DOCUMENT RESUME

ED 174 494 SE 028 638

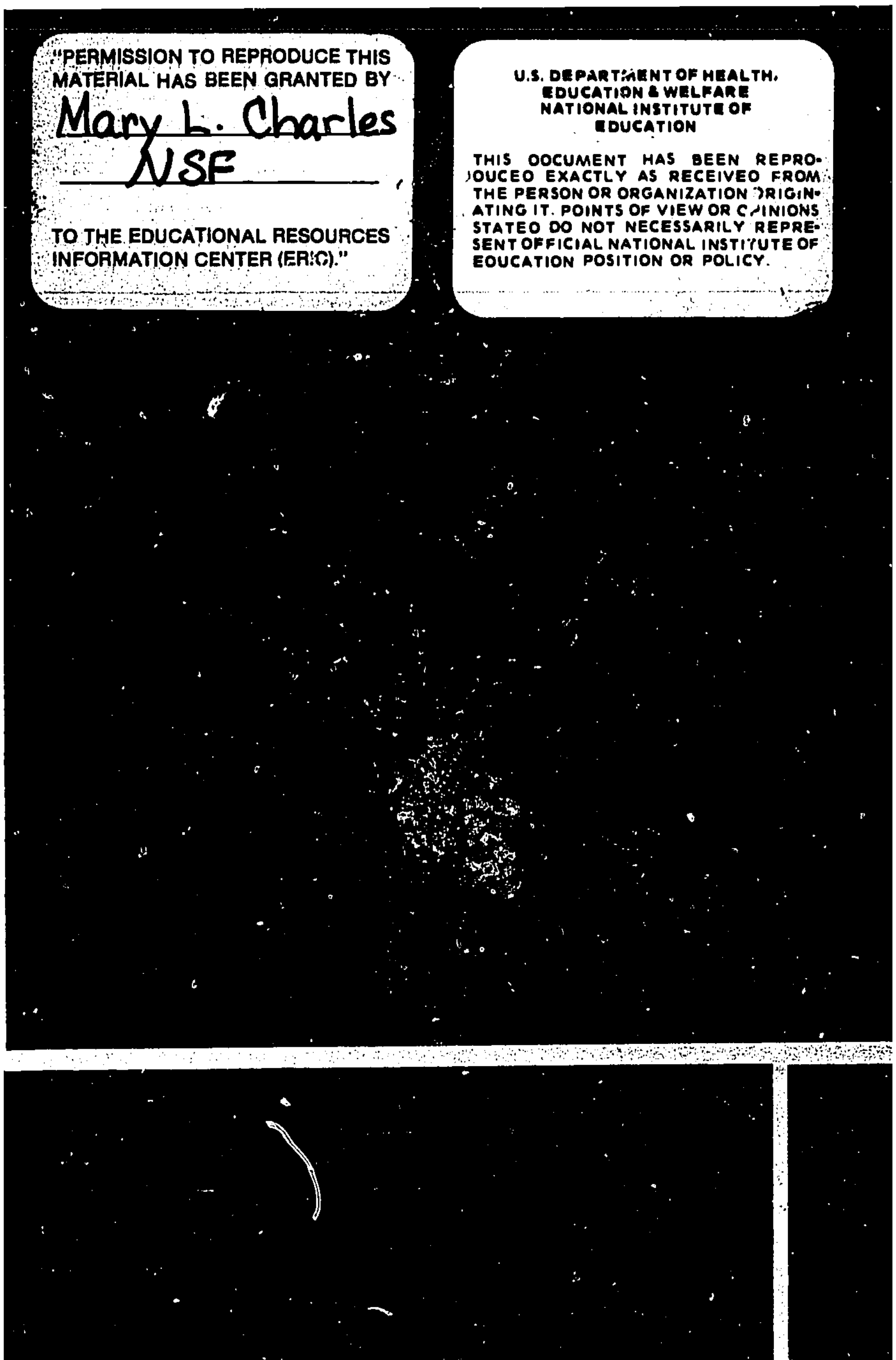
AUTHOR Stenberg, Warren; Walker, Robert J.
TITLE Calculus: A Computer Oriented Presentation, Part 1 [and] Part 2.
INSTITUTION Center for Research in Coll. Instruction of Science and Mathematics, Tallahassee, Fla.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE Jun 70
GRANT NSF-GY-3696
NOTE 1,165p.; Contains occasional light and broken type

EDRS PRICE MF09/PC47 Plus Postage.
DESCRIPTORS *Algorithms; *Calculus; *College Mathematics; *Computer Oriented Programs; Computers; Curriculum Development; Flow Charts; Higher Education; Mathematical Applications; *Mathematics Curriculum; Mathematics Education; Mathematics Instruction; *Mathematics Materials
IDENTIFIERS *Functions (Mathematics)

ABSTRACT

Parts one and two of a one-year computer-oriented calculus course (without analytic geometry) are presented. The ideas of calculus are introduced and motivated through computer (i.e., algorithmic) concepts. An introduction to computing via algorithms and a simple flow chart language allows the book to be self-contained, except that material on programming languages is excluded in order to allow the use of any language. Chapter topics include sequences, integrals, applications, functions, maxima, chain rule, derivatives, logarithmic and exponential functions, infinite series, and differential equations. (MP)

PART 1



INCIPAL AUTHORS

ERG AND ROBERT J. WALKER

CALCULUS

PART 1

by
WARREN STENBERG
University of Minnesota
ROBERT J. WALKER
Cornell University

with the assistance of

GEORGE H. ANDREWS
Oberlin College
LAURA COOPER
Florida State University
URSULA EBERHARD
Florida State University
RAY L. GRAHAM
Appalachian State University
PAUL OHME
Florida State University
EMILE ROTH
Florida State University
EDWARD F. WISHART
University of Nevada
ROBERT J. WISNER
New Mexico State University

A COMPUTER ORIENTED PRESENTATION

CRICISAM

An Experimental Textbook Produced by
THE CENTER FOR RESEARCH IN COLLEGE
INSTRUCTION OF SCIENCE AND MATHEMATICS

Financial support for this project has been provided by the National Science Foundation (Grant #GY-3696).

Center for Research in College Instruction
of Science and Mathematics
Room 212 Diffenbaugh
Florida State University
Tallahassee, Florida 32306

First Printing, 1968-1969 in five parts
Second Printing, May 1969 in five parts
Third, Revised Printing, June 1970

CRICISAM is an organization devoted to the improvement of undergraduate instruction in science and mathematics. Its governing board is:

The National Advisory Committee for this calculus project is:

The Local Coordinator at the grant site and host institution for this project is:

E. P. Miles, Director, Computing Center, Florida State University

The Director of CRICISAM is:

Guenter Schwarz, Florida State University

An expression of thanks is extended to the following persons who assisted in the preparation of this book:

G. J. Burks
Susan Carey
Kati B. Carus
Cynthia Corbin
Dorothea Driggs
Robert Ducharme
Tim Garland
J. B. Johnson
Susan McClellan
Judith Monette
Elizabeth Moore
Ilba L. Rosete
Linda Thompson
Patricia Vidzes

Special credit is given to Linda D. Thompson for her dedication in typing the manuscript and in coordinating activities concerned with the distribution of the text.

PREFACE TO THE THIRD REVISED PRINTING

Since the appearance of the second printing, this text has had a fairly extensive tryout in some 50 schools with 1,500 students involved. The teachers participating in this experimental use of the text have been most generous in supplying corrections and suggestions for improvement.

A great deal of feedback was gleaned from a three day conference, which was attended by nearly all the teachers using the book, and was held in Tallahassee, Florida March 23-25, 1970.

All the corrections supplied by the users have been made, but unfortunately time and resources have not allowed us to incorporate the suggestions for rewriting. Only a dozen or so pages of additional text appear in this printing.

However, most of the suggestions made by the teachers in addition to some after thoughts of the authors appear in the Teacher's Commentary. This commentary contains:

(a) suggestions for the handling of theory,

(b) related mathematics for interest of teachers or possible enrichment material,

(c) some additional flow charting to supplement Chapter 1,

(d) explanation of the rationale where treatment of topics is unusual,

(e) alternative or improved treatments of various topics which we were unable to incorporate in the text,

(f) supplementary exercises, problems and examples,

(g) a summary of the proceedings of the three day conference mentioned above.

It is expected that this commentary will be most valuable although no one teacher is expected to want to use or read all of it. It is suggested that the commentary be lightly scanned before teaching the course to obtain an idea as to what part of it might prove useful.

We finally remark that in response to overwhelming demand, the solutions of problems have been included in the student text instead of the separate answer booklet. The only exception to this are solutions which involve flow charts. These solutions appear instead in the Teacher's Commentary.

PREFACE TO THE FIRST PRINTING

This volume is the first part of a one year computer oriented calculus course (without analytic geometry). Considerable interest has been manifested in the impact of the computer in the calculus course, and several books have already appeared. This book goes much further than any of the others in the directions of introducing and motivating the ideas of calculus through computer (i.e., algorithmic) concepts.

Chapter 1 comprises an introduction to computing via algorithms and a simple flow chart language. The book is thus self-contained except that material on programming languages is excluded in order to allow the teacher to use FORTRAN, BASIC, ALGOL, PL/I, or any other programming language. In trying out an earlier version of these materials in the classroom, one of the authors found it expedient to teach FORTRAN by merely displaying a flow chart with the corresponding FORTRAN program alongside followed by two pages of explanation of peculiarities of the language (integer and real variables, etc.).

The authors have followed the algorithmic approach along the paths where it led us. This has resulted in a departure from the traditional ordering of some topics (e.g., sequences and integration treated before differentiation). It has also radically changed some of the proofs of theorems and in a few cases slightly modified the statement of theorems. For example, the form of the completeness axiom in Section 9 of Chapter 2 is quite unconventional.

It is natural to suppose that a computer approach to calculus would place more emphasis on heuristic than on rigor. While this may be the case with later efforts derived from this one, it is not so in this book. The authors regarded it part of their task to show that such a treatment could be made theoretically sound. Consequently, the course is somewhat more rigorous than may be appropriate for a beginning calculus course. Some of the more theoretical material has been placed at the ends of chapters, in starred sections, or in appendices to chapters. Some users may wish to de-emphasize this material, but it is hoped that they will give everything cautious trial to help determine whether the theory viewed in this new light becomes accessible to more students.

It is hoped that the dynamic or mechanistic character of the algorithmic approach will place concepts of calculus within the comprehension of a wider audience. Whether this hope will be justified only time and testing will tell. CRICISAM is anxious to obtain feedback and criticism in order to determine how the text might best be modified. We invite you to send your reaction to CRICISAM, Room 212 Diffenbaugh, Florida State University, Tallahassee, Florida.

The authors here express their appreciation for the efforts of Prof. E. P. Miles who supplied the impetus for bringing this project into existence and coordinating the computing facilities at Florida State University with the writing.

TABLE OF CONTENTS

VOLUME I

Chapter 0 — PRELIMINARIES

1	Introduction	1
2	Numbers	3
3	Inequalities	12
4	Absolute Value	17
5	Intervals	25
6	Functions	31
7	More on Functions	44
8	Composition and Inverses	57

Chapter 1 — BASIC COMPUTER CONCEPTS

1	Algorithms and Flow Charts	65
2	A Numerical Algorithm	73
3	A Model of a Computer	76
4	Input	89
5	Computer Memory	94
6	Arithmetic and Control Units of SAMOS	103
7	Machine Language	105
8	Odds and Ends	116
9	Iteration Boxes	120

Chapter 2 — SEQUENCES

1	Sequences of Approximations	125
2	Approximating Solutions of Equations	136
3	Problem Solution Using Sequences	142
4	Definition of Convergence	149
5	The Simplest Limit Theorems	165
6	Product and Quotient Theorems	172
7	The Squeeze Theorem	187

8	A Geometric Limit	194
9	Completeness	198

Chapter 3 AREA AND INTEGRAL

1	Area	207
2	An Algorithm for Area	219
3	Non-Monotone Functions	229
4	Integrals	234
5	Theory of Integration	241
6	Unicon Functions	252
7	Unicon Functions and Integrals	261
8	Formulas for Integrals	269
9	Tabulating the Sine and Cosine Functions	283
*10	The Unicon Modulus	290
A	Moduli for Combinations of Unicon Functions	298
B	Dispensing with Area	317

Chapter 4 APPLICATIONS OF INTEGRALS

1	Calculating Areas	325
2	Calculating Volume	337
3	Solids of Revolution	340
4	Another Characterization of the Integral	352

Chapter 5 LOCAL PROPERTIES OF FUNCTIONS

1	Speed and Limits	367
2	Limit Theorems and Continuity	384
3	Composition and Intermediate Value Theorems	398
4	Derivatives	407
5	The Derived Function	416
6	Derivative Theorems	424

Chapter 6 MAXIMA. THE MEAN VALUE THEOREM

1	An Example	439
2	The Maximum Theorem	441
3	Numerical Methods	455
4	The Mean Value Theorem	462
5	Approximations	470
6	Monotone Functions	479
*7	Existence of a Maximum	490

Chapter 7 THE CHAIN RULE

1	Differentiation of Composite Functions	495
2	Implicit Functions	509
3	Inverse Functions	519
4	Related Rates	533
5	Some Notations	539
6	Differentials	541
7	Parametric Equations	548
A	Theory of Inverse Functions	554

Chapter 8 THE CONNECTION BETWEEN DIFFERENTIATION AND INTEGRATION

1	Anti-derivatives	559
2	The Fundamental Theorem of Calculus	570
3	Second Form of Fundamental Theorem	578

CHAPTER 0

PRELIMINARIES

1. Introduction

We begin this book with an attempt to get our mathematical house in order so that the subsequent chapters will be more easily read. This involves some reminders of items from earlier courses, but then a bit of review never hurt anyone. It also involves some definitions and minor theorems. And finally, it involves some agreements on terminology and some conventions.

Students often do not read a mathematics book properly. When approaching a mathematics book, the reader should be well-armed with pencil and paper. The reading should be active, not passive: test every claim made by the author(s) with concrete examples of your own choice, try out suggested procedures, "doodle" with the ideas, and read with some doubt (for there may well be errors). When you've finished a paragraph or so, ask yourself if you understand, and if not, go through it again.

The problems are designed as challenges, to test your skill at manipulation, to check whether you understand the ideas, and in some cases to expand upon the material in the text. Make sure you can do them, obtaining what assistance is necessary from your instructor, your classmates, or from other books.

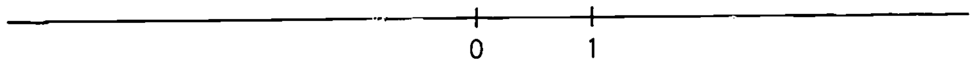
2. Numbers

In a certain vague sense, you are already pretty well aware of the numbers used in the study of calculus, and you are familiar with the arithmetic of those numbers. In your study of numbers, you started by learning how to count, and then progressed to more complicated numbers. We classify them as follows: integers, rational numbers, real numbers.

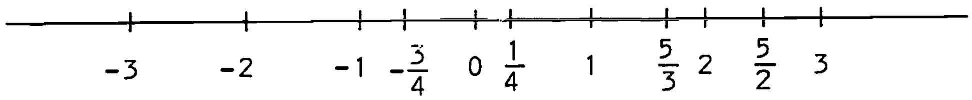
The integers (or whole numbers) are the numbers ..., -3, -2, -1, 0, 1, 2, 3, ... consisting of 0, the positive integers (1, 2, 3, ...), and the negative integers (-1, -2, -3, ...). An integer is even if it is divisible by 2, otherwise odd.

The rational numbers (or fractions) can be constructed by taking quotients of integers to obtain numbers like $\frac{3}{7}$, $-\frac{81}{4}$, $\frac{3}{9}$, $-\frac{12}{4}$, etc. In general, they are of the form $\frac{a}{b}$ where a and b are integers with the proviso, of course, that $b \neq 0$. Every integer n is a rational number since it may also be expressed as $\frac{n}{1}$. Notice that there is a smallest positive integer (namely, 1) but that there is no smallest positive rational number (Problem 2).

The integers and rational numbers have a commonplace geometric representation on a straight line. Once 0 and 1 have been selected, generally with 1 to the right of 0,



we merely take multiples of this unit length to obtain points corresponding to the integers and fractional parts of it for the rational numbers. A few have been plotted in this next scene



where the positive numbers are to the right of 0, the negative numbers to the left.

As we shall now see, though, these points do not "fill up" the line. That is to say, there are points on the line which do not correspond to integral or fractional multiples of the unit length. At 1, construct a right angle and form the triangle indicated in Figure 1-1, where the point P is one unit above the point corresponding to 1.

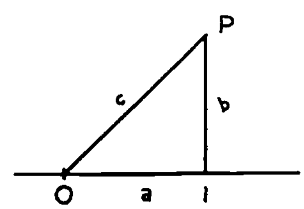


FIGURE 1-1

Thus, a=b=1, and by the Pythagorean Theorem (which states that $c^2 = a^2 + b^2$), we have $c^2 = 1^2 + 1^2 = 1 + 1 = 2$. Now if we rotate the line segment OP so that P falls on the base line, we can ask if it falls upon a rational number. If it does, then it will be at a distance c from 0, so that for

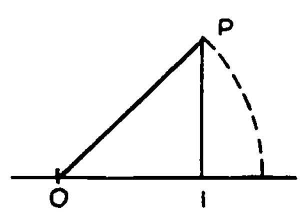


FIGURE 1-2

some rational number $\frac{x}{y}$, we have

$$c = \frac{x}{y}$$

We may as well (and will) assume that not both x and y are even, for otherwise we could cancel 2 from both the numerator and denominator. Now we have

$$c^2 = 2 = \frac{x^2}{y^2}$$

so that

$$2y^2 = x^2$$

Thus, x^2 is even, and by Problem 1(f), we conclude that x is even. Write $x = 2p$. Then

$$2y^2 = x^2 = (2p)^2 = 4p^2$$

so that

$$y^2 = 2p^2$$

Now y^2 is even, and Problem 1(f) demands that y be even. But look, we've concluded that both x and y are even, which contradicts our assumption. Thus, c cannot be written as a quotient of integers and is therefore not rational. The number c is of course $\sqrt{2}$, and we call it, quite naturally, an irrational number.

The real numbers include the rationals and fill in the number line by including as well all the irrationals. Every real number can be represented by decimals; for example,

$$\frac{4}{5} = 0.8$$

$$\frac{13}{7} = 1.857142857142857142...$$

$$-\frac{25}{6} = -4.16666...$$

$$\sqrt{2} = 1.41424...$$

$$\pi = 3.1415926...$$

The rationals, in this representation, either terminate (as with $\frac{4}{5}$) or repeat indefinitely (as with $\frac{13}{7}$ and $-\frac{25}{6}$). The nonrepeating infinite decimals characterize the irrationals. When we speak of numbers, we will simply mean real numbers; otherwise, we'll specify further by adjectives such as rational, whole, positive, nonzero, etc.

Notice that the real numbers and their basic operations satisfy the following properties (and you should notice the similarities):

Addition	Multiplication
1. If a and b are numbers, then so is a+b.	6. If a and b are numbers, then so is ab.
2. $a + (b+c) = (a+b) + c$	7. $a(bc) = (ab)c$
3. $a + b = b + a$	8. $ab = ba$
4. The number 0 has the property that $a + 0 = a$ for every number a.	9. The number 1 has the property that $a \cdot 1 = a$ for every number a.
5. Every number a has an "additive inverse," $-a$, so that $a + (-a) = 0$.	10. Every number a, except 0, has a "multiplicative inverse," a^{-1}, so that $a \cdot a^{-1} = 1$.

11. (This connects the operations of addition and multiplication) $a(b+c) = ab + ac$.

Any system of numbers with at least two elements satisfying these first eleven properties is called a field.

Order There are real numbers which exclude zero, called positive numbers satisfying

12. If a and b are positive numbers, then so are $a + b$ and $a \cdot b$.

13. If a is a number, then exactly one of the following statements is true:

(i) a is positive
(ii) $a = 0$
(iii) $-a$ is positive.

And finally, there is a property named after Archimedes called the Archimedean Axiom.

14. (Archimedean Axiom) If a is a real number, then there is a positive integer n which is greater than a.

(We prefer to present the Archimedean Axiom here although it properly should follow the definition of "greater than" in the next section.)

We will use these properties generally without comment or explicit reference, except that where their use is more unusual or crucial or subtle than in routine arithmetic, an appropriate comment will be included.

These properties do not characterize the real numbers. Notice, for example, that the system of rational numbers does satisfy all fourteen of them. However, an additional property, to be presented in Chapter 2, will serve with these to characterize the real numbers in the sense that the only set of numbers satisfying all fifteen properties is the system of real numbers -- whether presented as points on a line, infinite decimals, or whatever.

When referring to these properties, let's agree that "R7" will mean "property 7 of the system of real numbers." Just as an example of the sort of things you can prove about real numbers, we deduce two properties and show how they arise. They are theorems, but we present them as trivialities. The idea is that they don't really tell you anything about real numbers that you didn't already know. However, it's a good thing (or at least a comforting thing) to realize that there has been developed a system of axioms from which one can actually prove such things. Faith in numbers is all right, to a certain extent, but lest we succumb to the disease of numerology, it's wise once in a while to check that the properties we "know" about numbers can either be proved or must be accepted as axioms.

<u>Trivial Theorem 1</u>. $x \cdot 0 = 0$ for every real number x.

<u>Proof</u>: $x \cdot 0 = x \cdot (0 + 0)$ by R4 (with $a = 0$)

$= x \cdot 0 + x \cdot 0$ by R11

so that

$0 = x \cdot 0 + [-(x \cdot 0)]$ by R5

$= (x \cdot 0 + x \cdot 0) + [-(x \cdot 0)]$ by substitution

$= x \cdot 0 + [x \cdot 0 + (-(x \cdot 0))]$ by R2

$= x \cdot 0 + 0$ by R5

$= x \cdot 0$ by R4

<u>Trivial Theorem 2.</u> $(-x)(-y) = xy$ for any real numbers x and y.

<u>Proof:</u> (See that you can justify each step)

$$
\begin{aligned}
xy &= xy + 0 \\
&= xy + (-x)\cdot 0 \\
&= xy + (-x)[y + (-y)] \\
&= xy + [(-x)y + (-x)(-y)] \\
&= [xy + (-x)y] + (-x)(-y) \\
&= [x + (-x)]y + (-x)(-y) \\
&= 0\cdot y + (-x)(-y) \\
&= 0 + (-x)(-y) \\
&= (-x)(-y).
\end{aligned}
$$

The other usual rules for elementary computational arithmetic can be deduced in a similar manner.

PROBLEMS

1. Prove:

   (a) the sum of any two even integers is even.

   (b) the product of any two even integers is even.

   (c) the sum of any two odd integers is even.

   (d) the product of any two odd integers is odd.

   (e) 0 is even.

   (f) if a is an integer and a^2 is even, then a is even.

2. Prove that there is no smallest positive rational number.

3. Prove that $\sqrt{3}$ is irrational.

4. Check that the rationals satisfy the fourteen properties listed in this section.

5. Which of the fourteen basic properties are not satisfied by the set of integers; the set of even integers; the set of odd integers; the number 0 all by itself; the set of irrational numbers?

6. Explain why the square, x^2, of a real number x cannot be negative.

7. Prove the "invert and multiply" rule for dividing fractions.

8. What troubles arise if you attempt to interpret $\frac{1}{0}$ or

$\frac{0}{0}$ (or in general $\frac{a}{0}$ for any number a) as real numbers?

9. Johnny "proved" that 2 = 1 by this argument: Let a = b; then

$$a^2 = ab$$

so

$$a^2 - b^2 = ab - b^2$$

whence

$$(a+b)(a-b) = b(a-b)$$

and a cancellation of a - b from each side yields

$$a + b = b.$$

Now since a = b, we write 2b = b, so 2 = 1. Find the hole in this argument.

10. Prove the trivial theorem that $1 \neq 0$ in the field of real numbers.

3. Inequalities

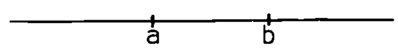
A large amount -- probably too large an amount -- of your early mathematical training has concerned equations, solving them, manipulating them, reducing them, etc. But for the necessary calculations in studying calculus, one must be able and willing to work with expressions that are not equal, or in some cases, not necessarily equal. This will enable us to make mathematically precise such loosely-worded expressions as "all numbers between 1 and 2," or "a is nearer to b than to c," or "the square roots of the positive integers become arbitrarily large." Furthermore, the language of inequalities is indispensable in discussions of numerical error.

We have stated in R12, R13, and R14 enough to proceed without delay. That is to say, just those properties about the notion of positive numbers will be sufficient for a systematic study of what can be done with numbers that are not equal. For example, one can prove (as Trivial Theorem 3) that 1 is a positive real number. We express this by writing $1 > 0$.

In general, $x > 0$ means that x is positive; and if $x - y > 0$, we write $x > y$ to express the fact that x is greater than y. Thus, $\pi > 3$, $0 > -1$, and $17 > -4$ for the respective reasons that $\pi - 3$, 1, and 21 are positive.

The ">" symbol turned around means "less than": $x < y$ means that x is less than y and is equivalent to $y > x$. So any statement concerning $>$ has a counterpart involving $<$.

Notice that, geometrically, $a < b$ means that a lies to the left of b on the number line, and this



is very often a useful picture which arises often in the chapters to follow.

Now there are three operational rules which could serve as Trivial Theorems 4, 5, and 6, and whose proofs are left to the Problem section. They are

If $x > y$ and $y > z$, then $x > z$

If $z > 0$ and $x > y$, then $xz > yz$

If $z < 0$ and $x > y$, then $xz < yz$

As with other statements, make sure you really understand the sorts of uses one makes of these rules. The first one allows us to write a string such as

$$a > b > c > d > e$$

without any misunderstanding and to pick as we wish from such a string any of several inequalities. For example, one can deduce that $b > e$. (Avoid writing such things as $a > b < c$, because it gives no information concerning any relationship between a and c.) The second rule is nice, allowing one to multiply an inequality by a positive number without changing its sense. The third rule forces one to change the sense of the inequality when multiplying both sides by a negative number. Thus, $2 < 3$, but $(-4)2 > (-4)3$.

The effects of subtracting and dividing are the substance of Problems 6, 7, and 8.

The symbol $\geq$ will be used to mean "greater than or equal to," and likewise $\leq$ will mean "less than or equal to." Thus, $8 < 17$ and $8 \leq 17$ are both true, as is $8 \leq 8$.

In many instances, you will have occasion to construct a string of relationships looking something like

$$\begin{aligned} A &= B \\ &\leq C \\ &< D \\ &\leq E = F \end{aligned}$$

where the capital letters will in some cases replace rather complicated-looking expressions and in other cases might be numbers. What we will generally be interested in is a relationship between what we started with, A, and what we ended with, F. Notice in this case that we can conclude $A \leq F$, but we can also conclude $A < F$, which is a "stronger" statement, thus generally more useful.

Probably for psychological reasons, you will see $<$ and $\leq$ more often than $>$ and $\geq$. This is undoubtedly because we've been taught to operate from left to right, both in reading and in plotting numbers. Thus, in thinking about several numbers, one generally starts with the one occurring first (i.e., the leftmost) on the line of real numbers. So, in writing those relationships, the $<$ symbol is somehow more "natural."

Concerning notation, you will find phrases such as, "... whenever a, b $\geq$ 0" or some such thing. This means both a and b are nonnegative.

PROBLEMS

1. Prove that 1 is a positive real number. (Hint: use Trivial Theorem 2 along with R12 and R13.) From this it follows that

$$\ldots -3 < -2 < -1 < 0 < 1 < 2 < 3 \ldots$$

2. Prove that if m, n, p, q are positive then

$$\frac{m}{n} > \frac{p}{q}$$

if and only if $mq > np$.

3. Prove Trivial Theorems 4, 5, and 6.

4. Prove that if $0 < x < 1$, then $x^2 < x$. On the other hand, what can you conclude if you're given that $x^2 < x$?

5. Prove that if $x > y$, then $x-z > y-z$.

6. Prove that if $x > y > 0$, then $\frac{1}{x} < \frac{1}{y}$.

7. Prove that if $x > y$ and $z > 0$, then $\frac{x}{z} > \frac{y}{z}$.

8. Where can $\geq$ (or $\leq$) replace $>$ (or $<$) in the inequality rules?

4. Absolute Value

We will often be concerned with the distance between two points on the line, and since points correspond to numbers, this concern is really with the "distance between x and y" for any pair of real numbers x and y. One could write $d(x,y)$ to denote the distance between x and y, but the notation $|x-y|$ is standard and fits in a framework of normal arithmetic; besides, you will soon recognize that it makes common numerical sense.

Notice first something trivial: the distance between x and y is the same as the distance between y and x. Thus,

$$|x - y| = |y - x|$$

We call $|x - y|$ the absolute difference of x and y. Notice further that the distance between x and 0 is $|x - 0|$, and the above equation yields

$$|x - 0| = |0 - x|$$

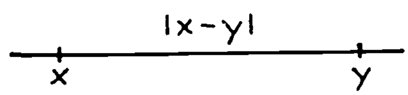
but $x - 0 = x$ and $0 - x = -x$, so

$$|x| = |-x|$$

Clearly, $|0| = 0$; and if $x \neq 0$, $|x| > 0$ because the distance between x and 0 is positive. Thus, $|x| \geq 0$ for all numbers x. We call $|x|$ the <u>absolute value</u> of x. Some examples are:

$$|3-7| = 4$$
$$|27| = 27$$
$$|7-3| = 4$$
$$|-27| = 27$$

and a picture is



Notice that we <u>could</u> have defined absolute value by

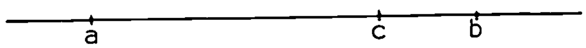
$$|x| = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -x & \text{if } x < 0 \end{cases}$$

and thus remain free of any appeal to geometry. Indeed, you may have learned this version of the definition.

What is the distance between 3 and -2? By our analysis, this is $|3-(-2)| = |3+2| = |5| = 5$. In general, $|x+y|$

represents the distance between x and -y. And it is natural to ask if there exist any relationships between $|x+y|$, $|x-y|$, $|x|$, and $|y|$ for arbitrary numbers x and y. There do, and the language of inequalities tells which ones.

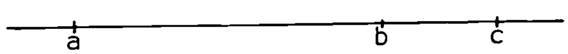
Consider numbers, a, b, and c as represented on the line. There are two cases: either c lies between a and b or it does not. If c lies between a and b, then



we have the equality

$$|a-b| = |a-c| + |b-c|$$

because the distance between a and b is the sum of the distance between a and c and the distance between b and c. If c does not lie between a and b,



then,

$$|a-b| < |a-c| + |b-c|$$

We can state in general, then, that

$$|a-b| \leq |a-c| + |b-c|$$

for any triple of numbers.

A series of replacements will yield some equivalent facts. First replace a-c by x, and replace b-c by y. Then since a-b = (a-c) - (b-c), we must replace a - b by x - y. The above statement then becomes

$$|x-y| \leq |x| + |y| .$$

Next, replace y by -z. Then we get

$$|x-(-z)| \leq |x| + |-z|$$

which is more simply written

$$|x+z| \leq |x| + |z|$$

This last statement is classically called the <u>triangle inequality</u>

Now in this last statement, replace x + z by w, and then z must get replaced by w - x and the triangle inequality becomes

$$|w| \leq |x| + |w-x|$$

which can be rewritten

$$|w-x| \geq |w| - |x| .$$

Also in the triangle inequality, we can replace x by u + v and z by -v so that x + z gets replaced by u, and we obtain

$$|u| \leq |u+v| + |-v| = |u+v| + |v|$$

so

$$|u+v| \geq |u| - |v| .$$

In all these recent calculations, notice that a, b, c, x, y, z, w, u, v are just numbers. We don't need all those symbols to write down an organized list of the basic relationships we've proved; using the fewest symbols, they are:

$$|a+b| \leq |a| + |b|$$
$$|a+b| \geq |a| - |b|$$
$$|a-b| \leq |a| + |b|$$
$$|a-b| \geq |a| - |b|$$

for any numbers a and b.

Our concern has been centered on absolute values of sums, and differences. What about products and quotients? An example will be instructive enough. Suppose $a = 7$ and $b = -12$. Then

$$|ab| = |7(-12)| = |-84| = 84$$

and $|a| = |7| = 7$, $|b| = |-12| = 12$, whence

$$|a|\cdot|b| = 7\cdot 12 = 84.$$

So in this case $|ab| = |a|\cdot|b|$. You can try other choices and see in general that the absolute value of a product is the product of absolute values. See problems 2 and 4.

In particular, $|a^2| = |a|^2 = a^2$, so that

$$|a| = \sqrt{a^2}$$

This is important to remember, because in some calculations, you may end up with the square root of a square of a number, and you must be careful about what you then conclude.

One last inequality: suppose $|x| < K$. Then this means that if $x \geq 0$, then $|x| = x < K$; and if $x < 0$, then $|x| = -x < K$, so $x > -K$. Thus, $|x| < K$ can alternatively be written

$-K < x < K$. In like manner, $|x| \leq K$ can be expressed by $-K \leq x \leq K$.

PROBLEMS

1. In the relations

$$|a + b| \leq |a| + |b|$$

$$|a + b| \geq |a| - |b|$$

$$|a - b| \leq |a| + |b|$$

$$|a - b| \geq |a| - |b|$$

decide in each case when the equality sign holds.

2. Prove that if $b \neq 0$ then

$$\frac{|a|}{|b|} = \left|\frac{a}{b}\right|$$

3. Determine which real numbers x satisfy:

(a) $|x + 2| < 3$

(b) $|x - 2| < 3$

4. Prove in general that

$$|ab| = |a| \cdot |b|$$

5. Intervals

A combination of the ideas of inequality and absolute value is handy in describing what are called intervals on the line of real numbers. Problem 3(b) of the last section demanded the determination of numbers x satisfying $|x - 2| < 3$.

There is both a geometric and an arithmetic approach to solving this problem and it's terribly important for future work that you learn to "see" the geometry and to carry out the corresponding arithmetic.

Geometrically, if x satisfies $|x - 2| \leq 3$, then just recall that $|x - 2|$ is the distance between x and 2. Thus, $|x - 2| \leq 3$ means that the distance between x and 2 cannot be more than 3. Thus, one can think of starting at 2 and proceeding in either direction for up to 3 units. Starting

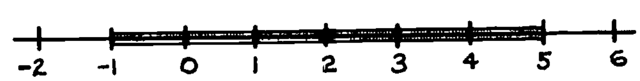


FIGURE 5-1

at 2, all the numbers up to and including 5 will satisfy the relation, along with all numbers down to and including -1. A picture is given in Figure 5-1.

Arithmetically, the statement $|x - 2| \leq 3$ can be analyzed by using the last paragraph of Section 4. The statement $|x - 2| \leq 3$ is equivalent to the chain of inequalities

$$-3 \leq x - 2 \leq 3.$$

Adding 2 all the way through yields

$$-1 \leq x \leq 5$$

which simply says that the number x must be between -1 and 5, inclusive.

In consequence, and by means of either the geometric or the arithmetic approach, we end up with what is called a <u>closed interval</u>. Its <u>center</u> is at 2 and its <u>radius</u> is 3.

If $|x - 2| < 3$, the same procedures yield

$$-1 < x < 5$$

and this set of points x is called an <u>open interval</u>. It also is said to have center 2 and radius 3.

In either of the cases

$$-1 \leq x < 5$$

or

$$-1 < x \leq 5$$

we call the associated interval on the line half open. It should be clear that neither of these intervals can be expressed simply by a single inequality involving absolute values.

There are of course infinite intervals

FIGURE 5-2

consisting of all points to one side of, and possibly including, a given point. The one in Figure 5-2 consists of all points to the right of 5 and includes 5. This can be expressed $x \geq 5$. The number 5 is excluded by $x > 5$.

The most common calculations you will be using with regard to intervals will be of the form

$$|A| < B.$$

You have seen from the above considerations that this is equivalent to writing

$$-B < A < B.$$

With intervals, there is a matter of notation. It is common to use $[-1,5]$ to mean the interval in Figure 5-1, i.e., all numbers x satisfying $-1 \le x \le 5$. In general

$[a,b]$ means all x for which $a \le x \le b$

$(a,b]$ means all x for which $a < x \le b$

$[a,b)$ means all x for which $a \le x < b$

(a,b) means all x for which $a < x < b$

$[a,\infty)$ means all x for which $a \le x$

(a,∞) means all x for which $a < x$

$(-\infty,a]$ means all x for which $x \le a$

$(-\infty,a)$ means all x for which $x < a$

$(-\infty,\infty)$ means all real numbers.

For each of the first four cases, the midpoint of the interval is its center, and half its length is the radius.

There is an aspect of intervals which involves averages. The term *average*, or *mean*, is ambiguous, for there are several ways in which those words are used. There is the *arithmetic mean* of two numbers, a and b. This is a number which when *added* to itself yields $a + b$, and that's of course $\frac{a+b}{2}$. There is also the *geometric mean* of two *positive* numbers a and b. This is a number which when multiplied by itself yields ab, and that's of course $\sqrt{ab}$. (By the way, $\sqrt{x}$ always means the positive number t for which $t^2 = x$.) A relationship between the arithmetic mean and the geometric mean of two positive numbers exists, and it can be viewed by means of intervals.

Let a and b be positive unequal numbers, and consider intervals of length a and b, placed side-by-side, as in Figure 5-3. At the midpoint M of the entire interval can be drawn a semicircle of radius $\frac{a + b}{2}$, the arithmetic mean of a and b. You can check to see that the right triangle depicted in Figure 5-5 has $\sqrt{ab}$ as one side and $\frac{a + b}{2}$ as the hypotenuse. Thus

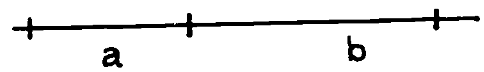


FIGURE 5-3

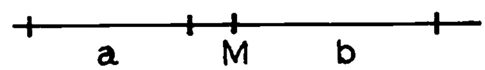


FIGURE 5-4

$$\sqrt{ab} < \frac{a + b}{2}$$

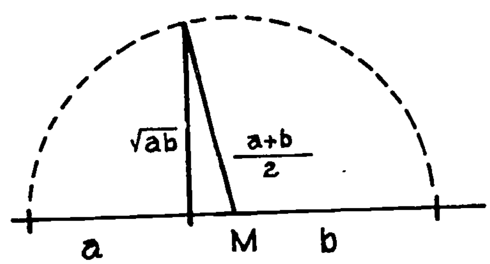


FIGURE 5-5

Problem 4 completes this issue by removing the demand that $a \neq b$. We can always conclude that

$$\sqrt{ab} \leq \frac{a + b}{2}$$

PROBLEMS

1. Determine and sketch all intervals (if any) corresponding to

   (a) $|x + 2| \leq 3$

   (b) $|x - 2| > 5$

   (c) $|x - 2| \leq 3$

   (d) $|x - 2| < 3$

   (e) $|x - 2| < -3$

   (f) $|x - 2| < |x - 3|$

   (g) $0 < |x-2| < 3$

   (h) $(x^5 + 3|x)^{100} (x-2) < 0$

2. Prove (by the Pythagorean Theorem) that Figure 5-5 is labeled correctly.

3. Prove the arithmetic mean - geometric mean inequality by considering

$$(\sqrt{a} - \sqrt{b})^2.$$

4. Prove, for a and b positive, that $\sqrt{ab} = \frac{a + b}{2}$ if and only if $a = b$.

6. Functions

In your previous school courses, you have dealt with functions of various sorts, particularly in studying logarithms, trigonometry, and so forth. Some reminders here should be of help in recalling the essence of the function concept.

The essence of this concept can be put in the nutshell of an example. Consider the sentence, "The area of a square is a function of the length of one of its sides." Think about this for a minute.

In the first place, that sentence begins with "The area," to denote that precisely one thing is being determined (by the length of a side). That is to say, a function embodies the idea of an "unambiguous designation." A given square has precisely one number associated with it which is designated as its area, and we're saying that this number is unambiguously designated in some manner by knowing the length of a side.

In the second place, the sentence under scrutiny does not talk about just one particular square. It concerns any square whatsoever, and this is the second basic idea embodied in the function concept. It isn't that a function makes just a single unambiguous designation, but rather that a function makes a bunch of unambiguous designations.

Thus, we could write A(S) to mean the area of square S; and if S has its sides each of length a, the well known formula for this whole bunch of unambiguous designations could be written

$$A(S) = a^2$$

to spell out (more specifically) that the area of any square is determined by the length of one of its sides.

If only certain squares were under consideration, a table would suffice. For example

length of side	1	2	3	5	7	11	13	17
area of square	1	4	9	25	49	121	169	289

One could think of a function as a collection or "set" of ordered pairs, wherein the first member unambiguously designates the second member. Thus, the idea of $A(S) = a^2$ could be written (a,a^2) where the first member designates the length of a side and the second member designates the area of the corresponding square.

Roughly speaking, then, a function assigns to each of certain numbers a uniquely-determined corresponding number. We will use letters, such as f, to refer to the functions. The symbol f(x) will designate the number corresponding to the number x. The number f(x) is called the value of the function f at x; or, sometimes, f evaluated at x.

We've already studied one function a little bit, the <u>absolute value function</u>. It assigns to each number x the number $|x|$. In signaling that one is about to work with this function, it is customary to write "Suppose $f(x) = |x|$..." or some such phrase.

In general, it is usually sufficient to simply reveal what f(x) is, for that carries implicitly the necessary information one needs about the correspondence under study. That is to say, rather than writing, "Consider the function which assigns to each number its square," we will be content with the shorter "Let $f(x) = x^2$." And <u>sometimes</u> you will see "the function x^2."

With each function, there is associated a set of numbers called the <u>domain</u> of the function. These are the numbers x for which we wish to study the functional values f(x). We say that f is defined <u>on</u> or sometimes (for geometric reasons) <u>over</u> its domain. For example, one might wish to study the behavior of $f(x) = |x|$ for $-1 \leq x \leq 1$. In such a case, the domain will be specified explicitly. If the domain is not so specified, then we assume the domain to be the set of all numbers for which $f(x)$ makes sense. With this agreement, the function $f(x) = \sqrt{x}$ with no domain specified automatically has as its domain $[0,\infty)$.

Given a function and its domain, there is another set of numbers, called the <u>range</u> of f. It is simply all the numbers $f(x)$.

Notice that we have not stated that $f(x)$ need be expressed by a formula, though that was indeed the case in the examples given. For instance, one can define a function f by stipulating that $f(x) = x^2$ if x is an integer and $f(x) = 0$ if x is not an integer. In this case, the domain of f is the set of all real numbers and the range is the set 0, 1, 4, 9, 16, ... of squares of integers.

You will need to become familiar with various forms of the word <u>map</u> and their mathematical usage, for they simplify many otherwise cumbersome phrases. One says that f maps its domain onto its range: for each number x in the domain of f, the function "sends x into $f(x)$" as suggested by

$$x \xrightarrow{f} f(x).$$

(Presumably this comes from map making wherein the mapmakers send Podunk onto an appropriately-placed dot on a piece of paper.) This arrow notation will be used from time to time.

Thus, x^2 maps the real numbers onto the nonnegative numbers; $\sqrt{x}$ maps the nonnegative real numbers onto the nonnegative real numbers; $|x|$ maps $[-1,1]$ onto $[0,1]$; $\sqrt{x}$ maps 9 onto 3 (you see, we also use this terminology for any part of the domain); $|x|$ maps $(-1,0)$ onto $(0,1)$.

Sometimes the "onto" terminology gets abandoned in favor of the looser "into." If f maps (any part of) its domain onto certain numbers S, then it maps those numbers <u>into</u> any set of numbers containing S. Thus, x^2 maps $(-1,1)$ into $[0,1]$, or into $[0, \infty)$ for that matter, since it maps $(-1,1)$ onto $[0,1)$. Again, $\sqrt{x}$ maps the positive integers into the positive real numbers, and $|x|$ maps the real numbers into the real numbers. The <u>into</u> terminology, then, admits a certain amount of sloppiness, but sometimes that's all that's needed in expressing our idea. You will note that all the functions we will deal with map selected parts of the real numbers into the real numbers. That is why they are called real - valued functions. A variable is a (generally unspecified) member of the domain of a function, and so our functions are real - valued functions of a real variable.

Functions can be "looked at" graphically, and certain insight about functions can often be gained in a geometric setting. Recall from your earlier training the coordinate system in the plane.

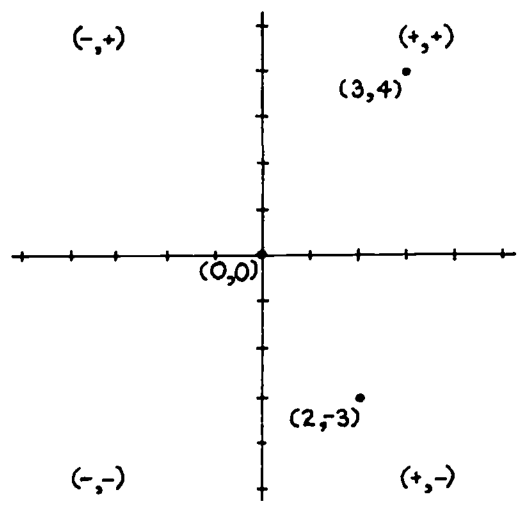


FIGURE 6-1

Every point has a name in terms of an "axis of abscissas" (often called the X-axis, a line of real numbers viewed horizontally) and an "axis of ordinates" (the Y-axis, a line of real numbers viewed vertically). The number 0 on each line is placed at a point called the origin, and positive numbers are to the right along the X-axis and extend upward along the Y-axis. Then the point of the plane which is reached by going three units to the right of the origin and then four units up is labeled, as in Figure 6-1, (3,4). The points (2, -3) and (0,0), the origin, also appear in Figure 6-1. A "general" point is often labeled (x,y), a reminder that each point in the plane is named by giving its abscissa first.

(You have by now noticed that the single symbol (3,4) can mean a point in the plane or an open interval. This is all right, though: a confusion never arises, because the context in which the symbol appears makes the meaning clear.)

In the coordinate plane, the <u>graph</u> of f is the set of <u>all</u> points of the form (x, f(x)) for x in the domain of f. Figure 6-2 illustrates a "general" point of the graph of some function f. The number x appears of course on the X-axis and the value of f at x appears on the Y-axis, and the pair (x, f(x)) shows up as a point in the plane. This means that the domain of f is a subset of the X-axis, and will in our work most often be an interval. The range of f is a subset of the Y-axis.

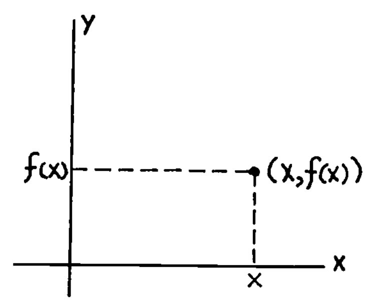


FIGURE 6-2

Since a general point in the plane has coordinates (x,y), and since a general point on the graph of f is given by $(x, f(x))$, the usage $y = f(x)$ has arisen and is often used in discussing functions, functional values, and graphs of functions.

For $f(x) = |x|$ over the interval $[-1,1]$, the graph is shown in Figure 6-3. Make sure that you both see it and "see" it. In this case, it was possible to display the entire graph. However, the situation is often complicated by dint of impossibility: you couldn't hope to graph $|x|$ for all real numbers x. In those cases, an incomplete picture fixing on the idiosyncracies of the function under study is pictured, and you must become adept at graphing functions to display those items of interest.

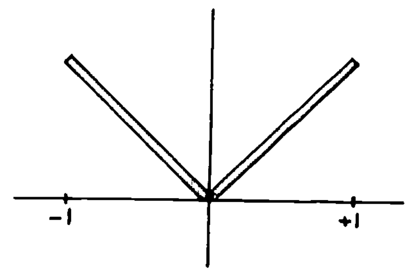


FIGURE 6-3

There are two "simplest" types of functions (simplest in the sense that calculating functional values constitutes absolutely no effort of any sort).

<u>Constant functions</u> are of the form $f(x) = k$ for a fixed number k. The domain is the set of all real numbers and the range is the single number k. For $f(x) = -2$ every point on its graph has the form $(x,-2)$ and is sketched in Figure 6-4.

The graphs of all the constant functions are of course parallel to the X-axis, and the X-axis itself is the graph of the constant function y = 0 . It is difficult to think of more to say in describing constant functions.

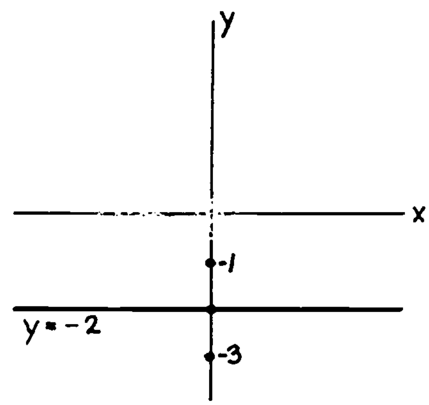


FIGURE 6-4

The identity function is of the form f(x) = x. This just maps each x onto itself, i.e., f "does nothing" and thus has the distinction of being the laziest of all functions. The domain is the set of all real numbers, and so is the range. Each point of the graph of the identity function is of the form (x,x), so y = x describes the geometric situation, shown in Figure 6-5.

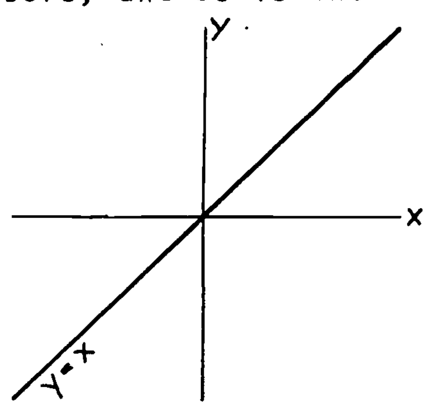


FIGURE 6-5

Functions, as well as numbers, admit of algebraic operations. If f and g are functions, then a new function called the sum of f and g and written f + g can be constructed. This is accomplished "pointwise": if x is a number, the value of f + g at x is obtained by adding the numbers f(x) and g(x). This means of course that x must be both in the domain of f and in the domain of g (hence in the intersection of those domains). Thus, the definition of the sum of two functions may be expressed in the equation

$$(f + g)(x) = f(x) + g(x).$$

Geometrically, the graph of the sum of two functions is obtained by summing ordinates at each point x on the X-axis. The geometric general rule is pictured in Figure 6-6.

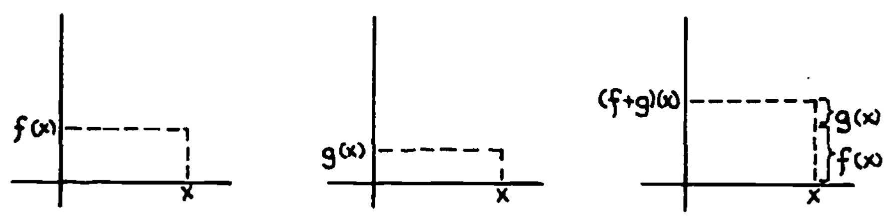


FIGURE 6-6

where the x is the same in each of the three parts. One can thus plot the graph of f + g by graphing both f and g on the coordinate plane and "eyeballing it" from there on out. Figure 6-7 depicts the graphs of the $f(x) = |x|$ and $g(x) = -2$ along with $(f + g)(x)$ which is also written $|x| - 2$. Figure 6-8 shows the sum of the functions $y = x$ and $y = |x|$, which of course coincide for $x \geq 0$,

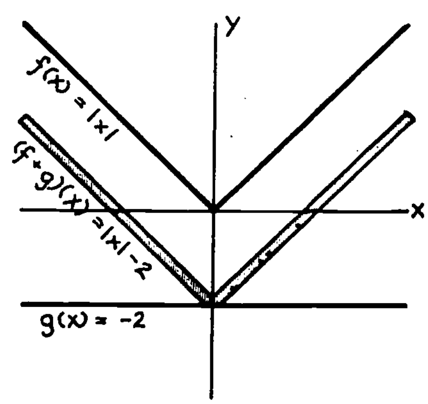


FIGURE 6-7

along with their sum $y = x + |x|$. Notice that this resulting function could also be given by

$$f(x) = \begin{cases} 0 \text{ for } x < 0 \\ 2x \text{ for } x \geqq 0 \end{cases}$$

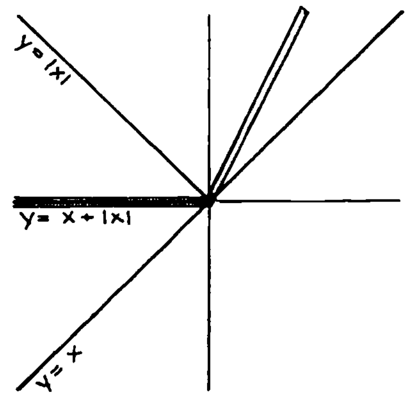


FIGURE 6-8

The product fg of two functions f and g is also defined pointwise. For each point x in the intersection of their domains, the value of fg at x is simply f(x)g(x). Thus

$$(fg)(x) = f(x)g(x).$$

With a little practice, an eyeballing procedure can be used to plot fg from the separate graphs of f and g. Figure 6-9 depicts (fg)(x) where $f(x) = |x|$ and $g(x) = -2$. Figure 6-10 does the same for $y = |x|$ and the identity function $y = x$,

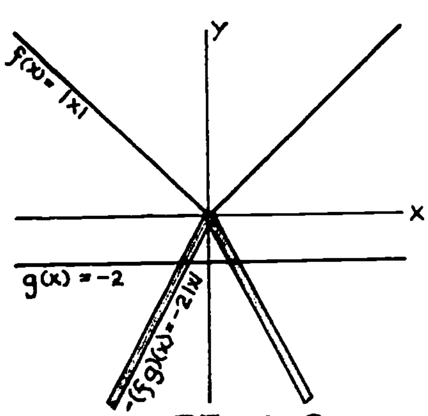


FIGURE 6-9

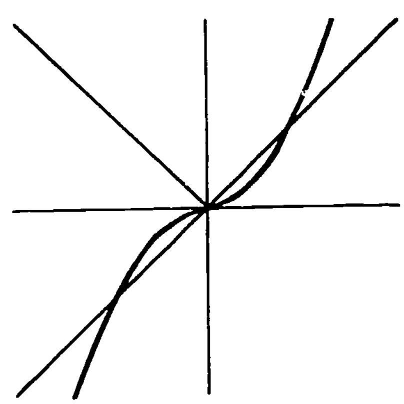


FIGURE 6-10

obtaining $y = x \cdot |x|$, this time without labeling anything. Make sure you "see" it.

Notice that this means for any function g, such expressions as 2g, πg, and in general kg for a fixed number k are well-defined. You simply take the product fg where $f(x) = k$, a constant function. Thus, -g makes sense, and so does f - g, the sum of f and -g. So differences of functions can be calculated also.

PROBLEMS

1. Sketch the graph of

   (a) $f(x) = x^2$ over $[-1,1]$

   (b) $y = x + 2$

   (c) $f(x) = x^2 + 2$

   (d) $f(x) = 2x$

   (e) $f(x) = 2x + x^2$

   (f) $y = x^3$

2. Write down an expression for $(f+g)(x)$ in each case and plot the graphs of f, g, and f + g.

   (a) $f(x) = x$ $\quad g(x) = x$

   (b) $f(x) = x$ $\quad g(x) = 2$

   (c) $f(x) = |x|$ $\quad g(x) = -2x$

   (d) $f(x) = x^3$ $\quad g(x) = |x|$

3. Repeat Problem 2 for $(fg)(x)$.

4. We have given meaning to 2f as the product of f and a constant function. How does this compare with f + f?

7. More on Functions.

A class of functions called polynomials is built up by starting with the two simplest types of functions, constant functions and the identity function, and constructing from them more complicated functions by the operations of sum and product of functions. For example

$$-2,\ x,\ 9$$

are polynomials; so are

$$-2 + x,\quad -2x,\quad -18 + 9x,\quad x^2,\quad 4x - 2x^3,$$

$$(x + 3)(2x - 7),\quad 3 + x(1 - x(\sqrt{2} - 2x))$$

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

It is evident that any expression of the form

$$(1) \qquad P(x) = c_0 + c_1x + c_2x^2 + \dots + c_nx^n ,$$

where $c_0, c_1, c_2, \dots, c_n$ are constants, is a polynomial. Conversely, it is easy to check that any polynomial can be written in the form (1); we merely have to multiply out, remove all parentheses, and collect like powers of x. Thus the last two of the above examples become

$$3 + x - \sqrt{2}x^2 + 2x^3,$$

$$0 + 0x - 10x^2 - 4x^3 + 7x^4 + x^5.$$

Because of this property it is customary to take equation (1) as the definition of a polynomial. Thus we define: A function P is a polynomial on a given domain if there are an integer n and constants $c_0, c_1, c_2, \ldots, c_n$ for which (1) is true for all values of x in that domain.

constants $c_0, c_1, c_2, \ldots, c_n$ are called the coefficients of the polynomial; in particular, c_k is the coefficient of the term $c_k x^k$, $k = 1, 2, \ldots, n$ and c_0 is the constant term. In writing a polynomial it is customary to omit terms with zero coefficients, so that the last example above would be written

$$-10x^2 - 4x^3 + 7x^4 + x^5.$$

The polynomial with all coefficients zero is written simply as 0.

Every polynomial except 0 has a degree, which is the greatest exponent of x appearing in the polynomial after terms with zero coefficients have been removed. Here we agree that $x = x^1$ and $c_0 = c_0 x^0$. The degrees of the 11 polynomials in the above example are respectively 0, 1, 0, 1, 1, 1, 2, 3, 2, 3, 5. Polynomials of the form ax + b, where a and b are numbers, are called linear (their graphs are straight lines) and are said to have slope a.

Functions in general (and polynomials in particular) may have zeros (also called roots). A zero of the function f is any number r for which f(r) = 0. The zeros of a function, then, appear on its graph as places where the graph touches or crosses the X-axis.

In addition to describing sums and products of functions, quotients may also be defined. If f and g are functions, then $\frac{f}{g}$ is defined, as you've probably already guessed, by

$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}$$

whenever this makes sense. Of course, then, the zeros of g are excluded from the domain of $\frac{f}{g}$. If P and Q are polynomials, then we call the function $\frac{P}{Q}$ a rational function. In working with quotients of functions, some fine distinctions must be made. For example, if $f(x) = x^2 - 4$ and $g(x) = x - 2$, then $\frac{f}{g}$ does not have 2 in its domain. Except for x = 2, however, this function is the same as h(x) = x + 2, which does have 2 in its domain.

The graph of $\frac{f}{g}$ may be sketched by viewing the separate graphs of f and g, and this again requires practice. The thing to remember is that large denominators produce small quotients and that small denominators produce large quotients. The function $\frac{1}{x}$ arises by choosing $f(x) = 1$ and $g(x) = x$, and its graph is displayed in Figure 7-1, this time without the separate graphs of f and g, which you should by now see through the eye of your mind. Other practice is provided in the problem section.

Now we present some new functions to play with. They are quite a bit different from the ones we've thus far faced.

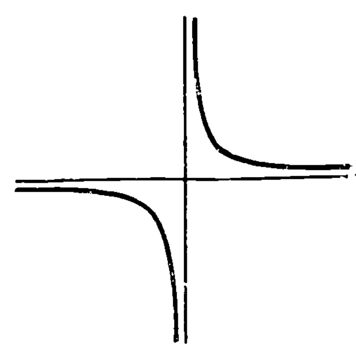


FIGURE 7-1

The greatest integer function arises as a result of realizing that every real number x can be expressed uniquely as

$$x = n + p,$$

where n is an integer and $0 \leq p < 1$. For example,

$$\frac{5}{2} = 2 + \frac{1}{2} \qquad -5 = -5 + 0$$

$$-\frac{5}{2} = -3 + \frac{1}{2} \qquad 0 = 0 + 0$$

$\sqrt{2} = 1 + 0.414...$ $\qquad \frac{1}{7} = 0 + \frac{1}{7}$

$\pi = 3 + 0.14159...$ $\qquad \frac{-1}{7} = -1 + \frac{6}{7}$

The greatest integer function maps each real number x onto n

$$x \longrightarrow n$$

and n is called the greatest integer in x; this integer is designated $[x]$. Thus, $[\frac{5}{2}] = 2$, $[-\frac{5}{2}] = -3$, $[\sqrt{2}] = 1$, and so on.

A partial graph of the greatest integer function is shown in Figure 7-2. Its "shape" is suggestive of steps, and this is just one example of what are known as step functions. Note that the domain of $f(x) = [x]$ is the set of all real numbers and the range is the set of all integers.

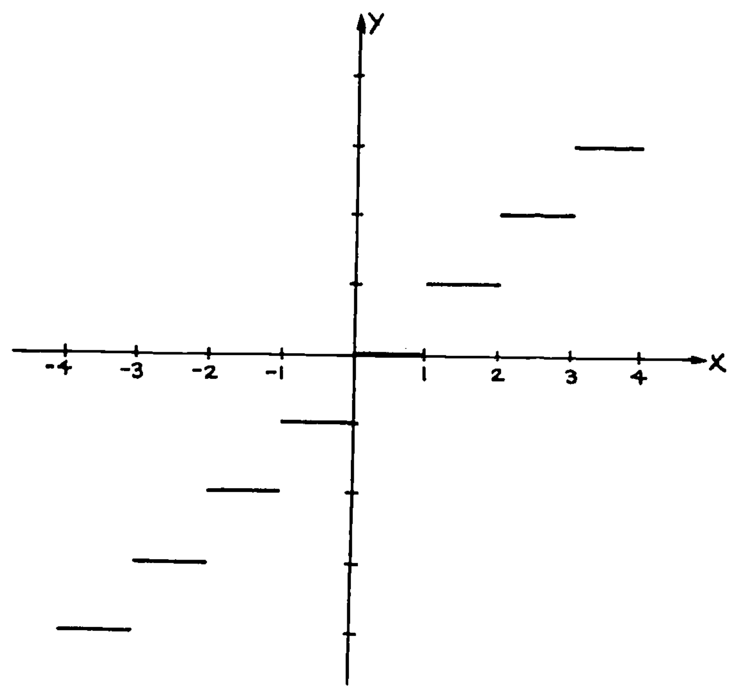


FIGURE 7-2

The fractional part function is closely associated with the greatest integer function: it maps each number x onto p in the aforementioned representation of x as n + p, where n is an integer and $0 \le p < 1$, so it's pictured

$$x \longrightarrow p$$

and p is called the fractional part of x. The notation for p is ((x)), so that $((\frac{5}{2})) = \frac{1}{2}$, $((-\frac{5}{2})) = \frac{1}{2}$, $((\sqrt{2})) = 0.414...$, and so forth.

For those who've worked with logarithm tables, [x] is called the characteristic of x, and ((x)) is called the mantissa of x.

Figure 7-3 shows a partial graph of the fractional part function f(x) = ((x)), and you will, from the definition of this function, note that the domain is the set of all real numbers and the range is [0,1).

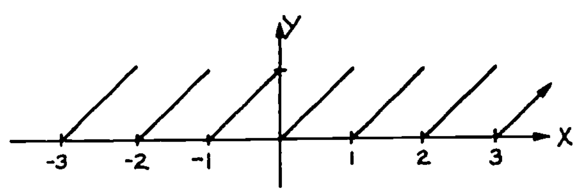


FIGURE 7-3

The circular functions require a reminder of the Pythagorean Theorem and its consequences in the coordinate plane. Suppose (a,b) and (c,d) are the coordinates of any two points in the plane. The shortest distance between those points may be obtained by considering, as in Figure 7-4, a right triangle with the segment between those points as the hypotenuse. Make sure you agree with all of the labels in that figure. Then the distance D between (a,b) and (c,d) is obtained from the Pythagorean equation.

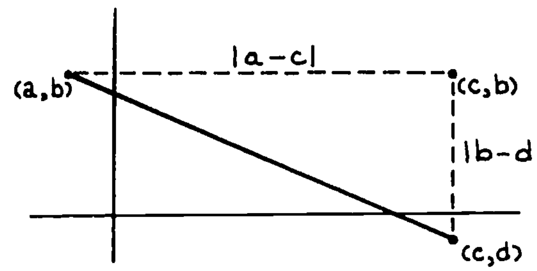


FIGURE 7-4

$$D^2 = |a-c|^2 + |b-d|^2$$
$$= (a-c)^2 + (b-d)^2$$

so that

$$D = \sqrt{(a-c)^2 + (b-d)^2}$$

Now let us consider a circle with center at (0,0) and radius 1. We call this <u>the unit circle</u>. This circle shown in Figure 7-5 consists of all points (x,y) at a distance 1 from (0,0), so that the distance formula yields

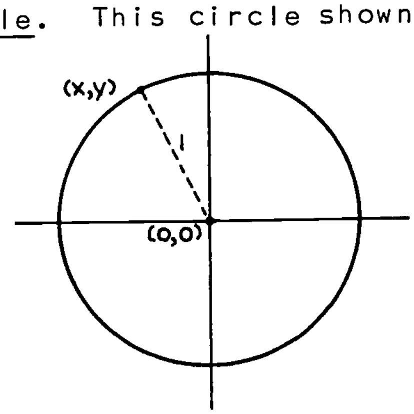


FIGURE 7-5

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

whereupon (squaring both sides) we obtain the following relationship between the first and second coordinates of all the points on the circle:

$$x^2 + y^2 = 1.$$

With these preliminaries out of the way, let u be any real number. Its distance from 0 is of course $|u|$. Imagine one end of a ruler being "planted" at the point (1,0) on the unit circle, and imagine further the ruler being "rolled along" the circle without sliding until the point corresponding to $|u|$ on the ruler is reached, with the agreement that if $u > 0$, then we proceed rolling counterclockwise, and if $u < 0$, we proceed clockwise along the circle. Figure 7-6 corresponds to a positive value of u. If $u = 0$, of course, we remain at (1,0).

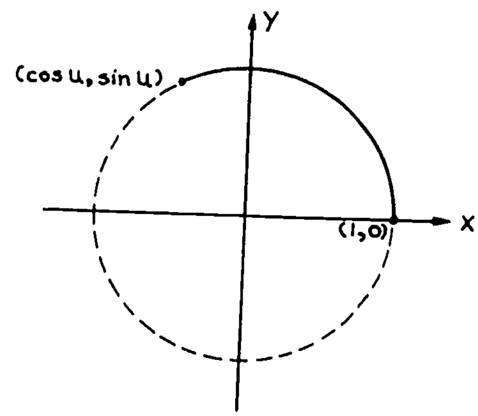


FIGURE 7-6

In any case, the point $|u|$ on the ruler ends up at some point on the circle, and we define cos u to be the abscissa of that point and sin u to be the ordinate.

The sine function is defined by $f(x) = \sin x$. Its domain is the set of all real numbers, and its range is the interval $[-1,1]$. You've memorized that the circumference of a circle is $2\pi r$, so the circumference of the unit circle is 2π. A

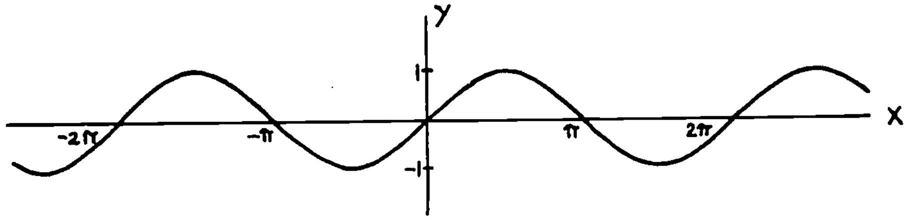


FIGURE 7-7

little thought about traversing the unit circle will be enough to convince you that a partial graph of the sine function is presented in Figure 7-7.

The <u>cosine function</u> is defined by $f(x) = \cos x$. Like the sine function, it maps the real numbers onto $[-1,1]$, and its graph is sketched in Figure 7-8.

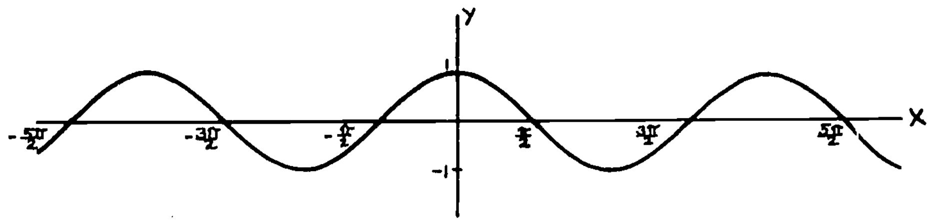


FIGURE 7-8

The <u>tangent</u>, <u>cotangent</u>, <u>secant</u>, and <u>cosecant</u> functions are defined by taking quotients of the two basic circular (or trigonometric) functions sine and cosine. Their abbreviation are suggestive enough to write the definitions as follows:

$$\tan x = \frac{\sin x}{\cos x} \qquad \cot x = \frac{\cos x}{\sin x}$$

$$\sec x = \frac{1}{\cos x} \qquad \csc x = \frac{1}{\sin x}$$

This presentation of the trigonometric functions as real-valued functions of a real variable departs from the more common presentation via angles, degrees, triangles, etc. However, all the usual trigonometric identities you learned earlier still hold, and you should use them freely.

We end this section with some language we will use to describe certain sorts of functional behavior.

A function which has the property that $a < b$ implies $f(a) \le f(b)$ is called *increasing*. Thus the identity function and the constant functions are increasing, and some more examples are

$$f(x) = x \qquad \text{on } [0,1]$$

$$f(x) = |x| \qquad \text{on } [0,\infty)$$

$$f(x) = [x]$$

$$f(x) = ((x)) \qquad \text{on } [0,1)$$

and it should be obvious that if a function is increasing on its domain, then it's increasing on any part of its domain. The sine and cosine functions are not increasing (except on selected subsets of their domains).

If $a < b$ implies that $f(a) < f(b)$ then the function f is called *strictly increasing*.

Similarly, if $a < b$ implies $f(a) \ge f(b)$, then f is *decreasing*; and the term *strictly decreasing* is reserved for those functions for which $a < b$ implies $f(a) > f(b)$.

Functions which are either increasing or decreasing are called *monotone*.

In general with regard to functions, situations will arise in which one might wonder if extreme cases are to be included, and the word "strictly" will be used to exclude extreme cases.

PROBLEMS

1. What is the degree of the polynomial

$$(x^2 + \sqrt{2}x + 1)(x^2 - \sqrt{2}x + 1) - x^4?$$

2. Prove that the degree of the product of two polynomials is the sum of their degrees, and that the degree of the sum is at most the larger of the two degrees.

3. Plot the graph of $\frac{f}{g}$ for each of the pairs of functions in Problem 2 of the previous section.

4. Show by example that there exist functions f, g, and h for which $gh = f$ but $\frac{f}{g} \neq h$. Is it, on the other hand, true that if $\frac{f}{g} = h$, then $gh = f$?

5. Prove that each real number x can be expressed in only one way as

$$x = n + p$$

where n is an integer and $0 \leq p < 1$. (Hint: show that if $x = m + q$ where m is an integer and $0 \leq q < 1$, then $n = m$ and $p = q$.)

6. What function is the sum of the greatest integer function and the fractional part function?

7. Sketch graphs of the other trigonometric functions.

8. Sketch the graphs of

(a) $f(x) = x + \sin x$

(b) $f(x) = [\sin x]$

(c) $f(x) = x \sin x$

(d) $f(x) = ((\sin x))$

8. Composition and Inverses

Thus far, we have carried out only algebraic operations with pairs of functions: $f+g$, $f-g$, $f \cdot g$, and $\frac{f}{g}$. One more operation needs to be discussed, and it is not algebraic. We present first an example.

Let $f(x) = x^2$ and $g(x) = x+1$. We evaluate f at 2 and then proceed to evaluate g at f(2): first, f(2) = 4 and g(4) = 5. In general,

$$g(f(x)) = f(x) + 1$$
$$= x^2 + 1.$$

We label the function thus obtained g(f), so that

$$(g(f))(x) = g(f(x))$$

and we call g(f) the composition of f and g. The f comes first in this phrase because the picture is

$$x \xrightarrow{f} f(x) \xrightarrow{g} g(f(x))$$

Notice that the composition of f and g is not necessarily the same as the composition of g and f. Indeed, the example we started with yields

$$(f(g))(x) = f(g(x))$$
$$= (g(x))^2$$
$$=(x+1)^2 = x^2 + 2x + 1.$$

The graph of f(g) can seldom be obtained from a hurried view of the graphs of the separate functions involved. This is because f(g) is evaluated at numbers which are in the range of g.

We give some examples:

(a) $f(x) = |x|$; $g(x) = -x$. Then
$(f(g))(x) = |-x| = |x| = f(x)$ and
$(g(f))(x) = -|x| = -f(x)$.

(b) $f(x) = \sin x$; $g(x) = x^2$. Then
$(f(g))(x) = \sin x^2$ and
$(g(f))(x) = (\sin x)^2$ which is usually written $\sin^2 x$.

(c) $f(x) = [x]$; $g(x) = ((x))$. Then
$(f(g))(x) = [((x))] = 0$ and
$(g(f))(x) = (([x])) = 0$.

(d) $f(x) = \sin x$; $g(x) = \frac{1}{x}$. Then

$(f(g))(x) = \sin \frac{1}{x}$ and

$(g(f))(x) = \frac{1}{\sin x} = \csc x$.

(e) $f(x) = x$; $g(x) = x^2 + \sin x + [x]$. Then

$(f(g))(x) = x^2 + \sin x + [x]$ and

$(g(f))(x) = x^2 + \sin x + [x]$.

Example (e) was selected in a complicated enough way that you should be convinced that any function composed with the identity function remains unchanged.

Another example is

$f(x) = 2x + 7$; $g(x) = \frac{1}{2}x - \frac{7}{2}$. Then

$g(f(x)) = x$, the identity function.

In general, if g is any function such that g(f) is the identity function, then g is called the <u>inverse</u> of f and we write f^{-1} to denote this inverse. Schematically, this means that for any x in the domain of f we have

$$x \xrightarrow{f} f(x) \xrightarrow{f^{-1}} x$$

Do not confuse $f^{-1}(x)$ with $\frac{1}{f(x)}$. In the last example, $f^{-1}(x) = \frac{1}{2}x - \frac{7}{2}$, but

$$\frac{1}{f(x)} = \frac{1}{2x + 7}$$

and these are clearly not the same functions.

Notice that not all functions have inverses. For example, $f(x) = 7$ has no inverse function. You should be able to see that the greatest integer function has no inverse.

It is very important that you give special attention to Problems 5, 6, and 7.

73

PROBLEMS

1. What are the zeros of the sine and cosine functions?

2. What are the zeros of $\sin \frac{1}{x}$?

3. Sketch the graph of $\sin \frac{1}{x}$.

4. Write an expression for the inverse of f where

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

   (b) $f(x) = x^2 + 1$

   (c) $f(x) = -x$

5. See if you can prove that if a function is strictly increasing, then it has an inverse. It's not very hard to do.

6. Does a strictly decreasing function have an inverse?

7. Let f be a strictly increasing function. Plot f and its inverse on the same coordinate axes, and compare them with the graph of the identity function. Do you notice anything?

Chapter I

BASIC COMPUTER CONCEPTS

by

A. I. Forsythe, T. A. Keenan, E. I. Organick, and W. Stenberg

ACKNOWLEDGEMENT: This chapter is borrowed, with permission of the publisher, John Wiley and Sons, Inc., New York from an introductory text by the above authors entitled

COMPUTER SCIENCE: A FIRST COURSE

This text in turn stems from an earlier effort entitled "Algorithms, Computation and Mathematics" published by the School Mathematics Study Group and from other SMSG publications

Chapter I

BASIC COMPUTER CONCEPTS

1. Algorithms and Flow Charts

A distinction should be made between the study of computers and the study of computing. The study of computers deals with the design of large complex networks of circuits and electronics that make up a computer. You will learn very little about that in this book. The subject of computing, on the other hand, deals with the organizing of problems so that computers can work them. As we shall see, this topic consists primarily of the study of algorithms--learning not only to understand but also to construct and improve them.

What is an algorithm? An algorithm is a list of instructions to carry out some process step by step. A recipe in a cook book is an excellent example of an algorithm. Here the preparation of a complicated dish is broken down into simple steps that every person experienced in cooking can understand. Another good example of an algorithm is the choreography for a classical ballet.

Here an intricate dance is broken down into a succession of basic steps and positions of ballet. The number of these basic steps and positions is very small, but by putting them together in different ways, an endless variety of dances can be devised.

In the same way, algorithms executed by a computer can combine millions of elementary steps such as additions and subtractions, into a complicated mathematical calculation. Also, by means of algorithms, a computer can control a manufacturing process or coordinate the reservations of an airline as they are received from ticket offices all over the country. Algorithms for such large scale processes are, of course, very complex, but they are built up of pieces as in the example we will now consider.

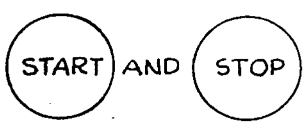
If we can devise an algorithm for a process, we will see that we can usually do so in many different ways. Here is one algor for the every-day process of changing a flat tire.

Algorithm for Changing a Flat Tire

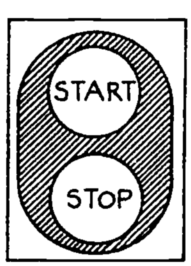
1. Jack up the car.
2. Unscrew the lugs.
3. Remove the wheel.
4. Put on the spare.
5. Screw on the lugs.
6. Jack the car down.

We could add a lot more detail to this algorithm. We could include the steps of getting the materials out of the trunk, positioning the jack, removing the hub-caps, loosening the lugs before jacking up the car, etc. For algorithms describing mechanical processes, it is generally necessary to decide how much detail to include. However, the steps we have listed will be adequate for getting across the idea of an algorithm. When we get to mathematical algorithms, we will have to be much more precise.

A <u>flow</u> <u>chart</u> is a diagram for representing an algorithm. In Figure 1-1, we see a flow chart for the flat tire algorithm. The



in the flow chart remind us of the buttons used to start and stop a piece of machinery.



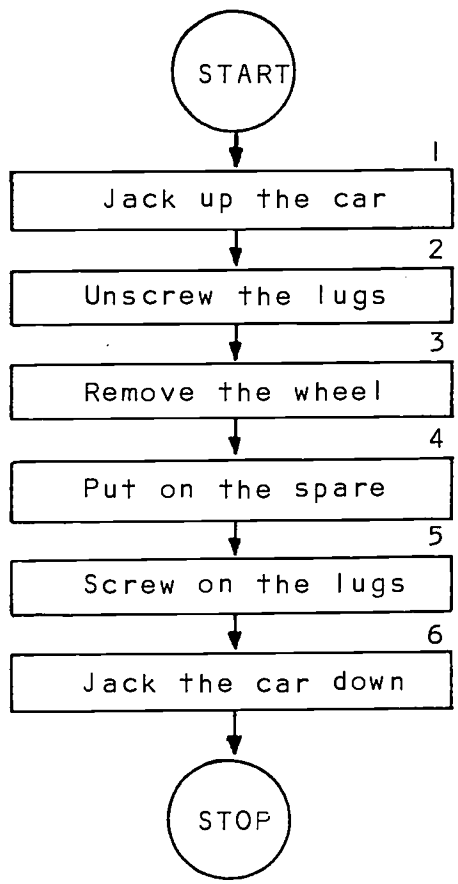
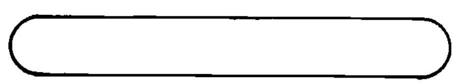


FIGURE 1-1

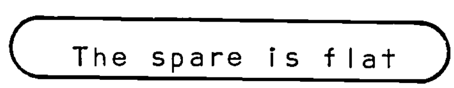
Each instruction in the flow chart is enclosed in a frame or "box". As we will soon see, the shape of the frame indicates the kind of instruction written inside. A ec-tangular frame indicates a command to take some action.

To carry out the task described by the flow chart, we begin at the start button and follow the arrows from box to box executing the instructions as we come to them.

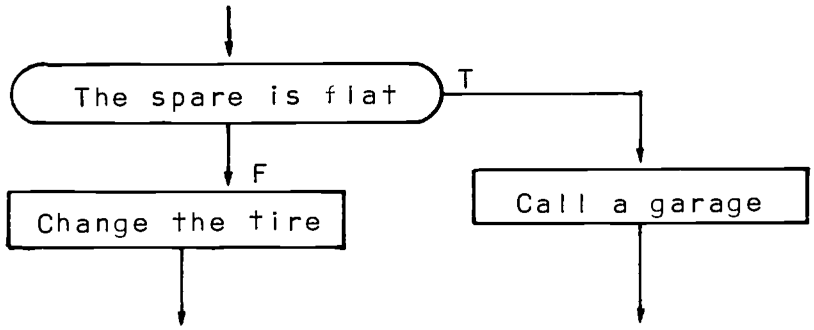
After drawing a flow chart, we always look to see whether we can improve it. For instance, in the flat tire flow chart, we neglected to check whether the spare was flat. If the spare <u>is</u> flat, we will not change the tire but will call a garage instead. This calls for a decision between two courses of action. For this purpose, we introduce a new shape of frame into our flow chart.



Inside the frame we will write an assertion instead of a command.

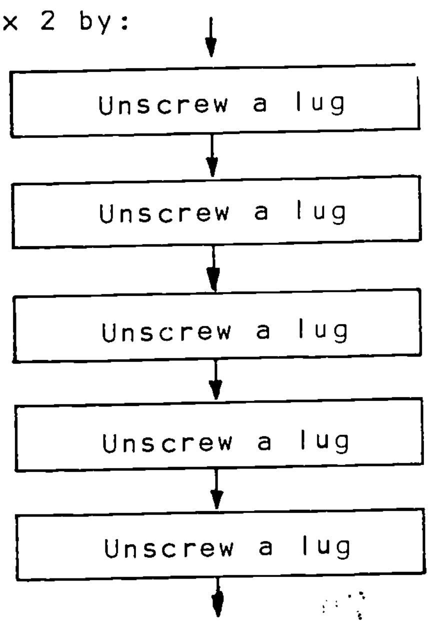


This is called a <u>decision box</u> and will have two exits, labelled T (for true) and F (for false). After checking the truth or falsity of the assertion, we choose the appropriate exit and proceed to the indicated activity.



Insetting this flow chart fragment into Figure 1-1, we obtain the flow chart in Figure 1-2.

There is still another instructive improvement possible. The instruction in box 2 of our flow chart in reality stands for a number of repetitions of the same task. To show the additional detail we could replace box 2 by:



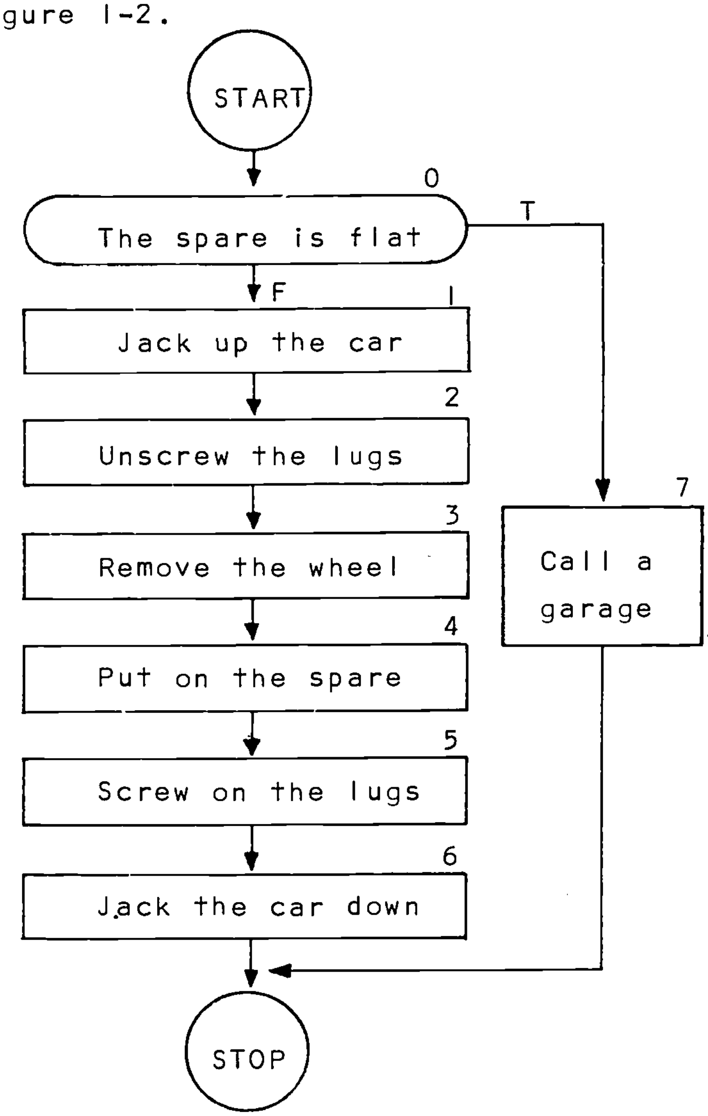
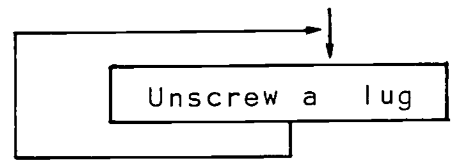
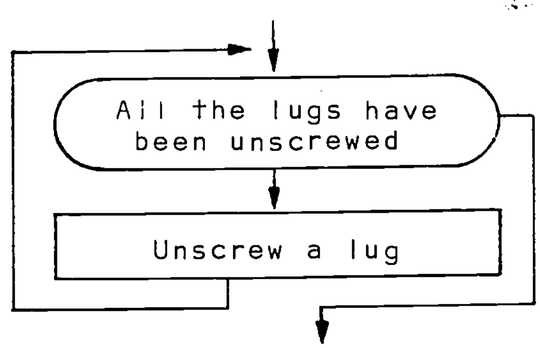


FIGURE 1-2

The awkwardness of this repeated instruction can be eliminated by introducing a loop.



As we leave the box, we find that the arrow leads us right back to repeat the task again. However, we are caught in an endless loop as we have provided no way to get out of the loop and go on with the next task. To rectify this situation, we again require a decision box, as follows:



Replacing box 2 of our flow chart with this mechanism and making a similar replacement for box 5, we get the final result shown in Figure 1-3.

Now that you have followed the development of the flat tire flow chart, try to devise one of your own. In the algorithm of the following exercise, you will probably discover some decisions and loops. There are many different ways of flow charting this algorithm, so probably many different looking flow charts will be submitted.

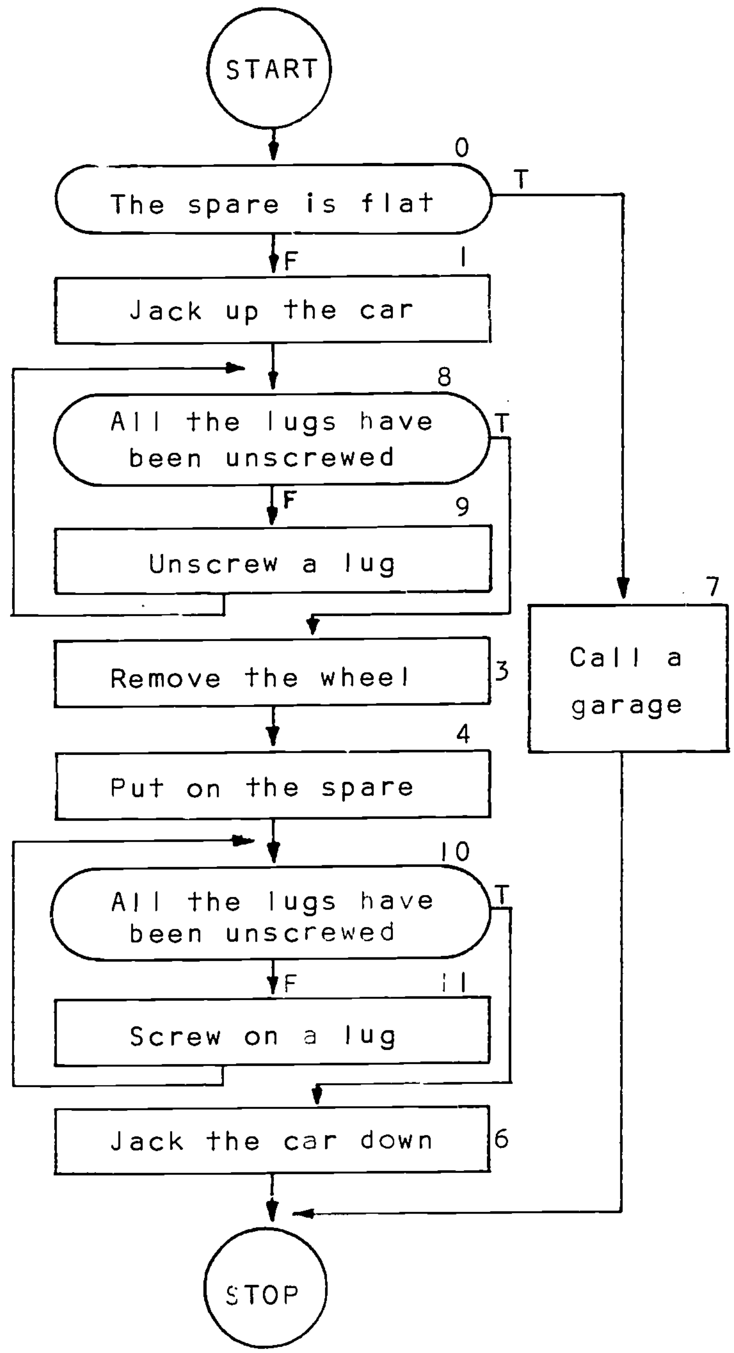


FIGURE 1-3

PROBLEMS

1. Prepare a flow chart representing the following recipe.

<u>Mrs. Good's Rocky Road</u>

Ingredients:

1 cup chopped walnuts

$\frac{1}{4}$ lb. block baker's chocolate

$\frac{1}{2}$ lb. marshmallows cut in halves

3 cups sugar

$\frac{1}{2}$ cup evaporated milk

$\frac{1}{2}$ cup corn syrup

1 tsp. vanilla

$\frac{1}{4}$ lb. butter

$\frac{1}{2}$ tsp. salt

Place milk, corn syrup, sugar, chocolate, salt in a four-quart pan and cook over high flame stirring constantly until mixture boils. Reduce to medium flame and continue boiling and stirring until a drop of syrup will form a soft ball in a glass of cold water. Remove from flame and allow to cool 10 minutes. Beat in butter and vanilla until thoroughly blended. Stir in walnuts. Distribute marshmallow halves over bottom of 10" square buttered baking pan. Pour syrup over marshmallows. Allow to cool 10 minutes. Cut in squares and serve.

2. A Numerical Algorithm

Now we are ready to look at an algorithm for a mathematical calculation. As a first example, we will take the problem of [illegible] terms of the Fibonacci sequence:

0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, ...

In this sequence, or list of numbers, the first two terms are given to be 0 and 1. After that, they are constructed according to the rule that each number in the list is the sum of the two preceding ones. Check that this is the case. Thus, the next term after the last one listed above is:

$$34 + 55 = 89.$$

Clearly, we can keep on generating the terms of the sequence, one after another, for as long as we like. But in order to write an algorithm for the process (so that a computer could execute it, for example), we have to be much more explicit in our instructions. Let's subject the process to a little closer scrutiny.

To the right is a table showing the computation of the Fibonacci sequence.

Next Latest Term	Latest Term	Sum
0	1	0 + 1 = 1
1	1	1 + 1 = 2
1	2	1 + 2 = 3
2	3	2 + 3 = 5
3	5	3 + 5 = 8
5	8	5 + 8 = 13
8	13	8 + 13 = 21

We can see that in each step the latest term becomes the new next latest term and the sum becomes the new latest term.

Let's construct a flow chart (Figure 2-1) for finding the first term to exceed 1000 in the Fibonacci sequence. After 64 steps which take us through the loop of flow chart boxes 2-5 fifteen times, we eventually emerge from box 3 at the T exit and proceed to box 6. This box is seen to have a different shape because it calls for a

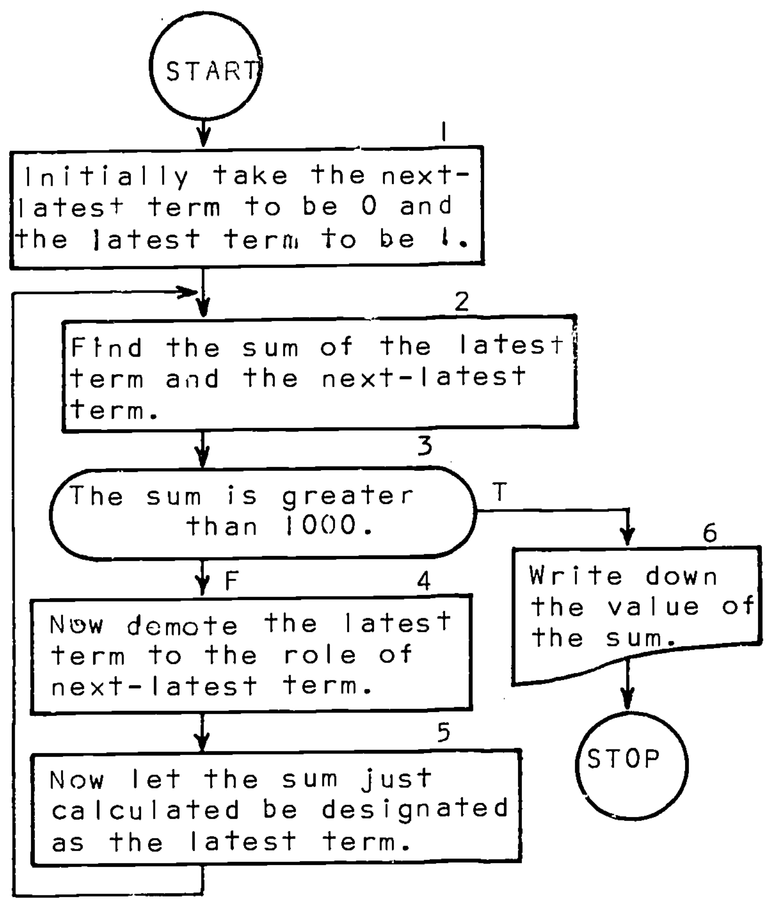
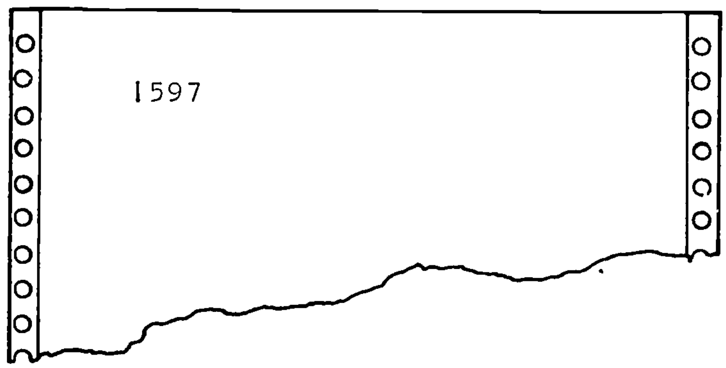


FIGURE 2-1

different kind of activity -- that of writing down our answer. The shape is chosen so as to suggest a page torn off a line printer, one of the most common computer output devices.



3. A Model of a Computer

The algorithm considered in the preceding section can be presented in much simpler notation which is at the same time more nearly ready to be given to a computer as a set of instructions. To do this, we need to introduce a conceptual model of how a computer works. This model is extraordinarily simple--childishly so, in fact. It is amazing but true that such a simple view of how a computer works is completely adequate for this entire course. We will present a more realistic picture of a computer in later sections of this chapter--but only to satisfy your curiosity, not because we have any real need of it.

Variables. In computing work, a variable is a letter or a string of letters used to stand for a number. In the formula

A = L × W

the letters A, L, and W are variables. In the formula

DIST = RATE × TIME,

DIST, RATE, and TIME are variables.

At any particular time, a variable must stand for one particular number called the value of the variable. Although at any time the value of a variable is one particular number, this value may change from time to time during a computing process. The value of a variable may change millions of times during the execution of a single algorithm.

In our conceptual model of a computer, we will associate with each variable a window box. On the top of each box the associated variable is engraved. Inside each box is a strip of paper with the present value (or current value) of the variable written on it.

Each box has a lid which may be opened when we wish to assign a new value to the variable. Each box has a window in the side so that we may read the value of a variable with no danger of altering the value. These window boxes constitute the memory of our computer. In Figure 3-1, we see the course of executing the Fibonacci sequence algorithm of the preceding section. Here NEXT

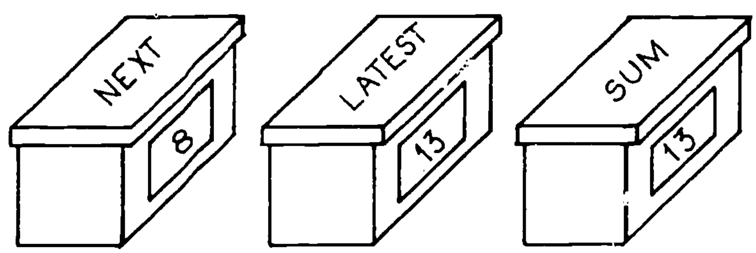
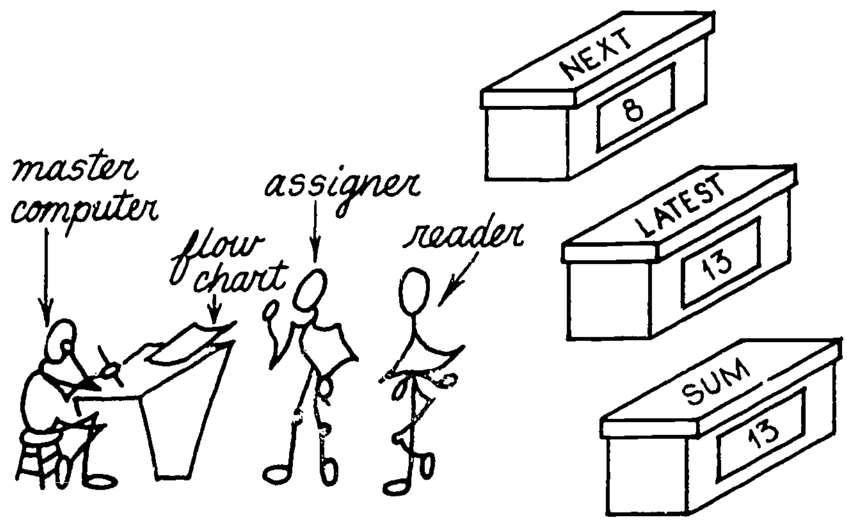


FIGURE 3-1

stands for "next to last term" and LATEST stands for "latest term".

<u>The Model and How It Works.</u> We visualize a computer as a room with a number of window boxes in it and a staff of three workers--the master computer and two assistants, the assigner and the reader. The master computer has a flow chart on his desk from which he gets his instructions, according to which he delegates certain tasks to his assistants. (In a real computer the tasks of these workers are performed by electronic circuits.)

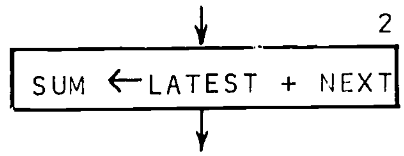


To see how this team operates, let us suppose that the computer is in the midst of executing the Fibonacci sequence algorithm of Figure 2-1. One of the instructions in this algorithm was:

2

Find the sum of the latest term and the next-latest term

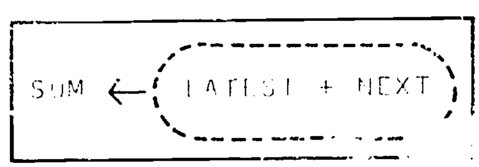
In a simplified flow chart notation, this instruction will take the form:



Inside this flow chart box, we find an <u>assignment statement</u>. Reading this statement aloud we would say, "Assign to SUM the value of LATEST + NEXT", or more simply, "Assign LATEST + NEXT to SUM". The left-pointing arrow is called the <u>assignment</u> operator and is to be thought of as an order or a command. Rectangular boxes in our flow chart language will always contain assignment statements and will therefore be called <u>assignment</u> <u>boxes</u>.

Now let's see what takes place when the master computer comes to this statement in the flow chart. We shall assume that the variables LATEST and NEXT (but not SUM) have the values seen in Figure 3-1.

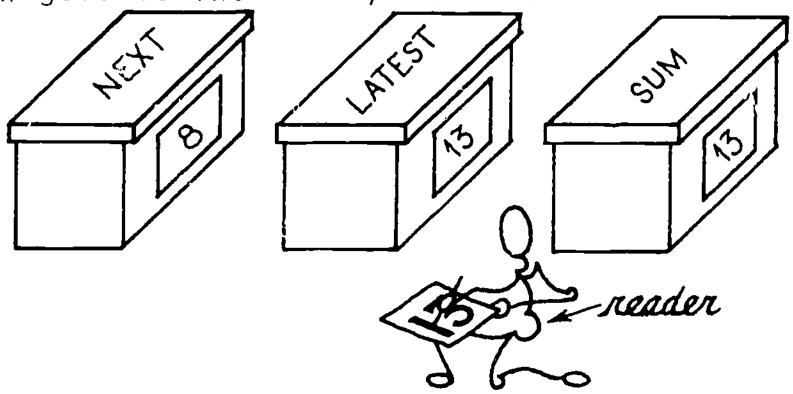
The computation called for in the assignment statement occurs on the right-hand side of the arrow, so the master computer looks there first.



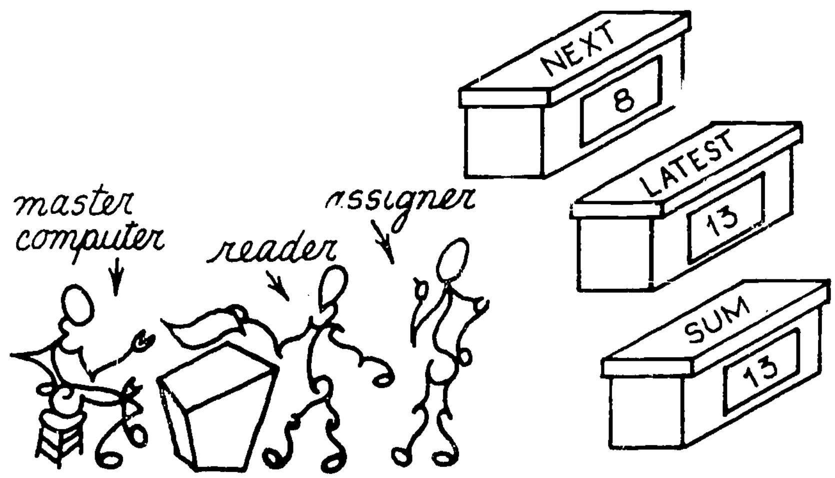
He sees that he must know the values of the variables

LATEST and NEXT so he sends the reader out to fetch these values from memory.

The reader then goes to the memory and finds the window boxes labeled LATEST and NEXT. He reads the values of these variables through the windows, jots the values down, and takes them back to the master computer.



The master computer computes the value of LATEST + NEXT using the values of these variables brought to him by the reader,



$$8 + 13 = 21.$$

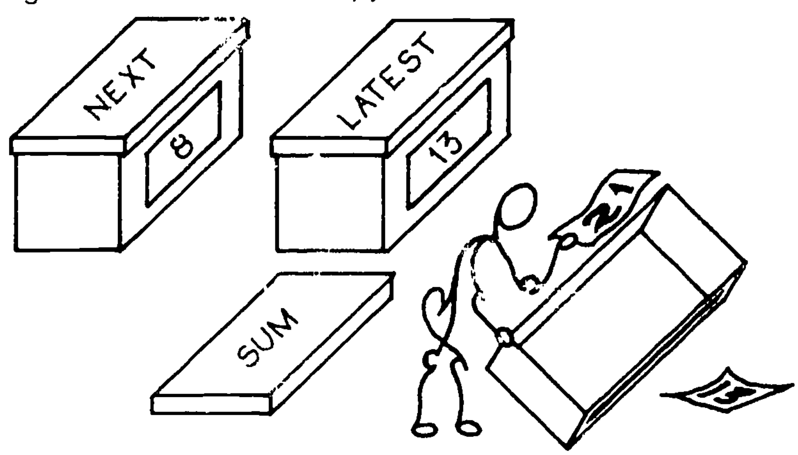
What does he do with this value?

The master computer now looks at the left-hand side of the arrow in his instruction.

SUM ← LATEST + NEXT

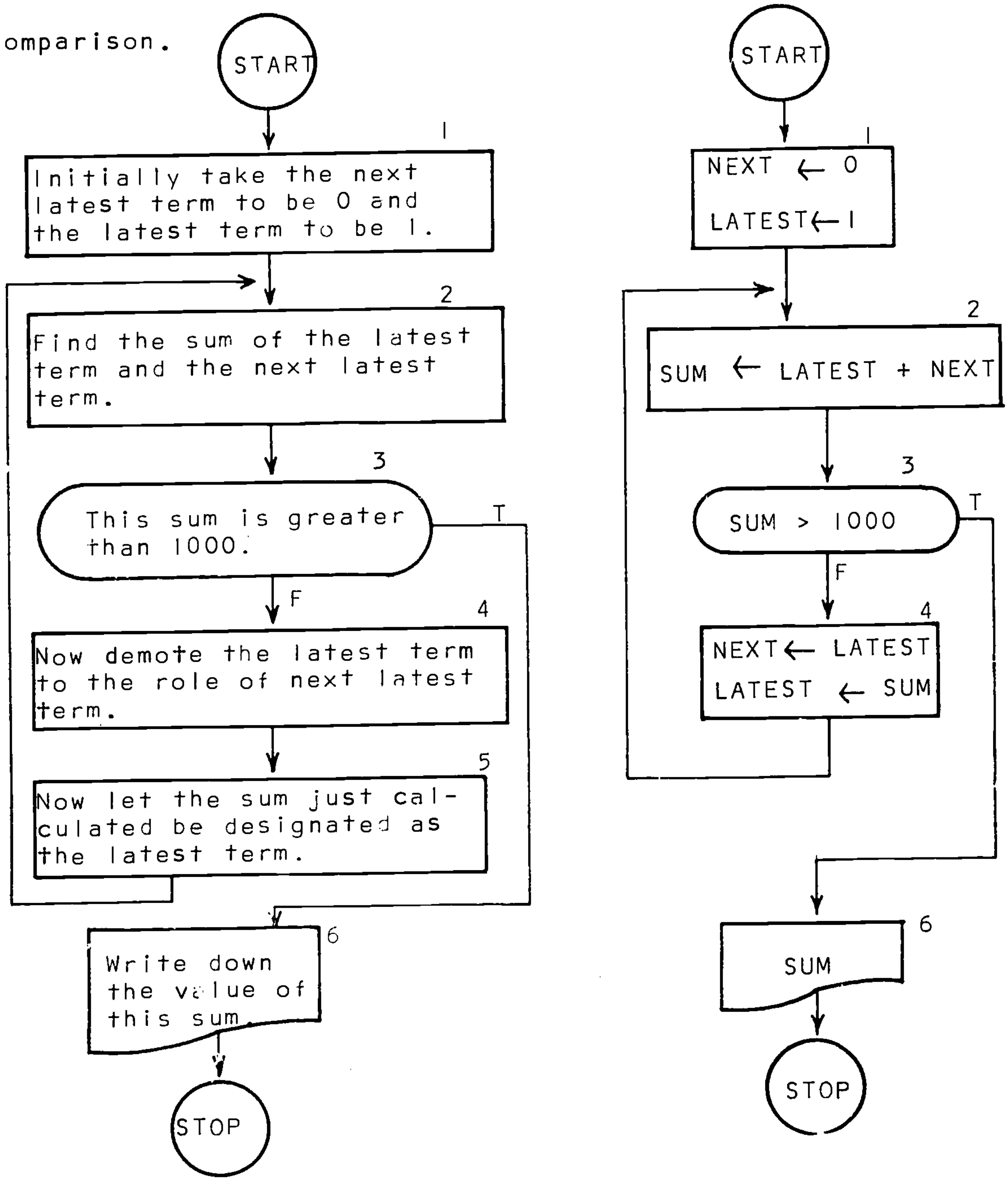
He sees that he must assign the computed value of LATEST + NEXT, namely, 21, writes "21" on a slip of paper, calls the assigner, and instructs him to assign this value to the variable SUM.

The assigner goes to the memory, finds the window box labeled SUM, and dumps out its contents. Then he puts the slip of paper with the new value in the box, closes the lid and returns to the master computer for a new task.



Recapitulating, we see that assignment is the process of giving a value to a variable. We say that assignment is <u>destructive</u> because it destroys the former value of the variable. Reading is <u>nondestructive</u> because the process in no way alters the values of any of the variables in the memory.

We present in Figure 3-2 the entire flow chart of Figure 2-1 in simplified flow chart language. The old and the new flow charts are placed side by side for easy comparison.



(a) Old

(b) New

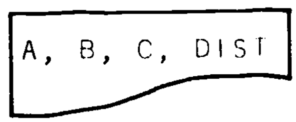
FIGURE 3-2

The translation requires very little explanation. In light of the foregoing explanation it should be obvious that the statement in box 1 is equivalent to the two statements in box 1 on the right. The new version of box 2 has been discussed in detail; box 3 is obvious.

We see that the two statements in boxes 4 and 5 of the old flow chart are compressed into one box, box 4 of the new flow chart. This is permissible whenever we have a number of assignment statements with no other steps in between. However, it is very important to understand that these assignment statements must be executed in order from top to bottom and not in the opposite order and not simultaneously. In fact, we should always think of a computer as doing just one thing at a time and the order in which things are done is generally extremely important.

You can see that the statements in box 4 involve no computation, but merely involve changing the values in certain window boxes. This sort of activity will occur very frequently in future flow charts.

In box 6 of the flow chart, we see written only the variable SUM. The shape of the box (called an output box) tells us that the value of the variable SUM is to be written down. If in some other algorithm we wished to write down the values of several variables, we would list these variables in an output box separated by commas, for example:



Tracing the ow Chart. To better understand what our flow chart in Figure 3-2(b) is doing, let us trace through it executing the steps as the master computer and his assistants would do them.

Tracing of the Flow Chart of Figure 3-2(b)

Step Number	Flow Chart Box Number	Values of Variables			Test	True or False
		NEXT	LATEST	SUM		
1	1	0	1	–		
2	2			1		
3	3				1 > 1000	F
4	4	1	1			
5	2			2		
6	3				2 > 1000	F
7	4	1	2			
8	2			3		
9	3				3 > 1000	F

Step Number	Flow Chart Box Number	Values of Variables			Test	True or False
		NEXT	LATEST	SUM		
10	4	2	3			
11	2			5		
12	3				5 > 1000	F
13	4	3	5			
14	2			8		
15	3				8 > 1000	F
16	4		8			
17	2			13		
18	3				13 > 1000	F
19	4	[illegible]	13			
20	2			21		
21	3				21 > 1000	F
22	4	13	21			
23	2			34		
24	3				34 > 1000	F
25	4	21	34			
26	2			55		
27	3				55 > 1000	F
28	4	34	55			
29	2			89		
30	3				89 > 1000	F
31	4	55	89			
32	2			144		

Step Number	Flow Chart Box Number	Values of Variables NEXT	LATEST	SUM	Test	True or False
33	3				144 > 1000	F
34	4	89	144			
35	2			233		
36	3				233 > 1000	F
37	4	144	233			
38	2			377		
39	3				377 > 1000	F
40	4	233	377			
41	2			610		
42	3				610 > 1000	F
43	4	377	610			
44	2			987		
45	3				987 > 1000	F
46	4	610	987			
47	2			1597		
48	3				1597 > 1000	T
49	6				1597	

In this trace, for ease of reading, the values of the variables are reproduced only when assignments are made to them. In between such steps, the values of the variables do not change and hence have the last previously recorded values. For example, in step 33 where we are working a test, the values of the variables are

NEXT = 55, LATEST = 89, SUM = 144.

In step 34, the values are

NEXT = 89, LATEST = 144, SUM = 144.

You can see that on step 48 in the execution of our algorithm, we finally leave box 3 by the true exit and pass on to box 6 where we output the answer, 1597, and stop.

The infantile simplicity of our conceptual model avoids and conceals certain pitfalls. There is a danger of thinking of assignment as being equality or substitution which it is not. (We'll have more to say about this later on.) This and other sources of confusion (such as the effect of a certain sequence of flow chart statements) can be cleared up by thinking in terms of our model which will always give the right answers.

In fact, the best way to get the ideas into your mind would be to make some window boxes and, with two other students, take the roles of master computer, assigner, and reader and work through a couple of algorithms as described in this section.

PROBLEMS

1. What would be the effect of changing the order of the two assignment statements in box 4 of Figure 3-2(b) so as to appear as seen at the right. Trace through the flow chart with this modification until you find the answer.

LATEST ← SUM
NEXT ← LATEST

2. (a) To compare the effects of the assignment statements

$$A \leftarrow B \quad \text{and} \quad B \leftarrow A$$

find the missing numbers in the table below.

Values before Execution of Assignment		Assignment to be Executed	Values After Execution of Assignment	
A	B		A	B
7	13	A ← B	?	?
7	13	B ← A	?	?

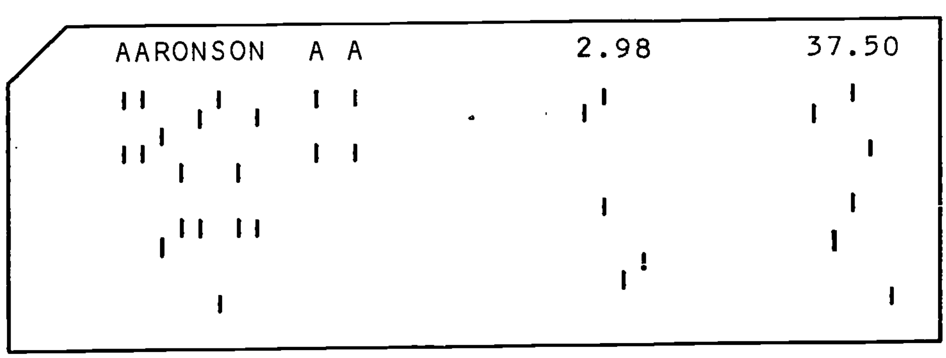
(b) In which of the two cases is it true that A = B after assignment?

(c) Are the effects of the two assignment statements the same or different?

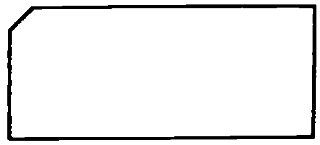
4. <u>Input</u>

Imagine that you are a bookkeeper in a large factory. You have records of the hourly rate of pay and the number of hours worked for each employee, and you have to calculate the week's wages. Of course, this can be done by hand; but assume there are nearly a thousand workers in the plant so that the job will be quite tedious. Naturally you prefer to have the computer do this task for you. That being the case, you will have to devise a flow chart to give the instructions to the computer.

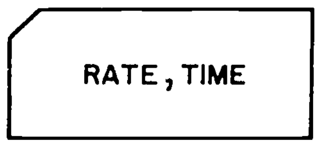
How will you get the hourly rates and the hours worked into the flow chart? Will you write them all in separately? If so, it will take a long time. The characteristic situation is that we have a stack of punch cards, one for each worker. According to a certain code of hole-patterns, each card is punched with the name of the worker, his hourly rate of pay, the number of hours worked, and perhaps some other things as well. Here is a sample of what such a card might look like



Now we will introduce a new shape of frame--the <u>input box</u>--into our flow chart language. The input box has this shape



to suggest a punch card. Inside the box will appear a single variable or a list of variables separated by commas.



When the above box is seen in a flow chart, it is interpreted as an instruction to the master computer to do these three things:

i) read two numbers from the top card in a stack of punch cards;

ii) assign these numbers respectively to the variables RATE and TIME; and

iii) remove this card from the stack.

We see that an input box is a command to make assignments, but this command is essentially different from that in an assignment box. In an assignment box, the values to be assigned are to be found in the computer's memory or are computed from values already in the computer's memory, whereas, with an input box the values to be assigned are obtained from outside the memory. No calculation may be called for in an input box.

In a real computer (not our conceptual one), the distinction between these two kinds of assignment shows up very sharply. The assignments called for in an input box usually involve some mechanical motion such as removing a card from a stack, while assignments called for in an assignment box are made by electronic pulses which move at nearly the speed of light, and hence much faster than input assignments.

Now, let's see how the input box is used in our hourly rate and payroll problem. Should we input the data from all the cards before we start our calculations? If so, we would need a tremendous number of window boxes in which to store all this data. Instead, we will calculate the wages after each card is read. A description of our process is:

1. Read the RATE and TIME from the top card in the stack and remove the card.
2. Multiply the RATE by the TIME to get the WAGE.
3. Output the values of RATE, TIME, and WAGE.
4. Return to step 1.

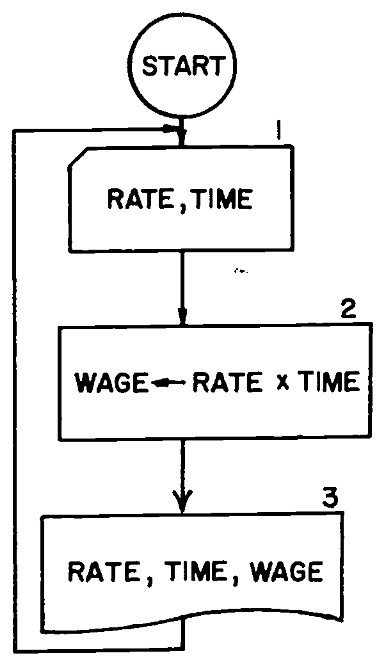


Fig. 4-1

This is realized in the flow chart of Figure 4-1. Each step in the above list appears in a similarly numbered box in our flow-chart except the 4th step.

That is represented by the arrow returning from box 3 to box 1.

You may wonder that the flow chart does not have a stop button. We assume as one of the functions of the input box, the duty of stopping the computation if the reading of another card is called for when the stack is empty.

PROBLEMS

1. Modify the flow chart of Figure 4-1 to provide for an overtime feature. All hours in excess of 40 are to be paid at time and a half. You will have to place a test somewhere in the flow chart to determine whether the worker actually put in any overtime. The formula by which his wages are computed will depend on the outcome of this test.

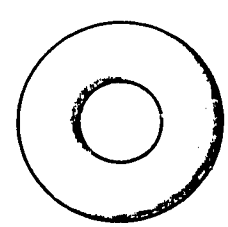
5. Computer Memory

Now we are ready to look at how our conceptual model of a computer can be realized in an actual machine. In this section and the next, we will discuss a prototype machine which we will call SAMOS. SAMOS is a prototype machine stripped down to the bare essentials. Some features of its operation are described in considerable detail while others are glossed over. The programming of SAMOS is described briefly in Section 6.

In order to study this book, it is only necessary that you should have a general idea of how a computer works. So we suggest that you read over the material in these two sections quite rapidly without attempting to master it. Just retain what sticks in your mind.

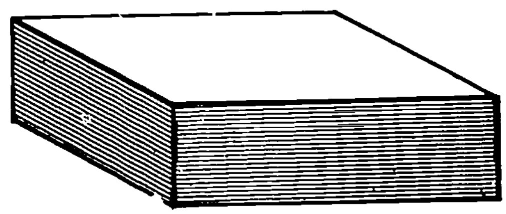
Cores

We will start with the memory. How are all those window boxes realized in actual practice? The memory of SAMOS is a rectangular

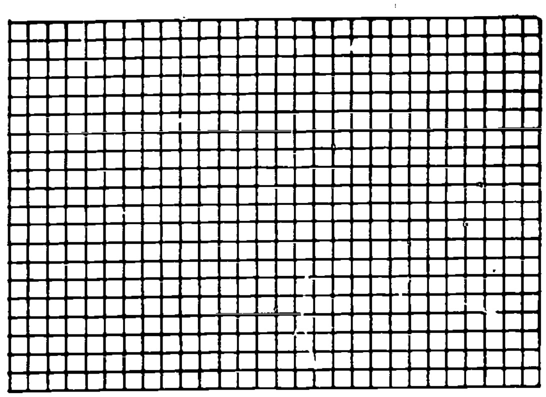


box. Inside the box, there is an arrangement of tiny magnetic doughnuts about $\frac{1}{20}$ of an inch in diameter. These doughnuts are called <u>cores</u>.

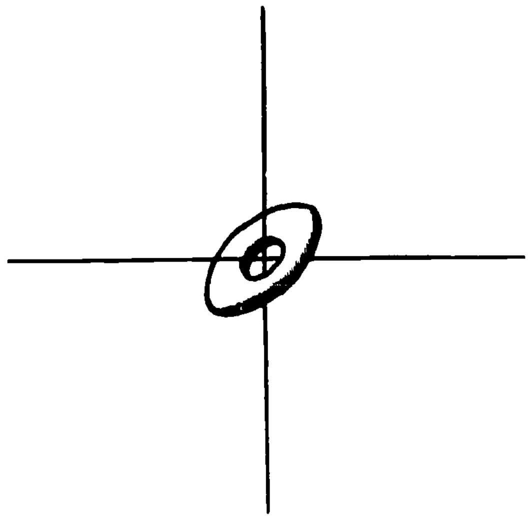
Our box is divided in 61 horizontal layers or trays called <u>core planes</u>.



On each of these layers wires are strung in two directions like the lines on a sheet of graph paper.

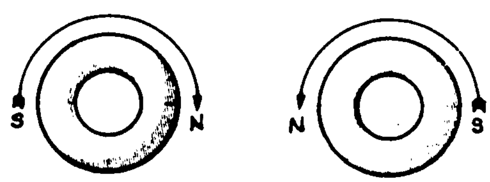


There are a hundred wires in each direction. At each point where two wires cross, the wires are threaded through a core like the thread passing through the eye of a needle.



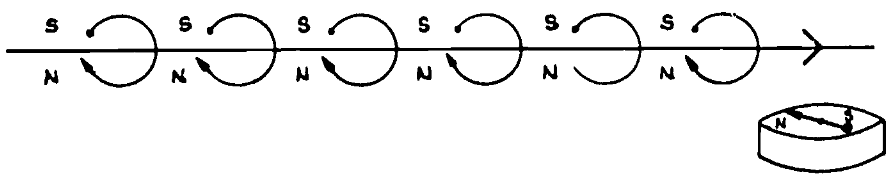
Since there are 100 × 100 crossings in each layer, we see that there are 10,000 cores in each core plane and hence 61 × 10,000 = 610,000 cores in the entire box.

These cores are capable of being magnetized in either the clockwise or the counter-clockwise sense.



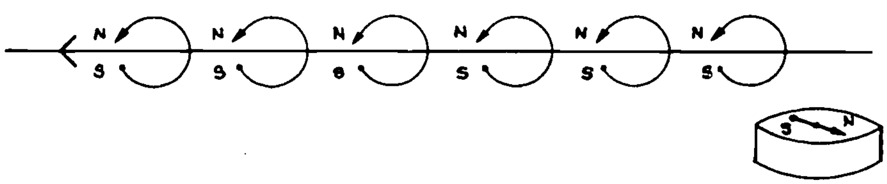
Because of this, the core can store information. We could think of clockwise magnetization as meaning "yes" and counter-clockwise as meaning "no". We will instead think of clockwise as standing for "0" and counter-clockwise for "1". In any event, the information contained in the direction of magnetization of a core is the smallest unit of information and is called a <u>bit</u> of information. We see that one core can store one binary digit 0 or 1, but a collection of cores can store a very large number. This we will discuss a little later on, after a digression to see how the cores get their magnetism.

First you must know that a pulse of electric current moving along a wire generates a magnetic field running around the wire, as depicted below.

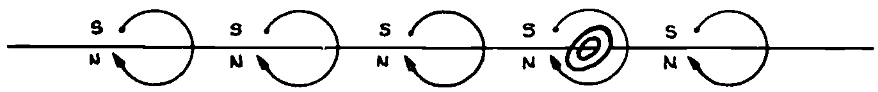


This field can be detected by a pocket compass. The strength of the magnetic field is strongest near the wire and dies away as we move further from the wire.

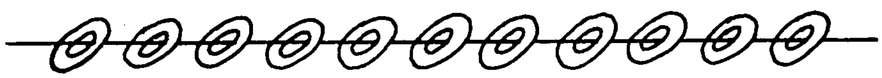
If the direction of the current is reversed, the direction of the magnetic field is also reversed.



Thus, when a pulse of current passes through a core, the core will become magnetized in one direction or the other, depending on the direction of the current.



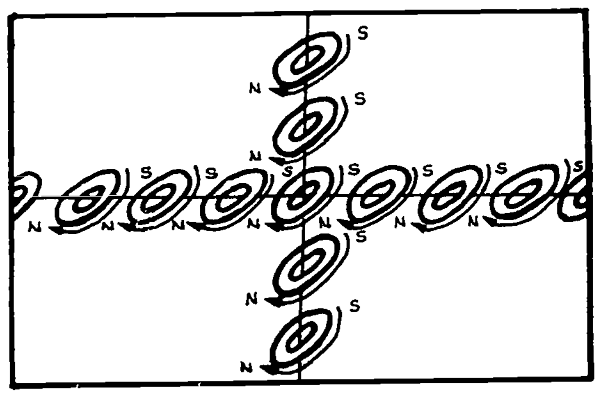
But how can we manage to magnetize just one core instead of the whole string of cores through which the



pulse passes? The answer lies in the magnetic properties of the material of which the core is made. With this material, if the pulse is too weak, then the direction of the magnetization of the core is not permanently altered. After the pulse of current has passed by, the core merely returns to its former magnetic condition, whatever that was.

On the other hand, if the current is strong enough, the core remains permanently magnetized in the sense established by the direction of the current, regardless of the former magnetic condition of the core. The situation is analagous to trying to throw a ball from the ground to the flat roof of a building. If you have enough power in your throw, the ball will land on the roof; otherwise it will bounce against the wall and fall back to the ground.

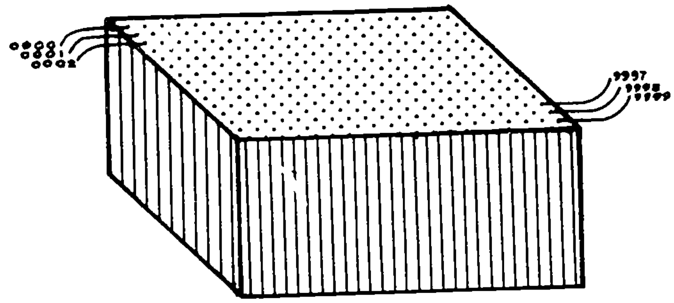
Now the strength of the pulses is carefully regulated so that one pulse is not sufficient to permanently magnetize a core but two pulses acting simultaneously will exceed the threshold strength and result in permanent magnetism. Thus, pulses passing along both wires shown below will permanently magnetize just the one core which is located where the wires cross.



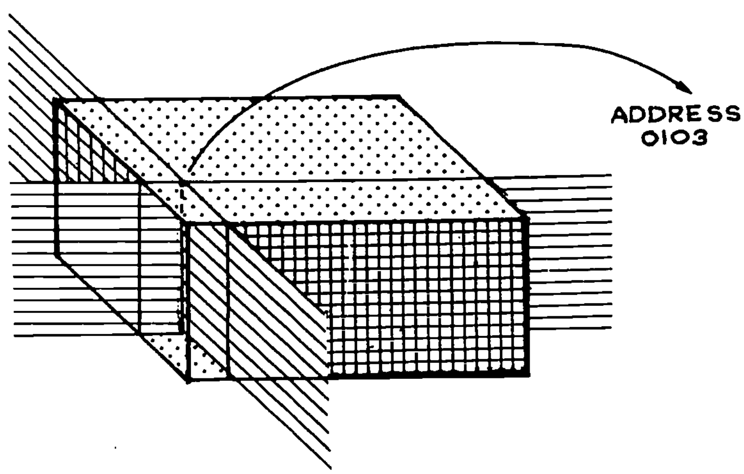
The Store

Now let's leave the individual core planes and consider the entire memory or store of the computer composed of the 61 core planes. Each vertical column of 61 cores constitutes a computer word. Thus, the memory of the computer is composed of 10,000 words. These words have addresses (like house numbers) which are 4-digit numbers from 0000 to 9999 by means of which we may refer to them. Each of the 10,000 dots on the top of the box is the top of a vertical column of 61 cores (or a word). The manner of assigning the addresses is indicated in the figure.

Each of these words corresponds to a window box in our conceptual model. Each variable in the flow chart will have a certain address. The word with that address will have in it a certain pattern of "bits" (directions of magnetization of its cores) representing the value of that variable. "Assigning a value to a variable" is effected by putting a certain pattern of bits into a word.



In more detail, when we said "the master computer tells the assigner to assign the value 1597 to the variable SUM", what actually takes place is this: The variable SUM is represented inside the machine by means of its address; suppose it is 0103. Now all the 61 x 2 = 122 wires passing through cores in the word addressed 0103 are energized with pulses of current in the proper direction so as to achieve the pattern of bits representing the number 1597.



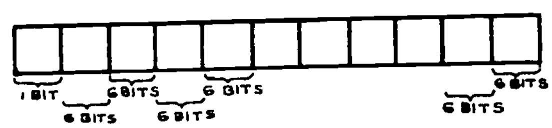
In a modern computer, this assignment process can be performed in 3/10 of a microsecond; that is, 3/10,000,000 of a second.

Characters

One obvious way of representing the number 1597 would be in the binary system as

1 1 0 0 0 1 1 1 1 0 1

preceded by a string of zeros to bring the total number of binary digits up to 61. But that isn't the way we'll do it. We want the words to operate on the decimal system rather than binary, and we would like to be able to store letters as well as digits. For this reason, we subdivide our 61 bit words into 11 characters as shown at right.

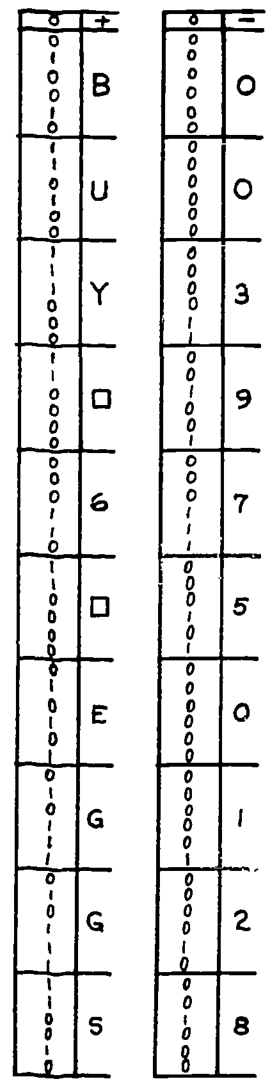


The first character is reserved for holding a sign, + or −. Here 0 stands for + and 1 for −. Each of the other characters consists of 6 bits. These characters can be used to store numbers or letters, according to the following code.

Char.	Code	Char.	Code	Char.	Code	Char.	Code
0	00 0000						
1	00 0001	A	01 0001	J	10 0001		
2	00 0010	B	01 0010	K	10 0010	S	11 0010
3	00 0011	C	01 0011	L	10 0011	T	11 0011
4	00 0100	D	01 0100	M	10 0100	U	11 0100
5	00 0101	E	01 0101	N	10 0101	V	11 0101
6	00 0110	F	01 0110	O	10 0110	W	11 0110
7	00 0111	G	01 0111	P	10 0111	X	11 0111
8	00 1000	H	01 1000	Q	10 1000	Y	11 1000
9	00 1001	I	01 1001	R	10 1001	Z	11 1001

We have used up only 36 of the 64 combinations available with a 6 bit code. This leaves 28 additional combinations for other special symbols such as +, $\geq$, etc. We introduce one of these right now, namely the blank space, ␣, which is coded as

1 1 0 0 0 0.



With this code, you can see that the 61 bit computer words displayed vertically at the right turn out to be

+	B	U	Y		6		E	G	G	S

and

-	0	0	3	9	7	5	0	1	2	8

From now on we will represent our computer words as strings of 11 characters instead of strings of 61 bits.

6. Arithmetic and Control Units of SAMOS

Now that we have seen how SAMOS's memory is structured we will consider how the memory is used in executing an algorithm.

Our computer has several other components besides the memory. These are shown in the block diagram in Figure 6-1.

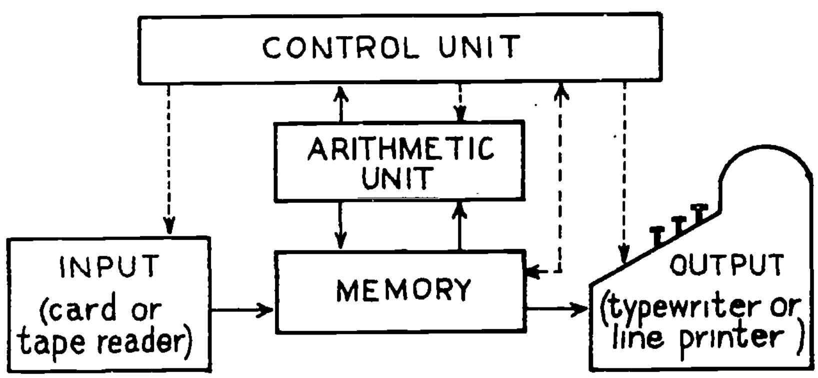


FIGURE 6-1

The solid lines indicate the directions in which values may be transferred. The dashed lines indicate the transferral of instructions or the exercise of control.

The control unit and the arithmetic unit perform the duties of the "master computer".

An important part of the arithmetic unit is the accumulator. This is a special computer word in which all arithmetic operations are performed. Furthermore, a simple assignment like

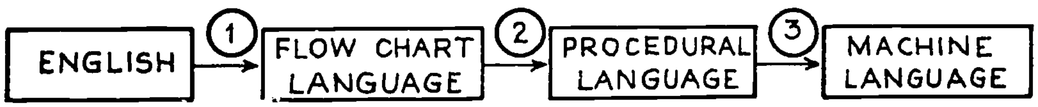
LATEST ← SUM

is carried out by first copying the value of SUM into the accumulator and then copying the value in the accumulator into the computer word belonging to the variable LATEST. The value of SUM is unchanged in this process. Note that values to be input or output do not pass through the accumulator but go directly in and out of memory.

Where does the control unit get such instructions as referred to in the last paragraph? These are also stored in the computer's memory. We will learn something about that presently.

7. Machine Language

Getting an algorithm into a form in which a machine can execute it involves several translations which may be depicted as follows:



You have already had a little experience with the first translation step. The second translation step is the process of translating a flow chart into a procedural language such as FORTRAN, ALGOL, MAD, or PL/I. Suffice it to say that this step is quite mechanical and can be performed by a person who has no idea what the algorithm is all about. The third translation process is completely mechanical and is done by the computer itself. This process is called compiling.

We don't need to know how compiling is done, but we do need to know the reason for doing it. Each make and style of computer has its own language--that is, its own set of instructions which it can understand. To avoid this tower of Babel in which a programmer would have to learn a new language for each machine with which he wished to communicate, the procedural languages were developed. These procedural languages constitute a kind of "Esperanto" which enables a programmer to communicate with many different machines in the same language.

The programmer merely prepares, say, a FORTRAN program on punched cards, feeds it into the computer which "compiles" a machine language program which is placed in the computer's memory.

In our SAMOS computer, these instructions will be placed in order in consecutively-addressed locations in memory starting with 0000. After the computer has executed an instruction, it will always look for the next instruction in the next address, except when there is a branching instruction telling it to go to a different address for the next instruction.

To see how this works, consider the following instruction seen in the Fibonacci sequence flow chart.

SUM ← NEXT + LATEST

FIGURE 7-1

In FORTRAN, this instruction would appear as:

SUM = NEXT + LATEST;

and in ALGOL:

SUM: = NEXT + LATEST

In the SAMOS language, these variables cannot be referred to by <u>name</u> but only by the <u>addresses</u> in memory associated with the variables. Suppose that NEXT, LATEST, and SUM have been given, respectively, the locations 0100, 0101, and 0102. Then in the SAMOS language, the above instruction would take the form or a sequence of three instructions:

+	L D A	0 0 0	0 1 0 0
+	A D D	0 0 0	0 1 0 1
+	S T Ø	0 0 0	0 1 0 2

Figure 7-2. SAMOS instructions for Figure 7-1

These instructions have the form of 11 character words, but the first character is not used here and neither are the 5tn, 6th, and 7th. The letters at the left of the instructions indicated the operation being performed, and the four-digit numerals at the right are addresses.

The letters LDA stand for "LoaD the Accumulator". The whole instruction means: "Copy the contents of the memory word addressed 0100 into the accumulator without altering the contents of address 0100." Clearly this is the function of the reader in our conceptual model. We will not go into the details of the electronics involved in carrying out this instruction. It is sufficient to know that when this pattern of bits in the instruction

0	L D A	0 0 0	0 1 0 0

is brought to the control unit, certain switches are thrown which allow a pulse of current to pass through the cores of the word 0100. The magnetized cores effect an alteration of the current which in turn permits a copy to be made.

The second instruction in Figure 7-2 means: "ADD the value in the word addressed 0101 to the value already in the accumulator and place the result in the accumulator." The third instruction means: "Copy (or STØre) the number in the accumulator into the word addressed 0102." Times vary from machine to machine, but in modern computers, the time required for carrying out such instructions will usually be less than $\frac{1}{1,000,000}$ of one second.

A Complete SAMOS Program. We are about ready to see how the entire flow chart for the Fibonacci sequence algorithm (repeated at the left) will emerge in SAMOS language. First, however, we must remark that in the SAMOS language we can never refer to a number directly but only to a memory address in which this number may be found. This even applies to constants. Thus, part of the compiling process will involve providing memory addresses for the constants (as well as the variables) appearing in the program. We assume that the addresses 0103, 0104, and 0105 have been set aside for the constants 0, 1, and 1000 appearing in the flow chart and that the proper values have already been put in the words with these addresses. The

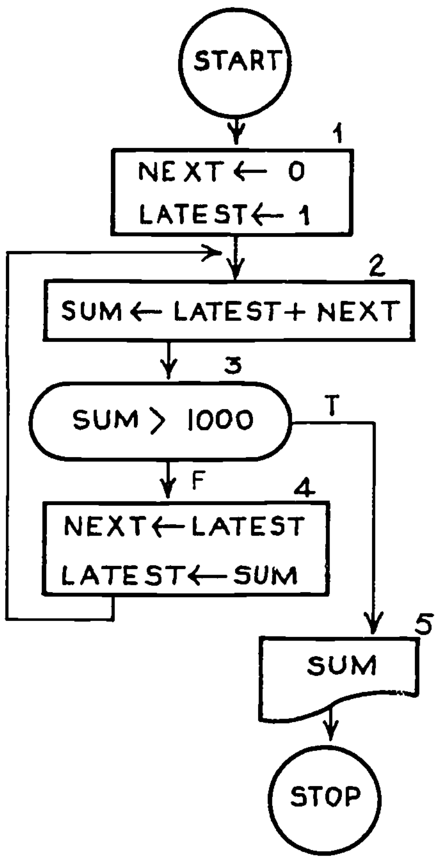


FIGURE 7-3

memory locations 0100, 0101, 0102 have been allocated for the variables NEXT, LATEST, SUM, but no values have been placed in these words. The state of the memory at the beginning of the execution of the SAMOS program for the Fibonacci algorithm is at the bottom of Figure 7-4. Operations not previously met will be explained in the discussion following this program.

MEMORY LOCATION (Address)	+ – 1	OPERATION 2 3 4	5 6 7	ADDRESS 8 9 10 11	FLOW CHART EQUIVALENT Character number
0 0 0 0		L D A		0 1 0 3	NEXT ← 0
0 0 0 1		S T Ø		0 1 0 0	
0 0 0 2		L D A		0 1 0 4	LATEST ← 1
0 0 0 3		S T Ø		0 1 0 1	
0 0 0 4		L D A		0 1 0 1	SUM ← LATEST + NEXT
0 0 0 5		A D D		0 1 0 0	
0 0 0 6		S T Ø		0 1 0 2	
0 0 0 7		L D A		0 1 0 5	SUM > 1000 (T →; F ↓)
0 0 0 8		S U B		0 1 0 2	
0 0 0 9		B M I		0 0 1 5	
0 0 1 0		L D A		0 1 0 1	NEXT ← LATEST
0 0 1 1		S T Ø		0 1 0 0	
0 0 1 2		L D A		0 1 0 2	LATEST ← SUM
0 0 1 3		S T Ø		0 1 0 1	
0 0 1 4		B R U		0 0 0 4	Arrow from flow chart box 4 to box 2
0 0 1 5		W W D		0 1 0 2	SUM
0 0 1 6		H L T			STOP
0 1 0 0					The variable NEXT
0 1 0 1					The variable LATEST
0 1 0 2					The variable SUM
0 1 0 3	+	0 0 0	0 0 0	0 0 0 0	The constant 0
0 1 0 4	+	0 0 0	0 0 0	0 0 0 1	The constant 1
0 1 0 5	+	0 0 0	0 0 0	1 0 0 0	The constant 1000

Figure 7-4. SAMOS Program for Fibonacci sequence algorithm

Discussion. The instructions in memory addresses 0004, 0005, and 0006 have already been discussed. Before looking at the other instructions, look first at memory locations 0100 through 0105 to see where your variables and constants are located.

From previous discussions, you see that the instruction in 0000 copies the value in 0103 (that is, the number 0) into the accumulator. Next, the instruction in 0001 copies the value in the accumulator into the word with address 0100. Together these steps are equivalent to assigning 0 to the variable NEXT. Similarly, the instructions in addresses 0002 and 0003 are equivalent to assigning the value 1 to the variable LATEST.

Remember that the control unit executes the instructions in order until it comes to a branching instruction. The first of these branching instructions is found in address 0009, reading

	B	M	I				0	0	1	5

The code BMI stands for "Branch on Minus". The whole instruction means, "If the number in the accumulator is negative, go to address 0015 for the next instruction, otherwise go on as usual to the next numbered address (0010)." We will see shortly that the number in the accumulator at this time is just

1000 - SUM

so that the number in the accumulator will be negative only in the case that

SUM > 1000.

In this case, the branching instruction sends us to address 0015 where we see the instruction

	W	W	D				0	1	0	2

which means, "Write the WorD in address 0102" which amounts to printing out the value of SUM.

Now why is it that when the instruction in address 000 is reached, the number in the accumulator is 1000 - SUM? Well, on looking at the instruction in address 0007, one sees that it instructs us to load the accumulator with the contents of 0105; that is, to put the number 1000 in the accumulator. The next instruction, that in 0008, tells us to "subtract the contents of 0102 from the accumulator and put the result in the accumulator." Since the contents of 0102 is just the value of SUM, this amounts to the placing in the accumulator.

You should be able to verify for yourself that the instructions in addresses 0010 through 0013 accomplish the assignments indicated in the right-hand column.

The instructions in memory address 0014 needs to be described.

	B	R	U				0	0	0	4

BRU stands for "BRanch Unconditionally". The meaning of the entire instruction is "Go back to memory address 0004 for the next instruction and continue in order from there." You can see this corresponds to the arrow from flow chart box 4 leading back to flow chart box 2 where we again repeat the assignment

SUM ← LATEST + NEXT

The instruction in 0016, of course, stands for HaLT and amounts to stopping the computing process.

You can best understand all this by tracing through the SAMOS program by hand, keeping a record of:

i) which instruction is being executed;

ii) the value in the accumulator;

iii) the values in the memory locations 0100, 0101, and

0102 (the values of NEXT, LATEST, and SUM).

Note that the contents of the instructions in addresses 0000 - 0016 are never altered, nor are the contents of the locations 0103 - 0105 (the constants 0, 1, and 1000).

PROBLEMS

1. Construct a list of SAMOS instructions for the flow chart of Figure 4-1. You will need two additional instructions. The first is

OPERATION							ADDRESS			
1	2	3	4	5	6	7	8	9	10	11
	R	W	D				1	0	0	5

which is an instruction to read a number from a card into the computer word addressed 1005.

The second is

	M	P	Y				1	0	2	3

which is an instruction to multiply the number in the accumulator by the number in address 1023 and put the result in the accumulator. (Of course, in the address part of these instructions we may put any address we wish.)

8. Odds and Ends

Only a few of the ideas we have learned about SAMOS need to be remembered.

Among the things to be remembered is the sequential nature in which the computer works, that is, the one-by-one steps in which the computer performs its tasks. The order in which the tasks are performed is just as important as what it does.

Another property of computers that we must understand is the finite word length. We have seen that SAMOS words consist of 10 characters and a sign so that the largest number representable in this coding system is

+ 9,999,999,999

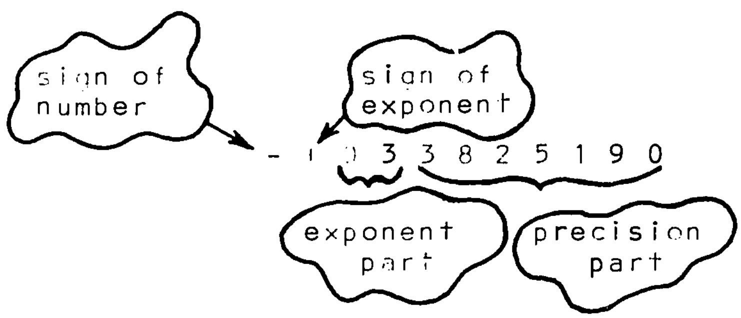
a rather large number but still finite.

You should be aware that there are other ways of coding numbers which allow us to work with numbers other than integers. One of the most common of these is floating point form which is similar to the so-called "scientific notation".

To see how this works, recall that any decimal numeral such as -382.519 can be expressed as

$$- .382519 \times 10^{3}$$

in which (right after the sign, if any) there is a decimal point followed by a non-zero digit multiplied by a suitable power of 10. We can code numbers in this way in SAMOS by reserving three characters for the exponent, thus having



Some examples of how numbers are coded in this system are shown in the table which follows:

NUMBER	FLOATING POINT FORM	SAMOS CODING		
3.1415926	$.31415926 \times 10^{1}$	+	+01	3141592
-273.14	$-.27314 \times 10^{3}$	-	+03	2731400
.0008761	$.8761 \times 10^{-3}$	+	-03	8761000
.73	$.73 \times 10^{0}$	+	+00	7300000
4	$.4 \times 10^{1}$	+	+01	4000000
1/3	$.333333333 \times 10^{0}$	+	+00	3333333
11/7	$.157142857 \times 10^{1}$	+	+01	1571428

By glancing at the table, we see that the eight-digit representation of π in the top of the left-hand column has to be chopped down to 7 digits of precision due to space requirements. The same holds true for 1/3 and 11/7 at the bottom of the table. Thus, we see that in a computer even such a simple fraction as 1/3 cannot be represented exactly, but only to a close approximation. This characteristic of "finite word length" presents important problems in computer work, which will be discussed in various places in the main text.

In this coding system, we can represent large numbers but we pay a price in giving up three places of precision. The largest floating point number representable is

+	+	9	9	9	9	9	9	9	9	9

which represents the number

999,999,900,000,000,000,000,000,000
000,000,000,000,000,000,000,000,000
000,000,000,000,000,000,000,000,000
000,000,000,000,000,000.

Similarly, there is a smallest positive number which can be represented, namely,

+	−	9	9	1	0	0	0	0	0	0

or

0.000 000 000 000 000 000 000 000 000
000 000 000 000 000 000 000 000 000
000 000 000 000 000 000 000 000 000
000 000 000 001

which is very small indeed.

In practice, most machines impose other restrictions which further limit the largeness and the smallness of the numbers which can be represented.

9. Iteration Boxes

Typical of the sort of thing we will have to do quite frequently in this course is adding up the reciprocals of the integers from 1 to 1000, indicated by the probably familiar "sigma" notation

$$\sum_{n=1}^{1000} \frac{1}{n} .$$

If you try to evaluate this sum by hand computation, you will soon find out why this problem had best be done on a computer.

For this algorithm, we need two variables, a variable n which successively takes on the values 1, 2, ..., 1000, and a variable SUM which keeps a running total of the sum of the reciprocals of the values of n. The rudimentary idea is illustrated in Figure 9-1.

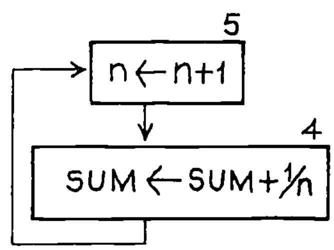


FIGURE 9-1

Of course we need a testing device in order to branch out of this loop when n exceeds 1000 as well as a means of assigning n its initial value 1. These additions are seen in Figure 9-2. The finishing touches to make this flow chart operational are: starting the variable SUM with a clean slate (SUM ← 0); and outputting the final value of SUM. The complete flow chart is seen in Figure 9-3.

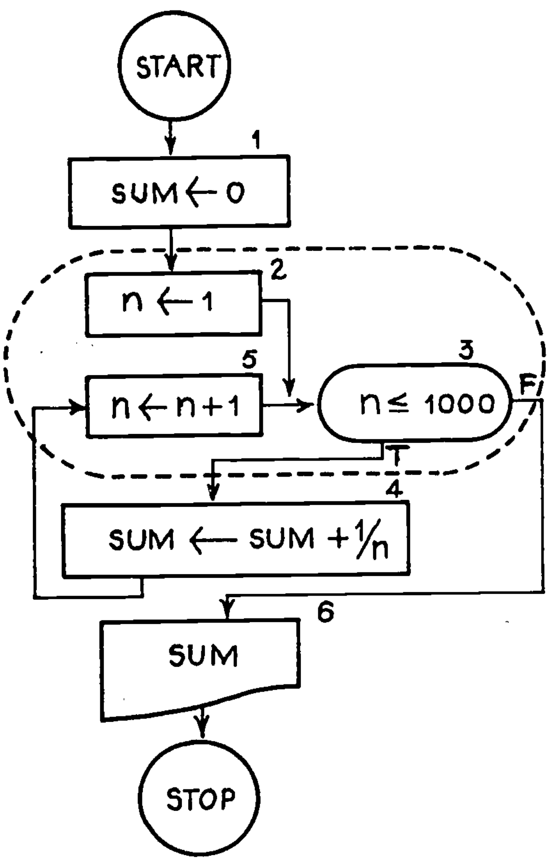


FIGURE 9 – 3

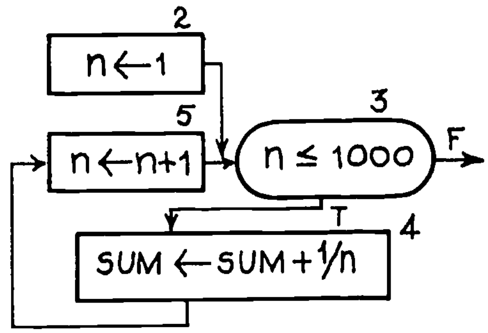


FIGURE 9 - 2

We see that the variable n acts as a sort of counter "controlling" the loop in this flow chart. This variable n:

i) starts with the value 1 (Flow chart box 2);

ii) and goes click, click, click in steps of 1 (box 5);

iii) through the value 1000 (box 3);

iv) executing the loop computation (box 4) at each step.

Loops controlled in this way by counters are of such frequent occurrence that we find it convenient to introduce a new kind of flow chart box to assume all these functions of the counter, initialization, testing, incrementation seen in Boxes 2, 3, and 5 of Figure 9-3. This three compartment flow chart box illustrated in Figure 9-5 is called an <u>iteration</u> <u>box</u>.

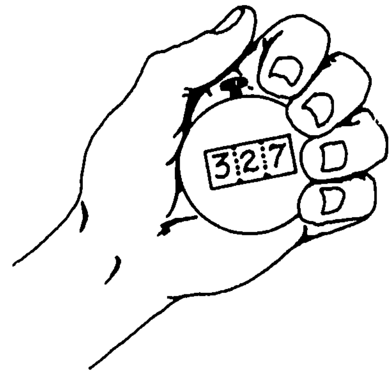


FIGURE 9-4

It is readily seen to be obtained by compressing together boxes 2, 3, and 5 of Figure 9-3. The flow chart of Figure 9-3 with this modification is seen in Figure 9-6.

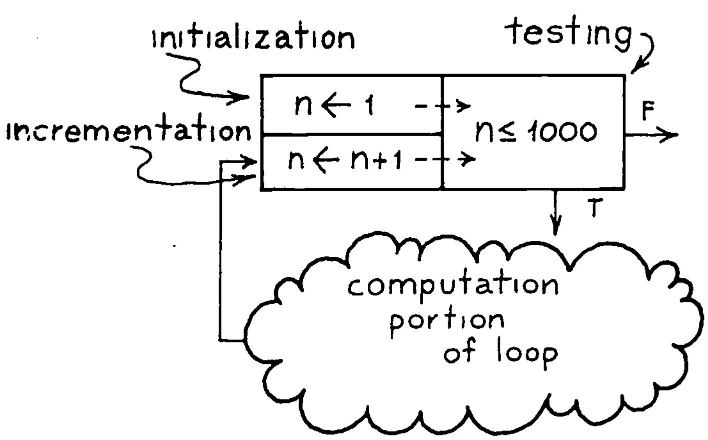


FIGURE 9-5

These iteration boxes are often very helpful in organizing our thinking in constructing flow charts. They are useful in many other contexts besides summation formulas, notwithstanding the fact that summation

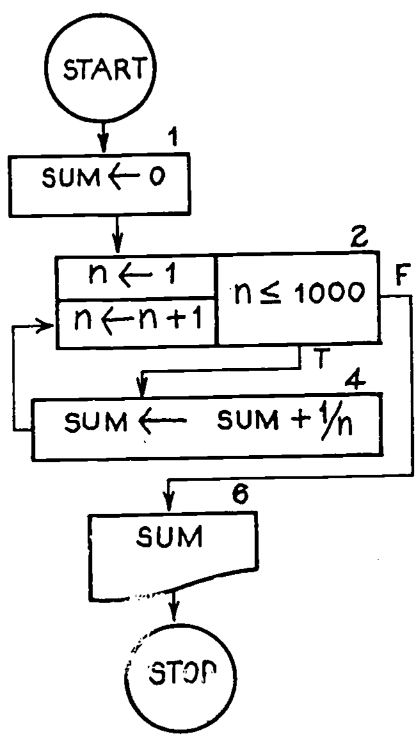


FIGURE 9-6

formulas will be among our best customers for iteration boxes.

The initial value of the variable n, the size of the step, and the relation in Box 2 may be chosen at will. Thus in Figure 9-7 we see a flow chart for calculating the product of the odd integers between 5 and 30.

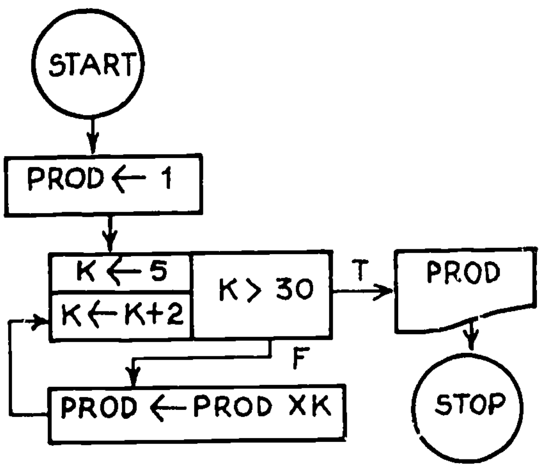


FIGURE 9-7

As a sample of use of iteration boxes, study the flow chart of Figure 9-8 which prints out every three digit number which is equal to the square of six more than the sum of its digits.

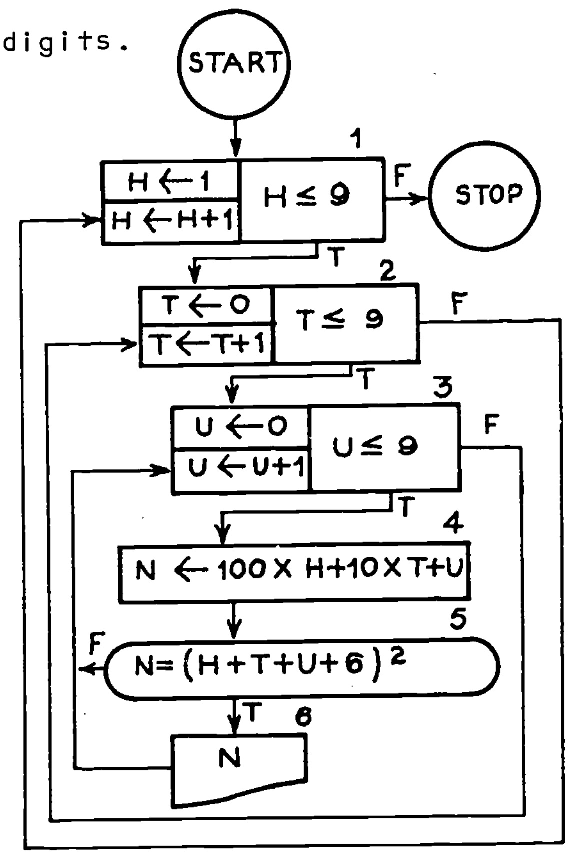


FIGURE 9-8

PROBLEMS

1. Using iteration boxes, write a flow chart and a computer program and run it for computing the exact values of $n!$ for $n = 1, 2, 3, \ldots, 9$.

2. Write and run a computer program for computing approximate values of $n!$ for all integer values of n from 1 to 100.

3. (a) Draw a flow chart for calculating $\sum_{n=1}^{12} 1$

   (b) Trace this flow chart by hand to find the value of the sum in (a)

   (c) Repeat (a) and (b) for the sum $\sum_{k=1}^{5} k^3$

4. (a) Draw a flow chart which will output n, $\sum_{k=1}^{n} k^3$, and $\left(\sum_{k=1}^{n} k\right)^2$ for each value of n from 1 to 50.

   (b) Write the program for part (a) and run it. Study your output and make a conjecture. Can you prove it?

5. In Figure 9-8, which flow chart boxes comprise

   (a) the loop controlled by box 3? box 2? box 1?

   (b) During the course of the algorithm, how many times do we enter box 1 from the top? box 2? box 3?

   (c) How many times will the test in box 5 be executed?

Chapter 2

SEQUENCES

1. Sequences of Approximations

There are many occasions in mathematics when we need to work with numbers which we cannot calculate exactly. When this happens, we must be content with approximations. You have experienced this fact in connection with finding square roots and finding areas of regions with curved boundaries.

A rather crude method for approximating, say, the square root of 2 would be to construct a square of side length 1 and then measure the length of the diagonal. If the figure is carefully drawn, an approximation can be obtained but only accurate to a few decimal places.

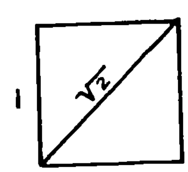


FIGURE 1-1

For a crude approximation of the area of a circle of radius 1, we could inscribe such a circle in a cardboard square. First weigh the square and then cut out the circular disc and weigh it. The area of the circle, π, is then calculated by

$$\frac{\pi}{4} = \frac{\text{weight of circle}}{\text{weight of square}}$$

Again, even with the most careful construction and the best scales, we could hardly expect a result accurate to more than three decimal places.

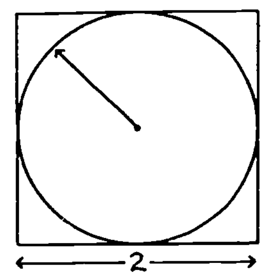


FIGURE 1-2

These methods of approximation, with their limited accuracy, are of no theoretical importance in mathematics. Mathematicians often require methods of approximation of unlimited accuracy. To this end, we look for a sequence or list of approximations with better and better accuracy, so that whatever accuracy may be demanded can be achieved by going sufficiently far down in our list. As an example, we will show how to construct such sequences for finding square roots.

You may have learned in your school days the "divide and average algorithm" for approximating square roots. This is a method for finding a sequence of approximations for square roots. It works like this:

To find the square root of a positive number a, we choose our first approximation to be any number we please, g_1, so long as it is positive. A number h_1 is computed by <u>dividing</u>

a by g_1,

$$h_1 = \frac{a}{g_1}$$

The next approximation, g_2, to the square root of a is obtained by <u>averaging</u> g_1 and h_1,

$$g_2 = \frac{g_1 + h_1}{2}$$

Now we iterate this process over and over, i.e.,

$$g_{n+1} = \frac{g_n + h_n}{2}$$

$$h_{n+1} = \frac{a}{g_{n+1}}$$

The sequence of g's generated in this way

$$g_1, g_2, g_3, g_4, \ldots$$

gets closer and closer to the square root of a. We can get numbers as close as we like to the square root of a. Why this is the case will be shown a little later on. For now, let us see how easy it is to write a computer program for this algorithm. The flow chart for this program is shown at right. This flow chart is not complete because we have not provided a stopping mechanism. This will be done later.

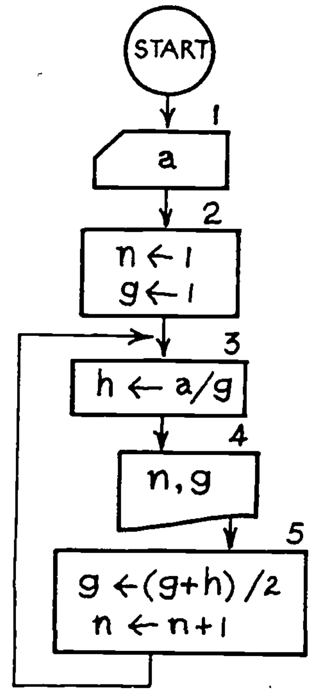


FIGURE 1-3

The purpose of the variable, n, is to number the lines of output. As we see, subscripted variables are not needed in the computer program. Instead, we let the variables g and h take on new values over and over again. We also note that for our first approximation (first value of g), we have arbitrarily selected the positive number 1.

Let's examine a few lines of output from this algorithm, ignoring the finite word length and concomitant round-off error obtained in computer calculations. We will make our calculations exact. We will take the input value of a to be 2. You should trace through the flow chart and verify the values listed below:

n	g
1	1
2	3/2
3	17/12
4	577/408
5	665857/470832
.	.
.	.
.	.

The square of the last value of g listed above is

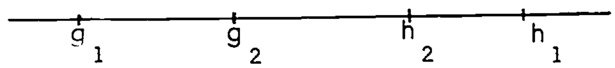
$$\frac{443365544449}{221682772224}$$

which exceeds 2 by $\frac{1}{221,682,772,224}$, that is, by less than 1 part in two hundred billion. Thus, g_5 is a quite good approximation of $\sqrt{2}$. We could repeat the process as many times as we like, getting more and more terms of the sequence ever closer to $\sqrt{2}$.

Let us see why this algorithm works. First note that if two pairs of positive numbers (g_1, h_1) and (g_2, h_2) have the same product, a, that is

$$g_1 h_1 = a = g_2 h_2,$$

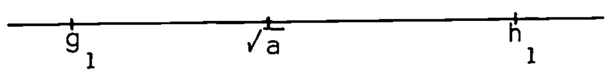
then the two members of one pair must lie between the two members of the other pair.



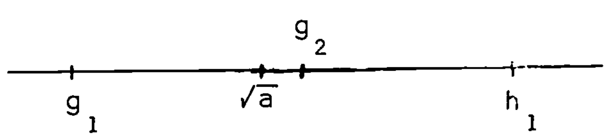
(If h_1 is larger than either g_2 or h_2, then g_1 must, in compensation, be smaller than either g_2 or h_2 in order for the products g_1h_1 and g_2h_2 to be the same.)

As a special case, $\sqrt{a}$ must lie between the members of any such pair. (Take $g_2 = h_2 = \sqrt{a}$.)

Since $h_1 = \frac{a}{g_1}$ by definition, we see that $g_1h_1 = a$, so that g_1 and h_1 lie on either side of $\sqrt{a}$.

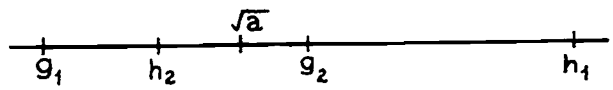


Now $g_2 = \frac{g_1 + h_1}{2}$ by definition and is therefore the midpoint of the interval joining g_1 and h_1.



The number h_2 is then determined by $h_2 = \frac{a}{g_2}$ so that $g_2h_2 = a$, whence by preceding remarks h_2 also lies between g_1 and h_1

and on the opposite side of $\sqrt{a}$ from g_2.

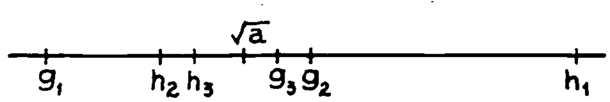


Furthermore, since g_2 is the midpoint of the segment joining g_1 and h_1, we can see that the length of the segment joining g_2 and h_2 is less than half that of the larger segment. That is,

$$|h_2 - g_2| < \frac{1}{2}|h_1 - g_1|$$

Now the same process is repeated to find g_3 and h_3. That is,

$$g_3 = \frac{g_2 + h_2}{2}, \quad h_3 = \frac{a}{g_3}$$



and we have

$$|h_3 - g_3| < \frac{1}{2}|h_2 - g_2| < \frac{1}{4}|h_1 - g_1|$$

Now we see that the length of the interval joining g_n and h_n is decreased by at least half each time n is increased by 1. Hence this interval "shrinks to a point" as n increases without bound. This point to which it "shrinks" is $\sqrt{a}$, the only number contained in all these intervals. The inequalities

$$0 \leq |g_n - \sqrt{a}| < |g_n - h_n| < \frac{1}{2^{n-1}}|g_1 - h_1|$$

show that the difference between g_n and $\sqrt{a}$ can be made as small as we like by choosing n large enough.

In light of this, we say that the sequence

$$g_1, g_2, g_3, g_4, \ldots$$

converges to $\sqrt{a}$. Also, the sequence

$$h_1, h_2, h_3, h_4, \ldots$$

converges to $\sqrt{a}$.

The trace of the flow chart of Figure 1-3 suggests that the terms of the sequence of g's converges to (or zeros in on) $\sqrt{a}$ much faster than by simply reducing the error by half at each step. This can be explained by means of the following calculation, where the first two equalities follow from the definitions of g_{n+1} and h_n and the remaining two from algebraic simplification:

$$g_{n+1} - \sqrt{a} = \frac{g_n + h_n}{2} - \sqrt{a} = \frac{g_n + \frac{a}{g_n}}{2} - \sqrt{a}$$

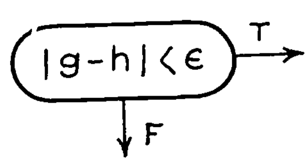
$$= \frac{g_n{}^2 - 2g_n \sqrt{a} + a}{2g_n} = \frac{(g_n - \sqrt{a})^2}{2g_n}$$

Thus, if g_n is such that the "error" $|g_n - \sqrt{a}|$ is less than 1/100, then $(g_n - \sqrt{a})^2$ is less than 1/10,000 so that the error in the next approximation $|g_{n+1} - \sqrt{a}|$ will be $\frac{1}{20{,}000g_n}$, which is less than 1/20,000 if g_n is greater than 1. And for the following term,

$$|g_{n+2} - \sqrt{a}| < \frac{\left(\frac{1}{20{,}000}\right)^2}{2g_{n+1}} < \frac{1}{800{,}000{,}000}$$

This shows that once the error starts getting small, it diminishes very rapidly indeed.

We cannot leave the square root problem until we provide a stopping mechanism for the algorithm given in Figure 1-3. In order to provide for this, we will input a number, ε, which is our maximum tolerance of error. (The Greek letter ε, pronounced "epsilon," is traditionally used for this purpose.) With each new computation of g and h, we will make the test



When we get a true answer to this test, we will know that g lies within ε of $\sqrt{a}$ since $\sqrt{a}$ is between g and h. Then we can terminate the computing process or go back for new values of a. In theory, this value of ε can be as small as you like. In practice, however, taking finite word length and round-off error in account, we must take ε sufficiently large so that the test will eventually be satisfied, thus avoiding an endless loop. How small ε may be safely taken will depend on the word length characteristics of your machine and/or your programming language as well as the size of the numbers a whose roots are being calculated. In other words, there is some minimum accuracy ε_1 beyond which we cannot hope to go, and we must take $\varepsilon \geq \varepsilon_1$.

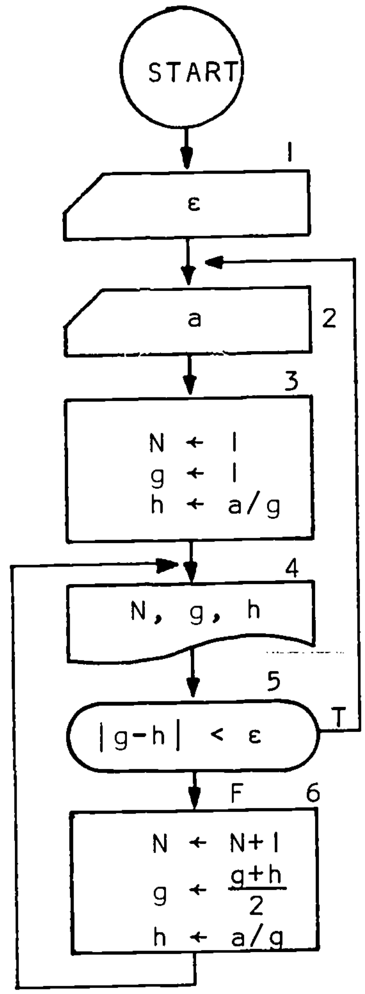


FIGURE 1-4

PROBLEMS

1. Write a program implementing the flow chart in Figure 1-4, and have the program approximate the square root of each of the following numbers: 1, 2, 20, .0002, 10^{10}. Read a value of 10^{-6} for ε.

2. Write a program to make a table of square roots. For each number, the program should obtain a sequence of approximations to the square root, but only the last approximation should be printed in each case. Your instructor will specify the size of your table.

2. Approximating Solutions of Equations

As you know, many mathematical problems may be reduced to the solving of equations. You know a formula for solving quadratic equations. Similar, but much more complicated, formulas exist for solving third and fourth degree equations. But for higher degree equations, no similar formulas can be found. For such equations and for equations of a nonalgebraic type like

$$\sin x = \frac{x}{2}$$

we will generally have to be satisfied with approximate solutions. Again we look for sequences of approximations by means of which we can get as near as we like to the true solution.

The method we present here for finding such sequences is very simple and at the same time one of the best for computer use. We first write our equation in the form

$$f(x) = 0.$$

For example, the equations

$$8x^3 = 6x + 1 \qquad \text{and} \qquad \sin x = \frac{x}{2}$$

would be expressed in the form

$$8x^3 - 6x - 1 = 0 \qquad \text{and} \qquad \sin x - \frac{x}{2} = 0.$$

Then the problem becomes that of finding a zero of the function f, that is a value of x for which $f(x) = 0$.

Graphically, this is the point at which the graph of the function, f, crosses (or touches) the X-axis.

Now if we have an interval $[L_1, R_1]$ such that the functional values at the end points, $f(L_1)$ and $f(R_1)$, have opposite signs, then the graph of the function must cross the X-axis somewhere between L_1 and R_1 (provided that the graph of f is an unbroken curve).

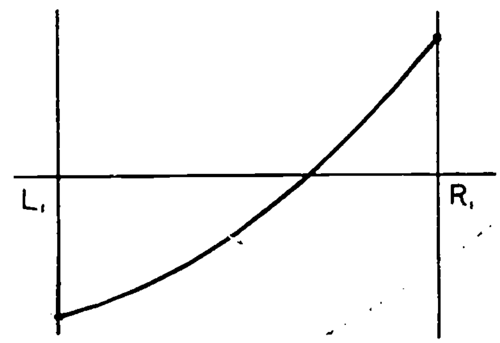


FIGURE 2-1

We look at the midpoint M_1 of the interval $[L_1, R_1]$

$$M_1 = \frac{L_1 + R_1}{2}$$

and consider three cases.

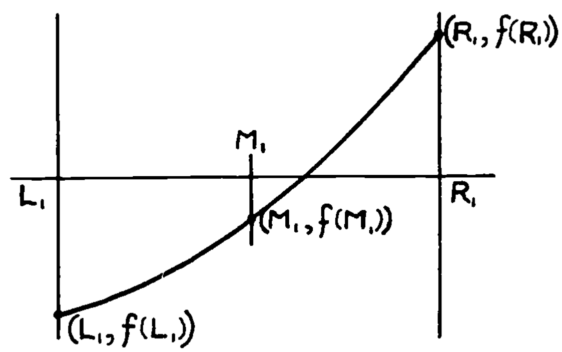


FIGURE 2-2

Case I. If $f(M_1) = 0$, then M_1 is a root.

Case II. If the sign of $f(M_1)$ is opposite to that of $f(L_1)$, then f will have a zero between L_1 and M_1. In this case, we let

$L_2 = L_1$ and $R_2 = M_1$

and repeat the process, finding M_2, etc.

<u>Case III.</u> If the sign of $f(M_1)$ is the same as that of $f(L_1)$, hence opposite to that of $f(R_1)$, then there is a zero between M_1 and R_1. In this case, we let

$$L_2 = M_1 \quad \text{and} \quad R_2 = R_1$$

In this way (unless we actually find a zero), the sequences

$$L_1, L_2, L_3, L_4, \ldots$$

and

$$R_1, R_2, R_3, R_4, \ldots$$

are constructed. A zero of the function is always located between L_n and R_n and moreover the length

$$R_n - L_n$$

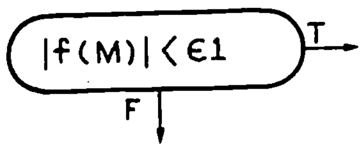
is reduced by half each time n is increased by 1. Thus, each of the numbers L_n and R_n can be made to differ from the zero by as little as we like. In other words, both of the sequences

$$L_1, L_2, L_3, \ldots \text{ and } R_1, R_2, R_3, \ldots$$

converge to the zero.

A flow chart for our algorithm is seen at the right of the next page. Again we see that the computer has no need of subscripted variables but instead prints out the successive values of L and R. The variable N merely numbers the lines of output.

The flow chart in Figure 2-3 is not quite ready to be converted into a computer program because we have not provided for a way to stop unless we actually hit a root. Also, because of round-off and finite word length, we should modify the test in box 7 to read



The final flow chart ready for translation into a computer program is shown in Figure 2-4. As in the previous example, we must have $\varepsilon \geq \varepsilon_1$. In most applications of this algorithm, one is interested only in the final value of M and not in the number of steps N or the intermediate values of L and R. In this case, the boxes N←1, N←N+1, and N,L,R can be eliminated.

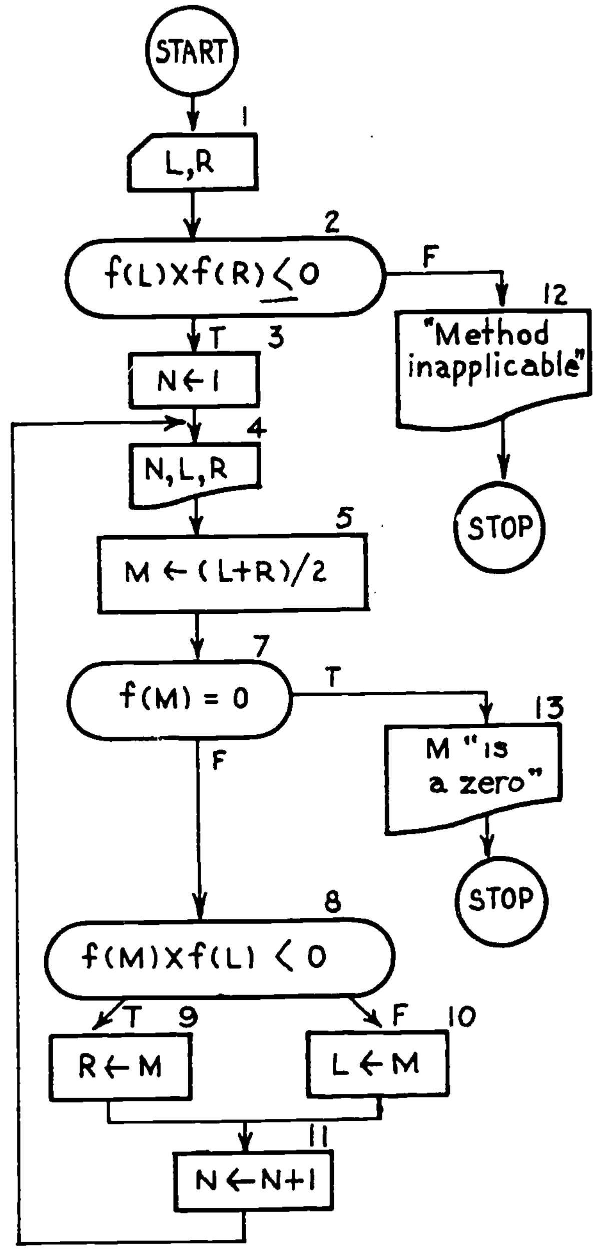


FIGURE 2-3

It is assumed that the input values of L and R are such that L < R. It is further assumed that f has opposite signs at L and R (i.e., that f(L)·f(R)<0).

It will be found that this sequence will not converge nearly as fast as the square root algorithm. Faster converging methods are possible for the problem at hand, but to use them it is necessary that the function satisfy extra conditions which in general are hard to verify. However, there are many important special cases where the faster methods are known to work and should be used. Such methods will be studied in a later chapter.

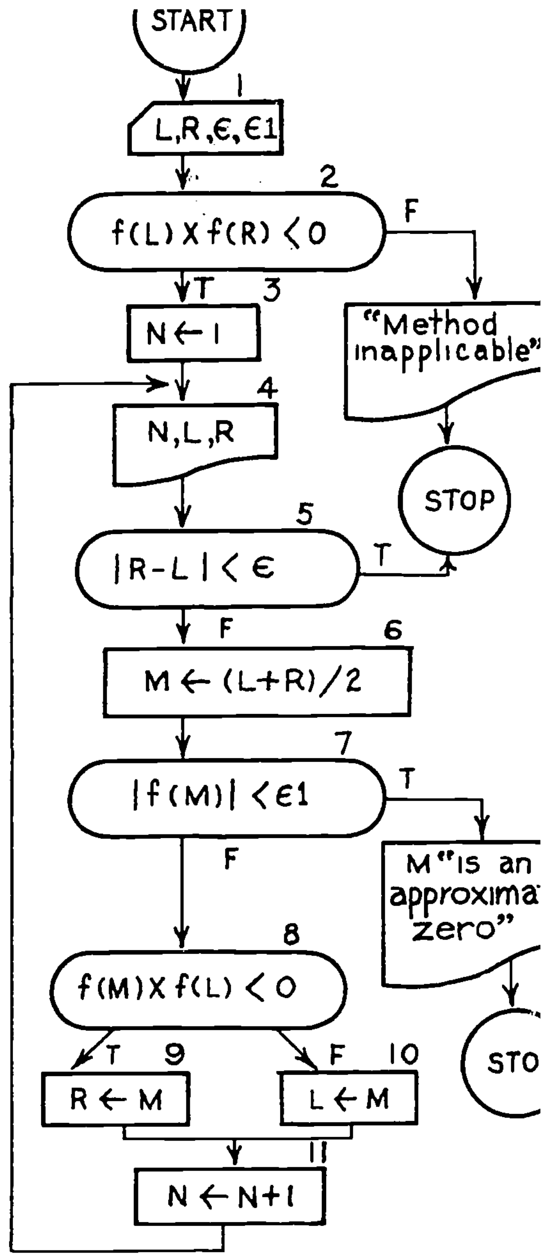


FIGURE 2-4

PROBLEMS

1. Draw a flow chart for a program which is to search for zeros of a function f in the following manner:

   (a) Read values for:

A	left end-point of an interval
B	right end-point of an interval
ε	error tolerance
N	a positive integer

   (b) For each integer K between 1 and N, apply the algorithm of this section to the sub-interval [L,R] of [A,B], where

   $$L = A + (K-1)\frac{B-A}{N} \text{ and } R = A + K\frac{B-A}{N}$$

2. Write the program described in Problem 1, and run it with the following functions f and numbers A, B, ε, and N:

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

   $A = -10, \quad B = 10, \quad \varepsilon = .01, \quad N = 40$

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

   $A = -10, \quad B = 10, \quad \varepsilon = .01, \quad N = 2$

   (c) $f(x) = x^4 - x^2$

   $A = -5, \quad B = 5, \quad \varepsilon = .01, \quad N = 20$

3. Use the program written in Problem 2 to approximate

   (a) the positive root of $x^2 - 2 = 0$

   (b) the positive root of $x^2 - 30 = 0$

   (c) the negative root of $x^5 + x + 1 = 0$

   In each case, let N=1 and choose appropriate values of A and B.

3. Problem Solution Using Sequences

Sequences also arise in other ways than in approximating numbers. We give here an illustration.

Example: A boy has a super-ball which when dropped will bounce back to $\frac{7}{8}$ of its original height. If the ball is dropped from a height of 5 feet and allowed to continue to bounce, what will be the total up and down distance it travels?

Solution: It is clear that the total up and down distance is 5 feet less than twice the total distance the ball travels downward. If we let d_n represent the total distance the ball falls before reaching the ground for the n^{th} time, we have, ignoring the diameter of the ball,

$$d_1 = 5$$
$$d_2 = 5 + 5\left(\frac{7}{8}\right)$$
$$d_3 = 5 + 5\left(\frac{7}{8}\right) + 5\left(\frac{7}{8}\right)^2$$
$$d_4 = 5 + 5\left(\frac{7}{8}\right) + 5\left(\frac{7}{8}\right)^2 + 5\left(\frac{7}{8}\right)^3$$
$$\vdots$$
$$d_n = 5 + 5\left(\frac{7}{8}\right) + 5\left(\frac{7}{8}\right)^2 + 5\left(\frac{7}{8}\right)^3 + \ldots + 5\left(\frac{7}{8}\right)^{n-1}$$

We have thus constructed a sequence. We will show that this sequence converges to a certain number which we will take as total distance the ball falls in infinitely many bounces. There may be some question as to whether the ball actually bounces infinitely many times. In the mathematical model of the problems we have selected, it is convenient to adopt the attitude that it does bounce infinitely many times. This gives answers close to actual experience. (The total distance the ball travels after the 100^{th} bounce is negligible for all practical purposes.)

In order to find the number to which our sequence converges, we express the term d_n in a different way. We illustrate this with d_6

$$d_6 = 5 + 5\left(\frac{7}{8}\right) + 5\left(\frac{7}{8}\right)^2 + 5\left(\frac{7}{8}\right)^3 + 5\left(\frac{7}{8}\right)^4 + 5\left(\frac{7}{8}\right)^5$$

$$\frac{7}{8}d_6 = 5\left(\frac{7}{8}\right) + 5\left(\frac{7}{8}\right)^2 + 5\left(\frac{7}{8}\right)^3 + 5\left(\frac{7}{8}\right)^4 + 5\left(\frac{7}{8}\right)^5 + 5\left(\frac{7}{8}\right)^6$$

Subtracting,

$$\frac{1}{8}d_6 = 5 + 0 + 0 + 0 + 0 + 0 - 5\left(\frac{7}{8}\right)^6$$

Thus,

$$d_6 = 8\left[5 - 5\left(\frac{7}{8}\right)^6\right]$$

Similarly,

$$d_n = 8\left[5 - 5\left(\frac{7}{8}\right)^n\right]$$

Now we can see what happens to d_n as n gets large; the value of $\left(\frac{7}{8}\right)^n$ gets closer and closer to zero so that the term $5\left(\frac{7}{8}\right)^n$

becomes insignificant. Thus, we see that the sequence $d_1, d_2, d_3, \ldots$ converges to the value of 40. The total up-and-down distance traveled by the ball will be $2(40) - 5 = 75$ feet.

(In this problem of the bouncing ball, we see an example of "mathematical modeling". Mathematics is not equipped to talk about nature directly; there is always a modeling process involved. In this case, we think of the position of the ball as being a point on a vertical line. You may if you wish think of this as the center of the ball. In this simplistic model, we ignore the difference in air resistance at different speeds, the deformation of the ball on hitting the ground, etc. Furthermore, we take a rule for the height of return of the ball observed in a certain range of heights and extend it to very small heights for which we are unable to make measurements. It is a moot question whether the ball actually bounces infinitely many times, but in our model we consider this to be the case. More sophisticated models of the bouncing ball are possible taking into consideration all of the phenomena mentioned above and others. But in any case, in using mathematics to describe real life occurrences some model is either tacitly or explicitly being used. The "correct" answers to such problems are considered to be those calculated by use of the model and not those obtained by performing the experiment and making measurements.)

PROBLEMS

1. Consider the sequence

$$d_1 = 1$$
$$d_2 = 1 + r$$
$$d_3 = 1 + r + r^2$$
$$d_4 = 1 + r + r^2 + r^3$$
$$\vdots$$
$$d_n = 1 + r + r^2 + r^3 + \ldots + r^{n-1}$$

(a) Compute $rd_4 - d_4$.

(b) Compute a new expression for d_4 by dividing your result in part (a) by $r-1$. (Assume $r \neq 1$.)

(c) Repeat steps (a) and (b) with n in place of 4.

(d) To what value does the sequence d_n converge if $|r|<1$?

(e) Can you determine whether or not the sequence d_n converges if $r \geq 1$?

2. Repeat Problem 1 with the sequence indicated below, where k is some number.

$d_1 = k$

$d_2 = k + kr$

$d_3 = k + kr + kr^2$

$d_4 = k + kr + kr^2 + kr^3$

$\vdots$

$d_n = k + kr + kr^2 + kr^3 + \ldots + kr^{n-1}$

3. For each sequence indicated, compute the number to which the sequence converges.

(a) $d_n = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots + \frac{1}{2^{n-1}}$

(b) $d_n = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots + \frac{1}{2^n}$

(c) $d_n = 10 + 2 + \frac{2}{5} + \frac{2}{25} + \ldots + \frac{2}{5^{n-2}}$

(d) $d_n = 7 - \frac{14}{3} + \frac{28}{9} - \frac{56}{27} + \ldots + 7 \cdot (-\frac{2}{3})^{n-1}$

4. (a) Show that the repeating decimal 0.232323... (that is, $\frac{2}{10} + \frac{3}{100} + \frac{2}{1000} + \frac{3}{10000} + \ldots$) is equal to $\frac{23}{99}$.

(b) Show that the repeating octal number 0.451451451... (that is, $\frac{4}{8} + \frac{5}{8^2} + \frac{1}{8^3} + \frac{4}{8^4} + \frac{5}{8^5} + \frac{1}{8^6} + \ldots$) is equal to

$$\frac{4 \cdot 8^2 + 5 \cdot 8 + 1}{8^3 - 1}$$

(c) Show that in the base B number system (for any integer B>1), the expression $0.c_1c_2\ldots c_n d_1 d_2 \ldots d_m d_1 d_2 \ldots d_m \ldots$ (where $d_1 \ldots d_m$ keeps repeating) represents a rational number.

Hint: The value of the first n+km digits is

$$C + \frac{1}{B^n}\left(\frac{D}{1} + \frac{D}{B^m} + \frac{D}{B^m} + \ldots + \frac{D}{B^{(k-1)m}}\right)$$

where

$$C = \frac{c_1}{B} + \frac{c_2}{B^2} + \ldots + \frac{c_n}{B^n} \quad \text{and} \quad D = \frac{d_1}{B} + \frac{d_2}{B^2} + \ldots + \frac{d_m}{B^m}$$

5. Suppose that Katonah loses a war against Nashville and agrees to pay reparations in perpetuity as follows: 1000 knashes the first year; each year after the first, the amount to be paid is $\frac{1}{6}$ of the amount for the preceding year. Suppose that fractions of knashes are minted in such a manner that each year it will be possible for the exact debt to be paid. How much will be paid during eternity?

6. The half-life of a radio-active isotope is the time in which half of any given quantity will decay. For example, the half-life of Strontium 90 is about 25 years. Thus, half of any quantity of Strontium 90 will decay in 25 years, three-fourths will decay in 50 years, seven-eights will decay in 75 years. Indeed, for any postive integer n, the fraction which decays in 25n years

will be

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots + \frac{1}{2^n}$$

For each of the following elements, calculate the amount of the element that will have decayed and the amount that will remain after the indicated time.

(a) 1 oz. Strontium 90 (half-life 25 years); 1000 years

(b) 3 oz. Strontium 90; 1000 years

(c) 3 oz. Strontium 90; 3000 years

(d) 1 oz. Rubidium 87 (half-life $(1.2)(10^{11})$ years); $(1.2)(10^{11})$ years

(e) 16 tons Rubidium 89 (half-life 15 minutes); 24 hours

(f) 1 oz. Beryllium 8 (half-life 10^{-10} seconds); 1 second (You may estimate the answer to part (f).)

(g) 1 lb. Radium 226 (half-life 1600 years); 16,000 years

7. Suppose that a bull breaks half of the remaining dishes in a china shop every 20 seconds. If the shop originally has 64000 dishes, how quickly must the bull be routed in order to save 1000?

4. Definition of Convergence

We have seen enough of the usefulness of sequences to warrant their systematic study. First we give a formal definition of a sequence.

> Definition. A sequence is a function whose domain is the set of positive integers. (We have not said what the range of the function must be. This is in fact quite arbitrary. But for most of our sequences in the present chapter, the range will be a set of real numbers.)

Thus in our sequence, a, of approximations for the square root of 2

n	a(n)
1	1
2	3/2
3	17/12
4	577/408
.	.
.	.
.	.

we have a(1) = 1, a(2) = 3/2, a(3) = 17/12, etc. However, in the case of sequences it is customary to depart from the ordinary function notation and use subscripts, writing

$$a_1 = 1,\ a_2 = \frac{3}{2},\ a_3 = \frac{17}{12},\ a_4 = \frac{577}{408},\ \text{etc.}$$

Sequences are usually defined by means of some formula, such as

$$b_n = \frac{3n^2 + 4}{2n^2 + 3} \qquad n = 1, 2, 3, \ldots$$

or by means of a recurrence relation by means of which terms of a sequence are defined in terms of their predecessors. An example of this is the above sequence for the square root of 2 in which the terms are defined by

$$a_1 = 1$$

$$a_n = \frac{a_{n-1} + \frac{2}{a_{n-1}}}{2} \qquad n = 2, 3, \ldots$$

Not all sequences converge. For example, the terms of the sequence

$$c_n = n^2 \qquad n = 1, 2, 3, \ldots$$

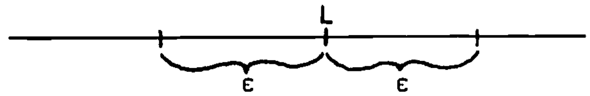
"go shooting off to infinity" while the terms of the sequence

$$d_n = (-1)^n \qquad n = 1, 2, 3, \ldots$$

oscillate between -1 and 1 and do not "zero in" on one particular value.

To be precise, it will be necessary to pin down the concept of convergence by giving a definition of it. In our earlier work on sequences of approximations, we said that a sequence converged to a number L provided that the sequence could be used to approximate L to any desired degree of accuracy. This is intuitive but rather vague. We can, however, give this statement a precise meaning. Before we talk about approximating to any desired degree of accuracy, let's talk about approximating to a given fixed degree of accuracy.

When we say that a sequence a_n, $n = 1,2...$, approximates a number L with accuracy ε, we mean that if we start computing terms of the sequence, we eventually get to the point where all the remaining terms will differ from L by less than ε. This condition can also be phrased geometrically, to wit: Consider an interval centered at L with radius ε (the radius of an interval is half its length). Now after n reaches a certain value, then for all higher values of n, the terms a_n will lie in this interval.



As an example, we will show that the sequence

$$a_n = \frac{n^2 + 1}{n^2} \qquad n = 1, 2, 3, \ldots$$

approximates the number 1 with accuracy 1/100. First of all, note that all terms of this sequence are greater than 1 since the numerator is greater than the denominator. On the other hand, by writing a_n in the form

$$a_n = 1 + \frac{1}{n^2}$$

we see that as the value of n is increased, the value of a_n decreases. Thus for all n greater than 10, we have

$$1 < a_n < a_{10} = 1.01$$

Therefore, after the 10^{th} term all succeeding terms will lie in an interval of radius 1/100 centered at 1. And this means that the sequence a_n, n = 1, 2, 3, ..., approximates 1 with an accuracy 1/100 according to our definition.

We must point out that there are many other numbers which this sequence approximates with this accuracy. For example, for all n greater than 10, we have seen that a_n lies between 1 and 1.01 and hence all these a_n lie between 1.005 - .01 and 1.005 + .01 (that is, between .995 and 1.015). Hence, this sequence approximates 1.005 with accuracy 1/100.

Similarly, the above sequence approximates 1 with accuracy 1/1,000,000 since for n > 1000, we have

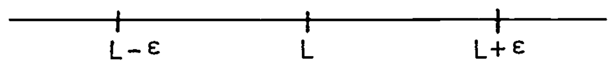
$$1 < a_n < a_{1000} = 1 + \frac{1}{1,000,000}$$

so that all terms after the 1000^{th} term lie within 1/1,000,000 of 1. (Note that this sequence does not approximate 1.005 with accuracy 1/1,000,000.)

Now we are ready to give a definition of convergence.

<u>Definition</u>. A sequence a_n, n = 1, 2, ..., <u>converges to</u> L provided that for every positive number ε, the sequence approximates L with accuracy ε.

This is what we meant earlier by our vague talk of a sequence approximating a number to any desired degree of accuracy. If this definition is interpreted geometrically, it says that for whatever positive ε we choose, we eventually reach the point where all further terms of the sequence lie in the interval centered at L with radius ε .



The above definition is generally worded in this somewhat more readily usable form:

<u>Definition</u>. A sequence a_n, $n = 1, 2, \ldots$, converges to L provided that for every positive number ε there is a number N so that for all values of n greater than N, a_n differs from L by less than ε.

We can easily show that the sequence

$$a_n = \frac{n^2+1}{n^2}, \quad n = 1, 2, 3, \ldots$$

actually converges to 1. Given an arbitrary positive number ε , we need only exhibit a positive integer N such that

$$\left| \frac{n^2+1}{n^2} - 1 \right| < \varepsilon$$

for all $n>N$. Now

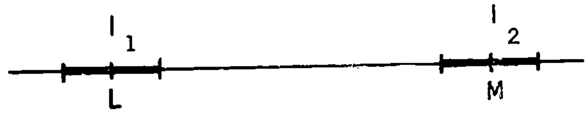
$$\frac{n^2+1}{n^2} - 1 = \frac{1}{n^2}$$

and $\frac{1}{n^2} < \varepsilon$ if $n^2 > \frac{1}{\varepsilon}$, that is if $n > \frac{1}{\sqrt{\varepsilon}}$. Thus, N may be taken to be any integer greater than $\frac{1}{\sqrt{\varepsilon}}$.

This shows that the sequence a_n, n = 1, 2, ..., approximates 1 with accuracy ε. And since ε was quite arbitrary except for being positive, we see therefore that for every positive number ε, the sequence a_n, n = 1, 2, 3, ..., approximates 1 with accuracy ε. In other words, the sequence converges to 1.

There are several theorems we should like to prove about convergence. The first of these is that a sequence cannot converge to two different numbers. For suppose $a_1, a_2, a_3, \ldots$ is a sequence and that L and M are two different numbers.

Choose small intervals I_1 and I_2 centered at L and M which do not intersect.



If the sequence converges to L, then after a certain point all the terms in the sequence will lie in the interval I_1 so that none of these terms will lie in I_2. Thus, the sequence cannot converge to M.

Hence, we see that the number to which a sequence converges (if the sequence converges at all) is unique. This number is called the limit of the sequence.

<u>Definition</u>. We denote by

$$\lim_{n\to\infty} a_n$$

the number to which the sequence $a_1, a_2, a_3, \ldots$ converges. If the sequence does not converge, we say that $\lim_{n\to\infty} a_n$ does not exist.

Thus, according to the previous example, where

$$a_n = \frac{n^2 + 1}{n^2}, \quad n, = 1, 2, 3, \ldots$$

we can say that $\lim_{n\to\infty} a_n = 1$ or, replacing a_n by $\frac{n^2 + 1}{n^2}$, we may write directly

$$\lim_{n\to\infty} \frac{n^2 + 1}{n^2} = 1.$$

You will be asked in the problems to prove that if r_n is a convergent sequence and if $r_n \geq a$ for all n, then $\lim_{n\to\infty} r_n \geq a$. It follows that if r_n is a convergent sequence and $r_n > a$ for all n, then $\lim_{n\to\infty} r_n \geq a$, but it need not be true that $\lim_{n\to\infty} r_n > a$. For example, $\frac{n^2+1}{n^2} > 1$ for all n, but we have just seen that $\lim_{n\to\infty} \frac{n^2+1}{n^2} = 1$. Similar statements can be made with $\leq$ and $<$ in place of $\geq$ and $>$.

It will also be a problem to prove that if x_n and y_n are convergent sequences and $x_n \geq y_n$ for all n, then $\lim_{n\to\infty} x_n \geq \lim_{n\to\infty} y_n$. The corresponding statement for $\leq$ is true, but the corresponding statements for $>$ and $<$ are subject to the reservations of the previous paragraph. Be certain to notice the hypothesis that x_n and y_n are convergent sequences. For example, $5+(-1)^n \geq \frac{n^2+1}{n^2}$ for all n, and $\lim_{n\to\infty} \frac{n^2+1}{n^2} = 1$, but it is not correct to conclude that $\lim_{n\to\infty} [5+(-1)^n] \geq 1$, because $\lim_{n\to\infty} [5+(-1)^n]$ does not even exist.

Another problem will be to prove that if $\lim_{n\to\infty} x_n = a$, then $\lim_{n\to\infty} |x_n| = |a|$. You should determine for yourself what can be said about the convergence of the sequence x_n if

$\lim_{n\to\infty} |x_n|=a.$

Example: An interesting example of a sequence defined by a recurrence relation is the Fibonacci sequence. In this sequence, the first two terms are 1 and each subsequent term is the sum of its two immediate predecessors. A few terms of the Fibonacci sequence are:

1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, ...

Thus, the terms a_n of the Fibonacci sequence are defined by the following conditions:

$$a_1 = 1$$
$$a_2 = 1$$
$$a_{n+1} = a_n + a_{n-1} \qquad n = 2, 3, \ldots$$

A flow chart for generating the terms of the Fibonacci sequence is shown in Figure 4-1. Here the variables ASUBN, ASNM1, and ASNP1 stand for a_n, a_{n-1}, and a_{n+1}, respectively. The use of MAX is to prevent

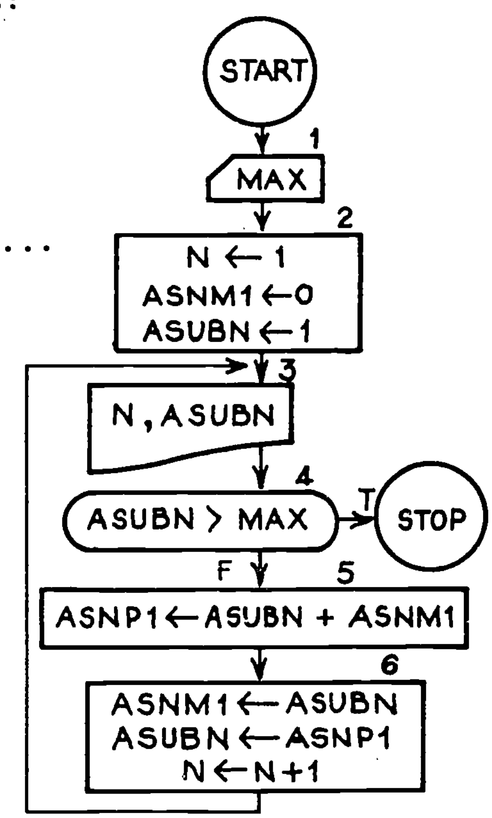


FIGURE 4-1

overflowing your word-length. The input value of MAX should not exceed half of the largest integer expressible in your machine.

The Fibonacci sequence obviously does not converge. The terms, increasing by at least 1 each time, eventually exceed any prescribed bound.

We call attention here once and for all to a principle which slightly generalizes the meaning of "sequence"; namely, if we alter the values of the first 53 terms of a sequence $a_1, a_2, a_3, \ldots$ this will not affect the convergence or divergence of the sequence nor the value to which it converges if it does converge. To see this, we need only require that the value of N in the definition be always taken greater than 53. Thus we may always be oblivious, in questions of convergence, to what happens to the first 53 terms, or even to whether they are properly defined. Of course, this principle is valid not only for 53 terms but also for 279 or 8967 or for any fixed finite number of terms.

For example we might define $a_n = \frac{1}{(n-3)(n-5)}$. Clearly this is meaningless for $n = 3$ and $n = 5$. We blithely ignore this fact in discussing the convergence of the sequence since there are only finitely many integers for which a_n is undefined. In light of these remarks one may, if he wishes, adopt a modified definition of sequence as a function whose domain is the set of all positive integers greater than some arbitrary number.

PROBLEMS

1. For each of the sequences indicated below, determine whether the sequence converges, and if it does, find the limit.

(a) $a_n = \frac{(-1)^n}{n}$

(b) $b_n = \frac{1}{\sqrt{n^2+1}}$ Hint: $\sqrt{n^2+1} > n$, so $\frac{1}{\sqrt{n^2+1}} < \frac{1}{n}$

(c) $c_n = \frac{1}{\sqrt{n^2-1}}$ if $n > 1$, and $c_1 = 14$.

(d) $d_n = \cos n\pi$

(e) $e_n = \cos 2n\pi$

(f) $f_n = \sqrt{4-\cos 2n\pi}$

(g) $g_n = 5$

2. Let $d_n = (-1)^n$. Answer each of the following questions and justify your answer.

(a) Does the sequence d_n approximate 2 with accuracy 10?

(b) Does the sequence d_n approximate -1 with accuracy $2 + \frac{1}{1,000,000}$?

(c) Does the sequence d_n approximate 1 with accuracy $\frac{5}{4}$?

(d) Does the sequence d_n approximate 0 with accuracy $\frac{5}{4}$?

(e) Does d_n converge?

3. The decimal expansion of $\frac{1}{7}$ is 0.142857142857142857....
Let a_n be the sequence of numbers obtained by keeping the first n digits of the expansion. That is,

$a_1 = 0.1$

$a_2 = 0.14$

$a_3 = 0.142$

etc.

(a) Does this sequence approximate 0.14 with accuracy 0.003?

(b) Does this sequence approximate $\frac{1}{7}$ with accuracy 0.003?

(c) Does this sequence converge to $\frac{1}{7}$?

4. Prove that each of the following sequences fails to converge.

(a) $a_n = n^2$

(b) $b_n = (-1)^n$

5. Affirm or deny each of the following statements and justify your reply.

(a) If $r_n \geq a$ for all n, and if r_n converges, then $\lim_{n\to\infty} r_n \geq a$.

(b) If $r_n > a$ for all n, and if r_n converges, then $\lim_{n\to\infty} r_n > a$.

(c) If r_n is a rational number for all n, and if r_n converges, then $\lim_{n\to\infty} r_n$ is a rational number.

(d) If $x_n \geq y_n$ for all n, and if both x_n and y_n converge, then $\lim_{n\to\infty} x_n \geq \lim_{n\to\infty} y_n$.

(e) If $x_n \geq y_n$ for all n, and if y_n converges, then x_n converges.

(f) If x_n and y_n are sequences which converge to the same limit m, then $\lim_{n\to\infty} z_n = m$, where z_n is the sequence $x_1, y_1, x_2, y_2, x_3, y_3, \ldots$

(g) If $\lim_{n\to\infty} r_n = m$, and if s_n is defined by $s_n = r_{n+1}$ then $\lim_{n\to\infty} s_n = m$.

(h) If $\lim_{n\to\infty} x_n = a$, then $\lim_{n\to\infty} |x_n| = |a|$.

(i) If x_n is a sequence and the sequence $y_n = |x_n|$ converges, then the sequence x_n converges.

6. Let T be defined for all positive numbers x by $T(x) = \frac{x+2}{x+1}$.

(a) Show that $T(x) > \sqrt{2}$ if $x < \sqrt{2}$, and $T(x) < \sqrt{2}$ if $x > \sqrt{2}$.

(b) Show that $|T(x) - \sqrt{2}| < (\sqrt{2} - 1)\,|x-\sqrt{2}|$.

Hint: Show that $\left|\frac{T(x) - \sqrt{2}}{x - \sqrt{2}}\right| = \frac{T(x) - \sqrt{2}}{\sqrt{2} - x} = \frac{\sqrt{2} - 1}{x + 1}$.

(c) Let a_n be the sequence defined by $a_1 = 1$, and $a_{n+1} = T(a_n)$ for $n \geq 1$. Show that a_n is a sequence of rational numbers which converges to $\sqrt{2}$.

(d) Calculate a_2, a_3, a_4, and a_5.

(e) Write a program to approximate $\sqrt{2}$ by calculating terms of the sequence a_n until two successive terms differ by less than 10^{-6}. Have the program print the number N of terms calculated, and the average of the last two terms. Compare the rate of convergence with the rate of convergence obtained by the program written for Problem 1 of Section 2-1.

The "arithmetic-geometric" mean M of two positive numbers a and b is defined as follows: suppose that $a \leq b$ and define a recursion formula by $a_1 = a$, $b_1 = b$, $a_{n+1} = \sqrt{a_n b_n}$, $b_{n+1} = \frac{a_n + b_n}{2}$. We get two sequences a_n and b_n and it will be shown later that both sequences converge to the same limit M. Write the appropriate flow-chart and program for computing an ε-approximation to M, and carry out the computation for $\varepsilon = .000005$ and the following pairs of values of a and b.

$a = 1, \ 1, \ 5, \ .001, \ 1 \ , \ 10$

$b = 2, \ 10, \ 6, \ 1000, \ 10^6, \ 10$

Draw some conclusions from the last three cases and prove your conclusions if you can. Is this an efficient algorithm, in the sense of giving an accurate result in few steps? Try some of the above cases for the smallest ε that you can use on your computer and see how many more steps are needed.

8. Modify the Fibonacci sequence flow chart so as to output a running total of the first n terms of the Fibonacci sequence, i.e.,

$$s_n = a_1 + a_2 + \ldots + a_n$$

Your output box should have the form

N, ASUBN, SSUBN

Write the program for this flow chart and run it. Can you spot a simple relationship between terms of the sequence $s_1, s_2, s_3, \ldots$ and terms of the sequence $a_1, a_2, a_3, \ldots$? Can you prove that this relationship always holds true? Hint: You may want to use mathematical induction.

9. Check for convergence:

(a) $47, 183, -10^{10}, 62.5, 1/2, 1/3, 1/4, \ldots, \frac{1}{n-3}, \ldots$

(b) $1, \frac{1}{2}, -\frac{1}{4}, 2, 1/8, -1/16, 3, 1/32, -1/64, 4, 1/128, \ldots$

5. The Simplest Limit Theorems

In the preceding section (Section 4), we had shown that

$$\lim_{n\to\infty} \frac{n^2 + 1}{n^2} = 1 .$$

We did this by means of the definition of convergence using ε and N, and it was a relatively tedious process. When the rule of formation of the terms of a sequence is fairly complicated, the ε and N process becomes positively painful.

Fortunately, there are theorems available which help us to avoid such unpleasant calculations. Among such theorems (as we will presently prove) are the sum and product theorems for limits:

$$\text{if } \lim_{n\to\infty} a_n = A \text{ and } \lim_{n\to\infty} b_n = B$$

$$\text{then } \lim_{n\to\infty} (a_n + b_n) = A + B$$

$$\text{and } \lim_{n\to\infty} (a_n b_n) = AB$$

The conclusion of the Sum Theorem could be written as

$$\lim_{n\to\infty} (a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n$$

which tells us that we may interchange the order of adding and taking the limit. The theorem is often verbalized as "the limit of the sum is the sum of the limits".

Returning to the example above, we could write:

$$\frac{n^2+1}{n^2} = \frac{n^2}{n^2} + \frac{1}{n^2} = 1 + \frac{1}{n^2}$$

Applying the Sum Theorem, we have

$$\lim_{n\to\infty} \frac{n^2+1}{n^2} = \lim (1 + \frac{1}{n^2}) = \lim_{n\to\infty} 1 + \lim_{n\to\infty} \frac{1}{n^2}$$

Since it is quite vious that

$$\lim_{n\to\infty} 1 = 1 \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n^2} = 0,$$

we have

$$\lim_{n\to\infty} \frac{n^2+1}{n^2} = 1 + 0 = 1.$$

This example provides us with an excel reason for wanting to have the Sum Theorem and similar theorems for products and quotients, etc., at our disposal. These theorems often enable us to decompose complicated limits into combinations of limits so simple as to be obvious on inspection.

The sum and product theorems themselves should seem quite obvious, for if n is large enough so that a_n is very close to A and b_n is very close to B, then $a_n + b_n$ ought to be very close to A + B, and $a_n b_n$ ought to be very close to AB. We are glad that this simple-minded way of looking at things is at hand to lend credibility to these theorems. But proofs are nevertheless necessary fo wo reasons:

1) the theorems are not all that obvious; the reasoning having been offered in their support will leave the

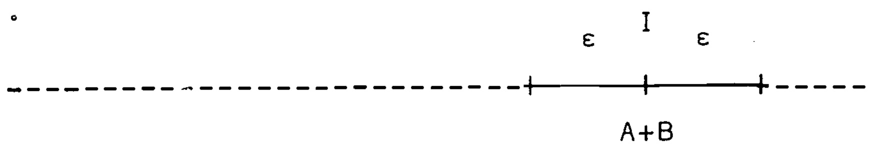
critical reader with an uneasy feeling in the pit of his stomach; and

2) we want to be sure that these theorems actually follow from the definition of convergence on which foundation we propose to erect a lofty edifice.

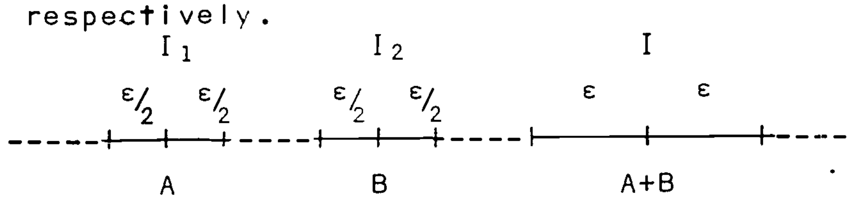
Now let's see how these proofs go.

<u>Theorem 1</u>. (Sum Theorem) If $\lim_{n\to\infty} a_n = A$ and $\lim_{n\to\infty} b_n = B$, then $\lim_{n\to\infty} (a_n + b_n) = A + B$.

<u>Proof:</u> Consider an interval I with center A+B and arbitrary radius ε.



Let I_1 and I_2 be intervals half as long as I, centered at A and B, respectively.



If for some value of n, a_n lies in I_1 and b_n in I_2, then we have

$$A - \varepsilon/2 < a_n < A + \varepsilon/2$$

$$B - \varepsilon/2 < b_n < B + \varepsilon/2$$

and adding

$$A + B - \varepsilon < a_n + b_n < A + B + \varepsilon$$

so that $a_n + b_n$ is in the interval I.

And since $\lim\limits_{n\to\infty} a_n = A$ and $\lim\limits_{n\to\infty} b_n = B$, we know that there are numbers N_1 and N_2 so that for $n > N_1$, a_n is in I_1 and for $n > N_2$, b_n is in I_2. Taking N to be the larger of N_1 and N_2, we see that when $n > N$, we have a_n in I_1 and b_n in I_2 so that, as seen above, $a_n + b_n$ is in I. Thus, the requirements of our definition are met and we can conclude that

$$\lim_{n\to\infty} (a_n + b_n) = A + B .$$

The next two theorems and their proofs are quite simple. The proofs will be left as exercises for the student.

Theorem 2. $\lim\limits_{n\to\infty} k = k$

Theorem 3. If $\lim\limits_{n\to\infty} a_n = A$, then $\lim\limits_{n\to\infty} (Ka_n) = KA$

Corollary 1. If $\lim\limits_{n\to\infty} a_n = A$ and $\lim\limits_{n\to\infty} b_n = B$, then $\lim\limits_{n\to\infty} (a_n - b_n) = A$

Proof: $\lim\limits_{n\to\infty} (-b_n) = \lim\limits_{n\to\infty} (-1)b_n = -1 \cdot B = -B$ by use of Theorem 3.

Thus, $\lim\limits_{n\to\infty} (a_n - b_n) = \lim\limits_{n\to\infty} (a_n + (-b_n)) = A + (-B) = A - B$ by Theorem 1.

Corollary 2. If $\lim\limits_{n\to\infty} a_n = A$, then $\lim\limits_{n\to\infty} (a_n - A) = 0$. The proof is left to the reader.

Corollary 3. If $\lim_{n\to\infty} a_n - A) = 0$, then $\lim_{n\to\infty} a_n = A$.

Again the proof is left to the reader.

PROBLEMS

1. Find two nonconvergent sequences whose sum converges.

2. Prove that if $\lim_{n\to\infty} x_n \geq a$ and $\lim_{n\to\infty} y_n \geq b$, then $\lim_{n\to\infty} (x_n + y_n) \geq a + b$.

3. Prove that if $\lim_{n\to\infty} x_n \geq a$ and $\lim_{n\to\infty} y_n \leq b$, then $\lim_{n\to\infty} (x_n - y_n) \geq a - b$.

4. Prove that if $\lim_{n\to\infty} x_n \geq a$ and $k \geq 0$, then $\lim_{n\to\infty} kx_n \geq ka$.

5. Prove that if $\lim_{n\to\infty} x_n \geq a$ and $k \leq 0$, then $\lim_{n\to\infty} kx_n \leq ka$.

6. Prove Theorem 2.

7. Prove Theorem 3.

8. Prove Corollary 2.

9. Prove Corollary 3.

10. Suppose a_n is a sequence which approximates A with accuracy ε and b_n is a sequence which approximates B with accuracy η. Prove that the sequence a_n+b_n approximates A+B with accuracy $\varepsilon+\eta$.

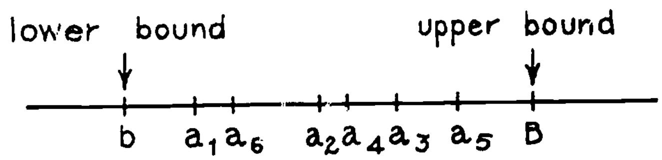
11. Suppose b_n is a sequence which approximates B with accuracy η. Prove that the sequence $-b_n$ approximates $-B$ with accuracy η.

12. Let a_n and b_n be as in Problem 10. Prove that the sequence $a_n - b_n$ approximates A-B with accuracy $\varepsilon+\eta$. Construct an example to show that $a_n - b_n$ need not approximate A-B with accuracy $\varepsilon-\eta$ even if $\varepsilon-\eta > 0$.

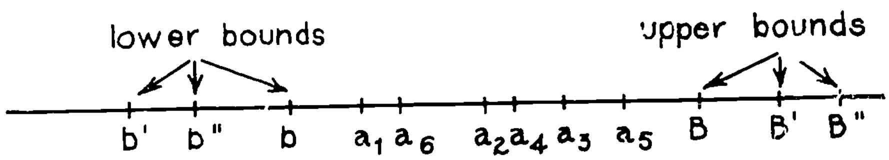
13. Let a_n be a sequence which approximates A with accuracy ε, and let k be some number. Prove that the sequence ka_n approximates kA with accuracy $|k|\varepsilon$.

6. Product and Quotient Theorems

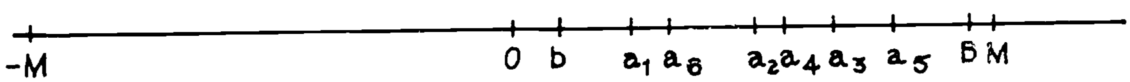
A sequence $a_1, a_2, a_3, \ldots$ is called <u>bounded</u> provided that all the terms of the sequence lie in some interval



The left and right end points of such an interval are called, respectively, lower and upper bounds for the sequence. It is clear that if we can find one such interval, we can find many.



Also it is quite clear that given a bounded sequence, one can find an interval centered at the origin in which all



the terms of the sequence lie. That is to say,

$$-M \leq a_n \leq M \qquad n = 1, 2, 3, \ldots$$

This can be expressed more concisely as

$$|a_n| \leq M \qquad n = 1, 2, 3, \ldots$$

We will use this form in our "official" definition of boundedness.

<u>Definition</u>. We say that a sequence $a_1, a_2, a_3, \ldots$ is <u>bounded</u> provided that there is a number M so that $|a_n| \leq M$ for all positive integers n. Such a number M is called "an upper bound for the absolute values of a_n."

We can see that the sequence

$$a_n = n \qquad n = 1, 2, 3, \ldots$$

is not bounded, for, no matter how large a number M may be chosen, we will be able to find n so that $a_r > M$.

On the other hand, we can see that the sequence

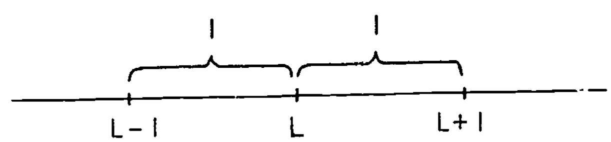
$$a_n = (-1)^n \qquad n = 1, 2, \ldots$$

(that is, the sequence whose terms have the values $-1, 1, -1, 1, -1, 1, \ldots$) is bounded since for this sequence $|a_n| = |(-1)^n| = 1$, so that the conditions of the definition are satisfied with M taken to be 1, or, for that matter, any number greater than 1.

This sequence, $a_n = (-1)^n$, $n = 1, 2, \ldots$, clearly fails to converge, and so it is evident that a bounded sequence does not necessarily converge. On the other hand, we have

<u>Theorem 5</u> If the sequence $a_1, a_2, a_3, \ldots$ converges, then it is bounded.

<u>Proof</u>: Denote the limit of the a_n by L. We can find a positive integer N so that for all values of n with $n \geq N$, we have a_n between L-1 and L+1.



Thus, for all $n \geq N$, we have

$$|a_n| \leq |L| + 1.$$

(This takes care of both the case that L is positive and the case that L is negative.) The number $|L|+1$ is now a candidate for the M in the definition. The only terms of the sequence whose absolute values could possibly exceed $|L|+1$ are

$$a_1, a_2, a_3, \ldots, a_{N-1}.$$

Since there are only finitely many of these, we can check them all out. We let M be the largest of the numbers

$$|L| + 1, \; |a_1|, \; |a_2|, \; |a_3|, \; \ldots, \; |a_{N-1}|$$

and then we will have $|a_n| \leq M$ for all positive integers n.

<u>Theorem 6</u>. If $\lim_{n\to\infty} a_n = 0$ and the sequence $b_1, b_2, b_3, \ldots$ is bounded, then the sequence

$$a_1 b_1, \; a_2 b_2, \; a_3 b_3, \; \ldots$$

converges to zero.

Note that we do not assume that the sequence $b_1, b_2, b_3, \ldots$ converges. Thus, for example, we could apply this theorem to prove that the sequence

$$\frac{1}{n} \sin n, \qquad n = 1, 2, 3, \ldots$$

converges since

$$\lim_{n\to\infty} \frac{1}{n} = 0$$

and sin n is bounded ($|\sin n| \leq 1$). This is true even though the sequence sin n, n = 1, 2, 3, ..., does not converge.

<u>Proof</u>: Le[illegible] 0. Let M be an upper bound for the absolute value of the b_n. Choose N so that for n > N, we have

$$-\frac{\varepsilon}{M} < a_n < \frac{\varepsilon}{M}$$

(in other words, so that $|a_n| < \frac{\varepsilon}{M}$).

Then for $n \geq N$, we have

$$|a_n| < \frac{\varepsilon}{M} \text{ and } |b_n| \leq M$$

so that

$$|a_n b_n| < (\frac{\varepsilon}{M})M = \varepsilon$$

or in other words, for $n \geq N$,

$$-\varepsilon < a_n b_n - 0 < \varepsilon ,$$

which is what is needed to show that the sequence $a_n b_n$, $n = 1, 2, 3, \ldots$, converges to 0.

Theorem 7. (Product Theorem) If $\lim_{n\to\infty} a_n = A$ and $\lim_{n\to\infty} b_n = B$, then $\lim_{n\to\infty} (a_n b_n)$ exists and is equal to AB.

Proof: Check that

$$a_n b_n = a_n B + a_n (b_n - B)$$

Since $b_n - B \to 0$ as $n \to \infty$ (by Corollary 2 of Theorem 3) and since the sequence a_n, $n = 1, 2, \ldots$, is bounded (by Theorem 5), we see by Theorem 6 that $\lim_{n\to\infty} a_n (b_n - B) = 0$.

Theorem 3 assures us that

$$\lim_{n\to\infty} a_n B = B \lim_{n\to\infty} a_n = BA$$

And now Theorem 1 (The Sum Theorem) assures us that $\lim_{n\to\infty} a_n b_n$ exists and is given by

$$\lim_{n\to\infty} a_n b_n = \lim_{n\to\infty} (a_n B) + \lim_{n\to\infty} a_n (b_n - B)$$

<u>Example</u>. Let x_n be the sequence $x_n = (4 - \frac{1}{n})(5 + \frac{1}{n})$, $n = 1, 2, 3, \ldots$. Then $\lim_{n\to\infty} x_n = \lim_{n\to\infty} (4 - \frac{1}{n}) \cdot \lim_{n\to\infty} (5 + \frac{1}{n})$ $= 4 \cdot 5 = 20$.

In order to prove the remaining two theorems, we need the following lemma whose proof we omit since it follows the same general lines as Theorem 5.

<u>Lemma</u>. If the sequence a_n, $n = 1, 2, \ldots$, converges to a number different from zero and has none of its terms equal to zero, then the sequence

$$\frac{1}{a_1}, \frac{1}{a_2}, \frac{1}{a_3}, \ldots$$

is bounded.

Theorem 8. (Reciprocal Theorem) If $a_1, a_2, \ldots$ converges to a number A different from zero and has none of its terms equal to zero then

$$\lim_{n\to\infty} \frac{1}{a_n} = \frac{1}{A}.$$

Proof: $\frac{1}{a_n} - \frac{1}{A} = \frac{A - a_n}{a_n A} = (A - a_n)(\frac{1}{a_n})(\frac{1}{A})$.

Since $A - a_n \to 0$ as $n\to\infty$ and since by our lemma $\frac{1}{a_n} \cdot \frac{1}{A}$ is bounded, then

$$\lim_{n\to\infty}\left(\frac{1}{a_n} - \frac{1}{A}\right) = 0$$

by Theorem 6.

Theorem 9. (Quotient Theorem) If $\lim_{n\to\infty} a_n = A$ and $\lim_{n\to\infty} b_n = B$ with $B \neq 0$ and none of the terms $b_n = 0$, then

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{A}{B}$$

Proof: $\lim_{n\to\infty} \frac{1}{b_n} = \frac{1}{B}$ by Theorem 8, and hence by the Product Theorem,

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty}\left(a_n \cdot \frac{1}{b_n}\right) = (A)\left(\frac{1}{B}\right) = \frac{A}{B}$$

Suppose that a_n is a sequence which converges to a number A. If f and g are functions such that

$$\lim_{n\to\infty} f(a_n) = f(A) \text{ and } \lim_{n\to\infty} g(a_n) = g(A) ,$$

then

$$\lim_{n\to\infty} (f(a_n) + g(a_n)) = f(A) + g(A) ,$$

and

$$\lim_{n\to\infty} (f(a_n)g(a_n)) = f(A)g(A).$$

If h is either a constant function or the identity function, we know that $\lim_{n\to\infty} h(a_n) = h(A)$. Thus if P is any polynomial function,

$$\lim_{n\to\infty} P(a_n) = P(A) .$$

Now if f and g are polynomial functions and R is the rational function $R(x) = \frac{f(x)}{g(x)}$, we have

$$\lim_{n\to\infty} R(a_n) = \frac{\lim_{n\to\infty} f(a_n)}{\lim_{n\to\infty} g(a_n)} = \frac{f(A)}{g(A)} = R(A),$$

providing that $g(A) \neq 0$ and $g(a_n) \neq 0$ for all n.

Example. Let c_n be the sequence defined by

$$c_n = \frac{(3 + \frac{1}{n})^2 + 1}{(3 + \frac{1}{n})^2 - 1}$$

Then

$$\lim_{n\to\infty} c_n = \frac{\lim_{n\to\infty} ((3 + \frac{1}{n})^2 + 1)}{\lim_{n\to\infty} ((3 + \frac{1}{n})^2 - 1)}$$

$$= \frac{\lim_{n\to\infty} (3 + \frac{1}{n})^2 + \lim_{n\to\infty} 1}{\lim_{n\to\infty} (3 + \frac{1}{n})^2 - \lim_{n\to\infty} 1}$$

$$= \frac{(\lim_{n\to\infty} (3 + \frac{1}{n}))^2 + 1}{(\lim_{n\to\infty} (3 + \frac{1}{n}))^2 - 1}$$

$$= \frac{3^2 + 1}{3^2 - 1} = \frac{10}{8}.$$

Example. Suppose $P(x) = 2x^3 - 4x^2 - 7x - 5$, and suppose a_n is a sequence which converges to 3. Then

$$\lim_{n\to\infty} (2(a_n)^3 - 4(a_n)^2 - 7a_n - 5) = 2(3)^3 - 4(3)^2 - 7(3) - 5$$

$$= 54 - 36 - 21 - 5 = -8$$

Example. Let a_n be any sequence which converges to 2 and for which $3(a_n)^2 - 4(a_n) - 2 \neq 0$ for any n. Then

$$\lim_{n\to\infty} \frac{(a_n)^2 + 1}{3(a_n)^2 - 4(a_n) - 2} = \frac{(2)^2 + 1}{3(2)^2 - 4(2) - 2} = \frac{5}{2}$$

Example: An interesting sequence can be obtained from the Fibonacci sequence. Let r_n be defined by the ratio

$$r_n = \frac{a_{n+1}}{a_n}$$

where a_n is the n^{th} term of the Fibonacci sequence. We will not now attempt to answer the question of whether the sequence $r_1, r_2, \ldots$ does converge, but we will show how to find the value it converges to if it does indeed converge. Using the definition of r_n and the recurrence relation for a_{n+1}, we have for $n \geq 2$, series converges absolutely. That is, $\sum a_n$ is absolutely convergent if $\sum |a_n|$ is convergent.

The alternating harmonic series is an example of a series that converges but does not converge absolutely. The following theorem shows that the reverse case is impossible.

Theorem 2. An absolutely convergent series is convergent.

Proof. Let $\sum |a_n|$ converge. Define two new series by

$$b_n = \begin{cases} a_n & \text{if } a_n \geq 0, \\ 0 & \text{if } a_n < 0, \end{cases} \qquad c_n = \begin{cases} 0 & \text{if } a_n \geq 0 \\ -a_n & \text{if } a_n < 0. \end{cases}$$

$b_n \geq 0$ and $b_n \leq |a_n|$, and hence, by Test 1, b_n converges. The same is true for $\sum c_n$. Hence $\sum (b_n - c_n) = \sum a_n$ converges. (Problem 1(a) of Section 2).

This theorem will sometimes tell us when a series containing negative terms converges but never when it diverges. For example, the series

$$1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} + \frac{1}{8} - \frac{1}{9} + \ldots$$

diverges, but none of our tests will give us this information. (See Problem 2).

From Theorem 2 and Test 1 we get another useful test for convergence.

__Ratio Test 1.__ If, for sufficiently large n,

$$a_n \neq 0 \text{ and } \left|\frac{a_{n+1}}{a_n}\right| \begin{Bmatrix} \leq M_1 < 1 \\ \geq M_2 > 1 \end{Bmatrix} \text{ then } \sum a_n \begin{Bmatrix} \text{converges} \\ \text{diverges} \end{Bmatrix}$$

__Proof.__ If for $n \geq N$ we have $\left|\frac{a_{n+1}}{a_n}\right| \leq M_1 < 1$ then $|a_n| \leq M_1|a_{n-1}| \leq M_1^2|a_{n-2}| \leq \cdots \leq M_1{}^{n-N}|a_N| = CM_1{}^n$, where $C = M_1{}^{-N}|a_N|$. Since $M_1 < 1$, $\sum CM_1{}^n$ converges and hence so does $\sum|a_n|$, and by Theorem 2 so does $\sum a_n$.

If for $n \geq N$ we have $\left|\frac{a_{n+1}}{a_n}\right| \geq M_2 > 1$ then, similarly, $|a_n| \geq CM_2{}^n$. Since $M_2 > 1$, a_n does not approach 0, and the series diverges by Theorem 1 of Section 2.

The following test is related to the one above in the same way that Comparison Test 3 is related to Comparison Test 2.

__Ratio Test 2.__ If $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = M$ then

$$\sum a_n \begin{Bmatrix} \text{converges} \\ \text{diverges} \end{Bmatrix} \text{ if } \begin{Bmatrix} M < 1 \\ M > 1 \end{Bmatrix}.$$

If M = 1 this test gives us no information.

Example 4, continued. Applying Ratio Test 2, we have

$$\lim_{n\to\infty} \left| \frac{(n+1)^5(.9)^{n+1}}{n^5(.9)^n} \right| = \lim_{n\to\infty} .9 \left(\frac{n+1}{n}\right)^5 = .9 \lim_{n\to\infty} (1 + \frac{1}{n})^5 = .9.$$

Since .9 < 1 the series converges.

Proof of Theorem 1. We need consider only the case of an increasing sequence. If the theorem is true for this case, and if $a_1, a_2, a_3, \ldots$ is a decreasing sequence with a lower bound B then $-a_1, -a_2, -a_3, \ldots$ is an increasing sequence with an upper bound -B and so has a limit L. Then $\lim_{n\to\infty} a_n = -L$.

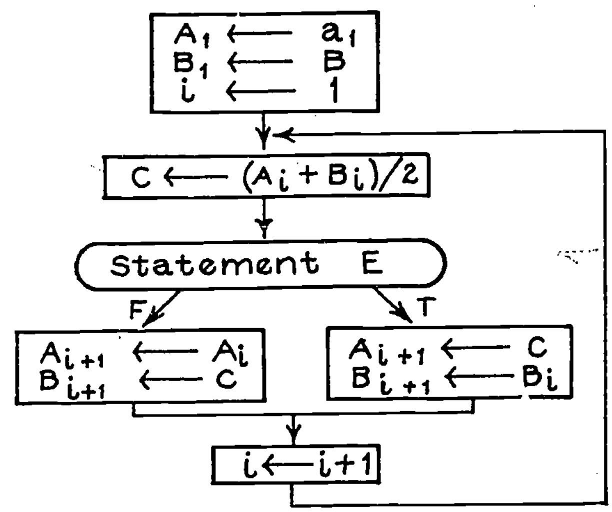


Figure 3-1

So let $a_1 \leq a_2 \leq a_3 \leq \ldots \leq a_n \leq \ldots \leq B$. The flow diagram in Figure 3-1 gives the

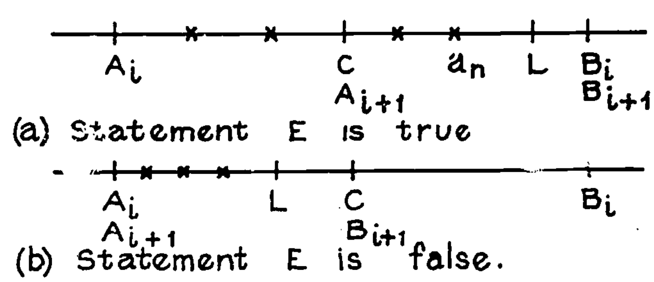


Figure 3-2

essense of the bisection process one follows to produce two sequences satisfying the conditions of the Completenes Axiom and determining $\lim_{n\to\infty} a_n$. Here Statement E is the following: There is an n for which $c \leq a_n \leq B_i$. The way this statement operates is shown in Figure 3-2. In each case the half of $[A_i, B_i]$ chosen for $[A_{i+1}, B_{i+1}]$ is the one containing the presumed limit L.

By the usual bisection argument the sequences $A_1, A_2, A_3, \ldots$ and $B_1, B_2, B_3, \ldots$ satisfy the conditions of the Completeness Axiom and so have a common limit L. We have only to show that L is the limit of $a_1, a_2, a_3, \ldots$.

Given $\varepsilon > 0$ we must find an N such that

$$|a_n - L| < \varepsilon \quad \text{whenever} \quad n > N.$$

Now the length of the interval $[A_i, B_i]$ is

$$B_i - A_i = (B_1 - A_1)/2^{i-1},$$

and so we can find an i for which $B_i - A_i < \varepsilon$. Since $B_1, B_2, \ldots$ is a decreasing sequence with limit L, $L \leq B_i$.

Similarly $L \geq A_i$, so L is in the interval $[A_i, B_i]$. By construction, there is an a_N in the interval $[A_i, B_i]$. If $n \geq N$ then also a_n is in $[A_i, B_i]$, for $a_n \geq a_N \geq A_i$ and $a_n \leq B_i$ since, by construction, each B_i is an upper bound of all the a_n. Figure 3-3 illustrates the relative positions of the various numbers.

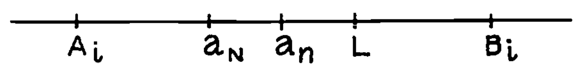


Figure 3-3

Since, for $n > N$, both a_n and L lie in an interval of length $< \varepsilon$ we have $|a_n - L| < \varepsilon$ as desired.

Like statement M in the proof of the existence of a maximum, in Section 6-7, Statement E is nonconstructive since it cannot, in general, be decided in a finite number of steps.

PROBLEMS

1. Prove Comparison Test 3.

2. Show that

$$1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} + \frac{1}{8} - \frac{1}{9} + \cdots$$

diverges. [Hint. Compare the partial sums $S_3, S_6, S_9, \ldots$ with the partial sums of the series $1 + \frac{1}{4} + \frac{1}{7} + \cdots$. Does this last series converge or diverge?]

3. Test the following series for absolute convergence, convergence, or divergence.

(a) $\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n^3}$

(b) $\sum_{n=1}^{\infty} (-1)^{n-1} \frac{n}{n+1}$

(c) $\sum_{n=1}^{\infty} (-1)^{n-1} \frac{n^2}{n^3+1}$

(d) $\sum_{k=3}^{\infty} (-1)^{k-1} \frac{1}{\sqrt[k]{2}}$

(e) $\sum_{k=1}^{\infty} (-1)^{k-1} \frac{k}{k^2 + 1}$

(f) $\sum_{k=0}^{\infty} (-1)^{k} \frac{1}{(k + 1)(k + 3)}$

(g) $\sum_{n=2}^{\infty} (-1)^{n} \frac{n^2}{e^n}$

(h) $\sum_{n=1}^{\infty} (-1)^{n-1} \left(\frac{1 + n}{1 + n^2}\right)^2$

(i) $\sum_{n=1}^{\infty} (-1)^{n-1} \frac{n}{3^n}$

(j) $\sum_{n=2}^{\infty} (-1)^{n} \frac{\log n}{n}$

(k) $\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{\sqrt{n}}$

(l) $\sum_{n=1}^{\infty} (-1)^{n-1} \frac{n^n}{(n + 1)!}$

4. We often think of a real number as an infinite decimal, e.g.

$$\pi = 3.14159265358979\ldots \quad .$$

In general, any positive number $A = N.a_1a_2a_3\ldots$, where N is a non-negative integer and each a_n has a value 0,1,2,...,8, or 9. What we mean by this is that

$$A = N + \frac{a_1}{10} + \frac{a_2}{10^2} + \frac{a_3}{10^3} + \cdots = N + \sum_{n=1}^{\infty} \frac{a_n}{10^n} .$$

(a) Prove that any such series converges to a value in $[N, N+1]$.

(b) What general statement can be made about the remainder after n terms; that is, about the error in <u>truncating</u> the number to n decimal places?

(c) To <u>round off</u> the number to n decimal places we add $5/10^{n+1}$ and truncate. What can be said about the remainder?

4. Infinite Series and Improper Integrals.

There is considerable similarity between infinite series

$$\sum_{k=1}^{\infty} a_k = \lim_{n\to\infty} \sum_{k=1}^{n} a_k$$

and improper integrals of the type

$$\int_a^{\infty} f(x)dx = \lim_{M\to\infty} \int_a^{M} f(x)dx.$$

To capitalize on this similarity we need the analog for functions of Theorem 1 of Section 3.

Theorem 1. A bounded monotone function on $[a,\infty)$, has a limit as $x \to \infty$.

The proof is the same as for the earlier theorem, with merely the substitution of x for n throughout. The same is true of Corollary 1 and the comparison tests.

Corollary 1. If $f(x) > 0$ for sufficiently large x then $\int_a^{\infty} f(x)dx$ converges if and only if $\int_a^{M} f(x)dx$ is bounded for all $M > a$.

We leave the statement and proofs of the comparison tests as an exercise.

<u>Example 1.</u> In Example 3 of Section 12-4 we encountered the integral

$$\int_0^\infty \frac{\sqrt[3]{t^8 + 4t^6 - 4t^5 - 4t^3 + 4t + 1}}{(1 + t^3)^2}\, dt$$

and made some vague statements about its convergence. We can now be good mathematicians and determine its convergence by comparing it with $\int_1^\infty \frac{1}{t^2}\, dt$, which converges. Using Comparison Test 3 we have

$$\lim_{t\to\infty} \left(\frac{\sqrt[3]{t^8 + \text{etc.}}}{(1 + t^3)^2}\right) \div \left(\frac{1}{t^2}\right)$$

$$= \lim_{t\to\infty} \frac{\sqrt[3]{t^{12} + 4t^{10} + \text{etc.}}}{(1 + t^3)^2}$$

$$= \lim_{t\to\infty} \frac{\sqrt[3]{1 + \frac{4}{t^2} - \frac{4}{t^3} - \text{etc.}}}{\left(\frac{1}{t^3} + 1\right)^3} = 3.$$

Hence the given integral converges.

Our main interest, however, is in the interaction of series and integrals, and for this the following theorem is fundamental.

Theorem 2. If f is a decreasing function then

$$(1) \qquad \sum_{k=n}^{m} f(k) - f(n) \le \int_n^m f(x)dx \le \sum_{k=n}^{m} f(k) - f(m)$$

and

$$(2) \qquad \int_n^m f(x)dx + f(m) \le \sum_{k=n}^{m} f(k) \le \int_n^m f(x)dx + f(n).$$

Proof. The inequalities

$$(3) \quad f(k+1) \le \int_k^{k+1} f(x)\,dx \le f(k)$$

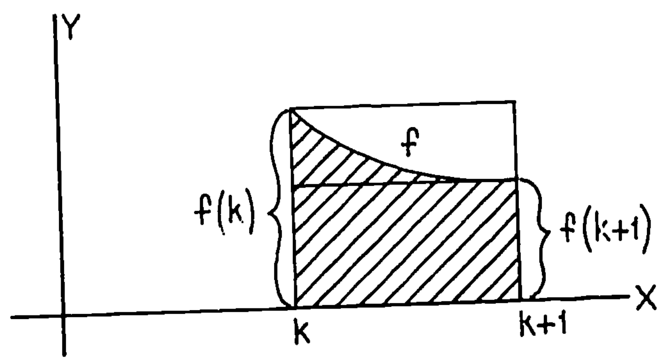


Figure 4-1

are obvious from Figure 4-1, since the middle term is the shaded area under the curve and the two bounds are the areas of the contained and the containing rectangles. In (3) let k have the successive values n, n + 1, ..., m - 1 and add the resulting inequalities. We get

$$\sum_{k=n+1}^{m} f(k) \le \int_n^m f(x)dx \le \sum_{k=n}^{m-1} f(k),$$

which is (1). To get (2) we solve the left-hand inequality and the right-hand inequality of (1) separately for $\sum_{k=n}^{m} f(k)$ and combine the two results as in (2).

Although this theorem is stated for any decreasing function our only interest, by virtue of Theorem 1 of Section 2, is in a decreasing function with limit 0. Such a function is necessarily positive.

Corollary 2. If for sufficiently large x, f(x) is a decreasing function with limit zero then $\int_a^\infty f(x)dx$ and $\sum f(k)$ either both converge or both diverge.

Proof. If $\sum f(k)$ converges then the rightmost sum of (1) remains bounded as $m \to \infty$. Hence $\int_n^m f(x)dx$ is bounded as $m \to \infty$ and by Corollary 1 the integral converges. If $\sum f(k)$ diverges then the leftmost sum of (1) is unbounded. Hence so is $\int_n^m f(x)dx$ and by Corollary 1 the integral diverges.

Example 2. The p-series. If $p \neq 1$,

$$\int_1^\infty x^{-p} dx = \lim_{m\to\infty} \frac{1}{1-p} x^{1-p}\Big|_1^m = \frac{1}{1-p} \lim_{m\to\infty} (m^{1-p} - 1).$$

The limit exists only if $1 - p < 0$, i.e. $p > 1$. For $p = 1$ we get $\lim_{m\to\infty} \log m$ which does not exist. Hence the p-series converges only for $p > 1$.

Example 3. We wish to compute $\sum_{k=1}^{\infty} k^{-2}$ to 5D accuracy. How many terms do we need? From (2) we have

$$\int_{n+1}^{m} x^{-2} dx + \frac{1}{m^2} \leq \sum_{k=n+1}^{m} k^{-2} \leq \int_{n+1}^{m} x^{-2} dx + \frac{1}{(n+1)^2},$$

or

$$\frac{1}{n+1} - \frac{1}{m} + \frac{1}{m^2} \leq \sum_{k=n+1}^{m} k^{-2} \leq \frac{1}{n+1} - \frac{1}{m} + \frac{1}{(n+1)^2}.$$

Now let $m \to \infty$. This gives

$$(4) \quad \frac{1}{n+1} \leq R_n \leq \frac{1}{n+1} + \frac{1}{(n+1)^2},$$

R_n being the remainder after n terms of $\sum_{k=1}^{\infty} k^{-2}$. This tells us that the remainder decreases quite slowly, like $\frac{1}{n+1}$. However, (4) also says that

$$-\frac{1}{2(n+1)^2} \le R_n - \frac{1}{n+1} - \frac{1}{2(n+1)^2} \le \frac{1}{2(n+1)^2} .$$

Hence if we add $\frac{1}{n+1} + \frac{1}{2(n+1)^2}$ to the n-th partial sum we get an approximation in error by at most $\frac{1}{2(n+1)^2}$. Using this, for 5D accuracy we need an n such that

$$\frac{1}{2(n+1)^2} < 5 \times 10^{-6},$$

which gives n = 233. In a sum of this magnitude roundoff error will not be serious but it must be allowed for. This will increase the value of n slightly.

Example 4. Although neither $\sum_{k=1}^{\infty} \frac{1}{k}$ nor $\int_1^{\infty} \frac{1}{x}dx$ converge, the difference of their "partial sums",

$$(5) \qquad S_n = \sum_{k=1}^{n} \frac{1}{k} - \log n$$

does converge as $n \to \infty$. This is evident from Figure 4-2.

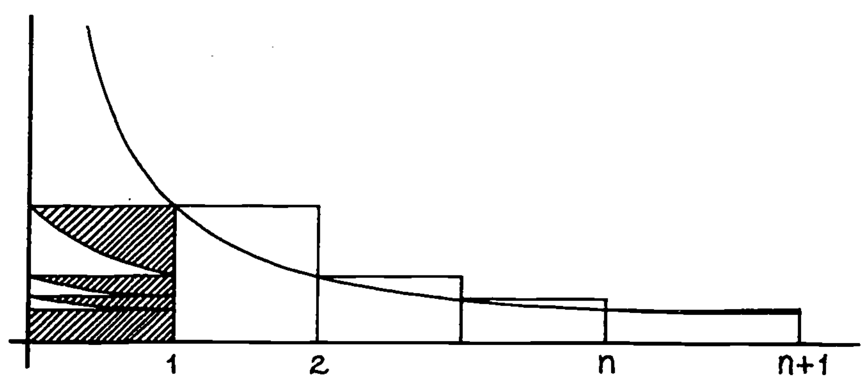


Figure 4-2

S_n is the area of the cross-hatched region, and as n increases this area decreases as pieces of the bottom rectangle are whittled away. Hence S_n, being a bounded decreasing sequence (bounded below by 0) has a limit. This limit, like π and e, crops up in a surprising number of places in mathematics. It is called <u>Euler's constant</u> and designated by the Greek letter γ (or sometimes by C or other symbols). Its value to 20D is

$$\gamma = 0.57721\ 56649\ 01532\ 86061.$$

Knowing γ, the best way to approximate $\sum_{k=1}^{n} \frac{1}{k}$ for large n is by replacing S_n in (5) by its limit γ. Thus

$$\sum_{k=1}^{100} \frac{1}{k} \approx \log 100 + \gamma \approx 5.18.$$

The error in such an approximation can be shown to be about 1/n.

This relation indicates the extremely slow growth of the partial sums of the harmonic series. To have $\sum_{k=1}^{N} \frac{1}{k} > 100$ we need, approximately, $\log N + \gamma > 100$, or $N > \exp(100 - \gamma)$. This is a very large number. As a computer exercise the smallest such value of N was computed exactly. It is

$$N = 1509\ 26886\ 22113\ 78832\ 36935\ 63264\ 53810\ 14498\ 59497.$$

PROBLEMS

1. Prove Comparison Test 1 for integrals: If $f(x) > 0$ and $g(x) > 0$ on $[a,\infty)$, and if $g(x)$ is unicon on $[a,M]$ for all $M > a$, then if

$$\left\{\begin{matrix} g(x) \le f(x) \\ \\ g(x) \ge f(x) \end{matrix}\right\} \text{ on } [a,\infty) \text{ and if } \int_a^\infty f(x)\,dx \left\{\begin{matrix} \text{converges} \\ \\ \text{diverges} \end{matrix}\right\},$$

so does $\int_a^\infty g(x)\,dx$.

2. State and prove Comparison Test 2 for integrals.

3. Determine, if possible, the convergence or divergence of the following integrals.

(a) $\int_1^\infty \frac{1}{x} e^{-x}\,dx$

(b) $\int_1^\infty \log x\, e^{-x}\,dx$

(c) $\int_1^\infty \sqrt{\frac{x+1}{x^2+1}}\,dx$

(d) $\int_0^\infty \frac{y}{\sqrt{y^3+1}}\,dy$

(e) $\int_{-\infty}^{\infty} e^{-z^2}\, dz$

(f) $\int_0^{\infty} \frac{1}{\sqrt{e^{2y}}}\, dy$

(g) $\int_1^{\infty} \frac{dx}{3x^2 - 1}$

(h) $\int_0^{\infty} \frac{dx}{\sqrt{8 + x}}$

(i) $\int_1^{\infty} \frac{dy}{y^{10}}$

(j) $\int_0^{\infty} \frac{dz}{2 + e^z}$

(k) $\int_1^{\infty} \frac{3x}{\sqrt{3x}}\, dx$

4. State and prove a comparison test for improper integrals of the type $\lim_{z \to a^+} \int_z^b f(x)dx$.

5. Show that $\int_0^1 x^{-p}dx$ converges for $p < 1$ and diverges for $p \geq 1$.

6. By comparison with integrals of the above type determine, if possible, the convergence or divergence of the following integrals.

(a) $\int_0^x \frac{e^t}{t}\,dt.$

(b) $\int_0^1 \frac{1}{\sqrt{t + t^2}}\,dt.$

(c) $\int_0^a \frac{1}{\sqrt{a^2 - x^2}}\,dx.$ [Hint. Let $u = a - x$.]

(d) $\int_0^1 \frac{1}{\sqrt{(1 - x^2)(1 - k^2x^2)}}\,dx, \quad k^2 < 1.$

(e) $\int_{0+}^1 \frac{\log x}{\sqrt{x}}\,dx.$

(f) $\int_0^{\pi/2} \tan 2x\,dx.$

7. Determine the convergence or divergence of the following series. Corollary 2 is not necessarily the best test to use.

(a) $\sum_{n=2}^{\infty} \frac{1}{n(\log n)^2}$

(b) $\sum_{n=1}^{\infty} \frac{n^3}{4^n}$

(c) $\sum_{n=1}^{\infty} 2 \sin \frac{\pi}{2^n}$

(d) $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n^2 + 2n}}$

(e) $\sum_{n=1}^{\infty} \log (3n + 1)$

(f) $\sum_{n=1}^{\infty} \log (1 + \frac{1}{n})$

(g) $\sum_{n=1}^{\infty} \arctan \frac{1}{n^2 + 1}$

(h) $\sum_{k=1}^{\infty} \frac{k + 2}{\log (k + 3)}$

(i) $\sum_{k=1}^{\infty} \frac{\sin k}{k^2}$.

8. In Section 11-8 we defined the gamma function, $\Gamma(x)$, by

$$\Gamma(x) = \int_0^{\infty} t^{x-1}e^{-t}dt.$$

(a) By comparison with $\sum n^{x-1}e^{-n}$ show that $\int_1^{\infty} t^{x-1}e^{-t}dt$ converges for all values of x.

(b) By comparison with $\int_0^1 t^{x-1}dt$ show that $\int_0^1 t^{x-1}e^{-t}dt$ converges for all $x > 0$.

(c) Hence show that $\Gamma(x)$ is defined for all $x > 0$.

(d) Show that $\Gamma(x + 1) = x\Gamma(x)$ for all $x > 0$.

(e) Hence show that the value of $\Gamma(x)$, for any $x > 0$, can be obtained from the value of $\Gamma(z)$ for a suitable z in $[1,2)$. Given that $\Gamma(1.5) \approx 0.88623$ find $\Gamma(.5)$ and $\Gamma(6.5)$.

9.

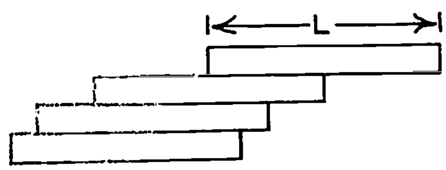


Figure A

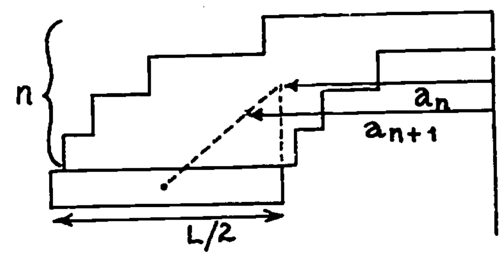


Figure B

I want to pile dominos, as in Figure A, so as to get as great an overhang as possible. The first (top) domino can be put with its center of gravity over the edge of the second one, so as to give an overhang of L/2. The c. of g. of these two can be put over the edge of the third; this gives an additional overhang of L/4. And so on. We can obviously do no better than this.

(a) Show that if the c. of g. of the first n dominos is a distance of a_n from the front edge (Figure B), then

$$a_{n+1} = a_n + \frac{L}{2(n + 1)} .$$

(b) What is the maximum overhang available with 28 dominos of length 2 in.? [Hint. Use Example 4.]

Ans. 3.89 inches.

(c) What can you say about the overhang if there is no restriction on the number of dominos?

5. Power Series.

We now return to the considerations of Section 1 but with a different approach. There we started with a function f and determined its Taylor series, of the form

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

Here we start with the series (1) and see whether it defines a function. A series of the form (1) is called a <u>power series</u>. Every Taylor series is a power series but the converse is not true.

Of course we are interested in whether or not (1) converges - more precisely, in the determination of those values of x for which (1) does converge. We see at once that (1) converges when $x = a$, for then the series is just

$$a_0 + 0 + 0 + 0 + \ldots = a_0.$$

To investigate other values of x we need a preliminary theorem.

<u>Theorem 1.</u> If (1) converges for $x = x_0$ it converges absolutely for any x for which $|x - a| < |x_0 - a|$. If (1) diverges for $x = x_0$ it diverges for any x for which $|x - a| > |x_0 - a|$.

Proof. If $\sum a_n(x_0 - a)^n$ converges then $|a_n(x_0 - a)^n| \to 0$ as $n \to \infty$. Hence for n sufficiently large $|a_n(x_0 - a)^n| \leq 1$. Now

$$\frac{|a_n(x - a)^n|}{|a_n(x_0 - a)^n|} = \left|\frac{x - a}{x_0 - a}\right|^n = r^n,$$

with $0 \leq r < 1$ since we are assuming $|x - a| < |x_0 - a|$. Hence for n sufficiently large $|a_n(x - a)|^n < r^n$, and by Comparison Test 1, $\sum|a_n(x - a)^n|$ converges since $\sum r^n$ converges. For the second half of the theorem, if $\sum a_n(x - a)^n$ converged so would $\sum a_n(x_0 - a)^n$ by the first half and Theorem 2 of Section 3. Hence $\sum a_n(x - a)^n$ must diverge.

We can now prove the basic theorem concerning the convergence of power series.

Theorem 2. For any power series of the form (1), one of the following is true:

(a) The series converges only for $x = a$;

(b) The series converges for all x;

(c) There is a positive number R such that the series converges if $|x - a| < R$ and diverges if $|x - a| > R$.

Proof. For simplicity we give the proof for the case $a = 0$, i.e. for

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

The general proof proceeds similarly.

That (a) and (b) can occur is shown by applying the Ratio Test to the two series $\sum_{n=0}^{\infty} n!x^n$ and $\sum_{n=0}^{\infty} \frac{x^n}{n!}$.

If neither (a) nor (b) is true, there must be two points, x_1 and y_1, with $x_1 \neq 0$, such that (2) converges for $x = x_1$ and diverges for $x = y_1$.
If $x_1 < 0$ replace it by $-x_1/2$;
if $y_1 < 0$ replace it by $-2y_1$.
Then by Theorem 1 the series (2) still converges for $x = x_1$ and diverges for $y = y_1$, and $0 < x_1 < y_1$. We now start the bisection process shown in Figure 5-1. This gives us an increasing sequence $x_1, x_2, \ldots$ and a decreasing sequence $y_1, y_2, \ldots$ with a common limit R, and such that (2) converges at each x_i and diverges at each y_i. If $|x| < R$ there is an $x_i > |x|$ and so by Theorem 1, $\sum a_n x^n$ converges. If $|x| > R$ there is a $y_i < |x|$ and by Theorem 1, $\sum a_n x^n$ diverges.

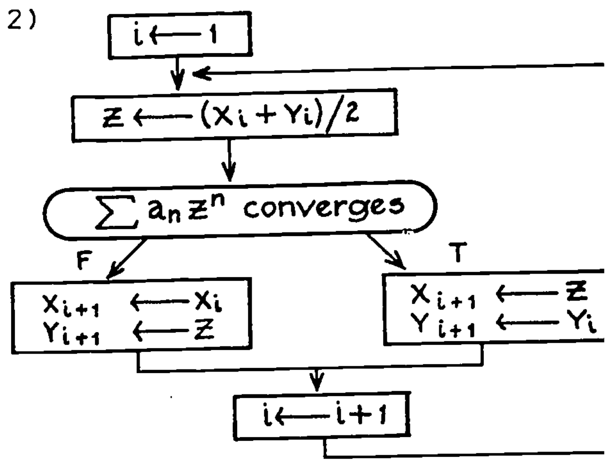


Figure 5-1

In case (c) the number R is called the radius of convergence of the series. Cases (a) and (b) can be included in this definition by allowing R to have the values 0 and ∞ respectively.

The algorithm of Theorem 2 is an impractical method of determining the radius of convergence, since the truth or falsity of the branch condition, $\sum a_n z^n$ converges, is difficult to determine. The following theorem is useful in many common cases.

Theorem 3. If $\lim_{n\to\infty} \left|\frac{a_n}{a_{n+1}}\right| = R \leq \infty$ then R is the radius of convergence of $\sum_{n=0}^{\infty} a_n(x - a)^n$.

Proof. Applying the Ratio Test to the series, we have convergence or divergence according as

$$\lim_{n\to\infty} \left|\frac{a_{n+1}(x - a)^{n+1}}{a_n(x - a)^n}\right| < 1 \text{ or } > 1,$$

or according as

$$\lim_{n\to\infty} \left|\frac{a_n(x - a)^n}{a_{n+1}(x - a)^{n+1}}\right| > 1 \text{ or } < 1.$$

This limit is $\frac{R}{|x - a|}$, and so we have convergence or divergence according as $R > |x - a|$ or $R < |x - a|$. This is just the condition that R be the radius of convergence.

Example 1. The Taylor series for log x about 1 is

$$(x - 1) - \frac{1}{2}(x - 1)^2 + \frac{1}{3}(x - 1)^3 - \dots .$$

Hence

$$\lim_{n\to\infty} \left|\frac{a_n}{a_{n+1}}\right| = \lim_{n\to\infty}\left(\frac{1}{n} \div \frac{1}{n + 1}\right) = \lim_{n\to\infty} \frac{n + 1}{n} = 1.$$

Hence $R = 1$, and the series converges for $|x - 1| < 1$ or $0 < x < 2$, which agrees with what we found at the end of Section 1.

Example 2. The Maclaurin series for sin x,

$$\sin x = x - \frac{1}{3!} x^3 + \frac{1}{5!} x^5 - \dots ,$$

has coefficients $0, 1, 0, -\frac{1}{3!}, 0, \frac{1}{5!}, \dots,$ and Theorem 3 obviously cannot be applied. However, we can put $z = x^2$ and write

$$\sin x = x(1 - \frac{1}{3!} z + \frac{1}{5!} z^2 - \dots),$$

and apply Theorem 3 to the series in z. We get

$$R = \lim_{n\to\infty}\left(\frac{1}{(2n + 1)!} \div \frac{1}{(2n + 3)!}\right) = \lim_{n\to\infty} (2n + 2)(2n + 3) = \infty$$

in agreement with Problem 4 of Section 1.

Neither Theorem 2 nor 3 tells us anything about convergence when $|x - a| = R$. Anything can happen here, as is shown by the following examples, each of which has $R = 1$:

(a) $\sum \frac{x^n}{n^2}$ converges at both 1 and -1;

(b) $\sum \frac{x^n}{n}$ converges at -1 but not at 1;

(c) $\sum x^n$ converges at neither 1 nor -1.

Frequently the behavior at the ends of the interval of convergence is of no great interest. If it must be determined the methods of Sections 2 to 4 are available.

For values of x within the interval of convergence the relation

$$f(x) = \sum_{n=0}^{\infty} a_n(x - a)^n$$

defines a function, since for each such x, f(x) has a definite value. The manipulation of these functions, for a given value of a, is particularly simple, being essentially the same as for polynomials. For simplicity we use a = 0 in the following discussion; in any case one can achieve this by introducing a new variable z = x - a and using power series in z.

The polynomials,

$$f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n,$$

$$g(x) = b_0 + b_1x + b_2x^2 + \dots + b_nx^n,$$

have the properties:

(a) $f(x) \pm g(x) = (a_0 \pm b_0) + (a_1 \pm b_1)x + \dots + (a_n \pm b_n)x^n;$

(b) $cf(x) = ca_0 + ca_1x + \dots + ca_nx^n;$

(c) $f(x)g(x) = a_0b_0 + (a_0b_1 + a_1b_0)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2 + \dots + (a_{n-1}b_n + a_nb_{n-1})x^{2n-1} + a_nb_nx^{2n};$

(d) $f'(x) = a_1 + 2a_2x + 3a_3x^2 + \dots + na_nx^{n-1};$

(e) $\int_0^x f(t)dt = a_0x + \frac{1}{2}a_1x^2 + \frac{1}{3}a_2x^3 + \dots + \frac{1}{n+1}a_nx^{n+1}.$

The corresponding properties of power series are stated in Theorem 4. Proofs of (a) and (b) follow from Problem 1 of Section 2 but (c), (d), and (e) are much harder and proofs will not be given here.

Theorem 4. Let

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \quad \text{and} \quad g(x) = \sum_{n=0}^{\infty} b_n x^n$$

have radii of convergence R_1 and R_2 respectively. Then

(a) $f(x) \pm g(x) = \sum_{n=0}^{\infty} (a_n \pm b_n)x^n$, $R = \min(R_1, R_2)$;

(b) $cf(x) = \sum_{n=0}^{\infty} ca_n x^n$, $R = R_1$;

(c) $$f(x)g(x) = a_0b_0 + (a_0b_1 + a_1b_0)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2 + \dots + (a_0b_n + a_1b_{n-1} + \dots + a_nb_0)x^n + \dots, \quad R = \min(R_1, R_2)$$

(d) $f'(x) = a_1 + 2a_2x + 3a_3x^2 + \dots + na_nx^{n-1} + \dots$, $R = R_1$;

(e) $$\int_0^x f(t)\,dt = a_0x + \frac{1}{2}a_1x^2 + \frac{1}{3}a_2x^3 + \dots + \frac{1}{n+1}a_nx^{n+1} + \dots, \quad R = R_1.$$

In each case R is the radius of convergence of the series.

Example 3. We have established that

(3) $$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots, \quad |x| < 1,$$

and so, setting $x = t^2$,

$$\frac{1}{1 + t^2} = 1 - t^2 + t^4 - t^6 + \dots , \quad |t| < 1.$$

Integrating from 0 to x we have, by (e),

$$\text{(4)} \qquad \arctan x = x - \tfrac{1}{3}x^3 + \tfrac{1}{5}x^5 - \tfrac{1}{7}x^7 + \dots , \quad |x| < 1.$$

Example 4. Starting again with (3) we get by differentiating and changing signs,

$$\frac{1}{(1 + x)^2} = 1 - 2x + 3x^2 - 4x^3 + \dots , \quad |x| < 1.$$

We can get the same result from (c) by multiplying the series in (3) by itself. It is often convenient to do this by the method used in elementary algebra for multiplying polynomials:

$$\begin{array}{rrrrrrr}
\frac{1}{1+x} = & 1 & - x & + x^2 & - x^3 & + \dots & \\
\frac{1}{1+x} = & 1 & - x & + x^2 & - x^3 & + \dots & \\
\hline
 & 1 & - x & + x^2 & - x^3 & + \dots & \\
 & & - x & + x^2 & - x^3 & + \dots & \\
 & & & x^2 & - x^3 & + \dots & \\
 & & & & - x^3 & + \dots & \\
 & & & & \dots & \dots & \\
\hline
\frac{1}{(1+x)^2} = & 1 & - 2x & + 3x^2 & - 4x^3 & + \dots &
\end{array}$$

From (d) of Theorem 4 we get the following important result.

Corollary 1. If the series

$$f(x) = \sum_{n=0}^{\infty} a_n(x - a)^n$$

has a radius of convergence $R > 0$ then

(a) All derivatives $f^{(n)}(x)$ exist for $|x - a| < R$;

(b) $a_n = \frac{1}{n!} f^{(n)}(a), \quad n = 0, 1, 2, \ldots ;$

(c) $\sum_{n=0}^{\infty} a_n(x - a)^n$ is the Taylor series of $f(x)$ about a.

Proof. By Theorem 4(d),

$$f'(x) = \sum_{n=0}^{\infty} na_n(x - a)^{n-1} = \sum_{n=1}^{\infty} na_n(x - a)^{n-1}, \quad |x| < R.$$

Applying Theorem 4(d) to this series gives

$$f''(x) = \sum_{n=1}^{\infty} n(n - 1)a_n(x - a)^{n-2} = \sum_{n=2}^{\infty} n(n - 1)a_n(x - a)^{n-2},$$

$|x| < R.$

And so on. In general

(5) $$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n - 1)\ldots(n - k + 1)a_n(x - a)^{n-k}, \quad |x| < R.$$

Putting $x = a$ in (5) gives

$$f^{(k)}(a) = k(k-1)\ldots1\cdot a_k = k!a_k,$$

which gives (b). (c) follows at once from the definition of a Taylor series.

This corollary tells us, among other things, that there is at most one way of expanding a function in a power series in $x - a$, namely, the Taylor series about a. Thus two power series that equal the same function must have their corresponding coefficients equal. This enables us to use the method of undetermined coefficients to compute terms of a Taylor series.

Example 5. To find terms of the Maclaurin series for the function

(6) $$f(x) = \frac{\cos x}{1 + e^x}$$

we assume that

$$f(x) = a_0 + a_1x + a_2x^2 + \ldots$$

and write (6) as

$$\cos x = (1 + e^x)(a_0 + a_1x + a_2x^2 + \ldots)$$

or

$$1 + 0x - \frac{1}{2}x^2 + 0x^3 + \frac{1}{24}x^4 + \dots$$

$$= (2 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{24}x^4 + \dots)(a_0 + a_1x + a_2x^2 + \dots)$$

$$= 2a_0 + (2a_1 + a_0)x + (2a_2 + a_1 + \frac{1}{2}a_0)x^2$$

$$+ (2a_3 + a_2 + \frac{1}{2}a_1 + \frac{1}{6}a_0)x^3 + \dots \quad .$$

Equating coefficients of powers of x gives

$$2a_0 = 1, \qquad a_0 = \frac{1}{2},$$

$$2a_1 + a_0 = 0, \qquad a_1 = -\frac{1}{4},$$

$$2a_2 + a_1 + \frac{1}{2}a_0 = -\frac{1}{2}, \qquad a_2 = -\frac{1}{4},$$

$$2a_3 + a_2 + \frac{1}{2}a_1 + \frac{1}{6}a_0 = 0, \qquad a_3 = \frac{7}{48},$$

and so on. Since we have formulas for the n-th terms of the two given series the above equations are the first few cases of a general recursion formula,

$$2a_n + \frac{a_{n-1}}{1!} + \frac{a_{n-2}}{2!} + \dots + \frac{a_0}{n!} = \begin{cases} 0 & \text{if } n \text{ is odd} \\ (-1)^{n/2}/n! & \text{if } n \text{ is even} \end{cases}$$

for determining a_n. Our present methods do not suffice to determine the radius of convergence of this series. From the theory of functions of a complex variable it can be shown that $R = \pi$.

Example 6. We wish to find the Taylor series about 1 of

$$f(x) = \begin{cases} \sqrt{\dfrac{\log x}{x - 1}} & \text{if } x \neq 1. \\ 1 & \text{if } x = 1. \end{cases}$$

To simplify the algebra we set $z = x - 1$ and find the Maclaurin series for

$$g(z) = f(1 + z) = \sqrt{\frac{\log (1 + z)}{z}} .$$

Proceeding as above we get

$$\frac{\log (1 + z)}{z} = 1 - \tfrac{1}{2}z + \tfrac{1}{3}z^2 - \tfrac{1}{4}z^3 + \ldots$$

$$= (a_0 + a_1 z + a_2 z^2 + \ldots)^2$$

$$= a_0^2 + 2a_0a_1 z + (2a_0a_2 + a_1^2)z^2$$

$$+ (2a_0a_3 + 2a_1a_2)z^3$$

$$+ (2a_0a_4 + 2a_1a_3 + a_2^2)z^4 + \ldots$$

Hence $a_0^2 = 1$ and since $f(0) = 1$, we must take $a_0 = 1$;

$$2a_0a_1 = -\tfrac{1}{2}, \qquad a_1 = -\tfrac{1}{4},$$

$$2a_0a_2 + a_1^2 = \tfrac{1}{3}, \qquad a_2 = \tfrac{13}{96},$$

and so on. Finally,

$$f(x) = 1 - \frac{1}{4}(x - 1) + \frac{13}{96}(x - 1)^2 + \dots + a_n(x - 1)^n + \dots ,$$

where the coefficients are defined by the recursion formula

$$\sum_{k=0}^{n} a_k a_{n-k} = (-1)^n/(n + 1), \qquad a_0 = 1.$$

You can see how this method might be used for a wide variety of problems. For instance, referring to Example 1 of Section 7-2, if we put

$$y = f(x) = \sum_{n=0}^{\infty} a_n(x - a)^n, \qquad a = 1, \qquad a_0 = .35$$

we can solve

$$x^3 + f(x)^3 = 3xf(x)$$

for $a_1, a_2, \dots,$ thereby obtaining a Taylor expansion of the implicit function discussed in that example.

Another application is to differential equations. The equation

(7) $\frac{dy}{dx} = x^2 + y^2, \quad y(0) = 1,$

for example, can be solved by putting

(8) $y = a_0 + a_1x + a_2x^2 + \dots , \quad a_0 = 1,$

and solving

$$a_1 + 2a_2x + 3a_3x^3 + \dots = x^2 + (a_0 + a_1x + a_2x^2 + \dots)^2$$

successively for $a_1, a_2, \dots$. In this type of problem the determination of the radius of convergence of the series can be extremely difficult, and in general it can only be approximated by numerical computation.

There is a serious objection to this method of getting power series if we wish to use the series to compute approximate values of the function; namely, we have no bounds for the remainder after n terms. For instance the series (8) is readily found to start

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

If we put $x = .2$, is

$$S_5 = 1 + .2 + (.2)^2 + \frac{4}{3}(.2)^3 + \frac{7}{6}(.2)^4 + \frac{6}{5}(.2)^5 \approx 1.256$$

a good approximation to y(.2), and if so, what is the maximum possible error? We have no easy way of answering these questions.

This does not mean that (9) is worthless, for there are uses for infinite series other than the computation of function values. One of these is the determination of limits, the topic of Section 10-4. Suppose, for instance, that we want to find

$$\lim_{x\to 0} \frac{1 - \cos x}{y(x) - e^x},$$

where y(x) is the solution of (7). Expanding each function in a Maclaurin series gives

$$\lim_{x\to 0} \frac{1 - (1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 - \dots)}{(1 + x + x^2 + \frac{4}{3}x^3 + \dots) - (1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \dots)}$$

$$= \lim_{x\to 0} \frac{\frac{1}{2}x^2 + \frac{1}{24}x^4 + \dots}{\frac{1}{2}x^2 + \frac{7}{6}x^3 + \dots}$$

$$= \lim_{x\to 0} \frac{\frac{1}{2} + \frac{1}{24}x^2 + \dots}{\frac{1}{2} + \frac{7}{6}x + \dots} = 1.$$

For limits as $x \to a$ this method is often the simplest one to use if the functions involved can be expanded in Taylor series about a.

Power series have many other interesting properties and applications, some of which are given in the problems. For the full development and understanding of the theory of power series it is necessary to allow the variable to assume values which are complex numbers. The related theory, the theory of functions of a complex variable, is one of the most interesting branches of mathematics.

PROBLEMS

1. Find the radius of convergence of $\sum_{n=1}^{\infty} a_n x^n$ for each of the following cases, if possible.

(a) $a_n = 2^n$

(b) $a_n = n^2$

(c) $a_n = \frac{(-2)^n}{n!}$

(d) $a_n = \frac{n^n}{n!}$

(e) $a_n = \frac{(n!)^2}{(2n)!}$

(f) $a_n = \frac{\sin n}{n^2}$

(g) $a_n = \frac{\log n}{n}$

(h) $a_n = \cos\left(\frac{n\pi}{3}\right)$

(i) $a_n = \sin\left(\frac{n\pi}{3}\right)$. [Hint. Try grouping terms.]

2. (a) Derive the "binomial series"

$$(1 + x)^m = 1 + mx + \frac{m(m-1)}{2} x^2 + \dots$$

$$+ \frac{m(m-1)\dots(m-n+1)}{n!} x^n + \dots$$

where m is any real number.

(b) Show that the radius of convergence of the binomial series is 1 except when m is a non-negative integer, in which case it is ∞.

3. Find the first four non-zero terms of the Maclaurin series of each of the following functions. If you can, also find the general term and the radius of convergence.

(a) $e^x \sin x$

(b) $\tan x$. [Hint. $\tan x = \dfrac{\sin x}{\cos x}$.]

(c) $\cosh x = \dfrac{e^x + e^{-x}}{2}$

$\sinh x = \dfrac{e^x - e^{-x}}{2}$

(d) $\log \cos x$. [Hint. Use $\int_0^x \tan t \, dt$.]

(e) $\sqrt{3 + \cos x}$. Ans. $2 - \frac{1}{8}x^2 + \frac{5}{768}x^4 + \frac{11}{18430}x^6 + \ldots$

(f) Solution of $y' = x + y$, $y(0) = 1$.

(g) Solution of $y' = x - y^2$, $y(0) = 0$.

Ans. $\frac{1}{2}x^2 - \frac{1}{20}x^5 + \frac{1}{160}x^8 - \frac{7}{8800}x^{11} + \ldots$.

4. The "sine integral" function Si(x) is defined by

$$Si(x) = \int_0^x \frac{\sin t}{t}\, dt,$$

(a) Find the Maclaurin series of Si(x).

(b) Compute Si(1) to 3D.

(c) Use the computer to compute Si(5) to 5D.

Ans. 1.54993.

(d) Write and run a program to tabulate Si(x) to 5D for x = 2(.1)7. Compare with page 242 of "Handbook of Mathematical Tables", Abramowitz and Stegun, Dover Publications.

(e) This table can also be computed by evaluating the integral by Simpson's rule. If this is done efficiently, i.e. without recomputing the whole integral for each value of x, which method do you think is most efficient. Give the reasons for your answer.

(f) Recompute the table using the Simpson rule and compare, if you can, the machine times needed for the two method.

5. The following theorem can be proved by methods somewhat beyond those of our text: If $\sum_{n=0}^{\infty} a_n$ converges then the function $f(x) = \sum_{n=0}^{\infty} a_n x^n$ is defined and continuous in $-1 < x \leq 1$. The important point is the continuity at $x = 1$.

(a) Apply this theorem to

$$f(x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$$

to prove that

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \log 2.$$

Check the answer of Problem 8, Section 2.

[Hint. Since $\log(1 + x) = f(x)$ for $x < 1$, and both functions are continuous at $x = 1$, we must have $\log(1 + 1) = f(1)$.].

(b) Prove that

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \frac{\pi}{4}.$$

(c) Using the method of Problem 8, Section 2, evaluate π to 2D. (But see Problem 7 below).

1002

6. Evaluate the following limits.

(a) $\lim_{x\to 0} \frac{\tan x - x}{\sin x - x}$

(b) $\lim_{x\to 0} \frac{\arctan x - x}{e^{x} - e^{-x} - 2x}$

(c) $\lim_{x\to 0} \frac{\log (1 + x)}{\log (1 - x)}$

(d) $\lim_{x\to 0} \frac{\sqrt{1 + x} - \sqrt{1 - x}}{2x}$

(e) $\lim_{x\to 0} \left(\frac{\sin x}{x^3} - \frac{\cos x}{x^2}\right)$

(f) $\lim_{x\to 0} \left(\frac{1}{x} - \frac{1}{\sqrt{x}}\right)$

(g) $\lim_{x\to 1} \frac{\sqrt{x} - 1}{\log x}$

(h) $\lim_{x\to \pi/2} (\tan x - \sec x)$

7. (a) Prove the identity

$$\arctan A + \arctan B = \arctan \frac{A + B}{1 - AB}$$

by taking tan of both sides.

(b) By successive application of this identity show that

$$4 \arctan \frac{1}{5} - \arctan \frac{1}{239} = \arctan 1 = \frac{\pi}{4} .$$

(c) Use this identity and the Maclaurin series for arctan x to compute π to 4D accuracy, using pencil and paper only.

This method was used by William Shanks in 1873 to compute π to 707 decimal places. Since the advent of the electronic computer π has been computed to more than 100,000 decimal places.

8. Associated with any sequence $a_0, a_1, a_2, \ldots$ there is a power series $a_0 + a_1 x + a_2 x^2 + \ldots$. Even though this power series may not converge for any x except zero we write

$$g(x) = \sum_{n=0}^{\infty} a_n x^n$$

and call g(x) the <u>generating function</u> of the sequence. Here is one of the many applications of generating functions.

(a) The Fibonacci numbers (see Section 2-4) are defined by

$$f_0 = 0, \quad f_1 = 1,$$

$$\text{(i)} \quad f_n = f_{n-1} + f_{n-2}, \quad n = 2,3,\ldots$$

Multiply (i) by x^n, sum from 2 to ∞, and reduce the result to

$$\text{(ii)} \quad g(x)(1 - x - x^2) = x,$$

where

$$\text{(iii)} \quad g(x) = \sum_{n=0}^{\infty} f_n x^n$$

is the generating function of the Fibonacci numbers.

(b) From (ii) derive by partial fractions

$$\text{(iv)} \quad g(x) = \frac{1}{\sqrt{5}}\left[\frac{1}{1 - ax} - \frac{1}{1 - bx}\right],$$

where $a = \frac{1 + \sqrt{5}}{2}$, $b = \frac{1 - \sqrt{5}}{2}$.

(c) Expand the two terms of the right-hand side of (iv) in series, and equate coefficients of like powers of x in the resulting series and (iii) to get, finally,

$$f_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right].$$

(d) Check the formula for $n = 0,1,2,3$.

(e) Prove that f_n is the integer closest to $\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n$, and that f_n is very nearly equal to this quantity for large n.

(f) Show that $f_{100} \approx 3.5 \times 10^{20}$.

(g) Criticize the derivation of the formula for f_n.

Chapter 14

DIFFERENTIAL EQUATIONS*

1. Numerical Solution.

In Section 9-2 we con idered some differential equations with initial conditions, of the form

$$(1) \qquad y' = f(x,y), \qquad y(x_0) = y_0.$$

For certain simple cases of the function f we were able to find a solution of (1), that is, a function y(x) satisfying the initial condition and such that

$$y'(x) = f(x,y(x))$$

for every x in some interval $[x_0, x_M]$. In this chapter we shall consider much more general cases of (1), discussing whether they actually have solutions, and, if so, how to determine these solutions either exactly or approximately.

To keep things simple at first we start off with an equation we know all about, namely

$$(2) \qquad y' = ky, \qquad y(0) = 1.$$

---

*Some of the material in this chapter is taken from "Engineering Mathematics" by Block, Cranch, Hilton, and Walker. Permission to use this material has been granted by the copyright owner, Cornell University, but its publication in this form is not endorsed by the copyright owner or the original authors.

We have seen that the unique solution to this equation is $y = e^{kx}$. Suppose we did not know this and we wanted to find the value of $y(x_1)$ where $x_1 = .1$. Since from (2) we can find $y'(0)$ we can use the linear approximation of $y(x)$,

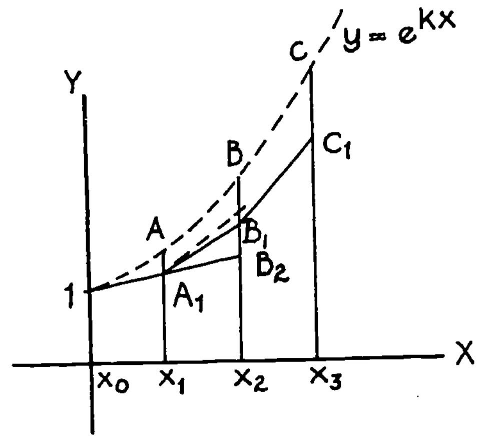


Figure 1-1

$$y(x) \approx y(0) + xy'(0),$$

to get

$$y(x_1) \approx Y_1$$

$$= y(0) + .1y'(0) = 1 + .1k.$$

In Figure 1-1, $y(x_1)$ is the ordinate of A and Y_1 the ordinate of A_1.

Now suppose we want $y(x_2)$, where $x_2 = 2x_1$. We could of course take

$$y(x_2) \approx 1 + .2k,$$

giving the point B_2, but it seems better to start with A_1 and take another step of x_1. Thus:

$y(x_2) \approx Y_2 = Y_1 + .1y'(Y_1) = Y_1 + .1kY_1.$

This gives us point B_1. Notice that the line A_1B_1 is not tangent to the curve $y(x)$, nor parallel to the tangent to $y(x)$ at A, but is tangent to the solution of $y' = ky$ that passes through A_1. (Dotted line in Figure 1-1).

The process can now be repeated to get $Y_3, Y_4, \ldots$ corresponding to points $C_1, D_1, \ldots$ that approximate points $C, D, \ldots$ on the true solution.

To investigate this process further it is convenient to introduce some notation. We assume that $x_{n+1} - x_n = h$ is constant, so that $x_n = x_0 + nh = nh$ in our present example. Then

(3) $Y_{n+1} = Y_n + hy'(Y_n) = Y_n + hkY_n, \quad Y_0 = 1.$

This equation gives a recursion formula for Y_n. Written in the form

$$Y_{n+1} = (1 + hk)Y_n, \quad Y_0 = 1,$$

we see that Y_n is multiplied by the constant $(1 + hk)$ at each step. Since its initial value is 1 we obviously have

$$Y_n = (1 + hk)^n.$$

Now we can write

$$Y_n = (1 + hk)^n = \left[(1 + hk)^{\frac{1}{hk}}\right]^{nhk} = \left[(1 + hk)^{\frac{1}{hk}}\right]^{kx},$$

and by Problem 7(a) of Section 10-4,

$$\lim_{hk \to 0} (1 + hk)^{\frac{1}{hk}} = e.$$

Hence for a fixed $x = nh$, as $h \to 0$ and $n \to \infty$, $\lim Y_n = e^{kx} = y(x)$. Thus we are assured that we can get as close an approximation as we wish, by taking h small enough.

We shall show that this happy conclusion applies to a very general class of equations of the form (1). Before proving this, however, we shall examine these equations from a geometric point of view.

PROBLEMS

1. The error in using Y_n as an approximation to $y(x)$ is

$$E = e^{kx} - (1 + hk)^{x/h} .$$

   (a) Regarding the right-most term as a function of h and using its linear approximation show that

$$E = \frac{1}{2}hk^2 xy(x).$$

   (b) The <u>relative error</u>, the ratio of the error to the true value, is in many cases more significant than the absolute error. Discuss the behavior of the two types of error in this problem, particularly as x increases with fixed h. The cases $k > 0$ and $k < 0$ must be distinguished.

2. Use the computer to determine E for various values of k, x, and h, with $kx = 1$. The value of $(1 + hk)^{x/h}$ is best obtained by successive squaring, using $h = x/2^N$. Does the linear approimation seem to hold pretty well? What happens for very small values of h?

2. <u>Graphical Solution</u>.

We consider the differential equation

(1) $y' = f(x,y)$,

assuming that the function f has enough continuity properties to make the following discussion meaningful.

At any point (x,y) at which f(x,y) is defined, (1) determines a direction, or more precisely, a slope at the point. The combination of point and slope is called a line element and is usually represented by a point with a short line segment through it. Figure 2-1 shows three line elements.

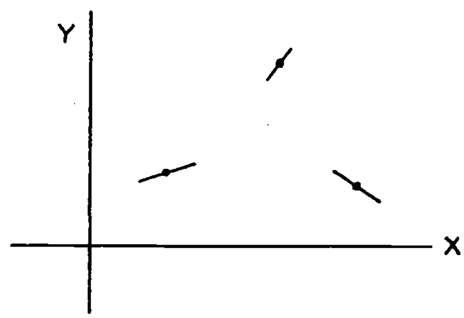


Figure 2-1

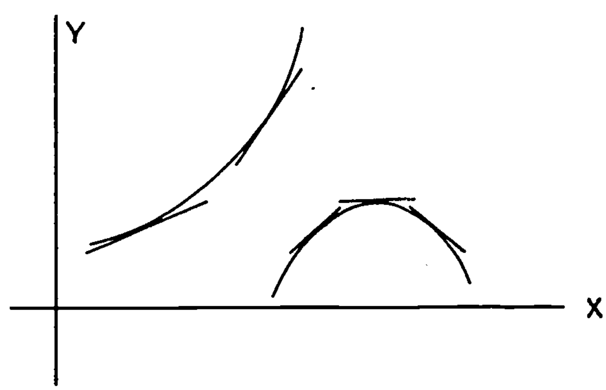


Figure 2-2

Any solution of (1) must be tangent to the line element at each of its points (Figure 2-2); and conversely, if we can find a curve that is tangent to the line element at each of its points then it determines a solution of (1). This property can be used to get some information about the solutions of (1).

To do this we first draw a large number of line elements, as in Figure 2-3, for the equation $y' = x - y^2$. This is somewhat of a chore if done by hand and is most easily accomplished by first drawing isoclines, curves along which the line elements have constant direction. These are obviously the curves $f(x,y) = m$ for various values of m. One of these is shown dotted in Figure 2-3.

A much pleasanter way to get the line element field is to use a computer with a good graphical output. Here it is easier to dispense with isoclines and just plot a large number of line elements on a rectangular grid.

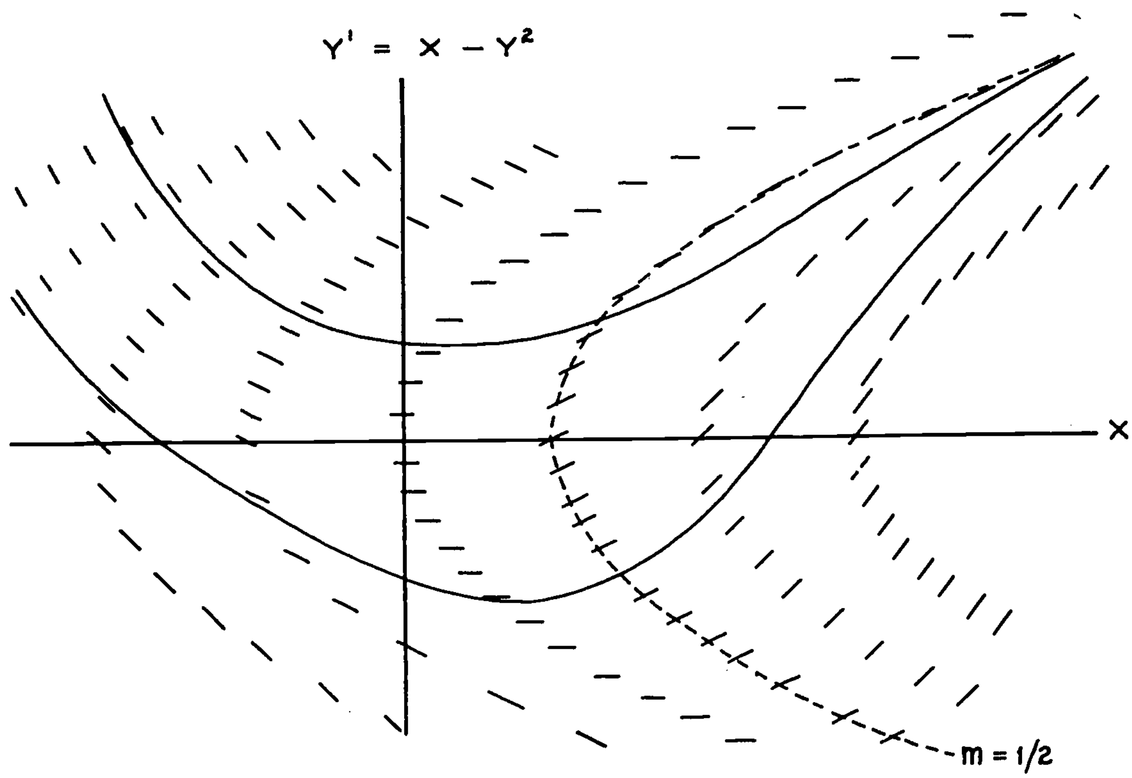


Figure 2-3

With a sufficient number of line elements one can fairly easily sketch in solutions of the equation. These give an idea of the general shape of the curves, their behavior with regard to local extrema and intervals of monotonicity, etc. In Figure 2-3 it is easy to deduce that the solutions approach the parabola $x - y^2 = 0$ as x increases but what happens as x decreases is not so obvious. In fact, each curve has a vertical asymptote (Problem 2).

This graphical approach is often useful in getting an idea of the shape of a solution before starting an elaborate analysis or computation to find it precisely. Knowing what to expect ahead of time is both a guide in the selection of a method of computation and a check on any serious errors that might occur.

PROBLEMS

1. Use line elements to sketch several solutions of each of the following equations. Make whatever comments you can about different types of solution of the same equation, local extrema, behavior for x increasing and decreasing, etc. Note that all local extrema occur on the isocline $f(x,y) = 0$.

   (a) $y' = x + y$ (d) $y' = \frac{x}{x^2 + y^2}$

   (b) $y' = \frac{x}{y}$ (e) $y' = \frac{5x + y}{x - 5y}$

   (c) $y' = -\frac{y}{x}$ (f) $y' = x + \frac{1}{y}$

2. We wish to show that a solution of $\frac{dy}{dx} = x - y^2$ has an asymptote as x decreases.

   (a) Setting $z = -x$, show that the above statement is the same as showing that $\frac{dy}{dz} = z + y^2$ has an asymptote as z increases.

   (b) Let $y(z)$ be a solution of $\frac{dy}{dz} = z + y^2$ and $w(z)$ of $\frac{dw}{dz} = w^2$, with initial condition $y(a) = w(a) = b$, $a \geq 0$, $b \geq 0$. Give an argument showing that $y(z) \geq w(z)$ for all $z > a$.

(c) Solve for $w(z)$, to get $w(z) = \dfrac{1}{a + \frac{1}{b} - z}$.

(d) Show that $y(z)$ has an asymptote as z increases.

3. Consider $y' = xy + 1$ for $x \geq 0$, and the solutions starting at $(0,b)$ for various negative values of b. Let C be the curve $xy = -1$ in the fourth quadrant.

(a) Show that a solution that crosses C eventually goes down rapidly in the fourth quadrant.

(b) Show that a solution that crosses the x-axis eventually goes up rapidly in the first quadrant.

(c) Show that there must be at least one curve that crosses neither the x-axis nor the curve C. [Hint. Use a bisection process.]

(d) We shall show later (Section 8, Problem 9) that there is exactly one such curve, through a point $(0,b_0)$. Locate b_0 as well as you can.

3. The Fundamental Theorem.

To discuss the solution of

(1) $y' = f(x,y), \quad y(x_0) = y_0$

we must first consider some properties of the function f. Suppose that f is defined in some open region R in the xy-plane. The adjective "open" means that the points on the boundary are not regarded as points of R. For example, R might consist of the points strictly inside a circle, or inside a rectangle.

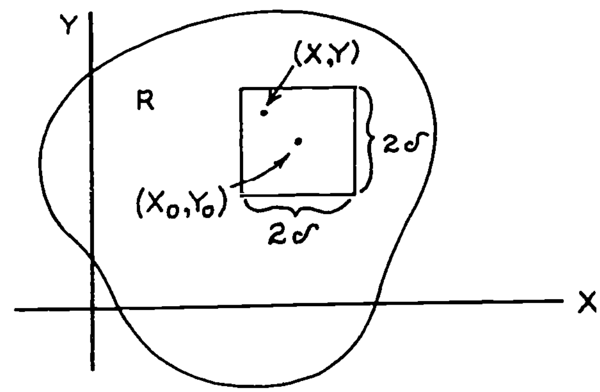


Figure 3-1

f is continuous in R if, given any point (x_0,y_0) in R and any $\varepsilon > 0$ there is a $\delta > 0$ such that $|f(x,y) - f(x_0,y_0)|$ whenever (x,y) is in R and $|x - x_0| < \delta$ and $|y - y_0| < \delta$. This is an obvious generalization of the definition of continuity for a function of one variable. Its geometrical significance is shown in Figure 3-1: Given a square of side 2δ with center at (x_0, y_0), the function values at the center and at any point inside the square will differ by less than ε.

f is <u>Lipschitzian in</u> y (see Section 3-10) in the region R if there is a number L such that

$$|f(x,y_1) - f(x,y_2)| \le L|y_1 - y_2|$$

for all pairs of points $(x,y_1),(x,y_2)$ in R. We usually prove that f is Lipschitzian by showing that $|f'_y(x,y)| < L$ in R, where $f'_y(x,y)$ designates the derivative of f with respect to y regarding x as a constant; i.e.

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

Proof is left to the reader. (Problem 1).

We can now state the fundamental existence and uniqueness theorem for differential equations.

<u>Theorem 1</u>. If f is continuous in a region R, then for any (x_0, y_0) in R there is an $H > 0$ such that (1) has a solution y(x) for $|x - x_0| < H$. If in addition, f(x,y) is Lipschitzian in y in the region R then the solution is unique.

The proof of this theorem is well beyond the level of this text.

Example 1. For the equation

(2) $y' = x + y, \quad y(0) = 1$

we can take R as the whole plane, since $f(x,y) = x + y$ is continuous for all values of x and y, and $f'_y(x,y) = 1$ is certainly bounded.

The solution is

$$y(x) = 2e^x - x - 1,$$

which extends indefinitely in both directions. Hence in this case we can take H as large as we please.

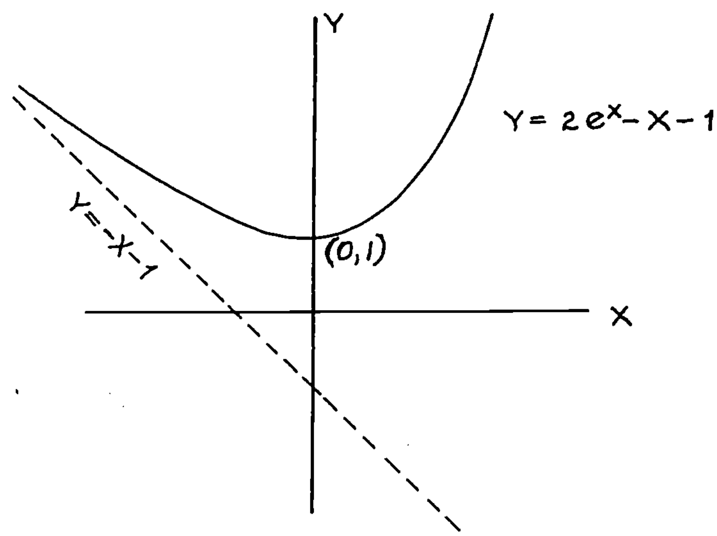


Figure 3-2

Example 2. Consider the equation

(3) $y' = \frac{4y^2 - 6y}{3x}, \quad y(1) = 1.$

The function is defined and continuous in each of the regions $x > 0, -\infty < y < \infty$ and $x < 0, -\infty < y < \infty$. Since our initial point (1,1) lies in the former region we use it. $f'_y = \frac{8y - 6}{3x}$ is bounded if $|y|$ is bounded above and x is bounded away from zero. So we must take R of the form

$$a < x < \infty, \quad |y| < M,$$

for some $a > 0$ and some $M > 0$.

The solution is

$$y = \frac{3}{2 + x^2},$$

curve ① in Figure 3-3. The curve goes indefinitely to the right but must stop when it hits the boundary of R at $x = a$. Hence $H = 1 - a$, where a can be arbitrarily small.

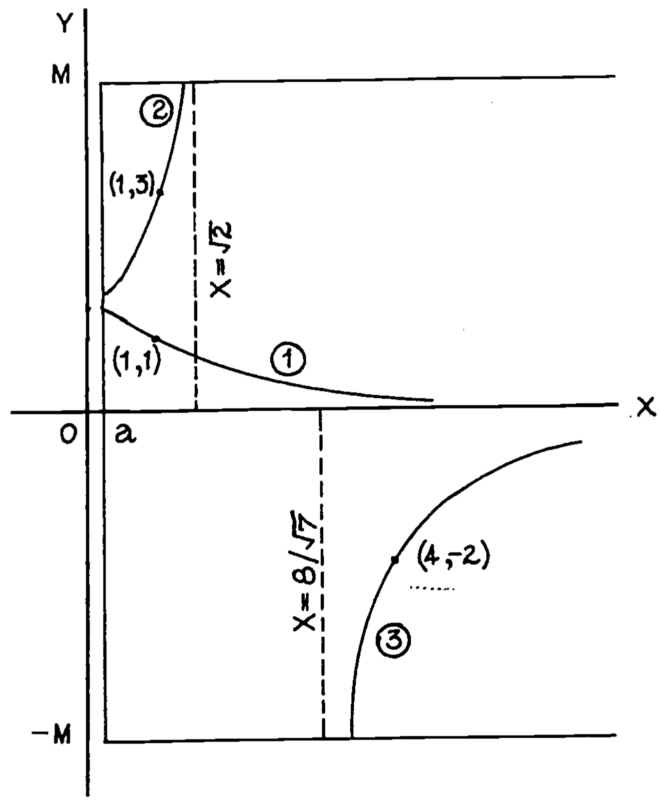


Figure 3-3

If we change the initial condition to $y(1) = 3$ we get curve ② with equation

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

This behaves quite differently, having an asymptote at $x = \sqrt{2}$, and hence having $H = \sqrt{2} - 1$.

Finally, curve ③ ,

$$y = \frac{96}{64 - 7x^2},$$

with initial condition $y(4) = -2$, has an asymptote on the left, at $x = 8/\sqrt{7} \approx 3$, and so $H = 4 - 8/\sqrt{7} \approx 1$.

Example 3. (See Problem 14 of Section 9-2).

(4) $\qquad y' = 3y^{2/3}, \quad y(0) = 0.$

Here f is continuous for all x and y but $f'_y = 2y^{-1/3}$ is unbounded near the x-axis. The fundamental theorem says that (4) has a solution but it may not be unique. In fact, any curve of the type shown in Figure 3-4 is a solution of $y' = 3y^{2/3}$. Thus there is an infinite number of such curves through any point on the x-axis.

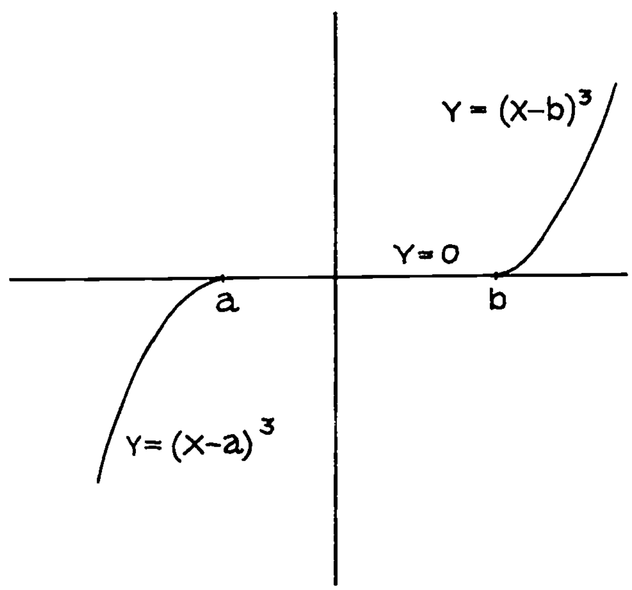


Figure 3-4

PROBLEMS

1. (a) Use the Mean Value Theorem to prove that if a function F has the property

$$L_1 \le F'(y) \le L_2$$

for all y in $[a,b]$, then for any y_1 and y_2 in $[a,b]$

$$F(y_1) - F(y_2) = k(y_1 - y_2),$$

with $L_1 \le k \le L_2$.

(b) Prove that if $|F'(y)| \le L$ for all y $[a,b]$ then

$$|F(y_1) - F(y_2)| \le L\,|y_1 - y_2|$$

for all y_1, y_2 in $[a,b]$.

2. Investigate carefully the solutions of $y' = -2\sqrt{y}$.

3. Consider the modification of Example 2:

$$f(x,y) = \begin{cases} \dfrac{4y^2 - 6y}{3x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

(a) Show that f is not continuous at any point (0,c).

(b) Show that $y' = f(x,y)$ has solutions over intervals (a,b) that include the value $x = 0$.

(c) Show that there are no solutions through (0,c) unless $c = 3/2$ or 0. For $c = 3/2$ there is an infinite number of solutions, for $c = 0$ there is one.

(d) Compare the above behavior with that of $y' = g(x,y)$ when

$$g(x,y) = \begin{cases} \frac{y}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

4. Euler's Numerical Method.

The numerical method that was introduced in Section 1 is known as Euler's method of solving differential equations. Given the equation with initial condition

(1) $\quad y' = f(x,y), \qquad f(x_0) = y_0,$

we choose a number h, generally small and positive, and define numbers $Y_0, Y_1, Y_2, \ldots, Y_N$ by

(2) $\quad Y_0 = y_0, \quad Y_{n+1} = Y_n + hf(x_n, Y_n), \quad n = 0, 1, \ldots, N-1,$

where $x_n = x_0 + nh$. The special case of Section 1 leads us to hope that in the general case Y_n will be an approximation to $y(x_n)$.

Example 1. $\quad y' = \dfrac{4y^2 - 6y}{3x}, \quad y(1) = 1.$

This case has been examined in Example 2 of the last section, so we know what to expect. Table 4-1 gives the computations for h = .5, N = 6, and Figure 4-1 shows the same data graphically. Even with such a large value of h the Y_n are not hopelessly bad approximations to the $y(x_n)$. Cutting down

x	Y	f(x,Y)	y(x)
1.0	1.000	-.667	1.000
1.5	.667	-.494	.706
2.0	.420	-.302	.500
2.5	.269	-.177	.364
3.0	.181	-.106	.273
3.5	.128	-.067	.211
4.0	.095		.167

Table 4-1

the size of h improves the approximation considerably, as one can see by comparing the values at $x = 2$ from Tables 4-1 and 4-2.

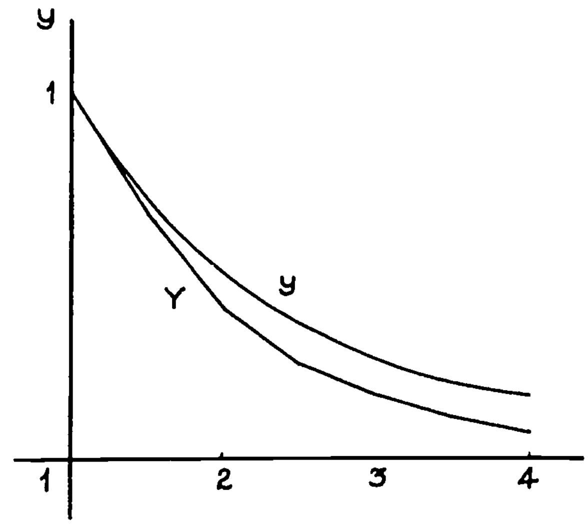


Figure 4-1

Now consider the same equation with the initial condition $y(1) = 3$. We saw in the earlier example that this has an asymptote at $x = \sqrt{2} = 1.414$ and hence the solution cannot be continued beyond this point. Nevertheless, the numerical "solution", as Table 4-3 shows, goes right on past $\sqrt{2}$ with no clear indication that its results are meaningless. The rather sudden jump in the value of Y at $x = 1.5$ does indicate possible trouble and suggests that we back up a bit and try a smaller value of h. But the computed value at $x = 1.4$ looks perfectly good even though it is hopelessly far off.

x	Y	f(x,y)	y
1.0	1.000	-.667	1.000
1.2	.867	-.578	.872
1.4	.751	-.535	.738
1.6	.644	-.455	.659
1.8	.553	-.388	.573
2.0	.475		.500

Table 4-2

This example shows the need for two things: first, a machine program to carry out the arithmetic involved in getting a numerical solution to any useful accuracy; and secondly, an error analysis that will tell us what value of h to use to get a given accuracy. We leave the first of these to the reader (Problems 2 & 3) and proceed to discuss the second.

x	Y	f(x,y)	y(x)
1.0	3.00	6.00	3.00
1.1	3.60	9.16	3.80
1.2	4.52	15.16	5.35
1.3	6.04	28.12	9.68
1.4	8.85	61.66	75.00
1.5	15.02		-----

Table 4-3

We assume that (1) satisfies the conditions of the Fundamental Theorem and has a unique solution y(x) for $x_0 \leq x \leq x_N$. We designate $y(x_n)$ by y_n. The error in the approximate solution given by (2) is then $E_n = y_n - Y_n$.

Now by the Extended Mean Value Theorem,

$$y(x_{n+1}) = y(x_n) + (x_{n+1} - x_n)y'(x_n) + \frac{1}{2}(x_{n+1} - x_n)^2 y''(\xi),$$

where $x_n < \xi < x_{n+1}$. This can be rewritten as

$$(3)\quad y_{n+1} = y_n + hf(x_n, y_n) + \frac{1}{2}h^2 y''(\xi).$$

Subtract from this the recursion formula

$$(4) \qquad y_{n+1} = Y_n + hf(x_n, Y_n),$$

and we get

$$(5) \qquad E_{n+1} = E_n + h[f(x_n, y_n) - f(x_n, Y_n)] + \tfrac{1}{2}h^2 y''(\xi).$$

The quantity $T_n = \frac{1}{2}h^2 y''(\xi)$ is called the truncation error, the error arising by cutting off all terms of (3) except those of first degree in h. To handle the expression in brackets we make the further assumption that all (x_n, Y_n) lie in the region R. Then by the Lipschitz condition,

$$f(x_n, y_n) - f(x_n, Y_n) = K_n(y_n - Y_n) = K_n E_n,$$

where $|K_n| \leq L$, the Lipschitz constant. Then (5) becomes

$$(6) \qquad E_{n+1} = (1 + hK_n)E_n + T_n .$$

Now, however, we must remember that we really do not compute Y_{n+1} exactly from (4), because of roundoff errors

in the computation. That is, we really have

$$Y_{n+1} = Y_n + hf(x_n, Y_n) + R_n,$$

where R_n is some unknown roundoff error, about which we can only say that $|R_n| \leq R$ for some small number R depending on the complexity of the function f, the word-length of the machine, etc. With this factor in the analysis (6) is replaced by

$$(7) \qquad E_{n+1} = (1 + hK_n)E_n + T_n - R_n.$$

To get bounds for E_n from (7) we must have bounds for the quantities K_n, T_n, and R_n. We have already seen that $|R_n| \leq R$. Assume that for all x in $[x_0, x_N]$ we have $|y''(x)| \leq M$; then $|T_n| \leq \frac{1}{2}h^2M$. For K_n we use only an upper bound, $K_n \leq K$. We take care of the lower bound by assuming that h is small enough to make $1 + hK_n$ positive. (In fact, if $|hK_n| \geq 1$ the approximation is too poor to be of any value.) Under these conditions,

$$(8) \qquad |E_n| \leq \left(\frac{hM}{2} + \frac{R}{h}\right)\frac{1}{K}\left[(1 + hK)^n - 1\right].$$

The proof, which is not difficult but rather long, is given at the end of this section. The case K = 0 can be handled by taking limits as $K \to 0$ (see Problem 1).

The bracketed expression in (8) can be handled as in Section 1. For small values of hK we have approximately

(9) $$(1 + hK)^N \approx e^{NhK} = e^{Kx},$$

if $x = Nh$. We can therefore draw the following conclusions from (8) and (9):

1. If there is no roundoff, i.e. in the case of exact mathematical analysis, $\lim_{h\to 0}|E_N| = 0$ and $\lim_{h\to 0} Y_N = y(x)$. This proves the convergence of Euler's method under the conditions we have assumed.

2. If roundoff is present then the upper bound for $|E_n|$ becomes infinite as $h \to 0$. This does not mean that $|E_N|$ necessarily becomes very large but it admits the possibility. A more exact analysis, using equation (7) and the statistical distribution of the R_n, shows that $|E_N|$ does indeed become arbitrarily large as $h \to 0$.

3. For fixed h and variable x, the bound on $|E_N|$ grows like e^{Kx}. Here again, this does not mean that $|E_N|$ grows this fast, but if the K_n remain fairly close to K the growth is of this order of magnitude.

4. Roundoff and truncation error are of about equal significance when $h^2 = 2R/M$. For a typical situation we might have $R = 10^{-14}$, $M = 8$, in which case the critical value of h is 5×10^{-8}. Since this would require twenty million steps to go from x_0 to $x_0 + 1$ one is hardly likely ever to use so small a value of h. On the other hand, for $R = 10^{-8}$, $M = .5$, the critical value is 2×10^{-4}. This implies only 5000 steps per unit change in x and gives

$$|E_n| \leq \frac{1}{K}(e^{Kx} - 1) \times 10^{-4}.$$

For $K = 2$, $x = 1$ we get $|E_n| \leq 3.1 \times 10^{-4}$, which is only 3-place accuracy. If more accuracy is needed and there is no way of decreasing R - by going to another machine or by using multiple precision programming - Euler's method must be abandoned in favor of one more complicated but more accurate. There are literally dozens of such methods and more are invented every year. Any good book on Numerical Analysis will discuss several of the most important ones.

Example 2. $y' = x - y^2$, $y(0) = 0$, $y(2) = ?$

This is the equation discussed graphically in Section 2. We see from the discussion and Figure 2-3 that y(x) is an increasing function whose value at x = 2 is roughly between 1 and 1.4. To oet the value of M we need to know something about y"(x). y" is obtained by differentiating the differential equation with respect to x, thus:

$$y'' = 1 - 2yy' = 1 - 2y(x - y^2).$$

$y''(0) = 1$. Since y is increasing, $y' \geq 0$ and $y \geq 0$; hence y" decreases. A few trial points taken from Figure 2-3 are enough to convince one that y" never gets close to -1 for x in the range [0,2], and so we can take M = 1.

To get a value for K we use the result of Problem 1(a) of Section 3. Since $f_y'(x,y) = -2y$ is bounded by -2.8 and 0, we see that the values of K_n are similarly bounded and so we can take K = 0. Then Problem 1 of this section gives as the bound on $|E_n|$,

$$|E_n| \leq n(\frac{1}{2} h^2M + R) = x\left(\frac{h}{2} + \frac{R}{h}\right).$$

For x = 2 the error bound is then simply h, plus the roundoff contribution.

In Table 4-4, columns Y1 and Y2 give the values of Y at corresponding values of x, for h = .1 and h = .0005 respectively. By our results above, Y2 should be accurate to 3D; hence the next column, giving the differences of Y1 and Y2, is an estimate of the errors in Y1.

(The machine which produced this table has $R < 10^{-12}$, so the roundoff error is negligible.)

APPROXIMATE SOLUTIONS OF DY/DX = X-Y*Y

X	Y1 H=.1	Y2 H=.005	Y2-Y1	Y3 H=.00025	Y3-Y2
0.0	0.000000	0.000000	0.000	0.000000	0.000000
.1	0.000000	.004975	.005	.004987	.000012
.2	.010000	.019934	.010	.019959	.000025
.3	.029990	.044805	.015	.044842	.000037
.4	.059900	.079395	.019	.079444	.000048
.5	.099541	.123344	.024	.123403	.000059
.6	.14855[illegible]	.176079	.028	.176147	.000067
.7	.206344	.236794	.030	.236868	.000074
.8	.272036	.304645	.032	.304523	.000078
.9	.344683	.377780	.033	.377859	.000079
1.0	.422802	.455392	.033	.455468	.000076
1.1	.504925	.535788	.031	.535860	.000072
1.2	.589431	.617485	.028	.617550	.000065
1.3	.674688	.599090	.024	.599145	.000056
1.4	.759158	.779371	.020	.779417	.000046
1.5	.841534	.857321	.016	.857357	.000036
1.6	.92071[illegible]	.932175	.011	.932201	.000026
1.7	.995944	1.003418	.007	1.003435	.000017
1.8	1.066754	1.070765	.004	1.070775	.000009
1.9	1.132957	1.134130	.001	1.134133	.000003
2.0	1.194598	1.193530	-.001	1.193578	-.000002

TABLE 4-4

Actually the error in Y_1 at $x = 2$ is very small. On the other hand, the error at $x = 1$, .033, is in good agreement with the computed bound, which is $h/2 = .05$. The cause of the decrease beginning at about $x = 1.1$ is the change in the curvature of the solution near this point. (See Figure 4-2 and Problem 7 of Section 10-1.) At first the curve is convex, and Y_1 underestimates, as in Figure 4-1. Then the curve becomes concave and Y_1 overestimates, thereby gradually cancelling the previous errors. From this point on the curve is fairly flat, i.e. M is small, and the error build-up in Y_1 will be less than the above estimate.

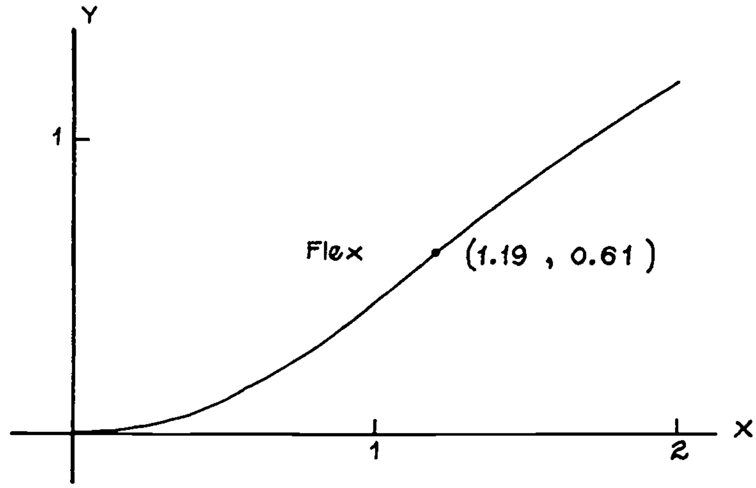


Figure 4-2

It is important to notice that the discussion in the above example is not mathematically rigorous. We have not <u>proved</u> that y(2) lies between 1.4 and 1, nor that -1 is a lower bound of y". This can be done in this simple example but in general the difficulties would be too great to be justified. Instead we usually proceed as we do for Simpson rule integration; we get a solution for a value of h hopefully small enough, repeat the process with a value of h half as big, and compare the results. This was done in columns Y2 and Y3 of Table 4-1. The difference, tabulated in the next column, indicate that the values are almost surely accurate to 3 places.

Proof of (8). We start with

$$(7) \qquad E_{n+1} = (1 + hK_n)E_n + T_n - R_n,$$

where

$$(10) \qquad 0 < 1 + hK_n \leq 1 + hK, \quad |T_n| \leq \tfrac{1}{2}h^2M, \quad |R_n| \leq R,$$

and $E_0 = 0$ since $Y_0 = y_0$. Consider quantities F_n defined by the recursion formula

$$(11) \qquad F_{n+1} = (1 + hK)F_n + \tfrac{1}{2}h^2M + R, \quad F_0 = 0.$$

It is easy to see that if $|E_n| \leq F_n$ then $|E_{n+1}| \leq F_{n+1}$. For, from (7) and (10),

$$\begin{aligned} |E_{n+1}| &\leq |1 + hK_n||E_n| + |T_n| + |R_n| \\ &\leq (1 + hK)|E_n| + \tfrac{1}{2}h^2M + R \\ &\leq (1 + hK)F_n + \tfrac{1}{2}h^2M + R \\ &= F_{n+1}. \end{aligned}$$

Since $|E_0| \leq F_0$ it follows that

(12) $$|E_n| \leq F_n$$

for all n.

To get (8) we have to solve (11). This equation is of the form

(13) $$F_{n+1} = aF_n + b,$$

where a and b are constants. We can simplify it still further by adding a suitable constant to F_n, i.e. let $G_n = F_n + c$. In terms of G_n, (13) becomes

$$G_{n+1} = aG_n - ac + b + c, \quad G_0 = c.$$

Taking $c = b/(a - 1)$ leaves

$$G_{n+1} = aG_n, \quad G_0 = b/(a - 1).$$

Now, obviously,

$$G_1 = aG_0, \; G_2 = aG_1 = a^2G_0, \; \ldots \; , \; G_n = a^nG_0,$$

and so the solution of (13) is

$$F_n = a^n \frac{b}{a - 1} - \frac{b}{a - 1} = \frac{b}{a - 1}(a^n - 1).$$

Finally, putting

$$a = 1 + hK, \quad b = \frac{1}{2}h^2M + R,$$

gives as the solution of (11),

$$F_n = \left(\frac{hM}{2K} + \frac{R}{hK}\right)\left[(1 + hK)^n - 1\right] \quad .$$

(8) then follows from (12).

PROBLEMS

1. (a) Show that taking limits in (8) as $K \to 0$ gives

$$|E_n| \le \left(\frac{hM}{2} + \frac{R}{h}\right) nh.$$

   (b) For $x = Nh$ discuss $|E_N|$ as h goes to zero with fixed x, and as x increases with fixed h.

2. (a) Write a flow diagram for the recursion process

$$Y_{n+1} = Y_n + hf(x_n, Y_n),$$

$$x_{n+1} = x_n + h,$$

   with initial values $Y_0 = y_0$, x_0. Output the successive values of x and Y.

   (b) Write a program from your flow diagram.

   (c) Test your program with the two cases of Example 1.

3. Modify your program in Problem 2 to output x and Y at every M steps rather than at every step. This enables you to use very small values of h without unduly depleting the forests of America.

4. Use your program in Problem 2 or 3 to make reasonable tabulations and graphs of the solutions of the following equations over the given integrals. Discuss any feature of the solution that seems unusual.

(a) $y' = \frac{5x + y}{x - 5y}$, $y(0) = -1$,

(i) for $0 \leq x \leq 1$, (ii) for $0 \leq x \leq 1.5$.

(Compare with Section 2, Problem 1(e)).

(b) $y' = \frac{\sqrt{y^2 - 1}}{x^2 - 1}$, $y(0) = 2$,

(i) for $-.9 \leq x \leq .5$. [Use negative values of h for $x < 0$.]

(ii) for $0 \leq x \leq .9$.

(Compare with Section 7, Example 2).

(c) $\frac{dy}{dt} = 2g(t) - \frac{1}{10}\sqrt{y}$, $y(0) = 0$, $0 \leq t \leq 1000$,

$$g(t) = \begin{cases} 1 - \cos \frac{t}{29} & \text{if } 2n \leq \frac{t}{182} \leq 2n+1 \\ 0 & \text{if } 2n+1 < \frac{t}{182} < 2n+2 \end{cases}, \quad n = 0, 1, 2, \ldots$$

(Compare with Section 5, Example 1).

5. Use your program to solve Problem 3(d) of Section 2.

6. (a) Use Euler's method with $h = .2$ to estimate $y(1)$ if

$$y' = xy + 1, \quad y(0) = -1.$$

Tabulate x, y, and y', and carry only two decimal places in your calculations.

(b) Use your values of x, y, y' to compute y'', and estimate values for M, K, and R.

(c) Determine the possible error in your approximation to $y(1)$.

7. (a) Solve $y' = xy + 1$, $y(0) = -1$, by expressing y as a Maclaurin series, as illustrated in Section 13-5.

(b) Compute $y(1)$ to two decimal places. Ans. -0.24

(c) Use the result of (b) to get the actual error in 5(a) to two places, and compare with the estimated possible error in 5(c). Is the latter a reasonable bound, much too pessimistic, or not a bound at all?

8. (a) Is Euler's method the best numerical way to find y(x) when given

$$y' = f(x), \quad y(x_0) = y_0?$$

Describe a better one.

(b) Flow chart and program the best method you can thir of.

(c) Solve

$$y' = \sqrt{1 + x^3}, \quad y(1) = 2,$$

to 5D accuracy for x = 1(.1)5

5. Applications.

A differential equation is apt to arise in a mathematical model of almost any problem involving continuously changing quantities. As is the case with all applications of mathematics the solution of such a problem involves the three steps of setting up the model, solving the equation, and interpretting the results. The rest of this chapter treats step 2. In this section we are mainly concerned with step 1, more particularly with the last half of step 1. The formulation of a model involves first the acceptance of some simplifying approximations to the true situation, and then the expression of the simplified picture in mathematical terms.

The examples and problems of Section 9-2 illustrate the two most common ways in which a differential equation is set up as a model of the approximation of a physical system. In one of these (Examples 3 and 4 and Problem 5 to 12) the derivative of a quantity enters directly, usually as a slope or a rate of change, and the differential equation results from a relation between this derivative, the quantity itself, and the independent variable. Motion problems, involving distance, velocity, and acceleration, are typical of this mode of formulation.

The other method, illustrated by Examples 1 and 2 of Section 9-2, makes no direct use of derivatives but analyzes the problem in terms of small changes of the variable, attempting so to formulate the situation that by letting these changes approach zero an equation involving one or more derivatives appears. This technique is more general than the former but is also less direct and often tricky to handle. We shall give examples of both methods.

In some of the examples and problems, notably Example 2 and its related problems, we consider the realism of the approximations made in step 1, by interpretting the solution of the differential equation back into physical terms. Results that are physically impossible indicate the need of changes in the simplifying assumptions.

Example 1. The Allukaw river, like the Nile, has a strongly seasonal flow, approximated by

$$f(t) = 10^8(1 + \sin \frac{t}{58}) \text{ cu ft/day} ,$$

t being measured in days from January 1. To level off the flow we build a dam 200 ft high, holding 2×10^{10} cu ft, to impound the water. The water runs out of an opening at the base of the dam he rate of $10^7\sqrt{y}$ t/day, y being the height of water in the dam. If we imagine the reservoir

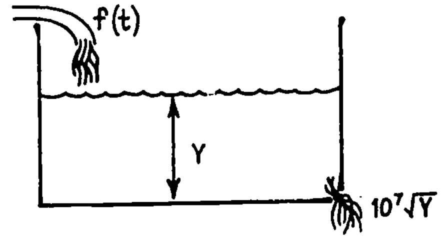


Figure 5-1

to behave simply like a tank, as in Figure 5-1, then the area of the base is $2 \times 10^{10}/200 = 10^8$ sq ft, and the volume of water in it is $10^8 y$. So

$$\frac{d}{dt}(10^8 y) = 10^8\left(1 + \sin \frac{t}{58}\right) - 10^7\sqrt{y} ,$$

or

$$\frac{dy}{dt} = 1 + \sin \frac{t}{58} - \frac{1}{10}\sqrt{y} .$$

A machine solution, using Euler's method with h = 1 but printing only every 20 steps, gives the curves in Figure 5-2. It was assumed that the reservoir was empty on January 1. No that the annual variation in flow has been reduced from 2×10^8 to $.3 \times 10^8$.

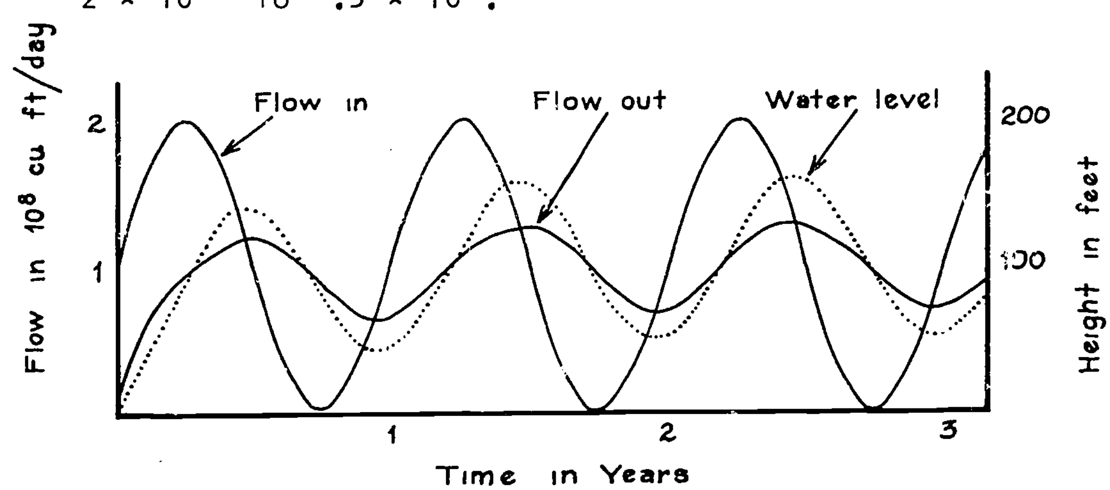


Figure 5-2

Example 2. R rabbits are running around in Australia. We make the following approximations regarding the change in R:

(a) The rabbits are spread uniformly throughout Australia;

(b) When two rabbits of opposite sex meet they produce r more rabbits;

(c) The average lifetime of a rabbit is b years.

From (a) the chance of a rabbit meeting a member of the opposite sex in a given time is proportional to R. Hence the total number of such meetings in a given time is proportional to R^2. By (b) the rate of increase by births is proportional to R^2. By (c), 1/b of the rabbits die each year. Hence, measuring time in years,

$$\frac{dR}{dt} = aR^2 - \frac{R}{b}.$$

The solution of this equation is analyzed in Problem 1.

This is of course a very crude model since our assumptions are grossly oversimplified. For one thing, we have neglected the food supply - If there are too many rabbits some of them starve. We can include this point by making b a decreasing function of R, for instance $b = c(1 - R/R_M)$, where R_M is an upper bound to the number of rabbits that could possibly live in Australia. If we want to get fancy we can let R_M be a function of t, to take care of climatic variation.

<u>Example 3</u>. A long rope of variable cross-section hangs vertically and supports a weight W at its lower end. (Figure 5-3). At distance x from this end let $A(x)$ be the area of cross-section of the rope, $\rho(x)$ the density of the material, and $S(x)$ the stress in the rope in force per unit area. Consider the forces acting on the portion of the rope between x and $x + \Delta x$. (Figure 5-4). Pulling upward at the top is the force

$$S(x + \Delta x)A(x + \Delta x)$$

of the stress at this point. Pulling downward is $S(x)A(x)$ plus the weight of the piece of

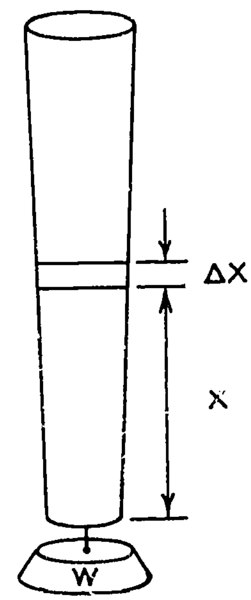


Figure 5-3

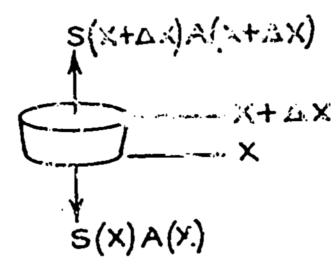


Figure 5-4

rope. This weight w is bounded,

$$g\rho(x_1)A(x_2)\Delta x \leq w \leq g\rho(x_3)A(x_4)\Delta x,$$

where $\rho(x_1)$, $A(x_2)$ and $\rho(x_3)$, $A(x_4)$ are the minima and the maxima of the density and the area functions in the interval $[x, x + \Delta x]$. Assuming that the functions ρ and A are continuous, it follows that there are x_5 and x_6 in $[x, x + \Delta x]$ such that

$$w = g\rho(x_5)A(x_6).$$

We must therefore have

$$S(x + \Delta x)A(x + \Delta x) = S(x)A(x) + g\rho(x_5)A(x_6).$$

Dividing by Δx and letting $\Delta x \to 0$ gives

$$\lim_{\Delta x \to 0} \frac{S(x + \Delta x)A(x + \Delta x) - S(x)A(x)}{\Delta x} = \lim_{\Delta x \to 0} g\rho(x_5)A(x_6)$$

or

(1) $$\frac{d(S(x)A(x))}{dx} = g\rho(x)A(x).$$

The initial condition is $T(0)A(0) = W$.

(a) If A is constant the equation becomes

$$S' = g\rho(x),$$

$$AS = W + \int_0^x Ag\rho(t)dt.$$

Thus the tension AS merely builds up with the weight of the rope below that point.

(b) For constant ρ, how should A vary so that S is constant? This is the "most economical" design, since S can be kept just below the breaking point. We get

$$SA' = g\rho A,$$

which integrates to

$$A = \frac{W}{S} e^{x/k}, \quad k = \frac{S}{g\rho}.$$

For steel, k ≈ 15,000 ft., so a vertical cable three miles long would have to have a cross-sectional area 2.7 times as large at the top as at the bottom.

(c) Suppose the rope is elastic, so that it stretches under tension. Assume that ρ is constant throughout the stretching and that A has the constant value A_0 in the unstretched state. The two states of a portion of the rope are shown in Figure 5-5. Hooke's

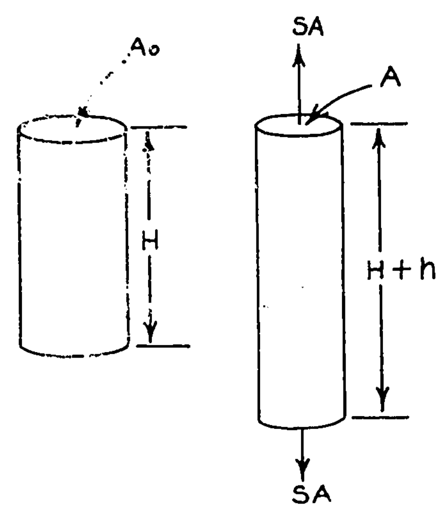


Figure 5-5

Law says that

$$h = cHSA$$

where c is a constant. Since we are assuming that the density is not changed, the volume must be the same in the two states; that is,

$$A_0H = A(H + h) = A(H + cHSA)$$

which gives

$$(2) \qquad A_0 = A(1 + cSA).$$

To combine this with (1) put it in the form

$$cSA = \frac{A_0}{A} - 1.$$

Then (1) gives

$$-\frac{A_0}{A^2} A' = cg\rho A,$$

or

$$-\frac{A'}{A^3} = \frac{cg\rho}{A_0} .$$

In equations involving physical quantities it is often convenient to work with "dimensionless" variables, defined as the quotient of two variables of the same physical dimensions. In the present case let us introduce the variable $y = A/A_0$. Replacing A in the above equation by $A_0 y$, we get

(3) $$-y^{-3}y' = k, \qquad k = cg\rho A_0.$$

Separating variables and integrating gives

$$y^{-2} - y(0)^{-2} = 2kx,$$

or

$$y = (y(0)^{-2} + 2kx)^{-1/2} .$$

To get (0) we combine (2) with $W = A(0)S(0)$. This gives

$$y(0) = \frac{A(0)}{A_0} = \frac{1}{1 + cW} .$$

So our final solution is

$$A = A_0(b^2 + 2kx)^{-1/2} ,$$

where $b = 1 + cW$ and $k = cg\rho A_0$.

Example 4. Newton's second law of motion says that the rate of increase of momentum of any body is equal to the force acting on the body. Consider a rocket whose mass and velocity at time t are M(t) and V(t), acted on by a force F (such as gravity, air resistance, ect.). At time t the momentum of the rocket is $M(t)V(t)$. At time $t + \Delta t$ this "body" has separated into two parts, the rocket at time $t + \Delta t$ and the portion of the fuel that was burned and ejected as gas in the interval Δt. The mass of the exhaust gas is

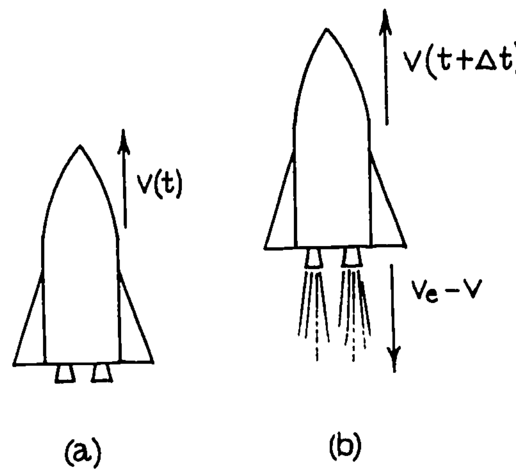


Figure 5-6

$$M(t) - M(t + \Delta t) = - \Delta M$$

and its velocity is $-(V_e - V)$, where V_e is the "exhaust velocity", a constant depending on the design of the rocket and the kind of fuel. More precisely, the momentum of the exhaust gas is bounded,

$$-(V_e - V(t_1)) \leq \text{Mom.} \leq -(V_e - V(t_2)),$$

where t_1 and t_2 are values of t in the interval $[t, t + \Delta t]$ at which $V(t)$ has minimum and maximum values. Since V is continuous there is a t_3 in this interval such that the momentum of the exhaust gas is $-(V_e - V(t_3))\Delta M$.

The increase of momentum in the interval Δt is then

$$M(t + \Delta t)V(t + \Delta t) - (V_e - V(t_3))\Delta M - M(t)V(t).$$

To get the rate of change of momentum we divide by Δt and let $\Delta t \to 0$; thus

$$\lim_{\Delta t \to 0} \left[\frac{M(t + \Delta t)V(t + \Delta t) - M(t)V(t)}{\Delta t} - (V_e - V(t_3)) \frac{\Delta(-M)}{\Delta t} \right]$$

$$= \frac{d}{dt} (M(t)V(t)) + (V_e - V(t)) \frac{dM}{dt} ,$$

since $t_3 \to t$ as $\Delta t \to 0$. By Newton's law we then have

$$MV' + M'V + V_e M' - VM' = F,$$

or

(4) $MV' = -V_e M' + F.$

There are several interesting special cases of this rocket equation.

(a) $F = 0$, i.e. motion in free space. We can write (4) as

$$M \frac{dv}{dt} = -V_e \frac{dM}{dt} ,$$

or

$$\frac{dM}{M} = -\frac{dV}{V_e} .$$

Assuming $M = M_0$ when $V = 0$ we get by integrating,

$$\log M - \log M_0 = -\frac{V}{V_e} ,$$

or

$$\frac{M_0}{M} = e^{V/V_e} .$$

M_0/M is the "mass ratio", the ratio of the initial to the final mass required to attain velocity V. For example, if $V_e = 10{,}000$ ft/sec and $V = 36{,}000$ ft/sec, the velocity needed to escape from earth's gravity, then $M/M_0 \approx 37$. In other words about 97% of the initial mass must consist of fuel. This large mass ratio is the reason why rockets cost so much.

(b) If the fuel is burned at a constant rate then $M = M_0 - ct$. The equation becomes

$$V' = \frac{cV_e + F}{M_0 - ct} .$$

For a rocket moving through the air at subsonic speed the resistance F is of the form $- kV^2$. The resulting equation is

$$V' = \frac{cV_e - kV^2}{M_0 - ct} .$$

Some conclusions that can be drawn in this case are treated in Problem 4.

PROBLEMS

1\. (a) Solve the rabbit equation

$$\frac{dR}{dt} = aR^2 - \frac{R}{b},$$

with inital condition $R(0) = R_0$, by the method of separation of variables introduced in Section 9-2.

Ans. $R = R_0\left[k - (k - 1)e^{t/b}\right]^{-1}$, $k = abR_0$.

(b) Show that if $R_0 < \frac{1}{ab}$ the rabbits die off, but if $R_0 > \frac{1}{ab}$ they become infintely numerous in a finite time.

2\. Using the computer, investigate the suggested equation

$$\frac{dR}{dt} = aR^2 - \frac{R}{c(1 - R/R_M)}.$$

(a) Using the values

$R_M = 3 \times 10^8 = 100/\text{sq mile}$;

$c = 5$, an average life of 5 years;

$a = 10^{-4}$, adjusted to try to get a reasonable answer;

$R_0 = 10^4$;

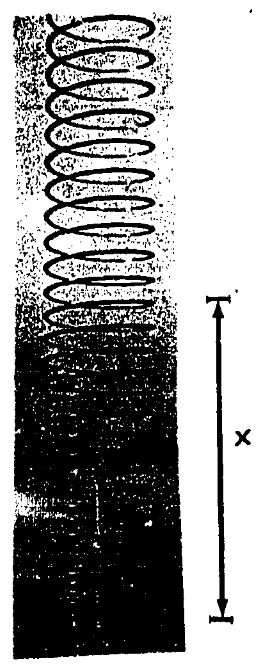
draw a graph of the growth for 100 years.

(b) Show that at the equilibrium situation the average life span of a rabbit, as given by $b = c(1 - R/R_M)$, is ridiculously small. To avoid this we must replace a by a function that is fairly constant until b gets down to, say, 1/2, and then decreases rapidly as b goes to zero. Try constructing such functions and run experimental trials with various parameters to see if you can get a realistic population growth.

3. A coiled spring (see figure) behaves like an elastic rope with the following changes:

(i) The quantity SA is replaced by the tension T;

(ii) Aρ is replaced by the linear density λ, mass per unit length.



Equations (1) and (2) then become

$$T' = g\lambda,$$

$$\lambda_0 = \lambda(1 + CT).$$

(a) Putting $y = \lambda/\lambda_0$ show that these give tho same equation (3), with suitable definition of k, and the same value of y(0).

(b) Suppose z is the length of an unstretched spring that stretches into the piece of length x in the stretched condition. Since mass is preserved in the stretching,

$$\lambda_0 z = \int_0^x \lambda(u)\,du.$$

Using the solution of (3), show that this gives

$$x = \frac{k}{2}z^2 + bz.$$

This is well illustrated with a Slinky, using $w = 0$. k is large enough, due to a large value of c, to spread the coil out nicely and give the effect pictured in the previous figure.

4. (a) Separate variables in the equation

$$V' = \frac{cV_e - kV^2}{M_0 - ct}$$

of Example 3(b) and solve with initial condition $V(0) = 0$ to get

$$(5) \quad V = \frac{bc}{2k} \frac{1 - \left(1 - \frac{ct}{M_0}\right)^b}{1 + \left(1 - \frac{ct}{M_0}\right)^b}, \quad b = 2\sqrt{\frac{kV_e}{c}} .$$

(b) Letting r be the mass ratio

$$r = \frac{M_0}{M} = \frac{M_0}{M_0 - ct},$$

reduce the above equation to the form

$$V = \frac{bc}{2k}\left(1 - \frac{2}{r^b + 1}\right)$$

and sketch the graph of V as a function of r.

5. If we neglect air resistance a rocket ascending vertically is acted on by a force $F = -Mg$.

(a) Set up the equation of motion if $M = M_0 - ct$, as in Example 3(b).

(b) Integrate the equation, assuming $V(0) = 0$.

Ans. $V = -V_e \log\left(1 - \frac{ct}{M_0}\right) - gt.$

(c) Integrate $s' = V$ to get the height s at time t.

(d) For a given mass ratio $r = \frac{M_0}{M}$ we have $t = \frac{M_0 - M}{c} = \frac{M}{c}(r - 1)$. Show that this gives

$$V = V_e \log r - \frac{gM}{c}(r - 1),$$

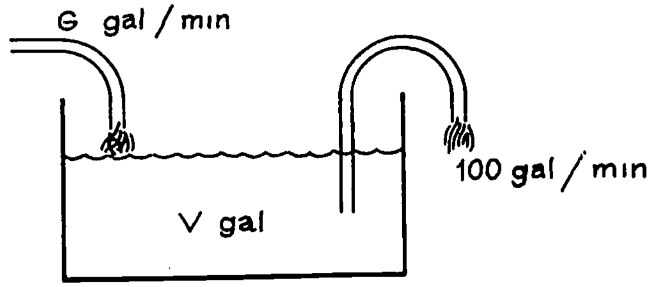
$$s = \frac{V_e M}{c}(r - 1 - \log r) - \frac{gM^2}{2c^2}(r - 1)^2.$$

With this velocity the rocket will coast to an additional height of $\frac{V^2}{2g}$. Show that the total height attained is

$$S = \frac{(V_e \log r)^2}{2g} - \frac{V_e M}{c}(r \log r - r + 1)$$

Show that $r \log r - r + 1$ is positive for $r > 1$ and hence that the maximum S is thus obtained by making c as large as possible. What are some limiting factors on the value of

6. A tank contains V gallons of salt solution. G gal/min of a solution containing .5 lb/gal is pumped in, and 100 gal/min of the solution in the tank is pumped out. The solution in the tank is stirred constantly and may be considered to be of uniform concentration. At the beginning of the process the tank contains 10,000 gal of fresh water. We want to know the concentration of the solution in the tank after t minutes.



(a) Assume G is constant, G = 100 gal/min. [Hint. Set up a differential equation for S, the amount of salt in the tank.] Ans. $C = \frac{1}{2}\left(1 - e^{-t/100}\right)$.

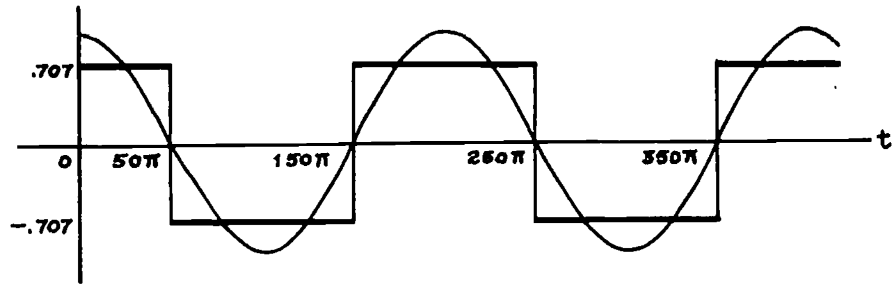
(b) Assume G fluctuates as given by

$$G = 100 + 50 \cos \frac{t}{100} .$$

[Hint. First find V as a function of t. The equation for S will have to be integrated numerically.]

Partial answer: V = 5000(2 + sin t/100).

(c) In (b) replace cos t/100 by the step function,



having the same root mean square, and solve the problem. Considering that either of these two cases might be used as an approximation to an oscillating input, has one of them any advantages over the other?

(d) Run solutions with .707 replaced by other constants and see if there is one that gives better agreement with the sine curve.

7. Two gasses combine to form a solid,

$$X + Y \rightarrow Z.$$

(a) Give an argument like that in Example 2 to show that

$$\frac{dx}{dt} = -axy,$$

where x and y are the concentrations of X and Y, and a is a constant. State any assumptions you make.

(b) Give an argument to show that $x - y$ is a constant p. By interchanging the roles of x and y, if necessary, we can assume $p \geq 0$.

(c) Solve the differential equation.

Partial answer: $x = \dfrac{p}{1 - (y_0/x_0)e^{-apt}}$ if $p > 0$.

8. In Problem 7 suppose that z is also a gas, which can spontaneously decompose into X + Y.

(a) Derive an equation

$$\frac{dx}{dt} = -axy + bz.$$

(b) Show that, with $x - y = p$, as before, we also have $x + z = q$.

(c) Reduce the differential equation to

$$\frac{dx}{x^2 + (p + r)x - rq} = -a\,dt, \quad r = b/a.$$

The quadratic polynomial in x must have one positive and one negative root. (Why?) Let α

and $-\beta$ be the roots and solve for x.

Ans. $x = \dfrac{\alpha + \beta k e^{-\gamma t}}{1 - k e^{-\gamma t}}$, $\gamma = a(\alpha + \beta)$, $k = \dfrac{x_0 - \alpha}{x_0 + \beta}$.

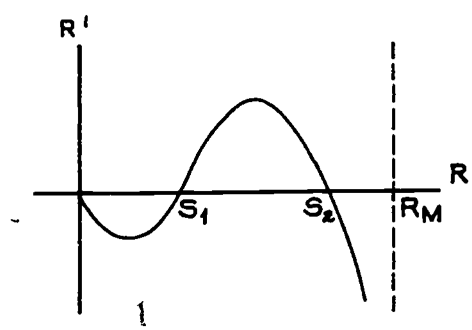
9. (a) The dynamic systems in Problems 7 and 8 approach "steady states" as $t \to \infty$. Show that the steady state can be obtained directly from the differential equation by setting $\frac{dx}{dt} = 0$. This is characteristic of steady states.

(b) Find the steady states of the simple rabbit equation in Example 2. Is a steady state necessarily one approached as $t \to \infty$? Discuss the notion of "stable" and "unstable" steady states.

(c) For the second rabbit equation,

$$R' = aR^2 - \frac{R}{c(1 - R/R_M)} ,$$

show that the graph of R' versus R has the shape shown. Hence show that for $S_1 < R_0 < R_M$ the population will approach the steady state



$R = S_2$, but for $0 < R_0 < S_1$ it will approach zero. What are the steady states, and which ones are stable?

10. Newton's law of cooling says that the transfer of heat from a body at temperature θ to its surroundings at temperature θ_s is proportional to $\theta - \theta_s$. If we assume the temperature of a body is uniform throughout and that the heat content is proportional to the temperature, this gives the equation

$$\frac{d\theta}{dt} = -k(\theta - \theta_s) .$$

The equation holds regardless of the sign of $\theta - \theta_s$.

(a) A pie is taken from an oven at a temperature of 350°F and set to cool in an atmosphere of 70°. In 45 minutes it is barely eatable, say 150°. When will it reach 90°? Ans. 94.8 min.

(b) An identical pie made at the same time is set in the draft of an oscillating fan which has the effect of multiplying k by a factor of $2 + \cos t/2$. When does this pie reach 90°? Ans. 48.7 min.

11. A body subject to Newton's law of cooling contains a small amount of radioactive material, which adds heat to the body at a constant rate. Show that this is equivalent to a non-radioactive body in a higher surrounding temperature.

12. Over a series of days the air temperature is approximately

$$\theta_s = 75 + 15 \sin 2\pi t.$$

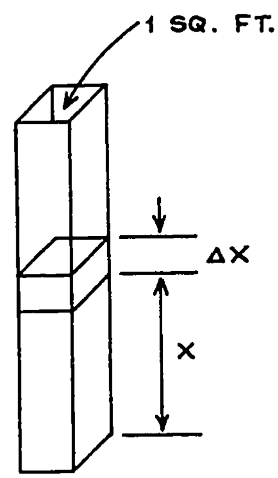
A closed car standing in the shade has a cooling coefficient of $k = 2$. Assuming that the car's temperature is 75° at $t = 0$, graph its temperature over a period of four days.

13. Consider a vertical column of atmosphere of 1 sq ft. cross-section. At height x above the ground let $\rho(x)$ be the density and $p(x)$ the pressure.

(a) Explain why

$$p(x) = p(x + \Delta x) + g\rho(x_1)\Delta x,$$

where x_1 is in $[x, x + \Delta x]$.



(b) Derive the differential equation

$$\frac{dp}{dx} = g\rho .$$

(c) If the atmosphere is at constant temperature then ρ is proportional to p. Solve for p as a function of x, given that p(0) = 15 lb/sq in. and $g\rho(0)$ = .08 lb/cu ft.

6. Systems of Equations.

In more elaborate problems than those considered in the last section differential equations can arise in forms other than the standard $y' = f(x,y)$. Consider, for example, the equation

$$(1) \qquad y'^2 + y^2 = x^2,$$

involving the independent variable x, a function $y = y(x)$, and its derivative $y' = y'(x)$. Equation (1) is equivalent to the combined statement

$$(2) \qquad y' = \sqrt{x^2 - y^2} \qquad \text{or} \qquad y' = -\sqrt{x^2 - y^2}\ .$$

Any curve that satisfies (2) at each point $(x,y(x))$ is a solution of (1): note that it is allowed to satisfy a different part of (2) at different points. A simpler example to see is

$$(3) \qquad y'^2 - (y + 1)y' + y = 0,$$

which gives by factoring

$$(4) \qquad y' - 1 = 0 \qquad \text{or} \qquad y' - y = 0.$$

The curve (Figure 6-1)

$$y = \begin{cases} x + 1 & \text{if } x \leq 0 \\ e^x & \text{if } x \geq 0 \end{cases}$$

satisfies (4), and hence (3), at all points. By combining the solid and dotted curves in the figure, four solutions satisfying f(0) = 1 can be obtained. This does not contradict the fundamental theorem since (3) is not in the form for applying this theorem. Each part of (4) is; but each part of (4) taken as a separate equation, has a unique solution through each point.

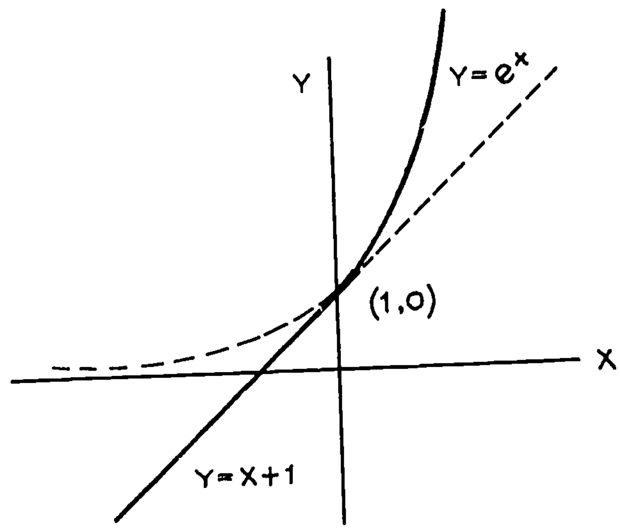


Figure 6-1

This example illustrates some of the complexities that can arise if we consider differential equations of the form $F(x,y,y') = 0$. In most cases, however, the routine, but tedious, procedure of the following example can be used.

Example 6-1. Solve

(5) $\cos y' + \frac{1}{2}y\, y' = x, \quad y(1) = 1.$

Here we cannot solve for y' explicitly, as in (2), but we can proceed by considering y' as an implicit function of x and y defined by the given equation. For $x = 1$, $y = 1$ equation (5) becomes

(6) $\cos y' = 1 - \frac{1}{2}y',$

which has solutions (Figure 6-2)

$y' = 0$, $y' \approx 1.11$, $y' \approx 3.70$.

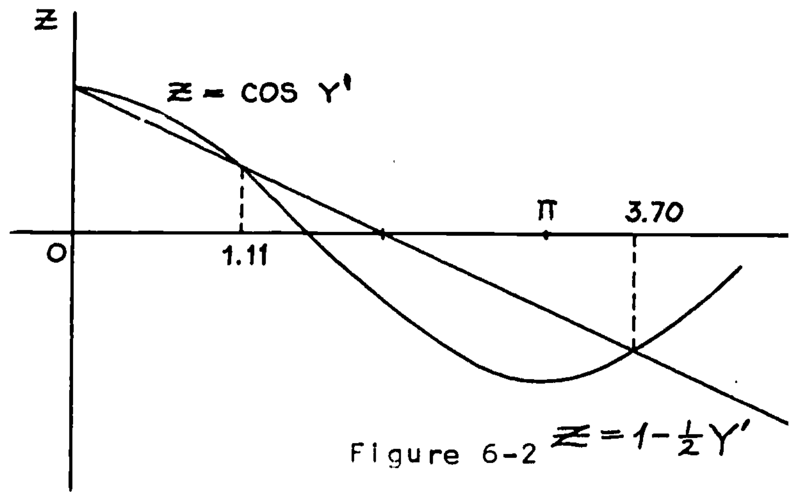


Figure 6-2

The last two values can be refined as much as needed by applying Newton's method to (6). The three values for y' imply that through the point (1,1) there pass three solutions of (5). Let us concentrate on the one with $y' = 1.11$.

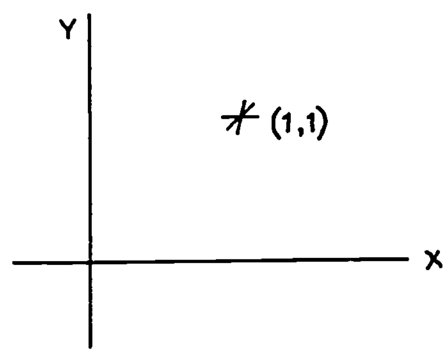


Figure 6-3

One step of Euler's method, with $h = .1$, gives

$x_1 = 1.1$, $y_1 = 1.111$, and

(7) $\cos y' = 1.1 - .617\, y'$.

Now (7) also has three solutions but we are interested in only one of them. For since we have assumed that y' is a continuous function of x, the value of y' at $x = x_1$ will be close to that at $x = x_0$, and so we want the solution of (7) that is close to 1.11. In other words, 1.11 should be (if our value of h is sufficiently small) a good first approximation in applying Newton's method to (7). Two applications of Newton's method gives $y' = 1.24$ correct to 2D. We can now get

$$y_2 = 1.11 + .1(1.24) = 1.23$$

and continue the process. Trouble can arise only when (5), as an equation in y', acquires two equal roots. At the next step they will (generally) separate again, and we do not know which one to follow. This is precisely the situation that causes the multiple solutions in the previous example (see Problem 1), and conditions outside the mere statement of the differential equations are needed to decide which path to take.

Of considerable more importance than the implicit equations, $F(x,y,y') = 0$, are the systems of equations. Here we have more than one unknown function and an equal number of equations. The standard form for a system of equations is analogous to that for a single equation, that is, the derivatives are expressed as functions of the variables. For example, for three equations in the three unknown functions $x(t)$, $y(t)$, $z(t)$, we have

$$
\begin{aligned}
x' &= f(t,x,y,z), \\
y' &= g(t,x,y,z), \qquad (8) \\
z' &= h(t,x,y,z).
\end{aligned}
$$

For a more general notation we can use a subscript notation, thus:

$$
\begin{aligned}
y_1' &= f_1(x,y_1,y_2,\ldots,y_n), \\
y_2' &= f_2(x,y_1,y_2,\ldots,y_n), \\
&\cdots\cdots\cdots\cdots \\
y_n' &= f_n(x,y_1,y_2,\ldots,y_n),
\end{aligned}
$$

or, more compactly,

(9) $y_k' = f_k(x,y_1,y_2,\dots y_n)$, $k = 1,2,\dots,n.$

For such systems of equations the fundamental theorem and Euler's method carry over in the simplest possible way. Thus we have

Theorem 1. If, in an $(n + 1)$-dimensional region R in $(x,y_1,y_2,\dots,y_n)$-space, each of the functions f_k in (9) is continuous and is Lipschitzian in each of the y_k, then for any point $(x_0,y_{1,0},y_{2,0},\dots y_{n,0})$ in R there is an $H > 0$ such that (9) has a unique solution $y_k(x)$, $k = 1,\dots,n$, in $|x - x_0| < H$ with $y_k(x_0) = y_{k,0}$, $k = 1,\dots,n.$

The proof of this theorem is essentially no more difficult than that of the simpler theorem in Section 3.

Euler's method is equally easy to generalize. For simplicity let us take equations (8). The basic recursion formula is

$$
(10) \quad
\begin{aligned}
x_{n+1} &= x_n + s\ f(t_n, x_n, y_n, z_n), \\
y_{n+1} &= y_n + s\ g(t_n, x_n, y_n, z_n), \\
z_{n+1} &= z_n + s\ h(t_n, x_n, y_n, z_n), \\
t_{n+1} &= t_n + s,
\end{aligned}
$$

with the initial values, t_0, x_0, y_0, z_0 being given. One important warning: in programming we generally omit the subscripts and use, for instance,

$$x \leftarrow x + s\ f(t,x,y,z)$$

instead of the first equation of (10). This is all right, but then to use

$$y \leftarrow y + s\ g(t,x,y,z)$$

for the second equation is wrong. The above assignment is equivalent to

$$y_{n+1} = y_n + s\ g(t_n, x_{n+1}, y_n, z_n),$$

which is not the given equation. The correct programming is

$$xI \leftarrow x + s \times f(t,x,y,z)$$

$$yI \leftarrow y + s \times g(t,x,y,z)$$

$$z \leftarrow z + s \times h(t,x,y,z)$$

$$x \leftarrow xI$$

$$y \leftarrow yI$$

$$t \leftarrow t + s \; .$$

The one phase of Euler's method that does not generalize easily is the error analysis. Although the same general conclusions can be obtained the analysis is much more complicated, involving techniques quite unsuitable for our present text.

So far we have considered only first order differential equations, that is, those in which only first derivatives appear. Higher order equations, or systems of equations, can be reduced to systems of first order equations by introducing extra unknown functions to stand for the lower order derivatives. Thus

(11) $y''' + 3xy'' - e^{x}y = \cos x$

is equivalent to the system

$$y' = u,$$
$$(12) \quad u' = v,$$
$$v' = -3xv + e^x y + \cos x.$$

The system

$$(13) \quad \begin{aligned} x'' &= f(t,x,y,x',y'), \\ y'' &= f(t,x,y,x',y'), \end{aligned}$$

which arises in considering the motion of a planet through a resisting medium, can be replaced by

$$x' = u,$$
$$y' = v,$$
$$u' = f(t,x,y,u,v),$$
$$v' = g(t,x,y,u,v).$$

Theorem I tells us what kinds of initial conditions are appropriate for such equations or systems. For (12) we would want values of $y(x_0)$, $u(x_0)$, $v(x_0)$ as initial conditions, corresponding to $y(x_0)$, $y'(x_0)$, $y''(x_0)$ for (11). Similarly (13) would need $x(0)$, $y(0)$, $x'(0)$, $y'(0)$, that is, the initial position and velocity of the planet.

For higher order equations, a different type of condition arises naturally. This is typified by the trivial equation

(14) $y'' = 0.$

The solutions of this are, of course, the lines $y = ax + b$, where a and b are arbitrary constants. The standard initial conditions

$$y(x_0) = y_0, \quad y'(x_0) = z_0,$$

amount to specifying a point on the line and the slope of the line. It is well known that a line can also be specified by two points; that is, that (14) has a unique solution satisfying

$$y(x_0) = y_0, \quad y(x_1) = y_1.$$

(The x_1, y_1 used here should not be confused with the x_1, y_1 occuring in Euler's method.) Such considerations lead to the important <u>two-point boundary value problem</u>, namely:

Given a region R in the xy-plane and a differential equation

(15) $y'' = f(x,y,y'),$

under what conditions will two points in R, (x_0, y_0) and (x_1, y_1), $x_0 < x_1$, determine a unique solution of (15) on $x_0 \le x \le x_1$ with $y(x_0) = y_0$, $y(x_1) = y_1$?

This is a difficult and complicated problem even for relatively simple equations. The following example shows the kind of thing that can happen.

Example 2. $y'' = -y$, $y(0) = 0$, $y(x_1) = y_1$. It is easy to check that

$$y = a \sin x + b \cos x$$

satisfies the differential equation for any a and b. In Problem 6 of Section 7 it will be proved that these are the only solutions. To satisfy the condition $y(0) = 0$ we must have

$$0 = a \sin 0 + b \cos 0,$$

or

$$0 = b.$$

Then to satisfy the condition $y(x_1) = y_1$ we need

$$y_1 = a \sin x_1.$$

Case A: $\sin x_1 \neq 0$. Then $a = y_1/\sin x_1$ and there is a unique solution.

Case B1: $\sin x_1 = 0$, $y_1 \neq 0$. We require $y_1 = a \times 0$, which is impossible; hence, no solution.

Case B2: $\sin x_1 = 0$, $y_1 = 0$. Here we need $0 = a \times 0$, which is true for any value of a. Hence there is an infinite number of solutions.

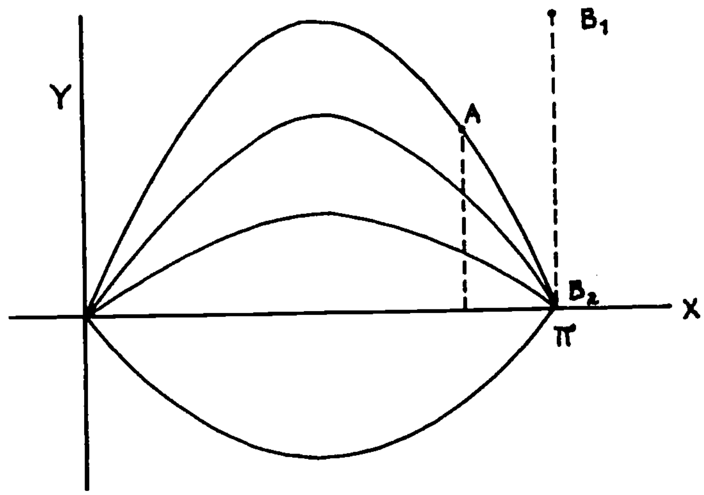


Figure 6-4

The situation is illustrated in Figure 6-4. The trouble in Case B lies in the fact that every solution that goes through (0,0) also goes through (π,0). Conditions under which situations like this occur, and the conclusions that can be drawn in these cases, constitute an important chapter in the theory of differential equations, but one that we cannot investigate here.

PROBLEMS

1. (a) Show that each solution of equation (3) has one of the forms:

(i) $y = x - a + 1,$

(ii) $y = e^{x-a}$

(iii) $y = \begin{cases} x - a + 1 & \text{if } x \le a \\ e^{x-a} & \text{if } x \ge a \end{cases}$

(iv) $y = \begin{cases} e^{x-a} & \text{if } x \le a \\ x - a + 1 & \text{if } x \ge a \end{cases}$.

(b) Show that for any point (x_0, y_0), with $y_0 \neq 1$, there is an $H > 0$ such that for $|x - x_0| < H$ there are two solutions of (5) through (x_0, y_0).

(c) Show that for any point $(x_0, 1)$ and for arbitrarily small $H > 0$ there are four solutions of (5) through $(x_0, 1)$ for $|x - x_0| < H$.

(d) Show that the exceptional points described in (c) are characterized by yielding double roots for y' as solutions of (5).

2. Set up a program, analogous to the one of Problem 2 or 3 of Section 4, to solve a system of m first order differential equations. Test it on the following, given the exact solutions.

(a) $y' = -yz, \quad y(0) = 1,$

$z' = 1/y^2, \quad z(0) = 0.$

Solve for $0 \leq x \leq 1$.

Solution: $y = \cos x$, $z = \tan x$.

(b) $x' = -x + y + z - t, \quad x(0) = 1,$

$y' = x - y + z - t, \quad y(0) = 2,$

$z' = x + y - z - t, \quad z(0) = 3.$

Solve for $0 \leq x \leq 2$.

Solution:

$$\left.\begin{matrix} x \\ y \\ z \end{matrix}\right\} = e^t + t + 1 + \left\{\begin{matrix} -e^{-2t} \\ 0 \\ e^{-2t} \end{matrix}\right.$$

3. Use the program of Problem 2 to solve the following.

(a) $y'' = -y, \quad y(0) = 0, \quad y'(0) = 1, \quad 0 \leq x \leq 8.$

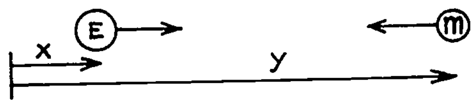
Solution: $y = \sin x$.

(b) $(1 - x^2)y'' - 2xy' + 20y = 0.$

$y(0) = 1, \quad y'(0) = 0, \quad 0 \leq x \leq .9.$

Solution: $y = \frac{1}{24}(3 - 30x^2 + 35x^4).$

(c)



$$x'' = \frac{Gm}{(y - x)^2}, \quad y'' = \frac{-GE}{(y - x)^2}, \quad x(0) = 0,$$

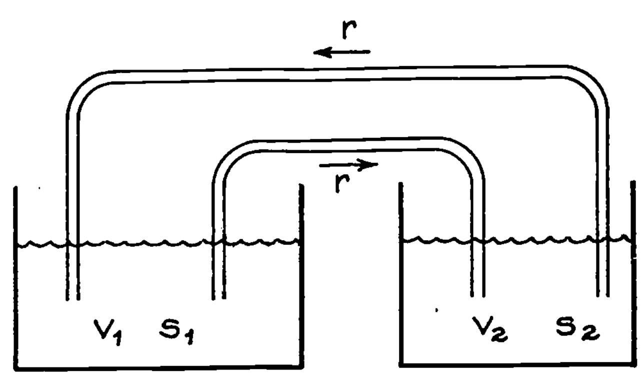
$y(0) = D, \quad x'(0) = y'(0) = 0, \quad G = 6.7 \times 10^{-11},$

$E = 6.0 \times 10^{24}, \quad m = 7.4 \times 10^{22}, \quad D = 3.8 \times 10^{8}$

In how many seconds will the moon hit the earth?

Ans. 4.1×10^5.

4. Two tanks, of capacities V_1 and V_2 gallons, are full of salt solutions. At $t = 0$ the first tank contains S_0 lbs of salt in solution, the second has pure water. It is desired to dilute the solution in the first tank by interchanging the contents of the tanks at the rate of r gal/min.



(a) Set up differential equations for S_1 and S_2, the amounts of salt in the two tanks at time t minutes. Assume perfect mixing in the tanks at all times.

(b) For the case

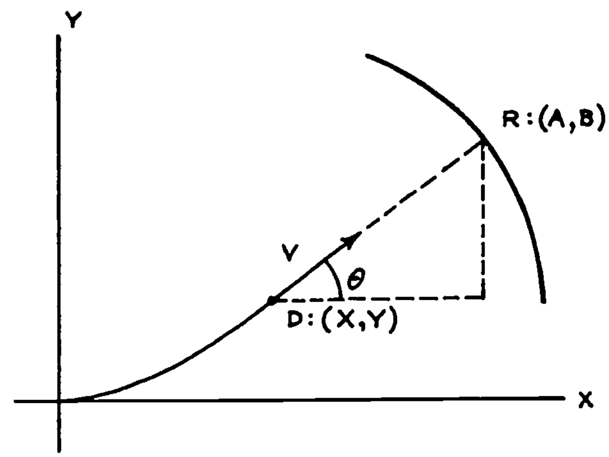
$$V_1 = 1000, \quad V_2 = 500, \quad S_0 = 500, \quad r = 60,$$

what are the concentrations in the two tanks at the end of 10 minutes? Ans. .361, .178.

(c) How long will it take to get the two concentrations within 1% of each other? Ans. About 28 min.

5. Modify Example 2 of Section 5 to include the presence of dingos, the wild dogs of Australia. Assume that dingos live exclusively on rabbits and make other appropriate assumptions to get differential equations governing the populations of rabbits and dingos. Solve these for various values of the constants and various initial values to see if you can get a fairly realistic model.

6. A dingo is chasing a rabbit by running directly towards it at any moment, with velocity V. If the dingo's coordinates are (x,y) and the rabbit's (A,B) then



$$\frac{dx}{dt} = V \cos \theta,$$

$$\frac{dy}{dt} = V \sin \theta.$$

where

$$\cos \theta = \frac{A - x}{D}, \quad \sin \theta = \frac{B - y}{D}, \quad D = \sqrt{(A-x)^2 + (B-y)^2}.$$

Supposing the rabbit runs in the circle

$$A = 1000 \cos \frac{t}{50} \text{ ft}, \quad B = 1000 \sin \frac{t}{50} \text{ ft},$$

and the dingo has velocity 30 ft/sec and is at (0,0) when t = 0, where and when does the dingo catch the rabbit? [Note. Because of the various errors in Euler's method you cannot expect D ever to become exactly zero.].

7. In a flu epidemic, out of N people let x be the number who have not been infected and y the number who are infected. The remaining N − x − y have been infected, have recovered, and are now immune to further infection.

(a) Justify the equations:

$$\frac{dx}{dt} = -axy,$$

$$\frac{dy}{dt} = axy - by,$$

where a and b are constants.

(b) Assume $y(0) = m > 0$, $x(0) = M > b/a$. Justify the following conclusions.

(i) y increases at first.

(ii) x always decreases, eventually becoming $< b/a$.

(iii) y is a maximum when $x = b/a$ and then decreases, approaching zero as $t \to \infty$.

(iv) x may approach a limit $x_\infty > 0$.

(c) Take N = 100, so that x and y are percentages, and M = 100 − m, and run solutions for various values of m, a, and b. In particular, try to see how x_∞ depends on these parameters.

7. Separation of Variables.

We return now to the first order equation $y' = f(x,y)$. Although numerical computations, using Euler's or some similar method, can get an approximate solution to essentially any such equation there are various reasons for finding analytic solutions, if possible. Perhaps the most important reason is that the form of the solution may tell us more about its behavior, particularly under changing initial conditions, than a mere table of values. Thus the knowledge that the solutions of $y' = x/y$ are the hyperbolas $y^2 - x^2 = c$ is useful information that is not likely to be obtained from numerical solutions.

The solving of $y' = f(x,y)$ is a generalization of the process of finding an indefinite integral, for the latter is simply the special case $y' = f(x)$. Thus we might expect to encounter all the difficulties, tricks, and special cases that were the subject of Chapter II, along with some new ones. This is indeed the case. Of all the various means that have been devised for solving a differential equation we shall consider in this chapter only two or three, suitable for many common and important cases.

The method of separating variables was introduced in Chapter 9 and has been used without comment earlier in the present chapter. Let us look at it a little more critically.

Example 1. $y' = \frac{1}{y}\sqrt{\frac{1-y^2}{1-x^2}}$, $y(0) = .5$.

We can separate variables, first writing y' as $\frac{dy}{dx}$, to get either

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

or

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

In the first form we must have $|x| < 1$, $|y| < 1$ and in the second, $|x| > 1$, $|y| > 1$. Seeing that our initial point satisfies the first pair of inequalities we must use (1). This integrates to

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

The value c is determined by the initial point to be $c = \sqrt{3}/2 \approx .866\ldots$. So, finally, solving (2) for y gives

(3) $y = \sqrt{1 - (c - \arcsin x)^2}$

the positive sign of the square root being chosen to give $y(0) = .5$.

The graph of (3) is given in Figure 7-1 for the maximum domain of x. However, this whole curve is not a solution of the differential equation. For it is evident from the original equation that $y' \geq 0$ for $y \geq 0$. Hence the piece BC of the curve must be excluded, and the domain of the solution (3) is only the interval

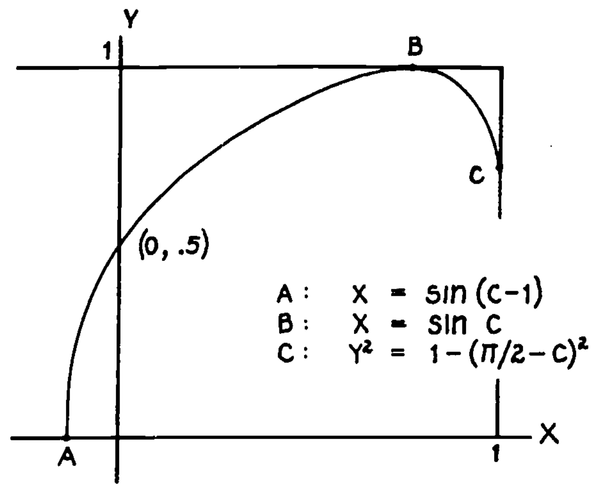


Figure 7-1

$$-\sin(1 - c) < x < \sin c,$$

or about

$$-.134 < x < .761.$$

In this case the restriction on the formal solution (3) was easy to deduce, but this kind of trouble may occur in subtle forms. One must always be most cautious when dealing with multiple valued expressions like square roots and inverse trig functions. The following theorem tells us how we can proceed in a general case.

<u>Theorem 1.</u> Let f be unicon in (a,b) and g in (c,d), and let $g(y) \neq 0$ for all y in (c,d). Let x_0 be in (a,b) and y_0 in (c,d). Let

$$F(x) = \int_{x_0}^{x} f(t)dt, \quad G(y) = \int_{y_0}^{y} g(t)dt.$$

Then

(a) F is defined in (a,b).

(b) G is defined and strictly monotone in (c,d).

(c) G has an inverse function H.

(d) $y(x) = H(F(x))$ is defined for all x for which F(x) is in the range of G.

(e) $G(y(x)) = F(x)$ for all such x.

(f) y(x) is the solution of

$$y' = \frac{f(x)}{g(y)}, \quad y(x_0) = y_0.$$

Proof. (a) Since x_0 and x are both in (a,b), the interval $[x_0,x]$ is contained in (a,b). Hence f is unicon on $[x_0,x]$ and $F(x)$ exists.

(b) Similarly, $G(y)$ exists. By the Fundamental Theorem of Calculus, $G'(y) = g(y) \neq 0$ and so $G'(y)$, being continuous, is either always positive or always negative. Hence G is strictly monotone, by Section 6-6.

(c) Every strictly monotone function has an inverse, Section 7-3.

(d) The range of G is the domain of H (Section 7-3).

(e) Basic property of the inverse function: $G(H(u)) = u$ for all u in the domain of H. Hence

$$G(y(x)) = G(H(F(x))) = F(x).$$

(f) Applying the chain-rule to $y(x)$, as defined in (d), gives

$$(4) \qquad y'(x) = H'(F(x))F'(x).$$

By Section 7-3,

$$H'(u) = \frac{1}{G'(H(u))}.$$

Hence

(5) $$H'(F(x)) = \frac{1}{G'(H(F(x)))} = \frac{1}{G'(y(x))} .$$

Since $F'(x) = f(x)$ and $G'(y) = g(y)$, (4) and (5) give us

$$y'(x) = \frac{1}{g(y(x))} f(x),$$

i.e., $y(x)$ is a solution of $y' = \frac{f(x)}{g(y)}$. Since, by definition, $G(y_0) = 0$, we have $H(0) = y_0$. Since, also, $F(x_0) = 0$, we have, finally

$$y(x_0) = H(F(x_0)) = H(0) = y_0.$$

The only reason for requiring $g(y) \neq 0$ was to insure (b). As long as G is strictly monotone the remaining conclusions follow. For instance, $g(y) = y^2$ on $(-1,1)$ is acceptable since

$$G(y) = \int_{y_0}^{y} t^2 dt = \frac{1}{3}y^3 - \frac{1}{3}y_0^3$$

is strictly increasing on $(-1,1)$.

To see how this theorem applies look at Example 1 again. We have

$$f(x) = \frac{1}{\sqrt{1 - x^2}}, \quad -1 < x < 1, \quad x_0 = 0,$$

$$g(y) = \frac{y}{\sqrt{1 - y^2}}, \quad 0 < y < 1, \quad y_0 = .5.$$

$y = 0$ is excluded since $g(y)$ must not be 0.

$$F(x) = \int_0^x \frac{dt}{\sqrt{1 - t^2}} = \arcsin x,$$

$$G(y) = \int_{.5}^y \frac{y\,dy}{\sqrt{1 - y^2}} = -\sqrt{1 - y^2} + c, \quad c = \sqrt{3}/2.$$

The inverse function H is thus

(6) $\qquad H(z) = \sqrt{1 - (c - z)^2}\,,$

the positive value being used since $H(z)$ must lie in the domain of G. The range of G is $c - 1 < G < c$, so this must be the domain of H. This is a restriction on (6), which otherwise could have the range $c - 1 < z < c + 1$. This restriction is what rules out the arc BC in Figure 7-1. We now get

$$y(x) = H(F(x)) = \sqrt{1 - (c - \arcsin x)^2}$$

as before, but with the restriction

1090

$$c - 1 < \arcsin x < c,$$

or

$$\sin(c - 1) < x < c,$$

giving the arc AB as the solution.

The conditions of Theorem 1 are sufficient to insure a solution but they may not be necessary. That is, there may be solutions that do not satisfy these conditions. The equation in Example 1, for instance, has the obvious solutions $y = 1$ and $y = -1$. The former can be combined with (3) to give the solution

$$y = \begin{cases} \sqrt{1 - (c - \arcsin x)^2} & \text{if } \sin(1 - c) < x < \sin c, \\ 1 & \text{if } \sin c \leq x < 1, \end{cases}$$

to give a solution in the interval $\sin(1 - c) < x < 1$. Since the two parts fit together with the same slope, 0, the differential equation is satisfied at this point.

One must always be on the lookout for these extra solutions. They usually lie along the boundaries of the regions in which the other solutions are defined.

Some of the advantages of the analytic solution (3) over a numerical solution can be appreciated by investigating the dependence of the solution on the initial point, or, equivalently, on the value of the constant c. In Figure 7-2, curves of the type AB are variations of the one obtained above for $c = \sqrt{3}/2$. The two extreme cases of this type go through the

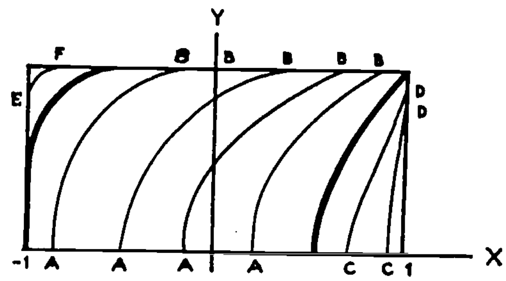


Figure 7-2

points (-1,0) and (1,1) and correspond to $c = 1 - \pi/2$ and $c = \pi/2$ respectively. For larger values of c, up to $c = 1 + \pi/2$ we get curves of type CD; and for $-\pi/2 < c < 1 - \pi/2$ curves of type EF. This kind of information can of course be obtained numerically by computing a large number of solutions for different initial values. This is apt to be time consuming and therefore expensive, especially if the critical curves through the corners are to be determined with some accuracy.

1002

The technique in Theorem 1 of using the integrals F and G with x_0 and y_0 as lower limits can be applied in general.

<u>Example 2.</u> $y' = \frac{\sqrt{y^2 - 1}}{(x^2 - 1)}$, $y(0) = 2$.

We obviously must have $x \neq \pm 1$, $|y| \geq 1$. Hence any solution lies in one of the six unshaded regions outlined by heavy lines in Figure 7-3. Considering the position of the initial point (0,2) we see that our solution must satisfy

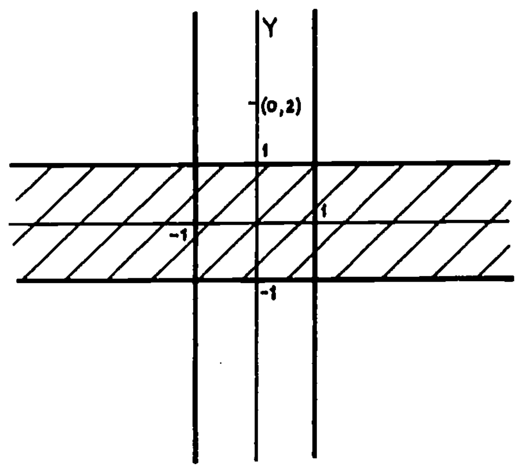


Figure 7-3

$$-1 < x < 1, \quad y \geq 1.$$

We get

$$F(x) = \int_0^x \frac{dt}{t^2 - 1} = \frac{1}{2}\log \left|\frac{t - 1}{t + 1}\right|\Bigg|_0^x = \frac{1}{2}\log \frac{1 - x}{1 + x},$$

the argument of the log function being chosen to be positive for the values of x in which we are interested. Similarly,

$$G(y) = \int_2^y \frac{dt}{\sqrt{t^2 - 1}} = \frac{1}{2} \log\left|t + \sqrt{t^2 - 1}\right|\Big|_2^y$$

$$= \frac{1}{2} \log(y + \sqrt{y^2 - 1}) - \frac{1}{2} \log(2 + \sqrt{3})$$

$$= \frac{1}{2} \log \frac{y + \sqrt{y^2 - 1}}{2 + \sqrt{3}} .$$

We wish to solve

$$\frac{1}{2} \log \frac{y + \sqrt{y^2 - 1}}{2 + \sqrt{3}} = \frac{1}{2} \log \frac{1 - x}{1 + x} .$$

This is obviously equivalent to

$$(7) \quad y + \sqrt{y^2 - 1} = a \frac{1 - x}{1 + x}, \quad a = 2 + \sqrt{3} .$$

The range of $y + \sqrt{y^2 - 1}$ is 1 to ∞ for $y \geq 1$, hence x must be restricted, if necessary, so that

$$1 \leq a \frac{1 - x}{1 + x} < \infty .$$

The right-hand inequality is automatic for $x > -1$; for $x + 1 > 0$ the left-hand inequality becomes

$$1 + x \leq a - ax,$$

$$(1 + a)x \leq a - 1$$

$$x \leq \frac{a - 1}{a + 1} = \frac{1 + \sqrt{3}}{3 + \sqrt{3}} = \frac{1}{\sqrt{3}} .$$

Hence x is restricted to $-1 < x \leq 1/\sqrt{3}$.

To solve (7) for y note that

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

Thus

$$y - \sqrt{y^2 - 1} = \frac{1}{a} \frac{1 + x}{1 - x} .$$

Adding this to (7), and doing a little algebra gives us as the solution

(8) $$y = 2 \frac{x^2 - \sqrt{3}x + 1}{1 - x^2} , \quad -1 < x \leq \frac{1}{\sqrt{3}} .$$

1095

Figure 7-4 shows what happens at $x = \frac{1}{\sqrt{3}}$. The curve (8) has a minimum point and is increasing for $1/\sqrt{3} < x < 1$, but the given differential equation obviously implies $y' \leq 0$. The dotted curve in the figure is a solution of

$$y' = -\frac{\sqrt{y^2 - 1}}{(x^2 - 1)}$$

obtained by taking the other determination of the square root. So here again it is the multiple valued quantity that cause the trouble.

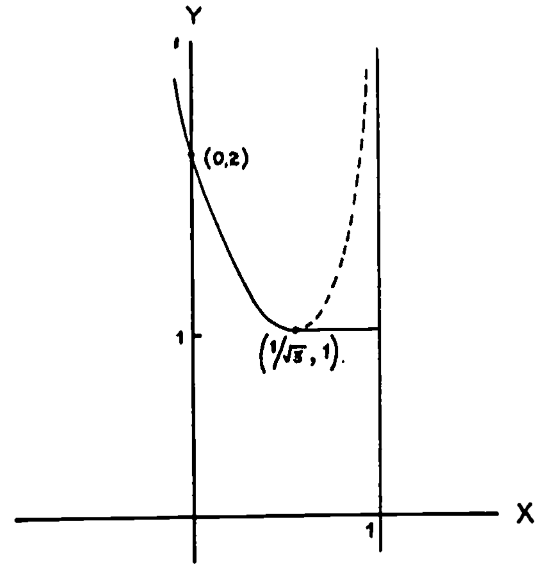


Figure 7-4

The solution can nevertheless be continued beyond $x = 1/\sqrt{3}$ since, as we have seen, $y = 1$ is a solution. Our solution is thus

$$y = \begin{cases} 2\,\dfrac{x^2 - \sqrt{3}\,x + 1}{1 - x^2} & \text{if } -1 < x < 1/\sqrt{3}, \\ 1 & \text{if } 1/\sqrt{3} \leq x < 1. \end{cases}$$

In Examples 1 and 2 we were able to solve explicitly for $y(x)$ in terms of elementary functions. This is not

to be expected in general. The following example exhibits two kinds of complications.

Example 3. $y' = (y^4 + y^2)e^{x^2}$, $y(0) = 1$.

Here we must avoid $y = 0$ but otherwise there are no restrictions. We get

$$\int_1^y \frac{1}{t^4 + t^2}dt = \int_0^x e^{t^2}dt.$$

The left-hand side integrates by partial fractions, and we get

$$(9) \qquad -\frac{1}{y} - \arctan y + 1 + \frac{\pi}{4} = \int_0^x e^{t^2}dt.$$

Here are our two troubles: First, the right-hand side cannot be integrated in elementary terms. We can express it as a power series, or we can use Simpson's rule to approximate it for a given x. Having done so, we then have the problem of solving (9) for y. This is another numerical process, using bisection or Newton's method. All in all, Euler's method, or some more accurate step-by-step method, would seem to be preferable.

PROBLEMS

1. Solve the following equations. Note that your answer is not necessarily wrong if it is not in the form given.

(a) $y' = x^2y$. Ans. $y = ce^{x^3/3}$.

(b) $y' = ay/x$. Ans. $y = cx^a$.

(c) $y' = ay^m/x^n$, $m \neq 1$, $n \neq 1$.

Ans. $y = \left[a \frac{1 - m}{1 - n} x^{1-n} + C\right]^{1/(1-m)}$.

(d) $y' = \cos x \cos y$. Ans. $y = \arctan \sinh(\sin x + c)$.

(e) $y' = e^{x+y}$. Ans. $y = -\log(c - e^x)$.

(f) $y' = \dfrac{y^2 + 1}{x}$. Ans. $y = \tan \log cx$.

(g) $y' = \dfrac{y^2 + 1}{x^2 + 1}$. Ans. $y = \dfrac{c + x}{1 - cx}$.

(h) $y' = \dfrac{y^2 - 1}{x^2 - 1}$. Ans. $y = \dfrac{c + x}{1 + cx}$.

2. The given solution of Problem 1(d) applies only to $-\pi/2 < y < \pi/2$. Find the solution with initial condition:

(a) $y(0) = 2\pi$.

(b) $y(0) = 3\pi/2$.

3. Consider the special case of Problem 1(c), $y' = -y^n/x^n$.

(a) For $n > 0$ what solution has this equation that is not covered by the given answer?

(b) For $n = 1/2$ find the special case of the given answer that satisfies $y(1) = 1$.

(c) Combine (a) and (b) to get the most complete solution with the given initial condition. What is the domain of this solution?

(d) Do (b) and (c) for the case $n = 2/3$. Warning: There are complications.

4. (a) What is the solution of Problem 1(h) not covered by the given answer?

(b) There are two points that behave peculiarly with respect to the solutions of this equation. What are they and what is the peculiar behavior?

5. Sometimes equations in which the variables cannot be separated can be reduced to separable form by an appropriate substitution. The substitution $y = ux$, replacing y by u, is often helpful. One common case in which it works is when

$$f(x, ux) = f(1, u),$$

f is then said to be "homogeneous of degree zero in x and y".

(a) Show that if f is homogeneous of degree zero and $y = ux$ then

$$\frac{du}{f(1,u) - u} = \frac{dx}{x} .$$

(b) Solve the following equations, at least to the point of expressing y as an implicit function of x.

(i) $y' = \frac{y}{x} + \exp\left(\frac{y}{x}\right).$

Ans. $y = -x \log(-\log cx)$.

(ii) $y' = \frac{y + x}{y - x}$.

Ans. $y = x \pm \sqrt{2x^2 + c}$.

(iii) $y' = \frac{5x + y}{x - 5y}$, (see Section 2, Problem 1(e)).

(iv) $y' = \frac{\sqrt{y^2 - x^2}}{x}$.

Ans. $\frac{y + \sqrt{y^2 - x^2}}{x^3} = C \exp \frac{y(y + \sqrt{y^2 - x^2})}{x^2}$

(c) What is the meaning of "homogeneous of degree zero" in terms of line elements?

6. To solve $y'' = -y$ (see Example 2 of Section 6) we use the form of y'' given in Problem 1 of Section 7-6:

$$y'' = \frac{dy'}{dx} = \frac{dy'}{dy}\frac{dy}{dx} = \frac{dy'}{dy}y' .$$

(a) Solve

$$\frac{dy'}{dy}y' = -y$$

to get $y' = c_0\sqrt{c_1{}^2 - y^2}$, $c_0 = \pm 1$.

(b) Solve

$$\frac{dy}{dx} = c_0\sqrt{c_1 - y^2}$$

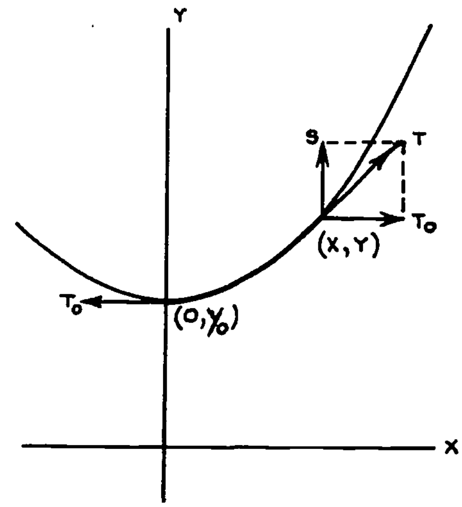
to get

$$y = c_1 \sin (c_0 x + c_2).$$

(c) Show that in all cases this is of the form

$$y = a \sin x + b \cos x.$$

7. A flexible rope of uniform linear density ρ is hanging loosely as in the figure. The weight of the portion between $(0, y_0)$ and (x, y) must be balanced by the upward component S of the tension T at (x,y), so



$$S = \int_0^x g\rho\sqrt{1 + y'(t)^2}\, dt.$$

Since the tension T is tangent to the curve we have

$$\frac{S}{T_0} = y'(x).$$

Hence

$$ay'(x) = \int_0^x \sqrt{1 + y'(t)^2}\, dt, \quad a = \frac{T_0}{g\rho}.$$

(a) How do we get from this equation to

$$ay''(x) = \sqrt{1 + y'(x)^2}\ ?$$

(b) Using $y'' = y' \frac{dy'}{dy}$, solve this equation for y' in terms of y. Show that the simplest form is obtained by choosing $y(0) = a$.

Ans. $y' = \frac{1}{a}\sqrt{y^2 - a^2}$.

(c) Solve the equation for y. [Hint. The easiest method is that of Problem 8, Section 11-3. Otherwise see Example 2.]

Ans. $y = a \cosh \frac{x}{a} = \frac{a}{2}\left(e^{x/a} + e^{-x/a}\right)$. This curve is known as a <u>catenary</u>.

8. Linear Equations.

A differential equation of the form

(1) $y' = p(x)y + q(x)$

is said to be linear. Linear equations, and their generalizations to higher order equations and systems of equations, have properties that make them especially useful in examining the behavior of certain kinds of physical systems. We shall return to this aspect later. For the present we study the solutions of (1).

Theorem 1. Let p and q be unicon in (a,b), let x_0 be any point in (a,b), and let y_0 be arbitrary. Then (1) has a unique solution in (a,b) satisfying $y(x_0) = y_0$.

Proof. Let r be a function, to be determined, positive and unicon on (a,b). Since r(x) is never zero, (1) is equivalent to

(2) $ry' - rpy = rq.$

We wish to choose r so that the left-hand side of (2) is the derivative of ry. Since

$$(ry)' = ry' + r'y,$$

this will be the case if

(3) $$r' = -pr.$$

Thus we have to solve (3) for r. Since r(x) is never zero on (a,b) we can separate variables, getting

$$\frac{dr}{r} = -p\ dx.$$

This gives

$$\log r = -\int_{x_0}^{x} p(t)dt + c,$$

or

$$r = e^{-s(x)}$$

where

$$s(x) = \int_{x_0}^{x} p(t)dt + c.$$

Since we are only after <u>some</u> function r, c can be taken to be any value we wish. We choose it to be 0 so as to make $s(x_0) = 0$.

Equation (2) now becomes

$$(e^{-s(x)}y)' = e^{-s(x)}q(x),$$

from which we get

$$y = e^{s(x)}\left[\int_{x_0}^{x} e^{-s(t)}q(t)\,dt + c\right]$$

Putting $x = x_0$ gives $c = y_0$.

This proof not only establishes the theorem but gives a formula for the solution, namely

$$(4)\quad \begin{aligned} y &= y_0 u(x) + u(x)\int_{x_0}^{x} \frac{q(t)}{u(t)}\,dt\,, \\ u(x) &= e^{s(x)}\,, \quad s(x) = \int_{x_0}^{x} p(t)\,dt. \end{aligned}$$

However, it is advisable to use the above derivation rather than trust one's memory of the formula.

If an initial point is not given we can take x_0 to be any convenient point in (a,b) and regard y_0 as the arbitrary constant of integration.

Example 1. $y' = y + e^{2x}$.

$p(x) = 1$ and $q(x) = e^{2x}$ are unicon for all x so we need not worry about bounds for x. Since x_0 is not given take it to be zero. Then

$$s(x) = \int_0^x 1\,dt = x.$$

$$r(x) = e^{-x}.$$

Equation (2) is

$$e^{-x}y' - e^{-x} = e^{-x}e^{2x}$$

or (always check this step)

$$(ye^{-x})' = e^{x}.$$

Integrating gives

$$ye^{-x} = c + e^{x}$$

or, finally,

$$y = ce^{x} + e^{2x}.$$

Example 2. $y' = \frac{x}{x^2 - 1} y + 1.$

Here $p(x) = \frac{x}{x^2 - 1}$ is unicon on any interval that does not contain 1 or −1. Let us take the interval (−1,1) and choose $x_0 = 0$. Then

$$s = \int_0^x \frac{t\, dt}{t^2 - 1} = \frac{1}{2}\log|t^2 - 1|\Big|_0^x = \frac{1}{2}\log(1 - x^2).$$

$$r = e^{-s} = (1 - x^2)^{-1/2}$$

$$((1 - x^2)^{-1/2}y)' = (1 - x^2)^{-1/2}$$

$$(1 - x^2)^{-1/2}y = \arcsin x + c$$

$$y = \sqrt{1 - x^2}\arcsin x + c\sqrt{1 - x^2}.$$

If we take the interval $(1,\infty)$, with $x_0 = 2$, we get

$$s = \frac{1}{2}\log\frac{x^2 - 1}{3},$$

$$r = \sqrt{\frac{3}{x^2 - 1}}$$

$$y = c\sqrt{x^2 - 1} + \frac{1}{2}\sqrt{x^2 - 1}\,\log(x + \sqrt{x^2 - 1}).$$

Example 3. $y' = xy - x^3$, $y(0) = 2$.

Here $p(x)$ and $q(x)$ are unicon over any interval. Using (4) for brevity, we have

$$s = \int_0^x t\,dt = x^2/2,$$

$$u = e^{x^2/2},$$

$$y = 2e^{x^2/2} - e^{x^2/2}\int_0^x t^3 e^{-t^2/2}dt.$$

The substitution $z = -t^2/2$ reduces the integral to

$$2\int_0^{-x^2/2} ze^z dz = 2(z - 1)e^z\Big|_0^{-x^2/2} = -(x^2 + 2)e^{-x^2/2} + 2.$$

Hence

$$y = 2e^{x^2/2} + x^2 + 2 - 2e^{x^2/2}$$

or

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

One can run into the same kind of trouble that was illustrated in Example 3 of Section 7.

Example 4. $y' = y + \sqrt{x}$, $y(1) = 1$.

We get

$$s = x - 1, \quad u = e^{x-1},$$

$$y = e^{x-1} - e^{x-1} \int_1^x e^{-t+1} \sqrt{t}\, dt$$

$$= e^{x-1} - e^x \int_1^x e^{-t} \sqrt{t}\, dt.$$

The last integral is not expressible in elementary form, so recourse must be had to numerical integration.

The situation is more annoying if p(x) cannot be integrated in elementary form. Then two numerical integrations are needed.

We turn now to the properties of linear equations of use in applications. Many a present-day device is what is called a "black box"; that is, a mechanism - using this term in the loosest fashion to include electrical, hydraulic, acoustic, etc. devices - that has an input q and an output y that are functions of time. Thus for a TV set the input is radio waves and the output is light and sound waves. For an automobile (in the simplest case) the input is displacement of the gas pedal and the output is angular velocity of the wheels. In these and other similar cases the relation between the input and output is much more complicated than can be represented by a single differential equation but many of the basic principles remain the same. For a TV set, for example, it is absolutely essential that the performance be linear to a very high degree of accuracy.

We consider, then, the situation illustrated in Figure 8-1, where q and y are functions of time t and are related by the

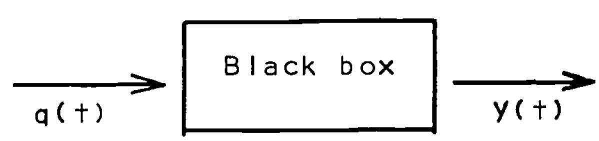


Figure 8-1

equation

(5) $$f(t)y' + g(t)y = q(t), \quad t \geq t_0 .$$

Here f(t) and g(t) are two functions of t determined by the internal structure of the black box. t_0 is some starting time, usually chosen to be 0. Our problem is to investigate the form of the function y(t) and in particular to see how it depends on the input and on the initial value y_0.

Consider first the case with zero input. A solution of

$$f(t)y' + g(t)y = 0$$

is called a null function of the equation. Equations (4) show that, omitting the identically zero null function, all null functions are multiples of one another, being simply multiples of

$$u(t) = \exp\left(\int_{t_0}^{t} - \frac{g(s)}{f(s)}\, ds\right) .$$

Note that the null functions depend only on f and g, that is, only on the structure of the black box.

It now follows that any solution of (5) can be expressed as any other solution plus a null function. For if y_1 and y_2 are two solutions, i.e.

$$f(t)y_1' + g(t)y_1 = q(t),$$

$$f(t)y_2' + g(t)y_2 = q(t),$$

then, subtracting

$$f(t)(y_1 - y_2)' + g(t)(y_1 - y_2) = 0.$$

That is, $y_1 - y_2 = z$ is a null function, and $y_1 = y_2 + z$.

Hence if we have one null function of the black box and if we have one output for a given input then we can get all possible outputs by simply adding multiples of the null function. In Example 3, for instance, if one were lucky or clever enough to notice the simple solution $y = x^2 + 2$ he could get any solution by adding a multiple of the null function $y = e^{-x^2/2}$. Which multiple is to be added is determined, of course, by the initial value $y(t_0)$.

So far we have only considered a single input. When we come to consider the outputs corresponding to several different inputs the basic property is the following, known as the <u>Principle of Superposition</u>:

Let $y_k(t)$ be an output corresponding to the input $q_k(t)$, $k = 1,2,\ldots,m$, and let $c_1,c_2,\ldots,c_m$ be any constants. Then

$$y(t) = \sum_{j=1}^{m} c_j y_j(t)$$

is an output corresponding to the input

$$q(t) = \sum_{j=1}^{m} c_j q_j(t).$$

It is left to the reader (Problem 1) to express this in terms of equations and to prove it by simple substitution of the output into the equation.

This principle finds its most common applications in the <u>static</u> case, the case in which the structure functions of the black box, $f(t)$ and $g(t)$, are constant. The differential equation is then

$$a y' + by = q(t)$$

and $y = e^{kt}$, $k = -b/a$, is a null function. Our first observation is the following:

The output with zero input grows without bound if $k > 0$, is constant if $k = 0$, and tends to zero if $k < 0$. Since it is essentially impossible to adjust a mechanism so that k is exactly zero the middle category is of little importance. In the case $k < 0$ the behavior of the black box is said to be <u>damped</u>; we set $h = -k$ and call h the <u>damping coefficient</u>. If $k > 0$ the behavior is <u>negatively damped</u>. This case can occur only when there is a source of energy within the box.

When we come to the problem of finding a solution for a given input the principle of superposition plays a vital role. If, for example,

$$q(t) = 6t^2 - 5 + 3 \sin 4t - 2e^{-6t}$$

we need only find solutions for

$$q = t^2,\ 1,\ \sin 4t,\ e^{-6t}$$

and then combine these solutions with the corresponding constant multipliers. The above terms are typical of those that appear in practical problems. For one thing, we can approximate many functions by partial sums of a power series, that is, by sums of powers of t. It is also possible to approximate many functions by sums of multiples of sin nt and cos nt for different values of n. Thus even if the input is quite complicated we may be able to find an output, approximately, by approximating the input by terms of the form x^n, sin nx, cos nx. e^{nx} terms are not so common but they are just as easily handled, so we include them in the following discussion.

The method we use is the "guessing" method, or "method of undetermined constants" of Section 11-6.

Case 1. q(t) is a polynomial of degree n. We observe that the solution

$$(6) \qquad \frac{1}{a}e^{kt}\int_{t_0}^{t} e^{-kz}q(z)\,dz$$

will also be a polynomial of degree n. (Problem 2). Letting this polynomial be

$$y = a_0 + a_1t + a_2t^2 + \dots + a_nt^n$$

we substitute in

$$ay' + by = q(t),$$

equate coefficients of corresponding powers of t, and solve successively for a_n, a_{n-1}, ... , a_0.

Example 5. $y' + 3y = x^2 - 1$,

Put $y = a + bx + cx^2$. Then

$$b + 2cx + 3a + 3bx + 3cx^2 = x^2 - 1.$$

Equating coefficients gives

$$3c = 1, \qquad c = 1/3,$$
$$2c + 3b = 0, \qquad b = -2/9,$$
$$b + 3a = -1, \qquad a = -7/27.$$

So

$$y = \frac{1}{27}(-7 - 6t + 9t^2).$$

Case 2. An input of the type $A \sin nt + B \cos nt$, has an output $a \sin nt + b \sin nt$. (Problem 2).

Example 6. To solve $y' + 3y = \cos 2t$ one sets $y = a \sin 2t + b \cos 2t$, to get

$$2a \cos 2t - 2b \sin 2t + 3a \sin 2t + 3b \cos 2t = \cos 2t.$$

Equating coefficients of sin and cos gives

$$3a - 2b = 0,$$
$$2a + 3b = 1,$$

from which

$$a = \frac{2}{13}, \quad b = \frac{3}{13}.$$

Hence $y = \frac{1}{13}(2 \sin 2t + 3 \cos 2t)$.

Case 3. (Proof left to reader: Problem 2).

$ay' + by = e^{ht}$ has a solution

$$y = \begin{cases} e^{ht}/(ah + b) & \text{if } h \neq k, \\ \frac{1}{a}te^{kt} & \text{if } h = k. \end{cases}$$

Example 7. Solve

$$y' + 3y = x^2 - 1 + 5 \cos 2t - 2e^{t} + 3e^{-3t}, \quad y(0) = 1.$$

Using the results of Example 5 and 6, and Case 3, we get as a solution of the equation

$$y_1(t) = \frac{1}{27}(-7 - 6t + 9t^2) + \frac{5}{13}(2 \sin 2t + 3 \cos 2t)$$
$$- \frac{2}{4}e^{t} + 3te^{-3t}.$$

The general solution is

$$y = y_1 + ce^{-3t}.$$

Putting in the initial condition gives

$$1 = -\frac{7}{27} + \frac{15}{13} - \frac{1}{2} + c, \qquad c = \frac{425}{702} \cdot$$

So, our answer is

$$y = \left(\frac{425}{702} + 3t\right) e^{-3t} - \frac{1}{2}e^{t} + \frac{1}{27}(9t^2 - 6t - 7)$$

$$+ \frac{5}{13}(2 \sin 2t + 3 \cos 2t).$$

PROBLEMS

1. Prove the Principle of Superposition.

2. (a) Prove that the expression (6) is a polynomial of degree n if $q(t)$ is a polynomial of degree n. [Hint. Prove it for a single term and then combine terms.]

   (b) Prove that (6) is of the form $\alpha \sin nt + \beta \cos nt$ if $q(t) = A \sin nt + B \cos nt$.

   (c) Prove that if $q(t) = e^{ht}$ then (6) is

$$\begin{cases} e^{ht}/(ah + b) & \text{if } h \neq k \\ \frac{1}{a}\, t\, e^{kt} & \text{if } h = k \,. \end{cases}$$

3. Solve each of the following:

   (a) $y' = \frac{y}{x} + x.$ Ans. $x^2 + Cx.$

   (b) $y' = xy + x.$ Ans. $Ce^{x^2/2} - 1.$

   (c) $y' = y \tan x + \sin x.$ Ans. $C \sec x - \frac{1}{2} \cos x.$

   (d) $y' = -y \sec x + \cos x.$

   Ans. $(\sec x - \tan x)(C + x - \cos x).$

(e) $y' = \frac{y}{x + 1} - 1$. Ans. $(C - \log(x + 1))(x + 1)$.

(f) $y' = \frac{xy}{x + 1} - 1$. Ans. $(Ce^{x} + x + 2)/(x + 1)$.

(g) $y' = y \log x + x^{x}$. Ans. $x^{x}(1 + Ce^{-x})$.

4. The equation

$$y' = \frac{2y}{x} + \frac{x^2}{(x^2 + 1)^{3/2}}$$

has one solution that remains bounded as $x \to \infty$. Find it, and show that $\lim_{x \to \infty} y(x) = -\frac{1}{2}$.

5. Solve each of the following.

(a) $y' - 2y = e^{3x} + 3$. Ans. $y = Ce^{2x} + e^{3x} - 3/2$.

(b) $y' - 3y = -3 \sin 3x$. Ans. $y = Ce^{3x} + \frac{1}{2}(\sin 3x + \cos$

(c) $y' = x^2 - y$. Ans. $y = Ce^{-x} + x^2 - 2x + 2$.

(d) $y' = y + 2x + 3 + 4e^{x} + 5 \sin x + 6 \cos x$.

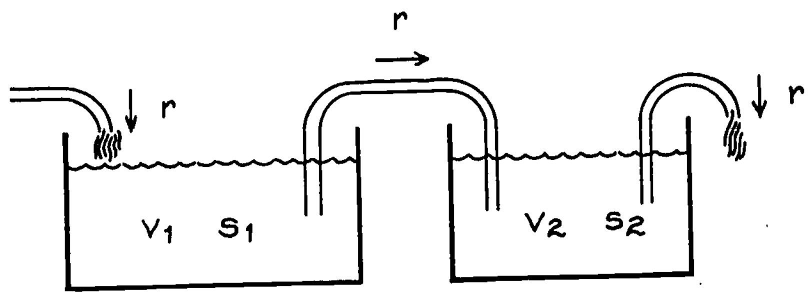
(e) $y' + y = \cos x + \sin x$. Can you see a solution?

6. Get the solutions of the equations in Problem 4 that satisfy the given initial conditions.

(a) $y(0) = 0$. (b) $y(0) = 2$.

(c) $y(1) = 2$. (d) $y(0) = 1$.

7. Two tanks of volumes V_1 and V_2, are connected as shown, and salt solution



flows through them at r gal/min. Let S_1 and S_2 be the amount of salt in the two tanks. The flow into the first tank is fresh water, and initially there is fresh water in the second tank. When is the salt content of the second tank a maximum?

Ans. If $V_1 \neq V_2$, $t = \dfrac{\log V_1 - \log V_2}{\frac{r}{V_2} - \frac{r}{V_1}}$.

If $V_1 = V_2$, $t = \dfrac{V_1}{r}$.

This problem illustrates how the output of one linear process can be used as the input of another. This is the principle of multi-stage amplifiers in radio sets.

8. An equation of the form

$$y' = p(x)y + q(x)y^n$$

is called a Bernoulli equation.

(a) Show that a substitution of the form $u = y^{1-n}$ gives a linear equation in u.

(b) Solve each of the following

(i) $y' = y + xy^2$

(ii) $y' = y + x\sqrt{y}$

(iii) $y' = \frac{x^2 + y^2}{y}$.

9. (Refer to Section 2, Problem 3).

(a) Solve $y' = xy + 1$, $y(0) = b$.

(b) Show that the answer can be written in the form

$$\left(b - b_0 - \int_x^\infty e^{-t^2/2}dt\right) e^{x^2/2} ,$$

where

$$b_0 = -\int_0^\infty e^{-t^2/2}dt.$$

(c) Show that for large x the solution grows like $e^{x^2/2}$ or $-e^{x^2/2}$ depending on whether $b > b_0$ or $b < b_0$.

(d) Investigate the behavior, as $x \to \infty$, of the solution with $b = b_0$.
[Hint. $y = -\int_x^\infty e^{-(t^2-x^2)/2}dt$. Use the fact that $t^2 - x^2 = (t + x)(t - x) \geq 2x(t - x)$ for $t \geq x$.

The value of b_0 can be approximated from (b) by the methods of Section 11-9. By other methods one can prove that $b_0 = -\sqrt{\pi/2}$.

10. Consider a damped linear system represented by

$$\frac{dy}{dt} = -hy + q(t).$$

We are interested in the long-time behavior and so can ignore the null function Ce^{-ht} which dies out.

(a) Fill in the gaps in the following argument.

If $q(t) = \sin(at + b)$, $a > 0$, then

$$y(t) = \frac{1}{a^2 + h^2}\left[h \sin(at + b) - a \cos(at + b)\right]$$

$$= \frac{1}{\sqrt{a^2 + h^2}}\left[\frac{h}{\sqrt{a^2 + h^2}} \sin(at + b) - \frac{a}{\sqrt{a^2 + h^2}} \cos(at + b)\right] = A \sin(at + c)$$

where $A = 1/\sqrt{a^2 + h^2}$, $c = b - \arctan \frac{a}{h}$.

(b) If $q(t) = \sum_{n=1}^{N} B_n \sin(n\omega t + b_n)$, $\omega > 0$, write an expression for the corresponding output $y(t)$.

The input in (b) is <u>periodic</u> of period $2\pi/\omega$, i.e. $q(t + 2\pi/\omega) = q(t)$. The answer to (b) tells us that the output is also periodic, with the same period. Since most periodic phenomena, such as sound waves, radio waves, tides, mechanical vibrations, etc., can be approximated by trigonometric sums we see that the property of periodicity is preserved by a linear differential equation.

11. To solve $y' = p(x)y + q(x)$, $y(x_0) = y_0$, by power series, first if $x_0 \neq 0$ replace x by $t + x_0$, to give

(7) $\frac{dy}{dt} = f(t)y + g(t)$, $y(0) = y_0$.

Assume we know the Maclaurin expansions of f and g,

$$f(t) = \sum_{n=0}^{\infty} a_n t^n, \quad g(t) = \sum_{n=0}^{\infty} b_n t^n.$$

It can be proved that if the radii of convergence of these two series are R_1 and R_2 then the solution of (7) has a Maclaurin expansion whose radius of convergence is $\geq \min(R_1, R_2)$. Let this series be

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

(a) Show that

$$c_0 = y_0,$$

$$c_1 = b_0 + a_0 c_0,$$

$$c_2 = \frac{1}{2}[b_1 + (a_0 c_1 + a_1 c_0)],$$

$$c_3 = \frac{1}{3}[b_2 + (a_0 c_2 + a_1 c_1 + a_2 c_0)],$$

and so on.

(b) Find 5 terms of the series solution of

$$y' = y \cos x + \sin x, \quad y(0) = 1.$$

Ans. $y = 1 + x + x^2 + \frac{1}{6}x^3 - \frac{1}{8}x^4 + \dots$.

12. (a) Draw a flow chart for the algorithm of Problem 11(a), assuming the a_n and b_n are stored.

(b) Adapt it to use recursion formulas to compute a_n and b_n when they are needed. Note that b_n is used only once and need not be stored whereas each a_n and c_n is used in all later steps.

(c) Program the algorithm for Problem 11(b) and find the first 20 terms of the series.

(d) Find the first 40 terms of a series solution of

$$y' = \frac{y}{\sqrt{x}} + 1, \quad y(1) = 0.$$

Use it to approximate $y(.1)$. How accurate do you think your value is?

[First derive $(1 + t)^{-1/2} = 1 - \frac{1}{2}t + \frac{1 \cdot 3}{2 \cdot 4}t^2 - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}t^3 + \ldots$]

13. A second order linear differential equation has the form

$$f(t)y'' + g(t)y' + h(t)y = q(t).$$

(a) What would we mean by a null function?

(b) Prove that if z_1 and z_2 are null functions so is $az_1 + bz_2$ for any constants a and b.

(c) Prove that any two solutions differ by a null function.

(d) Prove the Principle of Superposition.

(e) For the special case

$$y'' - 2y' - 3y = \sin t$$

(i) Find two null functions of the form e^{kt}.

Ans. $k = -1, 3$.

(ii) Find a solution of the form $a \sin t + b \cos t$.

Ans. $\frac{1}{10}(-2 \sin t + \cos t)$.

(iii) Find the solution satisfying the initial conditions $y(0) = 0$, $y'(0) = 1$.

Ans. $40y = 11e^{3t} - 15e^{-t} - 8 \sin t + 4 \cos t$.

ANSWERS

Volume II

Chapter 9

1-8, page 607

a) ae^{ax}, b) $2e^{x} + e^{2x}$, c) $2x^3e^{x^2} + 2xe^{x^2}$, d) $\frac{1}{x + 3}$

e) $\frac{2e^{x}}{(e^{x} + 1)^2}$, f) $\frac{1 - \log t}{t^2}$, g) $\frac{z}{\sqrt{z^2 - 1}} e^{\sqrt{z^2-1}}$,

h) $3e^{3x}\left[\frac{1}{\log(e^{x} - 1)} - \frac{e^{3x}}{[\log(e^{3x} + 1)]^2(e^{3x} + 1)}\right]$,

i) $2e^{-2x}(\cos 2x - \sin 2x)$, j) $3^{x}x^2(x \log 3 + 3)$,

k) $\log x$, l) $\cot x$, m) $\sec x$, n) $\frac{1 + x(x^2 + 1)^{-1/2}}{x + \sqrt{x^2 + 1}}$,

o) $\frac{2}{e^{x} - e^{-x}}$, p) $\frac{2x \log(1 + x^2)}{1 + x^2}$, q) x^2e^{x},

r) $e^{ax}(a \sin bx + b \cos bx)$, s) $25e^{3x} \cos 4x$.

1-9, page 609

a) $y' = \left(\frac{2}{x} + \frac{x}{1 + x^2} - 2\right) y$,

b) $y' = \left(\frac{1}{x + 1} + \frac{2}{x + 2} + \frac{3}{x + 3} + \frac{4}{x + 4}\right)y$,

c) $y' = (-1 + \cot x - 2 \tan 2x - \frac{2}{x} - 3 \cot 3x + 4 \tan 4x)y$,

d) $y' = \frac{y(x \log y - y)}{x(y \log x - x)}$.

Chapter 9

1-10, page 609

a) Min at $P(x,y) = (-1, \frac{1}{e})$,

b) Min at $P(x_1, y_1) = (0,0)$

Max at $P(x_2, y_2) = (2, \frac{4}{e^2})$

c) Max at $P(x,y) = (0,10)$,

d) y is not defined at $x = 0$, no extrema.

e) Min at $P(x,y) = (\frac{1}{e}, -\frac{1}{e})$,

f) Max at $P(x,y) = (e, \frac{1}{e})$

1-11, page 609

a) $-\frac{1}{5}e^{-5x} + c$, b) $\frac{1}{3}e^{x^3} + c$, c) $\log|x - 1| + c$

d) $\frac{t^2}{2} + \log|t| + c$, e) $\frac{1}{2} \log|t^2 + 1| + c$,

f) $\log|e^x + 1| + c$, g) $\log|e^x + 1| + c$,

h) $\frac{1}{2}x - \frac{1}{8} \log|2 + 3e^{4x}| + c$, i) $\arctan e^z + c$,

j) $\log|\sin x| + c$, k) $-\log|\cos x| + c$, l) $-\frac{1}{e} + e$,

m) $-\frac{1}{e^2} + 1$, n) $-\frac{1}{2}\left(\frac{1}{e^{25}} - 1\right)$, o) $\log 3$,

p) $\frac{\pi}{4} - \frac{1}{2} \log 2$, r) does not exist, s) does not exist.

<u>1-13, page 612</u>

a) cosh x, sinh x, $\frac{1}{\cosh^2 x}$, $-\frac{1}{\sinh^2 x}$,

b) $\int \sinh x \, dx = \cosh x + c$

$\int \cosh x \, dx = \sinh x + c$

$\int \tanh x \, dx = \log|\cosh x| + c$

$\int \cosh x \, dx = \log|\sinh x| + c$

Chapter 9

<u>1-16, page 612</u>

log 2 = .693174

log 3 = 1.098657

log 5 = 1.6095

log 7 = 1.94598565

log 11 = 2.397984

log 13 = 2.5650518

<u>2-1, page 626</u>

a) $s = 16t^2 + 5t + 100$

b) $r^3 = -3 \cos \theta + 1 + \frac{3}{2}\sqrt{2}$

c) $y^2 = x^2 + 2x + 9$

d) $y^2 = x^2 + 1$

e) $y = \log(c - e^x)$

f) $z = (\frac{1}{3} + \sqrt{t} + 1)^2$

<u>2-2, page 626</u>

a) $a \neq 0$, $ay + b = ke^{ax} + c$, b) $a = 0$, $y = bx + c$

<u>2-3, page 627</u>

a) 2

b) Twice the mass of earth.

<u>2-4, page 627</u>

a) 1981

b) 1993

Chapter 9

<u>2-5, page 627</u>

a) $e^{-.1386t}$, b) $e^{-69.3174}$

<u>2-6, page 627</u>

$-12 = xy - 8y$

<u>2-7, page 628</u>

$y = e^{-kx}$

<u>2-8, page 628</u>

$ke^{y} = x^{3}$

<u>2-9, page 628</u>

$r = -\frac{1}{4k\pi t}$

<u>2-10, page 628</u>

a) .09956 lb/gal.

b) 105 min.

<u>2-11, page 629</u>

13.7 min.

Chapter 10

1-1, page 642

a) convex, b) concave, c) convex for $x \geq 2$, concave for x

d) convex and concave, e) convex, f) concave, g) convex

h) convex, i) neither

1-2, page 642

a) Proof: $h'' = f'' + g'' \geq 0$

b) False. $f(x) = 0$, $g(x) = x^2$, $h(x) = -x^2$

c) False. $f(x) = -1$, $g(x) = x^2$, $h(x) = -x^2$

1-3, page 643

a) Yes, b) No

1-4, page 643

a) e^x is convex, $y = x + 1$ is tangent at $(0,1)$, so $e^x \geq x +$

b) log x is concave, $y = x + 1$ is tangent at $(1,0)$,

so $\log x \leq x + 1$

Chapter 10

<u>1-6, page 643</u>

a) flex at $(-1,2)$, b) flexes at $(0,4)$, $(\frac{3}{2}, -\frac{17}{16})$,

c) flexes at $(\pm\frac{1}{\sqrt{3}}, \frac{3}{4})$, d) flexes at $(\pm\sqrt{3}, \pm\frac{\sqrt{3}}{4})$,

e) flexes at $(2, \frac{2}{e^2})$, f) flexes at $(\pm.66, \pm.22)$,

g) $\cot x = \frac{x}{2}$, h) no flex

<u>2-1, page 653</u>

a) Output

1, .75000, .66667, .04167

2, .70833, .67933, .01450

4, .69383, .69122, .00131

EXCESSIVE ROUNDOFF 8, .69383, 100131

<u>2-2, page 653</u>

a) .69314, b) 19.625

<u>2-4, page 655</u>

a) .375, b) T = .38935, E = -.01435

c) M = .36794, E = -.00706, d) A_1 = .32865, E = -.00369

e) A_2 = .37508, E = -.00008

Chapter 10

3-2, page 663

a) $x_0 = 2$

$x_1 = 1.54$

$x_2 = 1.522$

$|E| \leq .002$

b) $x_0 = 1$

$x_1 = .75$

$x_2 = .73$

$|E| = .005$

3-3, page 663

a) $x_{n+1} = x_n - \dfrac{x_n^2 + c}{2x_n} = x_n - \dfrac{x_n}{2} + \dfrac{c}{2x_n} = \dfrac{1}{2}(x_n + \dfrac{c}{x_n})$

b) $x_0 = 3$, $x_1 = \frac{1}{2}(3 + \frac{8}{3}) = 2.8333$, $x_2 = \frac{1}{2}(\frac{17}{6} + \frac{48}{17}) = 2.8284$

$|E| \leq .0003$

3-4, page 664

a) Any number $N \neq 0$ can be written as $c \times 10^{2m}$, where $.1 \leq c \leq 10$. Then $\sqrt{N} = \sqrt{c} \times 10^m$

b) Program and run for c = .1(.1)1(1)10, counting the number of steps to get 8 place accuracy. One more step will be needed for 15 place accuracy.

Chapter 10

3-5, page 664

$$x_{n+1} = x_n - \frac{c - 1/x_n}{1/x_n^2} = x_n - cx_n^2 + x_n = x_n(2 - cx_n)$$

3-7, page 664

a) Use $a = 0$, $b = 1$, b) Use $a = 0$, $b = 1$,

c) $a = 2.5$, $b = 3$ for large root

$a = 2$, $b = 1.5$ for smaller root

d) Two roots are 2 and 4. For the third root put $z = -x > 0$ and write as $\log z = -\frac{\log 2}{2} z$. Take $a = .5$, $b = 1$.

4-4, page 680

a) $-\frac{3}{2}$, b) 0, c) 0, d) $\frac{1}{2}$, e) $-\infty$, f) 0, g) 0,

h) 0, i) $\frac{1}{2}$, j) $= \lim_{x\to 1-} \frac{\log(1 - x)}{\frac{1}{\log x}} = \lim_{x\to 1-} \frac{\frac{1}{1 - x}}{-\frac{1}{\log^2 x} \cdot \frac{1}{x}}$

$$= \lim_{x\to 1-} (-x) \cdot \lim_{x\to 1-} \frac{\log^2 x}{1 - x}$$

$$= -1 \cdot \lim_{x\to 1-} \frac{\frac{1}{x} \log x}{-1} = 0.$$

4-7, page 682

a) e, b) e^x, c) 1, d) ∞

1-1, page 690

a) $\frac{1}{4}x^4 - x^3 + \frac{3}{2}x^2 - x + c$, b) $\sqrt{x}(\frac{2}{3}x - 2) + c$, c) $-\frac{2}{3}\cos 3\theta + c$

e) $\frac{1}{\sqrt{5}} \arctan \frac{x}{\sqrt{5}} + c$, d) $-\frac{1}{2} e^{-2x} + c$, f) $\frac{1}{2\sqrt{5}} \log \left|\frac{x - \sqrt{5}}{x + \sqrt{5}}\right| + c$,

g) $4 \log |x + \sqrt{x^2 + 3}| + c$, h) $\tan \theta + c$, i) $\sec \theta + c$

2-1, page 696

a) $-\frac{1}{4} e^{-2x^2} + c$, b) $\frac{1}{2} \log |1 + \sin 2x|$ c,

c) $-\cos \log t + c$, d) $\sqrt{x^2 + 1} + c$, e) $\frac{1}{2} \log |x^2 + \sqrt{x^4 + 1}| + c$

f) $\log(e^x + e^{-x}) + c$, g) $\frac{1}{3}\left(\log |x^3 + 1| + \frac{1}{x^3 + 1}\right) + c$,

h) $-\arctan \cos x + c$, i) $\frac{1}{8} \sin 4x - \frac{1}{20} \sin 10x + c$,

j) $\frac{1}{8} \sin 4x + \frac{1}{20} \sin 10x + c$, k) $-\frac{1}{2} \cos x - \frac{1}{10} \cos 5x + c$,

l) $\log |4x^2 - 4x - 3| + \frac{1}{4} \log \left|\frac{2x - 3}{2x + 1}\right| + c$,

m) $\log |1 + \tan x| + c$, n) $\frac{1}{\sqrt{2}}$ l $\left|x + 1 + \sqrt{x^2 + 2x + \frac{3}{2}}\right| + c$,

2-1, page 696 - con't.

o) $\arcsin\left(\frac{2}{5}x - 1\right) + c$, p) $\frac{1}{4}(\log x)^4 + c$,

q) $x^2 - x + \frac{1}{16}\, 4\left[\log |2x^2 + x| + 19 \log \left|\frac{2x}{2x + 1}\right|\right] + c$,

r) $\frac{1}{3}x^3 - x^2 + 3x - 4 \log |x + 1| - \frac{2}{x + 1} + c$,

s) $\arctan \sqrt{x^2 + 4x + 3} + c$, t) $\frac{1}{3} \log |2 + 3 \tan x| + c$,

u) $\sqrt{x^2 + 2x} - \arctan \sqrt{x^2 + 2x} + c$, v) $\frac{1}{3} \tan^3 x + c$,

w) $\tan x + \frac{1}{3} \tan^3 x + c$, x) $- \cot x - \csc x + c$,

y) $-2 \cot \frac{x}{2} - x + c$, z) $\frac{1}{\sqrt{5}} \arctan (\sqrt{5} \tan x) + c$

2-2, page 698

$\frac{1}{8}\left(\cos 2x - \frac{1}{2} \cos 4x - \frac{1}{3} \cos 6x + \frac{1}{4} \cos 8x\right) + c$

2-5, page 698

c) $\int \tan^8 x\, dx = \frac{1}{7} \tan^7 x - \frac{1}{5} \tan^5 x + \frac{1}{3} \tan^3 x - \tan x + x + c$

$$\int \tan^9 x\, dx = \frac{1}{8} \tan^8 x - \frac{1}{6} \tan^6 x + \frac{1}{4} \tan^4 x - \frac{1}{2} \tan^2 x$$

$$- \log|\cos x| + c\ .$$

Chapter 11

3-1, page 710

a) $\frac{1}{2}\log\left|\frac{x^2}{1+x^2}\right| + c$, b) $\frac{1}{6}\log\left|\frac{\sqrt{x^2-9}-3}{\sqrt{x^2-9}+3}\right| + c$,

c) $-\frac{1}{2}\log\left|\frac{e^{2x}}{e^{2x}-1}\right| + c$, d) $\frac{1}{2}\sec 2x + c$,

e) $\frac{1}{3}(2x-1)^{1/2}(x+10) + c$, f) $-\frac{1}{a}\cos(ax+b) + c$,

g) $\frac{2}{3}(x+5)\sqrt{x-1} + c$, h) $\frac{1}{3\sqrt{2}}\arctan\sqrt{\frac{x^3}{8}-1} + c$,

i) $\frac{3}{80}(5x^2+3)(x^2-1)^{5/3} + c$, j) $x + c$,

k) $2(\sqrt{x} - \log(1+\sqrt{x})) + c$, l) $\frac{3}{28}(x+1)^{4/3}(4x-3) + c$,

3-2, page 711

Easily integrable for odd integers.

3-3, page 712

a) $-\frac{1}{10}\cos^5 2x + c$, b) $\frac{1}{4}\tan^2 2x + c$,

c) $2\sin x - \frac{1}{2}\log\left|\frac{1+\sin x}{1-\sin x}\right| + c$,

d) $\frac{1}{105}(5\cos^2 3x - 7)\cos^5 3x + c$,

3-3, page 712 - con't.

e) -s $\frac{1}{3}\sec^3 x + c$, f) $\frac{1}{3}\tan^3 x + c$,

g) $-\frac{2}{b^2}(a + b\cos x - a\log|a + b\cos x|) + c$,

h) $= -\frac{1}{2}\log|1 - \tan^4 x| + c$

3-5, page 712

Hint: $\int \frac{\sin 2x + a\sin x}{\cos 2x + a\cos x}\,dx = \int \frac{(2\cos x + a)}{2\cos^2 x - 1 + a\cos x} \cdot \sin x\,dx$

3-6, page 713

a) $-\frac{1}{3}(4 - x^2)^{1/2}(x^2 + 8) + c$, b) $-\frac{x}{2}\sqrt{4 - x^2} + 2\arcsin\frac{x}{2} + c$

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

d) $\sqrt{x^2 - 1} + \log|x + \sqrt{x^2 - 1}| + c$,

e) $\left(\frac{3 - x}{2}\right)\sqrt{4 - (1 + x)^2} + 3\arcsin\left(\frac{1 + x}{2}\right) + c$,

f) $(x - 2)/4\sqrt{x^2 - 4x + 8} + c$, g) $(x - 1)/3\sqrt{2x^2 + 2x - 1} + c$

h) $-\frac{\sqrt{2ax + x^2}}{ax} + c$

Chapter 11

3-9, page 715

a) $x - \tan \frac{x}{2} + c$, b) $\log \left| 1 + \tan \frac{x}{2} \right| + c$,

c) $\frac{1}{\sqrt{2}} \log \left| \frac{1 + \tan \frac{x}{2} + \sqrt{2}}{1 + \tan \frac{x}{2} - \sqrt{2}} \right| + c$, d) $\frac{1}{2} \log \tan \frac{x}{2} - \frac{1}{4} \tan^2 \frac{x}{2} +$

e) $\log \left| \frac{\sin x}{1 + \sin x} \right| + c$, f) $\sqrt{2} \log \left| 1 + \tan \frac{x}{2} \right| + c$,

g) $\log \left| \frac{A + B}{A - B} \right| + c$

$A = \sqrt{4 - 2 \cos x - 2 \cos^2 x} - \sqrt{3} (1 + \sin x - \cos x)$

$B = \sqrt{2} (1 - \cos x)$

h) $a^2 > 1$: $\frac{2}{\sqrt{a^2 - 1}} \arctan \frac{(a - 1) \tan \frac{x}{2}}{\sqrt{a^2 - 1}} + c$

$a^2 < 1$: $\frac{1}{\sqrt{1 - a^2}} \log \left| \frac{(1 - a) \tan \frac{x}{2} + \sqrt{1 - a^2}}{(1 - a) \tan \frac{x}{2} - \sqrt{1 - a^2}} \right| + c$

$a^2 = 1$: $\csc x - a \cot x + c$

3-10, page 715

See Problem 5-2, Page 719

Chapter 11

3-11, page 716

b) α) $a^2 > b^2 + c^2$: $\frac{2}{d} \arctan \frac{(a - c) \tan \frac{x}{2} - b}{d} + k$

$d = \sqrt{a^2 - b^2 - c^2}$

β_1) $a^2 < b^2 + c^2$, $a \neq c$: $\frac{1}{d} \log \left| \frac{(a - c) \tan \frac{x}{2} - b - d}{(a - c) \tan \frac{x}{2} - b + d} \right| + k$

$d = \sqrt{b^2 + c^2 - a^2}$

β_2) $a = c$, $b \neq 0$: $\frac{1}{b} \log |b \tan$ [illegible] $a| + k$

γ_1) $a^2 = b^2 + c^2$, $a \neq c$: $\frac{2}{b - (a - c) \tan \frac{x}{2}} + k$

γ_2) $a = c$, $c \neq 0$, $b = 0$: $\frac{1}{a} \tan \frac{x}{2} + k$

4-1, page 726

a) $- x \cos x + \sin x + c$, b) $-e^{-x}(x + 1) + c$,

c) $x \arcsin 2x + \frac{1}{2}\sqrt{1 - 4x^2} + c$, d) $x \tan x - \log |\sec x| + c$,

e) $3x(\tan 3x - 3x) - \log |\sec 3x| +$ [illegible]$_2x^2 + c$,

f) $(x - \frac{1}{2}) \arcsin \sqrt{x} + \frac{1}{2} \sqrt{x - x^2}$, g) $e^{x}(x^2 - 2x + 2) + c$,

h) $x[(\log x)^2 - 2 \log x + 2] + c$, i) $4\sqrt{x^2 + 2}\,(3x^4 + 4) + c$,

j) $\frac{1}{2} (x^2 \arctan x - x + \arctan x) + c$,

4-1, page 726 - con't.

k) $\frac{e^{ax}}{a^2 + b^2}(a \cos bx + b \sin bx) + c,$

l) $\frac{1}{2}[(x^2 + 1) \sin 2x - x \cos 2x] + c,$

m) $\frac{1}{3}\sqrt{4 + x^2}(x^2 - 8) + c,$

n) $\sqrt{1 + x^2} \arctan x - \log |x + \sqrt{1 + x^2}| + c,$

o) $x \log |x^2 + 1| - 2(x - \arctan x) + c,$

p) $2e^{\sin x}(\sin x - 1) + c,$ q) $2(\sin \sqrt{x} - \sqrt{x} \cos \sqrt{x}) + c,$

r) $\frac{1}{2} e^{x^2}(x^2 - 1) + c,$ s) $-\frac{1}{x + 2} \log x + \frac{1}{2} \log \left|\frac{x}{x + 2}\right| + c$

4-2, page 727

$\frac{1}{2}[\sec x \tan x + \log |\sec x + \tan x|] + c$

4-3, page 727

$\frac{e^x}{2}[x(\sin x - \cos x) + \cos x] + c$

5-1, page 737

a) $-\frac{1}{2}(2x + 3)^{-1} + c$, b) $-x + \frac{1}{7}[\frac{11}{3} \log|3x + 2| - \frac{5}{2} \log|2x - 1|] + c$,

c) $\frac{1}{2}x^2 + 2x + \frac{1}{4} \log |x + 1| + \frac{27}{4} \log |x - 3| + c$,

d) $\frac{1}{x - 1} - \log |x - 1| + \log |x - 2| + c$,

e) $\frac{1}{20} \log \left|\frac{x + 2}{x - 2}\right| + \frac{1}{30} \log \left|\frac{x - 3}{x + 3}\right| + c$,

f) $x - 6 \log |x - 1| + 11 \log |x - 2| + c$,

g) $\frac{1}{4}[5 \log |x - 1| - \log |x + 1| + \frac{2}{x - 1}] + c$,

h) $\frac{1}{2}x^2 - 2 \log (x^2 + 4) + \frac{3}{2} \arctan \frac{x}{2} + c$,

i) $2 \log |x + 3| + \frac{1}{3} \log \left|\frac{x - 1}{x + 2}\right| + c$, j) $x - \log \left|\frac{(x + 3)^4}{(x - 4)^3}\right| + c$,

k) $-\log |x - 1| + 2 \log |x - 2| + c$,

l) $\frac{1}{25} \left[2 \log \left|\frac{x^2 + 4}{(x - 1)^2}\right| + \frac{10}{1 - x} + \frac{19}{2} \arctan \frac{x}{2} \right] + c$,

m) $\frac{1}{8} \left[\log (x^2 + 2) + 7 \log |x^2 + 6| + \frac{1}{2\sqrt{2}} \arctan \frac{x}{\sqrt{2}} + \frac{1}{2\sqrt{6}} \arctan \frac{x}{\sqrt{6}}\right] + c$,

5-1, page 737 - con't.

n) $\frac{1}{2}[3 \log |x + 1| - 14 \log |x + 2| + 13 \log |x + 3|] + c,$

o) $\frac{3}{2} \frac{x - 1}{x^2 + 2x + 2} + \frac{1}{2} \log |x^2 + 2x + 2| - \frac{5}{2} \arctan (x + 1) + c,$

p) $\frac{1}{2} \log \left|\frac{(x - 1)^3(x + 2)^4}{x + 3}\right| + c,$

q) $\frac{1}{1 - x} + 2 \log |x - 1| + \frac{1}{2} \log |x \quad + x + 1| + \frac{1}{\sqrt{3}} \arctan \frac{2x+1}{\sqrt{3}} +$

r) $\frac{3}{x^2 + 1} + 3 \log |x^2 + 1| + c,$ s) $\frac{1}{2} \log \left|\frac{x^2}{x^2 + 1}\right| + c,$

5-2, page 738

a) $x + \frac{2}{1 + \tan \frac{x}{2}} + c,$ b) $\frac{1}{2} \log \left|\frac{\tan^2\frac{x}{2} - 2 \tan \frac{x}{2} - 1}{\tan^2\frac{x}{2} + 2 \tan \frac{x}{2} - 1} - x\right| + c,$

c) $\frac{\sec^2\frac{x}{2} \tan \frac{x}{2}}{(1 - \tan^2\frac{x}{2})^2} - \log \left|\frac{1 - \tan \frac{x}{2}}{1 + \tan \frac{x}{2}}\right| + c,$

6-1, page 744

a) $\frac{1}{2\sqrt{2}} \left(\log |\sqrt{2} x + \sqrt{2x^2 - 1}| - \frac{\sqrt{2} x}{\sqrt{2x^2 - 1}}\right) + c,$

b) $\frac{1}{6} (x^3 - \frac{1}{6}[x \cos 6x + \frac{x^2}{2} \sin 6x - \frac{1}{36} \sin 6x]) + c$

<u>6-1, page 744</u> - con't.

c) $\frac{1}{4}\left(\frac{(2x - 1)^{3/2}}{3} + 2(2x - 1)^{1/2} - (2x - 1)^{-1/2}\right) + c,$

d) $\frac{1}{12}\left(\sin^5 2x \cos 2x - \frac{\sin 4x}{4} + \frac{\sin 8x}{32} + \frac{3}{4}x\right) + c,$

e) $\frac{10x + 2}{-18\sqrt{2x^2 - x - 1}} + \frac{1}{2\sqrt{2}} \log \sqrt{2x^2 - x - 1} + \frac{x\sqrt{2}}{2} - \frac{1}{4\sqrt{2}} + c,$

f) $\frac{1}{8}\left(\log |2x - 1| - \frac{2}{2x - 1} - \frac{1}{2(2x - 1)^2}\right) + c,$

g) $e^{2x}\left[\frac{1}{74} (\sin 3x \quad 6 \cos 3x) \sin^3 3x\right.$

$\left. + \frac{27}{5} ((\sin 3x - 3 \cos 3x) \sin 3x + \frac{9}{2})\right] + c,$

h) $\frac{1}{24}\left(\frac{x}{(2x^2 - 1)^2} - \frac{3x}{2(2x^2 - 1)} + \frac{3}{4\sqrt{2}} \log \left|\frac{2x - \sqrt{2}}{2x + \sqrt{2}}\right|\right) + c,$

i) $\frac{1}{54}((\tan^2 3x + 7 \log |\cos 3x + 1|) + \frac{2}{3}x \tan 3x - x^2 + c,$

j) $\frac{1}{4} \arctan \frac{x}{2} - \frac{x}{2(x^2 + 4)} + c$

<u>6-3, page 745</u>

$(x^4 \quad x^3 + 12x^2 + 23x + 23)e^{-x} + c$

6-4, page 745

$$\frac{1}{a^2 + b^2} e^{ax} [(ac + bd)\sin bx + (ad - bc)\cos bx] + c$$

6-5, page 745

$$p = \frac{(ac + d)}{e}, \quad r = \frac{ad - c}{e}, \quad q = \frac{-a^2c - 2ad + c}{e^2}$$

$$s = \frac{-a^2d + 2ac + d}{e^2}, \quad e = a^2 + 1$$

6-6, page 745

$$e^{hx}[(px^2 + qx + r)\sin kx + (lx^2 + mx + n)\cos kx] + c$$

7-1, page 761

a) $3 \log 2 - \frac{5}{2} \approx -.42$, b) $\frac{\pi}{4} - \frac{1}{2} \log 2 \approx .43845$,

c) $\frac{1}{9} (12\sqrt{2} \log 2 - 8\sqrt{2} + 4) \approx .394$, d) $\log 2 \approx .6931$,

e) 0, f) $\frac{2}{35} (1 + 6\sqrt{2}) \approx .537$, g) $\frac{1}{2} (-e^{-\pi} + 1)$

7-3, page 762

a) πab, b) $\frac{4}{3}\pi ab^2$

Chapter 11

8-1, page 773

a) 3, b) divergent, c) π, d) divergent, e) $\dfrac{\omega}{\omega^2 + h^2}$

f) divergent, g) 0, h) divergent

8-2, page 773

$\frac{\pi}{2}$

8-3, page 774

a) $\frac{\pi^2}{4}$, b) divergent

8-4, page 774

$\frac{\pi}{2a}$, $\frac{2\pi}{a^2}$

8-5, page 774

No. Since $\lim_{x \to 0} x \log x = 0$

$$\int_0^1 x \log x \, dx = -\frac{1}{4}$$

8-6, page 774

a) 1 × 5280 ft/lb.
b) 800 × 5280 ft/lb.
c) 3930 × 5280 ft/lb.
d) 4000 × 5280 ft/lb.

9-2, page 792

b) n = 130

9-5, page 793

$\approx .04$

Chapter 12

1-1, page 820

$\frac{2}{5}\pi$

1-2, page 820

$\frac{\pi}{7}$

1-3, page 820

$2\pi^2$

1-4, page 820

$2\pi^2 + 4\pi$

1-5, page 820

a) $\frac{2401\pi}{80}$, b) $\frac{5\pi}{14}$

1-8, page 821

a) $\frac{\pi}{4}$, b) 4π

1-9, page 822

a) $c > \frac{1}{2}$, $c > 2$

2-1, page 830

b) $\frac{Gm\rho L}{D(D + L)}$. No, (c) No.

2-2, page 831

a) $G\rho^2 \log \frac{(D+L_1)(D+L_2)}{D(D+L_1+L_2)}$

3-1, page 834

28.080π ft/lb

3-2, page 834

50000π

3-3, page 834

350000 ft/lbs

3-4, page 834

a) $\frac{1}{2}\pi r^2 L^2 \rho$ ft/lbs, b) $\frac{3}{8}\pi r^2 L^2$ ft/lbs

3-5, page 834

$8333333.33325\rho\pi$ ft-lbs.

4-2, page 843

a) 16.43, b) 3.82, c) $2\pi^2 a$,

d) $\sqrt{1 + a^2} - \log|1 + \sqrt{1 + a^2}| + \log|a| - 1.032$,

4-2, page 843 - con't.

e) $\sqrt{e^{2c} + 1} - \log \left| 1 + \sqrt{e^{2c} + 1} \right| + c - 1.032$

4-4, page 843

b) $S \to \infty$, C has no length.

4-6, page 845

c) $L_{1/2} = 1 + \frac{1}{\sqrt{2}} \log \left| 1 + \sqrt{2} \right| \approx 1.62$

5-2, page 853

a) $2/\pi$, b) 0, c) $1/2$, d) $\pi/4$, e) $\log 2 - \frac{1}{2}$,

5-3, page 854

a) $\frac{2}{3}a^2$, b) $\frac{\pi}{4}a$, c) a

5-4, page 854

a) $\frac{2}{\pi}a^2$, b) $\frac{2}{\pi}a$, c) $\frac{4}{\pi}a$

5-5, page 854

a) $\frac{1}{\sqrt{2}}$, b) $\frac{1}{\sqrt{2}}$, c) $\sqrt{\frac{3}{8}}$, d) $\frac{1}{2}\sqrt{1 + \pi}$,

e) $\frac{2}{3\sqrt{3}} (\log^2 8 - 2 \log 8 + 2)^{1/2} \approx .57$

6-1, page 866

a) $6\ \dfrac{e^2 - 3}{e^2 - 1} \approx 4.1$,   b) 6

6-2, page 866

a) $\bar{x} = \frac{3}{5}$, $\bar{y} = \frac{3}{8}$,   b) $\bar{x} = 0$, $\bar{y} = \frac{3}{5}b$,   c) $\bar{x} = \frac{\pi}{2}$, $\bar{y} = \frac{\pi}{8}$,

d) $\bar{x} = \frac{1}{4}$, $\bar{y} = -1$,   e) $\bar{X}$ does not exist, $\bar{y} = \frac{1}{8}$

6-3, page 866

Area is infinite.

6-4, page 867

Area is infinite, therefore centroid not defined.

6-5, page 867

$\bar{x} = \pi a$, $\bar{y} = \frac{5}{6}a$

6-6, page 867

$\bar{x} = \frac{4a}{3\pi}$, $\bar{y} = \frac{4b}{3\pi}$

6-9, page 867

a) Origin at lower left corner

$\bar{x} = 2.8$, $\bar{y} = 5.1$,   b) $\bar{x} = 5$, $\bar{y} \approx 7.0$

c) $\bar{x} = \bar{y} \approx 4.5$,   d) $\bar{x} = 10$, $\bar{y} \approx 4.9$

e) Origin at upper left corner

$\bar{x} = \dfrac{a}{6 - \frac{3}{2}\pi}$, $\bar{y} = -\bar{x}$,   f) $\bar{x} \approx 2.9$, $\bar{y} \approx -5.6$

Chapter 12

6-12, page 869

a) P(1/2,1/2): barycentric coordinates $m_1 = m_2$, $m_3 = 2m_2$

P(1,0): barycentric coordinates $m_1 = m_2$, $m_3 = 0$

P(0,1): barycentric coordinates $m_1 = m_2 = 0$, m_3 arbitrary

P(x,y): barycentric coordinates $m_1 = \frac{m_3(2 - x - 2y)}{2y}$

$$m_2 = \frac{m_3 x}{2y}$$

b) barycentric coordinates	xy coordinates
(1,1,1)	$(\frac{2}{3}, \frac{1}{3})$
(1,2,3)	$(\frac{2}{3}, \frac{1}{2})$
(0,1,1)	$(1, \frac{1}{2})$
(1,0,0)	(0, 0)
(a,b,c)	$\left(\frac{2b}{a+b+c}, \frac{c}{a+b+c}\right)$

Chapter 13

1-4, page 882

a) .001, b) .013, c) 4×10^{-5}, d) 4×10^{-6}

1-6, page 883

a) $-\int_0^x \frac{t^n}{1-t}dt$, d) $x = \frac{1}{2}$, $n = 13$, No.

1-7, page 884

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

b) $\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\left(x - \frac{\pi}{4}\right) - \frac{\sqrt{2}}{4}\left(x - \frac{\pi}{4}\right)^2$, $a_n = \frac{\sqrt{2}}{2n!}(-1)^{\left[\frac{n}{2}\right]}$

c) $e^a + e^a(x - a) + \frac{1}{2}e^a(x - a)^2$, $a_n = \frac{e^a}{n!}$,

d) $2 + x - 2x^2$, $a_n = 0$ if $n > 4$; $4 - 15(x + 1) + 28(x + 1)^2$,

e) $x + \frac{1}{3}x^3 + \frac{2}{15}x^5$,

f) $3 + \frac{1}{6}(x - 9) - \frac{1}{216}(x - 9)^2$, $a_n = \frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\left(\frac{1}{2}-2\right)\cdots\left(\frac{1}{2}-n+1\right)}{n!}3^{1/2-n}$

g) $\log 4 + \frac{1}{4}(x - 4) - \frac{1}{32}(x - 4)^2$, $a_n = \frac{(-1)^{n-1}}{n4^n}$,

1-7, page 884 - con't.

h) $2 - \frac{1}{4}x^2 - \frac{5}{64}x^4$, i) $1 + x + \frac{1}{2}x^2$, j) $x + \frac{1}{6}x^3 + \frac{3}{40}x^5$

1-9, page 886

a) .947, $n = 5$, b) .17, $n = 6$

1-10, page 886

e) The Maclaurin series is $0 + 0x + 0x^2 + 0x^3 + \dots$. It converges for all x, its sum is 0 and thus converges to $f(x)$ only for $x = 0$.

1-11, page 887

Hint: Let $u = t - x$ and $x = 1$.

2-3, page 900

a) $\lim_{n\to\infty} \frac{n^3}{3^n} = 0$, b) None, c) None, d) $\sum_{n=1}^{\infty} \frac{1}{\sin n}$ diverges

2-4, page 901

a) $n = 200001$, if there is no roundoff, b) 9, c) 7, 13,

d) 4

Chapter 13

2-8, page 902

a) .90, b) .176

2-9, page 904

d) Hint: Use n = 9, recursive formula a ← −a × (n −1)/(n + 4).

3-3, page 916

a) conv. absolutely, b) div., c) conv., d) div.,

e) conv., f) conv. absolutely, g) conv. absolutely,

h) conv. absolutely, i) conv. absolutely, j) conv.,

k) conv., l) div.

4-3, page 926

a) conv., b) conv., c) div., d) div., e) conv.,

f) conv., g) conv., h) div., i) conv., j) conv.,

k) div.

4-6, page 927

a) div., b) conv., c) conv., d) conv., e) conv., f) div

Chapter 13

4-7, page 928

a) conv., b) conv., c) conv., d) div., e) div.,

f) div., g) conv., h) div., i) conv. absolutely.

4-8, page 929

e) $\Gamma(.5) = 1.77246$, $\Gamma(6.5) = 387.92$

4-9, page 930

b) 3.87 in., c) it is unlimited

5-1, page 949

a) $\frac{1}{2}$, b) 1, c) ∞, d) $\frac{1}{e}$, e) 4, f) our present methods do not suffice to determine the radius of convergence of this series, g) 1, h) 1, i) 1

5-3, page 950

a) $x + x^2 + \frac{1}{3}x^3 + \frac{1}{20}x^5 + \ldots$, b) $x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \frac{17}{315}x^7 + \ldots$,

c) $1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \ldots$, $\sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}$, $R = \infty$,

$x + \frac{x^3}{3!} + \frac{x^7}{7!} + \frac{x^9}{9!} + \ldots$, $\sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!}$, $R = \infty$.

5-3, page 950 - con't.

d) $-\frac{1}{2}x^2 - \frac{1}{12}x^4 - \frac{1}{45}x^6 - \frac{17}{2520}x^8 - \cdots ,$

f) $1 + x + x^2 + \sum_{n=3}^{\infty} \frac{2}{n!}x^n$

5-4, page 951

a) $x - \frac{1}{3\cdot 3!}x^3 + \frac{1}{5\cdot 5!}x^5 + \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)(2n+1)!}$

b) .946, c) 1.54993, d) flow chart, see next page.

5-6, page 953

a) -2, b) 1, c) -1, d) $\frac{1}{2}$, e) $\frac{1}{3}$, f) ∞,

g) $\frac{1}{2}$, h) 0

Chapter 14

4-7, page 990

a) $y(x) = -1 + x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{2\cdot 4}x^4 + \frac{1}{3\cdot 5}x^5 - \frac{1}{2\cdot 4\cdot 6}x^6 + \cdots$

b) $y(1) = -.236$, c) $|E| \leq .05$

4-8, page 991

a) Simpson's rule, or even the trapezoidal rule, is better.

5-5, page 1008

a) $V' = \frac{cV_e}{M_0 - ct} - g$, b) $V = -V_e \log\left(1 - \frac{ct}{M_0}\right) - gt$,

c) $S = \frac{V_e M_0}{cr}(r - \log r - 1) - \frac{1}{2}gt^2$

5-10, page 1013

a) 94.8 min., b) 48.7 min.

5-13, page 1014

c) $p = 14e^{x/4.7}$ if x is measured in miles.

6-4, page 1030

a) $\dfrac{dS_1}{dt} = -\dfrac{rS_1}{V_1} + \dfrac{rS_2}{V_2}$, $\dfrac{dS_2}{dt} = \dfrac{rS_1}{V_1} - \dfrac{rS_1}{V_2}$,

b) .361, .178,   c) about 28 min.

6-6, page 1032

Flow Chart, see next page.

7-2, page 1048

a) $y = \arctan \sin h \sin x + 2\pi$,

b) $y = \dfrac{3\pi}{2}$, for $\sec y + \tan y$ becomes infinite.

7-3, page 1049

a) $y = (-x^{1-m} + c)^{\frac{1}{1-m}}$, $y = 0$,   b) $y = (-x^{1/2} + 2)^2$,

c) $y = \begin{cases} (4 - x^{1/2})^2 & \text{if } 0 \leq x < 4 \\ 0 & \text{if } x > 4 \end{cases}$,   Domain $0 \leq x < \infty$.

d) There is an infinite number of solutions of the form

$$y = \begin{cases} (2 - x^{1/3})^3 & \text{if } 0 < x < 8 \\ 0 & \text{if } 8 \leq x \leq b \\ (b^{1/3} - x^{1/3})^3 & \text{if } x > b. \end{cases}$$

Chapter 14

7-4, page 1049

a) The solutions are $y = 1$, $y = -1$

b) $P_1(x,y) = (1,1)$

$P_2(x,y) = (-1,-1)$

All the curves $y = \frac{c + x}{1 + cx}$ go through $(1,1)$ and $(-1,-1)$, but they have all different slopes at these points.

8-4, page 1070

$y = -x^2 + \frac{x^3}{\sqrt{x^2 + 1}}$, $\lim_{x \to \infty}\left(-x^2 + \frac{x^3}{\sqrt{x^2 + 1}}\right) = -\frac{1}{2}$

8-5, page 1070

d) $y = ce^x + 4xe^x - 2x - 5 + \frac{1}{2}(\sin x + 11 \cos x)$,

e) $y = \sin x$

8-6, page 1070

a) $c = \frac{1}{2}$, b) $c = \frac{3}{2}$, c) $c = 3$, d) $c = \frac{23}{2}$

8-8, page 1072

b) (i) $= \frac{1}{ce^{-x} + x - 1}$, (ii) $y = (ce^{x/2} - x - 2)^2$,

8-9, page 1072 - con't.

(iii) $y = \sqrt{ce^{2x} - x^2 - x - \frac{1}{2}}$

8-9, page 1072

a) $y = e^{x^2/2} \quad b + \int_0^x e^{-t^2/2}dt$

8-12, page 1075

Flow Chart, see next page.

INDEX VOLUMES I - II

Absolute convergence,157,911
Absolute value,17,33
Acceleration,419,594
Algorithm,65
  for area,219 ff.
  for Fibonacci sequence,
    108,158
  numerical,73
  for partial fractions,800,
    801
Alternating series,895
Anti-derivatives,562 ff.
Approximation
  of extrema,455 ff.
  of function value,470 ff.
  of integrals,211 ff.,
    645 ff.,776 ff.
  of limit of sequence,125 ff.
  of root of equation,136,
    656 ff.
  of solution of differential
    equation,958,975 ff.
Area,207,325 ff.
Arithmetic units,103
Assignment box,79
Assignment statement,79
Axiom(s)
  Archimedean,7
  completeness,202
  of real numbers,6

Barycentric coordinates,868
Bounded function,670,919
Bounded sequence,173,905

Catenary,1053
Center of area,861
Center of gravity,856
Center of mass,856
Centroid,856
Chain rule,499
Circle, unit,50
  area,125,211,347,703

Coefficient
  damping,1064
  of polynomial,45
Comparison test,907,909,
    910
Completeness axiom,202
Composition theorem,399
Computer,65
  concepts,65
  language of,105 ff.
  memory of,77,94 ff.
  model of,76 ff.
  word,95
Compiling,105
Control unit,103
Convergence
  of improper integrals,
    766
  rate of,661
  of series,889
  of sequences,132,149,
    153,154
Continuity,386,387
Continuous functions,
    386 ff.,968
Coordinate(s),37
  barycentric,868
  system,36
Cores,95
Curve
  length of,837
  plane,835
  slope of,410
Cycloid,549
Cylindrical shells,812 ff.

Decision box,68
Definite integrals,582,746
Derivative(s),408 ff.
  chain rule for,499
  of composite functions,
    495 ff.
  formulas,418,426,428,429,
    427,598,599,603
  of implicit functions,509

Differential(s),541 ff.
Differential equations,
    621,957 ff.
  first order systems,1020
  higher order,1023
  linear,1054 ff.
Differentiation,407 ff.
  explicit,509
  implicit,509
  logarithmic,605
Duhamel's theorem,828

Error,471
  in approximation,219
  bound,473,649,661,778,782
  roundoff,118,438,979
  truncation,978
Euler's constant,925
Exponential function,597 ff.
Extended mean value theorem,
    464

Floating point,116
Flow chart,67 ff.
  for integrals of convex
    functions,651
  for maxima,458
  for Newton's method,662
  for root of an equation,
    140
  for Simpson's rule,784
  for trapezoidal rule,226
Formulas for volumes and
    areas,449
Fraction,3
Function(s),31 ff.
  circular,49
  composition of,57
  concave,634
  constant,38,542
  convex,633
  continuous,386,968
  decreasing,53
  derived,416
  domain of,33
  elementary,686
  exponential,597 ff.
  gamma,769 ff.
  generating,954
  graph of,37
  greatest integer,47
  identity,39
  increasing,53
  inverse trigonometric,
    525 ff.
  linear,46
  logarithm,593 ff.
  monotone,53,479 ff.
  null,1061
  piecewise monotone,229
  polynomial,44
  range of,34
  rational,46
  roots of,46
  strictly decreasing,53
  strictly increasing,53
  unicon,254
  weight,850
Fundamental existence and
    uniqueness theorem,969
Fundamental theorem of
    calculus,575,579

Gamma function,769 ff.
Generalized mean value
    theorem,673
Greatest integer function,
    47

Identities,trigonometric,
    275,695
Inequalities,12,13
  triangle inequality,20
Input,89
Input box,90
Integer,13

Integral
approximation of,221 ff., 645 ff.,776 ff.
definition of,235,317
definite,582,746
elliptic,834
exponential,785
generalization of,823 ff.
improper,764
indefinite,582
tables,740
Integration,234 ff.
by partial fractions,729
by parts,717
by substitution,566,699 ff.
Intermediate value theorem, 403
Interval,25 ff.

Law of the Mean, see Mean Value Theorem
Length of a curve,837
L'Hospital's Rule,666
Limit(s)
evaluation of,666 ff.
infinite,670
left-hand,668
of a function,374 ff.
of a sequence,156
Line element,962
Linear differential equation, 1054 ff
Lipschitz(ian)
coefficient,293
condition,293
function,293,431,969
Local maximum tests,466,485, 486
Local maximum point,443
approximation of,455
Logarithm,593 ff.

Machine Language,105
Maclaurin expansion,873
series,879
Mass ratio,1003
Maximum theorem,441
Mean
arithmetic,28
geometric,28
value,846,848
Mean value theorem,462
applications,470
for integrals,851
Modulus, unicon,290 ff.

Newton's method,656 ff.
Null function,1061
Numbers,3 ff.

Open region,968
Ordinates,36

Pappus theorem,864
Parametric equations,548
Partial fractions,729 ff.
Partial sum,888
Piecewise monotone functions,229
Polynomial,44
Power series,932
Principle of superposition, 1062
Pythagorean theorem,4

Radius of convergence,935
Rate of change,533
Ratio tests,912
Rational numbers,3
Real numbers,5
Related rates,533
Remainder
in Taylor expansion,873
in series,893

Rolle's theorem,463
Round-off error,118,438,979

Samos,94,103 ff.
Separation of variables,
1035
Sequence,125 ff.,149
bounded,173
convergence of,132,153,
154,156
Fibonacci,158
limit of,156
Series,879,888 ff.
absolute convergence,911
alternating,895
convergence,889
divergence,889
expansion of functions,
879 ff.
geometric,908
harmonic,890
Maclaurin,879
p-series,908
power,932
Taylor's,879
Set,631
concave,631
convex,631
Simpson's rule,784 ff.
Slope of a curve,410
Solids of revolution,340,
809 ff.
Speed,367
Squeeze theorem,189,394
Step function,48
Support,633
Systems of differential
equations,1020

Tangent line,407
equation of,410
slope of,410

Taylor's series,879
theorem,465
Trapezoidal rule,266,778
Trigonometric functions,
51 ff.
Trigonometric identities,
275,695
Truncation error,978

Variable,36,76
Velocity,368
Volume,337
formulas for,449
of solid of revolution
340,809

Weight functions,850
Weighted average,187
Work,360,832

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

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

Now if $\lim_{n\to\infty} r_n = L \neq 0$, we have

$$\lim_{n\to\infty} r_n = \lim_{n\to\infty} \left(1 + \frac{1}{r_{n-1}}\right) = 1 + \lim_{n\to\infty} \frac{1}{r_{n-1}} = 1 + \frac{1}{\lim_{n\to\infty} r_{n-1}}$$

or

$$L = 1 + \frac{1}{L} .$$

(Do you see why $\lim_{n\to\infty} r_{n-1}$ must also equal L?) This equation $L = 1 + \frac{1}{L}$ can be rewritten in the form $L^2 - L - 1 = 0$. Solving this quadratic equation, we have

$$L = \frac{1 \pm \sqrt{5}}{2}$$

Since the limit must obviously be positive, we can throw away the negative solution, so that the only possible value for $\lim_{n\to\infty} r_n$ is $\frac{1 + \sqrt{5}}{2}$.

Example. Suppose a_n is the sequence defined by $a_1 = 3$, and $a_{n+1} = a_n^2 - 2$ for $n \geq 1$. If $\lim_{n\to\infty} a_n = A$, then

$$A = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} ((a_n)^2 - 2) = A^2 - 2.$$

Solving the equation $A = A^2 - 2$, we obtain $A = 2$ or $A = -1$. The first few terms of the sequence a_n are 3, 7, 47, 2207; the sequence does not seem to be converging to 2 or -1. In fact, all the terms are ≥ 3, because $a_1 = 3$, and if $a_n \geq 3$, then $a_{n+1} = (a_n)^2 - 2 \geq 3^2 - 2 = 7 > 3$. Therefore, the sequence a_n does not converge to 2 or -1, and hence does not converge at all.

PROBLEMS

1. Find the limit, if it exists, of each of the following sequences.

(a) $a_n = 2 + \frac{3}{n}$

(b) $b_n = \dfrac{1 - (\frac{7}{8})^n}{1 - \frac{7}{8}}$

(c) $c_n = (4 - \frac{1}{n})^2 - 3(4 - \frac{1}{n})$

(d) $d_n = n^2 - (n^2 - 1)$

(e) $e_n = \dfrac{2^{n-1} - 1}{2^n}$

(f) $f_n = \dfrac{3 - \frac{2}{n^2}}{4 + \frac{3}{n}}$

(g) $g_n = (-1)^n + \dfrac{1}{2^n}$

(h) $h_n = \dfrac{n^2}{n^2 + 1} (1 - \dfrac{2}{n^1} + \dfrac{2}{2^n})^3$

2. Exhibit two nonconvergent sequences whose product converges.

3. Prove that if $\lim_{n\to\infty} x_n \geq a \geq 0$ and $\lim_{n\to\infty} y_n \geq b \geq 0$, then $\lim_{n\to\infty} x_n y_n \geq ab$.

4. Prove that if $0 \leq \lim_{n\to\infty} x_n \leq a$ and $\lim_{n\to\infty} y_n \geq b > 0$, then $\lim_{n\to\infty} \dfrac{x_n}{y_n} \leq \dfrac{a}{b}$, providing y_n is never 0.

5. Prove that if $\lim_{n\to\infty} x_n = 0$ and if $|y_n| \leq a$ and $|z_n| \geq b > 0$ for all n, then $\lim_{n\to\infty} \frac{x_n y_n}{z_n} = 0$.

6. Find sequences x_n and z_n such that $\lim_{n\to\infty} x_n = 0$ and $z_n > 0$ for all n, and $\lim_{n\to\infty} \frac{x_n}{z_n} = 1$. (Compare Problem 4.)

7. Suppose that x_n is a sequence of nonzero numbers such that for every number K there is an integer N such that $|x_n| > K$ for all $n > N$. Prove that $\lim_{n\to\infty} \frac{1}{x_n} = 0$.

8. Modify the flow chart for the Fibonacci sequence in Section 4 to provide for the output of the value of r_n. Output box should have the form shown at the right. Write the program and run it. See whether the terms r_n actually seem to be converging to $\frac{1 + \sqrt{5}}{2}$. (Since in any earlier program you computed $\sqrt{5}$ to a large number of decimal places, it will be an easy hand calculation to compute $\frac{1 + \sqrt{5}}{2}$ to compare with your computer output in this problem.)

```
N, ASUBN, RSUBN.
```

9. Suppose K is a number and a_n is a sequence defined by $a_1 = K$, and $a_{n+1} = (a_n)^2 - 2$ for $n \geq 1$. For each of the following values of K, evaluate the first five terms

of the sequence a_n. In each case, determine whether the sequence converges, and if it does, find its limit.

(a) $K = 2$

(b) $K = -1$

(c) $K = -3$

(d) $K = 0$

(e) $K = \frac{-1 + \sqrt{5}}{2}$

10. Prove that the sequence sin n does not converge. Hint: Assume that $\lim_{n\to\infty} \sin n = L$. Use the identity

$$\sin (n + 1) = \sin n \cos 1 + \cos n \sin 1$$

to show that $\lim_{n\to\infty} \cos n$ exists and that, if $M = \lim_{n\to\infty} \cos n$, then $L = L \cos 1 + M \sin 1$. Use the identities

$$\sin 2n = 2 \sin n \cos n$$

and $\sin^2 n + \cos^2 n = 1$ to obtain other equations relating L and M. Finally, show that the three equations are contradictory.

7. The Squeeze Theorem

Before going on with our limit theorems, we will discuss a theorem that has nothing to do with limits but which is needed in the proof of the corollaries to the "squeeze" theorem, which does involve limits.

The Weighted Average Theorem. If $r, s \geq 0$ with $r + s = 1$, then $ra + sb$ lies between a and b. (Here we use the word 'between' to include a and b themselves.)

When r and s satisfy these conditions the expression, $ra + sb$ is called a convex combination or weighted average of a and b. For the common special case where $r = s = \frac{1}{2}$, we have $ra + sb = \frac{1}{2}a + \frac{1}{2}b = \frac{a + b}{2}$, the ordinary average or arithmetic mean of a and b, which of course lies between a and b. The theorem in the general case, for all its simplicity, is frequently useful.

Proof: If $a \leq b$, then

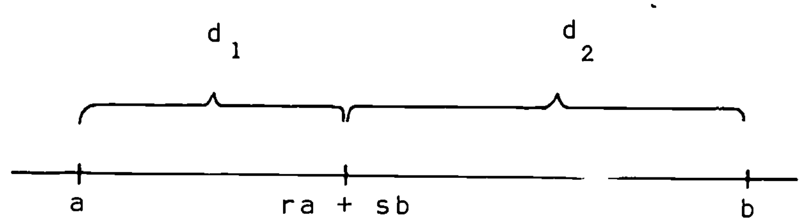
$$a = ra + sa \leq ra + sb \leq rb + sb = b.$$

If $a \geq b$,

$$a = ra + sa \geq ra + sb \geq rb + sb = b,$$

and the proof is complete.

Actually, we can say quite a bit more about $ra + sb$ than is actually contained in the theorem. Assume $a \leq b$. The fact is that the point $ra + sb$ divides the segment $[a,b]$ in the ratio of s to r. That is, if d_1 and d_2 are the distances illustrated below



then, using the relation $r + s = 1$ in the forms $s = 1 - r$ and $r = 1 - s$, we have

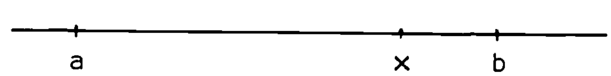
$$d_1 = (ra+sb) - a = sb - (1-r)a = sb - sa = s(b-a)$$

$$d_2 = b - (ra+sb) = (1-s)b - ra = rb - ra = r(b-a)$$

so that

$$\frac{d_1}{d_2} = \frac{s(b-a)}{r(b-a)} = \frac{s}{r}$$

It is further true that every number x between a and b can be expressed as a convex combination of a and b.



We leave it to you to check that choosing

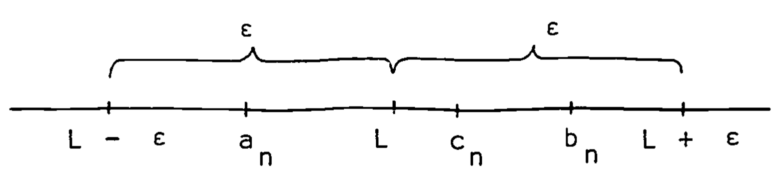
$$r = \frac{b - x}{b - a} \qquad s = \frac{x - a}{b - a}$$

yields $r \geq 0$, $s \geq 0$, $r + s = 1$, and

$$ra + sb = \frac{b - x}{b - a}\, a + \frac{x - a}{b - a}\, b = x.$$

<u>Theorem 10</u>. (Squeeze Theorem) Suppose that $\lim_{n\to\infty} a_n = L = \lim_{n\to\infty} b_n$. Further suppose that c_n lies between a_n and b_n for $n = 1, 2, \ldots$; Then the sequence $c_1, c_2, c_3, \ldots$ converges and $\lim_{n\to\infty} c_n = L$.

<u>Proof:</u> Let $\varepsilon > 0$. We know that we can find N_1 so that for $n > N_1$, a_n will lie within a distance ε of L. Similarly, we can find N_2 so that for $n > N_2$, b_n will lie within a distance ε of L. Letting N be the larger of N_1 and N_2, we can see that for $n > N$, both a_n and b_n lie within a distance ε of L, whence c_n lying between a_n and b_n will also lie within a distance ε of L.



<u>Example</u>. Consider the sequence c_n defined by $c_n = 2^{\frac{1}{n}}$

By the binomial formula,

$$(1+\frac{1}{n})^n = 1 + n(1)^{n-1}(\frac{1}{n}) + \ldots \geq 2.$$

Taking the n^{th} root, we obtain $(1+\frac{1}{n}) \geq 2^{\frac{1}{n}}$. Also $2^{\frac{1}{n}} \geq 1$. Since $\lim_{n\to\infty} (1 + \frac{1}{n}) = 1$ and $\lim_{n\to\infty} 1 = 1$, it follows by the Squeeze Theorem that $\lim_{n\to\infty} 2^{\frac{1}{n}} = 1$.

Corollary 1. If $r_n \geq 0$ and $s_n \geq 0$ and $r_n + s_n = 1$ holds for all integers n and $\lim_{n\to\infty} a_n = L = \lim_{n\to\infty} b_n$, then

$$\lim_{n\to\infty} (r_n a_n + s_n b_n) = L.$$

[It should be noted that we do not assume that the sequences $r_1, r_2, \ldots$ and $s_1, s_2, \ldots$ converge.]

Proof: Letting

$$c_n = r_n a_n + s_n b_n$$

we see by the conditions on r_n and s_n that c_n is a convex combination of a_n and b_n and hence c_n lies between a_n and b_n. Since a_n and b_n converge to the same value L, then by the Squeeze Theorem, so also does c_n. That is,

$$\lim_{n\to\infty} c_n = \lim_{n\to\infty} (r_n a_n + s_n b_n) = L.$$

Corollary 2. If $\lim_{n\to\infty} a_n = L = \lim_{n\to\infty} b_n$ and $p_n \geq 0$, $q_n \geq 0$ and for no value of n are p_n and q_n both zero, then

$$\lim_{n\to\infty} \frac{p_n a_n + q_n b_n}{p_n + q_n} = L.$$

Proof: Let $r_n = \frac{p_n}{p_n + q_n}$ and $s_n = \frac{q_n}{p_n + q_n}$, so that $r_n \geq 0$, $s_n \geq 0$, and $r_n + s_n = 1$ whence the conclusion follows from the theorem.

PROBLEMS

1. Prove that $\lim_{n\to\infty} nr^n = 0$ if $|r| < 1$. Hint: If $r \neq 0$, we have, setting $\varepsilon = \frac{1}{|r|} - 1$,

$$0 < |nr^n| = \frac{n}{(1+\varepsilon)^n} < \frac{n}{1 + n\varepsilon + \frac{n(n-1)}{2}\varepsilon^2}$$

for all $n > 2$.

2. Prove that $\lim_{n\to\infty} n^k r^n = 0$ if $|r| < 1$ and k is a positive integer.

3. Let r be a number and let d_n be the sequence defined by

$$d_1 = 1$$
$$d_2 = 1 + 2r$$
$$d_3 = 1 + 2r + 3r^2$$
$$\vdots$$
$$d_n = 1 + 2r + 3r^2 + \ldots + nr^{n-1}.$$

(a) Calculate $rd_4 - d_4$.

(b) Calculate $rd_n - d_n$.

(c) Obtain a new expression for d_n by dividing the result of part (b) by $r-1$, assuming $r \neq 1$.

(d) Prove that $\lim_{n\to\infty} d_n$ exists if $|r| < 1$.

(e) If $|r| < 1$, what is the value of $\lim_{n\to\infty} d_n$?

4. Let r be a number and let d_n be the sequence

$$d_1 = a_1$$
$$d_2 = a_1 + a_2 r$$
$$d_3 = a_1 + a_2 r + a_3 r^2$$
$$\vdots$$
$$d_n = a_1 + a_2 r + a_3 r^2 + \ldots + a_n r^{n-1}$$

where $a_1, a_2, a_3, \ldots$ is the Fibonacci sequence.

(a) Calculate $rd_n - d_n$.

(b) Prove that $\lim_{n\to\infty} a_n r^n = 0$ if $|r| < \frac{1}{L}$, where

$$L = \lim_{n\to\infty} \frac{a_{n+1}}{a_n}.$$

(c) Calculate the limit of the result of part (a), assuming that $\lim_{n\to\infty} d_n$ exists and that $|r| < \frac{1}{L}$.

(d) Calculate $\lim_{n\to\infty} d_n$ assuming that $\lim_{n\to\infty} d_n$ exists and that $|r| < \frac{1}{L}$.

5. Use the Squeeze Theorem to show that $\lim_{n\to\infty} \sqrt[n]{2^n + 3^n} = 3$.

8. A Geometric Limit

Right now, we are mainly interested in the limit to be developed in this section as an example of the use of the Squeeze Theorem. Later on, we shall see that this limit is of very basic importance.

Suppose that $a_1, a_2, a_3, \ldots$ is a sequence of positive numbers converging to zero. Let

$$b_n = \frac{\sin a_n}{a_n}, \quad n = 1, 2, 3, \ldots$$

What can we say about the sequence $b_1, b_2, b_3, \ldots$? Does it converge? If so, to what value? The Quotient Theorem cannot be applied since $\lim_{n\to\infty} a_n = 0$.

This puzzler becomes accessible to us by the application of the Squeeze Theorem to a pair of inequalities which we will derive from geometric considerations. In this

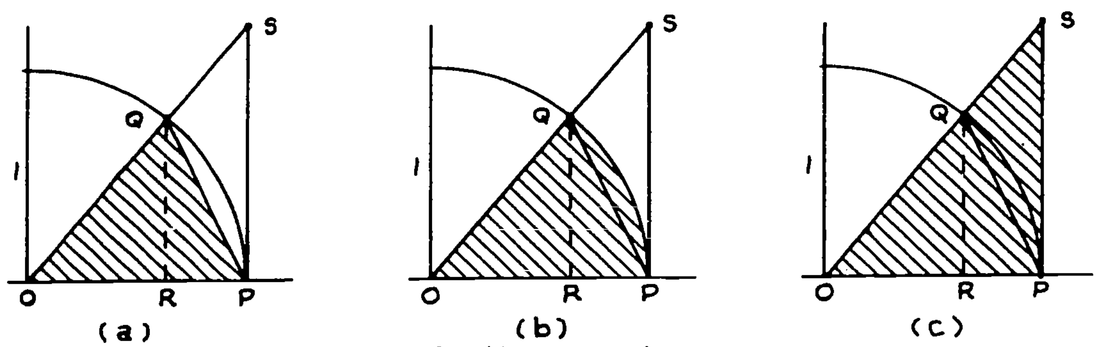


FIGURE 8-1

derivation, we need $a_n < \frac{\pi}{2}$, but since the a_n's converge to zero, this will be true if we restrict ourselves to sufficiently large values of n.

From the inclusion relations (see Figure 8-1)

$$\Delta OPQ \subset \text{sector } OPQ \subset \Delta OPS,$$

we conclude that

$$\text{area } \Delta OPQ \leq \text{area sector } OPQ \leq \text{area } \Delta OPS.$$

By simple trigonometric considerations, each of these areas is expressible in terms of x, the radian measure of angle POQ, where $0 < x < \frac{\pi}{2}$:

area $\Delta OPQ = \frac{\sin x}{2}$; area sector $OPQ = \frac{x}{2}$; area $\Delta OPS = \frac{\tan x}{2}$.

Therefore, the last inequality can be reexpressed as

$$\frac{\sin x}{2} \leq \frac{x}{2} \leq \frac{\tan x}{2}$$

Now the geometry has done its duty and we resort to simple algebraic manipulation to bring this inequality into a more usable form.

Multiplyi ı by the positive number $\frac{2}{\sin x}$ gives

$$1 \leq \frac{x}{\sin x} \leq \frac{1}{\cos x}$$

and taking reciprocals,

$$1 \geq \frac{\sin x}{x} \geq \cos x.$$

Thus, when n is large enough that $a_n < \frac{\pi}{2}$, we have

$$1 \geq \frac{\sin a_n}{a_n} \geq \cos a_n .$$

Since $\cos a_n$ = length of $\overline{OR}$, evidently $\lim_{n\to\infty} \cos a_n = 1$.

Also $\lim_{n\to\infty} 1 = 1$. Therefore by the Squeeze Theorem,

$$\lim_{n\to\infty} \frac{\sin a_n}{a_n} = 1.$$

This "proof" rests on certain facts concerning circles, areas, and arc lengths which you learned in your high school trigonometry course. In high school, these facts were supported by heuristic reasoning and appeal to intuition. Consequently, we cannot regard the proof given here as being entirely rigorous. Nevertheless, we will accept the results derived here until we are in a position to establish them firmly.

PROBLEMS

1. Suppose that a_n is a sequence of nonzero numbers such that $\lim_{n\to\infty} a_n = 0$. Evaluate each of the following limits, and justify your answers.

(a) $\lim_{n\to\infty} \frac{\sin a_n}{a_n}$ (Hint. What is $\frac{\sin(-x)}{-x}$ if $x \neq 0$?)

(b) $\lim_{n\to\infty} \frac{1 - \cos a_n}{a_n^2}$ (Hint. Use the formula $1 - \cos 2x = 2 \sin^2 x$)

(c) $\lim_{n\to\infty} \frac{1 - \cos a_n}{a_n}$

(d) $\lim_{n\to\infty} \sin a_n$

(e) $\lim_{n\to\infty} \tan a_n$

(f) $\lim_{n\to\infty} \frac{\sin ka_n}{a_n}$, where k is any constant. (Hint For $k \neq 0$, consider the sequence $b_n = ka_n$.)

(g) $\lim_{n\to\infty} \frac{\tan a_n - \sin a_n}{a_n^3}$. (Hint. Write in terms of $\sin a_n$ and $\cos a_n$ and express in terms of previously-determined limits.)

9. Completeness

Recall from Chapter 0 that a function f is said to be increasing if $f(a) \leq f(b)$ whenever $a<b$. In particular, a sequence c_n is *increasing* if $c_i \leq c_j$ whenever $i<j$. On the other hand, the sequence c_n is *decreasing* if $c_i \geq c_j$ whenever $i<j$.

In examples in earlier sections, we considered the Fibonacci sequence

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \ldots$$

and a certain sequence of ratios

$$r_1, r_2, r_3, r_4, \ldots$$

derived from the Fibonacci sequence by the rule,

$$r_n = \frac{a_{n+1}}{a_n} \qquad n = 1, 2, 3, \ldots$$

We had shown that *if* the sequence $r_1, r_2, r_3, \ldots$ converges, then it converges to $\frac{1 + \sqrt{5}}{2}$. Our computer output gave a strong indication that this sequence does converge to this value, but we have not yet been able to prove it. We will take up that question now.

In order to determine whether the sequence $r_1, r_2, r_3, \ldots$ actually converges, let us consider the differences d_n defined by

$$d_n = r_{n+1} - r_n .$$

By the definition of r_n, we have

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

Calling the numerator of this fraction p_n and using the recurrence relation for the terms of the Fibonacci sequence, we have

$$\begin{aligned} p_n &= a_{n+2}\, a_n - a_{n+1}\, a_{n+1} \\ &= (a_{n+1} + a_n)\, a_n - a_{n+1}\, (a_n + a_{n-1}) \\ &= a_{n+1}\, a_n + a_n a_n - a_{n+1}\, a_n - a_{n+1}\, a_{n-1} \\ &= a_n a_n - a_{n+1}\, a_{n-1} \\ &= -(a_{n+1}\, a_{n-1} - a_n a_n) \\ &= -\, p_{n-1} \end{aligned}$$

Using this relation repeatedly, we see that

$$p_n = -p_{n-1} = p_{n-2} = -p_{n-3} = \cdots = \pm p_1 ,$$

so that all p_n's alternate in sign and their absolute values are all equal to that of p_1, which is easily computed to be

$$p_1 = a_3a_1 - a_2a_2 = (2)(1) - (1)(1) = 2 - 1 = 1.$$

Thus, the p_n's alternately take on the values +1 and -1. And now we have

$$d_n = r_{n+1} - r_n = \frac{(-1)^{n+1}}{a_n a_{n+1}}$$

Consequently, the differences d_n alternate in sign and decrease in magnitude. (The denominators $a_n a_{n+1}$ obviously increase as n increases.) This means that the values of r_n alternately oscillate to the right and left with ever decreasing oscillations (see Figure 9-1).

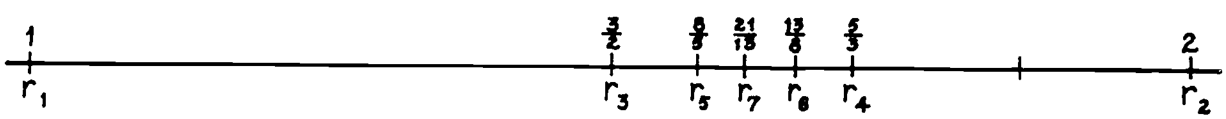


FIGURE 9-1

(Another way of saying this is that each r_n lies between its two immediate predecessors.) Furthermore

$$r_{n+1} - r_n = \frac{1}{a_n a_{n+1}}$$

which obviously approaches the limit 0. We can see that the r_n's with odd subscripts form an increasing sequence and the r_n's with even subscripts form a decreasing sequence, and the intervals

$$[r_1, r_2]$$
$$[r_3, r_4]$$
$$[r_5, r_6]$$
$$[r_7, r_8]$$
$$\vdots$$

are nested one within the other and shrink down to a point. Clearly, the sequence $r_1, r_2, r_3, \ldots$ converges.

Clear as the convergence of this sequence is, it cannot be proved. When these intervals shrink down to a point, there is nothing to guarantee us that there is a number associated with that point. This situation must be remedied by adopting some completeness axiom. (An axiom is a statement adopted without proof.) There are many statements which could be taken as completeness axioms with equivalent effect. The one most convenient for our purposes is the following:

Completeness Axiom. If $L_1, L_2, L_3, \ldots$ is an increasing sequence and $R_1, R_2, R_3, \ldots$ a decreasing sequence with $\lim_{n\to\infty} (R_n - L_n) = 0$, then the two sequences both converge to the same number.

This axiom assures us that the sequence $r_1, r_2, r_3, \ldots$ discussed above converges. It furthermore guarantees that the bisection process used in finding roots will always converge, even without the assumption that the function f has a root on any interval [c,d] for which f(c) and f(d) have opposite signs.

When this completeness axiom is adjoined to the Field, Order, and Archimedean axioms of Chapter 0, the development of our axiom system for the real numbers is finished. From these axioms, we can derive all the properties of the real number system. The completeness axiom constitutes the fundamental distinction between Algebra and Calculus. We will find that this axiom will be invoked over and over throughout the course to guarantee the existence of the basic concepts of calculus.

PROBLEMS

1. Let a sequence c_n be defined by $c_0 = 1$ and $c_{n+1} = \frac{2c_n + 5}{c_n + 2}$

   (a) If c_n converges to a number L, what is L?

   (b) Calculate c_1, c_2, c_3, c_4 as fractions and as decimal approximations to the nearest ten thousandth. Does this sequence seem to be converging to the value L calculated in part (a)?

   (c) Calculate c_{n+2} in terms of c_n and simplify.

   (d) Let $d_n = c_{n+1} - c_n$. Calculate $\frac{d_{n+1}}{d_n}$ in terms of c_n and simplify. (Use the formula $c_{n+1} = \frac{2c_n + 5}{c_n + 2}$ and (c).) Show that the values of d_n alternate in sign and that $\lim_{n\to\infty} d_n = 0$.

   (e) Explain how we can see that the sequence c_n does indeed converge.

   (f) How large must N be in order that $|c_n - L| < 10^{-9}$ for all $n > N$?

2. (a) Prove that $4xy \leq (x+y)^2$. (Hint: $(x-y)^2 \geq 0$.)

   (b) In the arithmetic-geometric algorithm of Problem 7, Section 4, prove that if $a_n \leq b_n$, then $a_{n+1} \leq b_{n+1}$.

   (c) Prove that if $a_n \leq b_n$, then $a_{n+1} \geq a_n$ and $b_{n+1} \leq b_n$.

   (d) Prove that if $a_n \leq b_n$, then $b_{n+1} - a_{n+1} \leq \frac{1}{2}(b_n - a_n)$. (Hint: Draw a picture.)

(e) Prove that $\lim_{n \to \infty} a_n$ and $\lim_{n \to \infty} b_n$ exist and are equal.

This proves the existence of the arithmetic-geometric mean M.

3. Suppose a sequence c_n is defined as follows:

$$c_1 = 1$$

$$c_2 = 1 - \frac{1}{2}$$

$$c_3 = 1 - \frac{1}{2} + \frac{1}{3}$$

$$c_4 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4}$$

$$\vdots$$

$$c_n = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots + (-1)^{n+1} \frac{1}{n}$$

We can form a new sequence R_n by choosing only the odd terms of the sequence c_n. Thus,

$$R_1 = 1$$

$$R_2 = 1 - \frac{1}{2} + \frac{1}{3}$$

$$R_3 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5}$$

$$\vdots$$

$$R_n = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} + \ldots + \frac{1}{2n-1}$$

Similarly, we can form a new sequence L_n choosing only the even terms of the sequence c_n.

(a) Using the Completeness Axiom, prove that the sequences R_n and L_n converge and that $\lim_{n\to\infty} R_n = \lim_{n\to\infty} L_n$.

(b) Prove that the sequence c_n converges and that $\lim_{n\to\infty} c_n = \lim_{n\to\infty} R_n = \lim_{n\to\infty} L_n$.

(c) Obtain a simple expression for $R_n - L_n$.

(d) What is the smallest value of n such that $R_n - L_n < \frac{1}{1000}$?

(e) Use the computer to approximate $\lim_{n\to\infty} c_n$ with an error of less than $\frac{1}{1000}$.

Chapter 3

AREA AND INTEGRAL

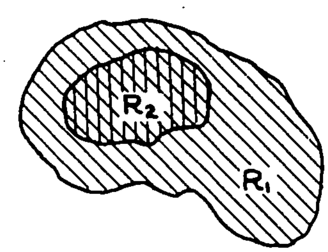
1. Area

There are five fundamental properties, all very natural, which form the basis for our development of the subject of area. These are:

I. Any bounded region in the plane has area, which is a nonnegative real number.

II. Congruent regions have the same area.

III. If the regions R_2 and R_1 are such that $R_2 \subset R_1$ then Area of $R_2 \leq$ Area of R_1.



IV. If a region R is decomposed into a number of non-overlapping parts, say R_1, R_2, R_3, then

Area of R = Area of R_1 + Area of R_2 + Area of R_3.

V. The area of a rectangular region is the length times the width.

We will not worry about units. We will always be working in the coordinate plane and will be using as our unit of length the unit used in constructing the coordinate system.

Knowing the area of a rectangle, we can immediately find the area of a right triangle. (See Figure 1-1.) In the rectangular region R, we see that a diagonal divides the rectangle into two right triangles which are congruent and therefore have equal areas.

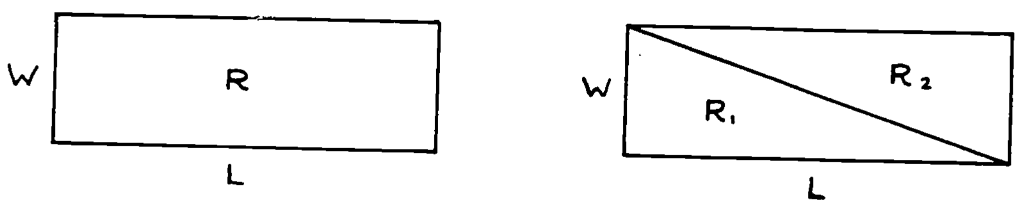


FIGURE 1-1

Thus

$$\text{Area of } R = 2(\text{area of } R_1)$$

so that

$$\text{Area of } R_1 = \frac{\text{area of } R}{2} = \frac{L \times W}{2}$$

This is the familiar fact that the area of a right triangle is half the product of the lengths of the legs.

We can further find the area of any triangle by a similar method. (See Figure 1-2.) We see that drawing an altitude

FIGURE 1-2

divides the triangle into two right triangles so that

$$\text{Area of } R = \text{Area of } R_1 + \text{Area of } R_2$$

$$= \frac{b_1h}{2} + \frac{b_2h}{2} = \frac{b_1h + b_2h}{2}$$

$$= \frac{(b_1 + b_2)h}{2} = \frac{bh}{2}$$

Thus the area of any triangle is half the base times the altitude. All this is of course very well known to you.

We can continue in this way to find the areas of polygonal regions, that is to say, regions whose boundaries are made up of line segments like the one shown in Figure 1-3. By drawing diagonals, such a region can always be decomposed into triangular regions whose area can be computed from the

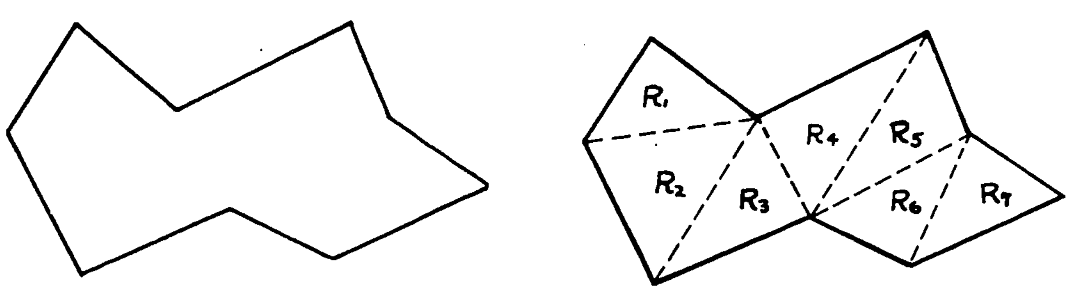


FIGURE 1-3

usual formula. And then the area of R is just the sum of the areas of the triangular regions.

The problem of computing the areas of regions whose boundaries are made up of line segments is now disposed of. We are ready to tackle one of the major problems of calculus-- the problem of finding the areas of regions with curved boundaries.

In calculus we usually work with the areas of regions having the configuration shown in Figure 1-4. That is to say, the regions are bounded on the bottom by the X-axis, on the sides by vertical lines, and on top by the graph of some function.

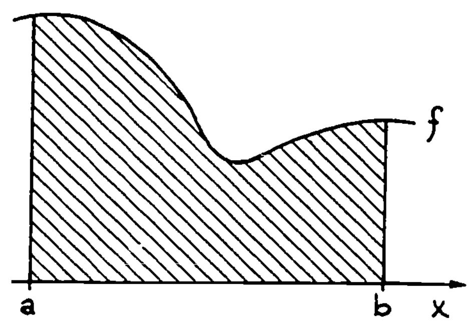
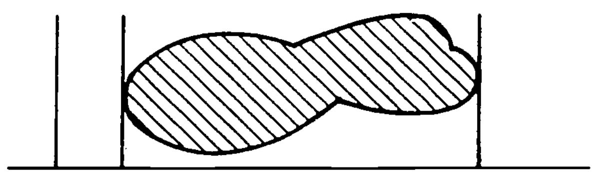
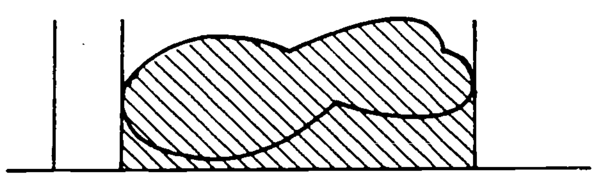


FIGURE 1-4

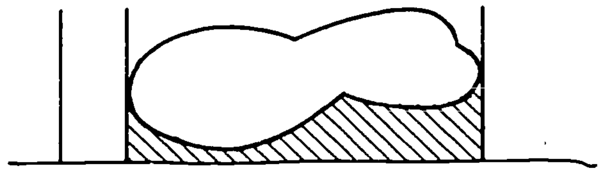
Thus, to find the area of a region shaped like this



we would first find this area:



then this one:



and then take the difference.

Although we are not in a position to find the exact area of any region with curved boundaries, still we can approximate such areas to any desired degree of accuracy.

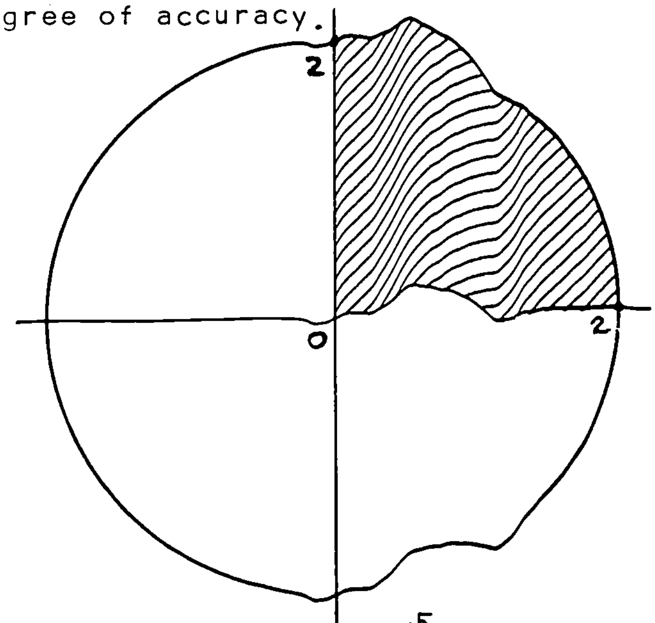


FIGURE 1-5

Let's see how this works out with a particular example. Let's compute the area of the quarter of the circle $x^2 + y^2 = 4$ depicted in Figure 1-5. You were taught in high school that the area of a circle with radius r is πr^2. Since the radius of the circle is 2, the area of the quarter circle is $\frac{\pi}{4} \cdot 2^2 = \pi$. As a check we will see how closely our computed area agrees with the well-known approximate values of π.

We see that the region is of the special configuration described above. It is bounded by the X-axis, the vertical lines $x = 0$, $x = 2$, and the graph of a certain function. We find a formula for this function by solving the equation $x^2 + y^2 = 4$ for y. Now

$$y = \pm \sqrt{4 - x^2}$$

So the function under consideration is given by

$$f(x) = \sqrt{4 - x^2}\ ,\ 0 \leq x \leq 2.$$

Although we cannot calculate the area exactly, we can approximate it by "rectangular configurations," a term we shall use to indicate regions composed of adjacent rectangles. In Figure 1-6, we have a region composed of three adjacent

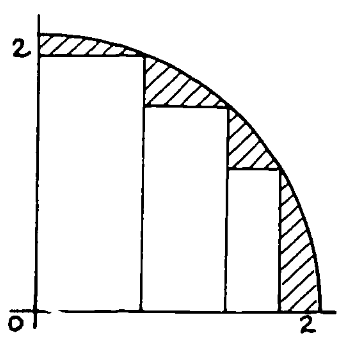


FIGURE 1-6

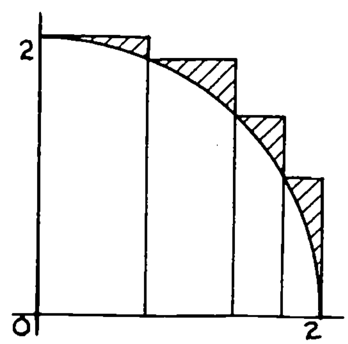


FIGURE 1-7

rectangles contained in the quarter circle. The area of the shaded part is the amount by which the area of the quarter circle exceeds that of the rectangular configuration.

In Figure 1-7, we see that the quarter circle is contained in a rectangular configuration composed of four rectangles.

The shaded area is the amount by which the total area of the rectangles exceeds that of the quarter circle.

Figure 1-8 shows Figures 1-6 and 1-7 superimposed. The shaded area represents the difference between the areas of the rectangular configuration in Figure 1-7 and that in Figure 1-6.

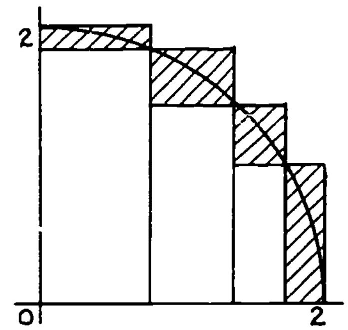


FIGURE 1-8

By repeating this process with a large number of rectangles, the shaded area can be made quite small. See Figure 1-9.

It looks as though, by using thinner and thinner rectangles, we can get closer and closer approximations to the area of the quarter circle. Let's now actually compute some of these rectangular configuration areas for our quarter circle.

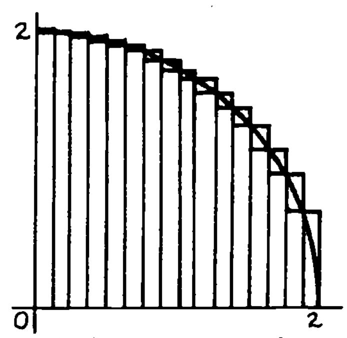


FIGURE 1-9

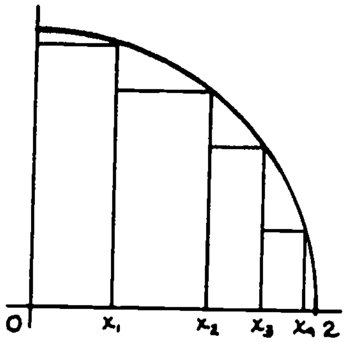


FIGURE 1-10

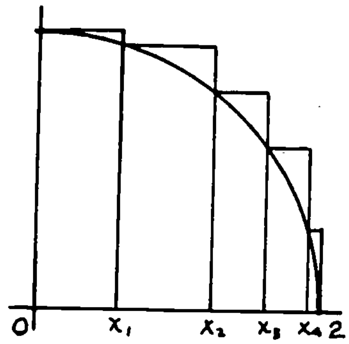


FIGURE 1-11

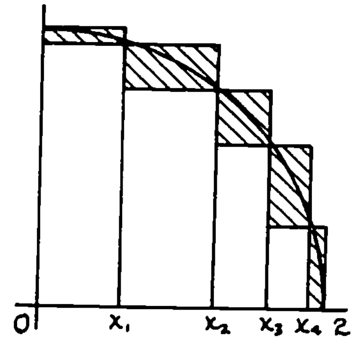


FIGURE 1-12

In Figures 1-10, 1-11, and 1-12, we have plotted points x_1, x_2, x_3, x_4, between 0 and 2, and we have sketched the corresponding inscribed and circumscribed rectangular configurations along with the superimposed rectangles, showing the difference of the areas.

The values of x_1, x_2, x_3, x_4, in the picture were chosen as,

$$x_1 = .56 \quad x_2 = 1.2 \quad x_3 = 1.6 \quad x_4 = 1.92$$

The areas of the four rectangles in Figure 1-10 are

$$(x_1 - 0)f(x_1),\ (x_2 - x_1)f(x_2),\ (x_3 - x_2)f(x_3),$$
$$(x_4 - x_3)f(x_4)$$

where

$$f(x) = \sqrt{4 - x^2}$$

224

Substituting in the values we obtain for the four areas

Rectangle	Width	Height	Area
1	.56 - 0 = .56	$\sqrt{4 - (.56)^2} = 1.92$	(.56)(1.92) = 1.0752
2	1.2 - .56= .64	$\sqrt{4 - (1.2)^2} = 1.6$	(.64)(1.6) = 1.024
3	1.6 - 1.2= .4	$\sqrt{4 - (1.6)^2} = 1.2$	(.4)(1.2) = .48
4	1.92- 1.6= .32	$\sqrt{4 - (1.92)^2} = .56$	(.56)(.32) = .1792
		Total	2.7584

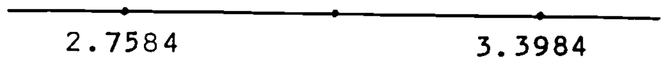
Similar computations yield for the sum of the areas of the five rectangles in Figure 1-11,

$$(x_1-0)f(0) + (x_2-x_1)f(x_1) + (x_3-x_2)f(x_2) + (x_4-x_3)f(x_3) + (2-x_4)f(x_4)$$

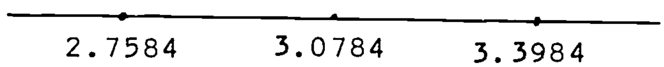
or

$$.56(2) + .64(1.92) + .4(1.6) + .32(1.2) + .08(.32) = 3.3984.$$

The true area, π, of the quarter circle lies between these two estimates. That is, π lies somewhere in the interval [2.7584, 3.3984]



and hence the distance between π and the midpoint of the interval cannot exceed half the length of the interval.



That fact may be expressed in the form

$$|\pi - 3.0784| \leq .32$$

Actually, our estimate, 3.0784, differs from π by less than .064 .

This average of our upper and lower estimates can be seen to be the sum of the areas of the four trapezoids and one triangle shown in Figure 1-13.

The area is again shaded and is seen to be much less than half the shaded area in Figure 1-12.

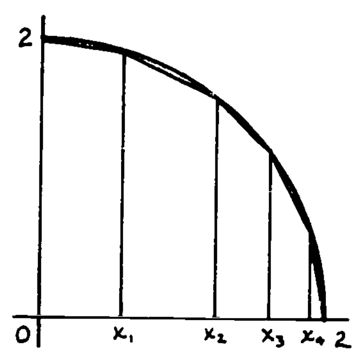


FIGURE 1-13

In this section, we have seen how our general principles regarding area have enabled us to find approximations of the areas of regions with curved boundaries. In the next section, we will improve our method so as to find sequences of approximations which converge to the actual area.

PROBLEMS

1. (a) Draw a graph of the function $y = x^2 + 5$.

   (b) Choose four points x_1, x_2, x_3, and x_4 between 0 and 3. Draw the five rectangles under the curve $y = x^2 + 5$ with bases $[0,x_1]$, $[x_1,x_2]$, $[x_2,x_3]$, $[x_3,x_4]$, and $[x_4, 3]$.

   (c) Compute the sum of the areas of the five rectangles drawn in (b).

   (d) Repeat parts (b) and (c), this time with rectangles above the curve $y = x^2 + 5$.

   (e) Estimate the area of the region in the first quadrant under the curve $y = x^2 + 5$ by averaging your results from parts (c) and (d).

   (f) Calculate the difference between your estimate made in (c) and the true value, which is 24.

2. Draw a flow chart for a program to do the computations like those in Problem 1 parts (c), (d), and (e) for a function f on an interval $[A,B]$. The program should read numbers x_1, x_2, ..., x_N which partition the interval $[A,B]$ into N + 1 parts $[A,x_1]$, $[x_1,x_2]$, ..., $[x_N,B]$. Assume that the function f is nonnegative and monotone on the interval $[A,B]$.

2. An Algorithm for Area

In order to devise an algorithm for calculating areas to any desired degree of accuracy, we need only make finer and finer subdivisions of our intervals. This leads to longer and more tedious calculations so that we would naturally prefer to have these calculations done by a computer.

The first step in developing our algorithm is to analyze the error. Suppose that f is a monotone function over an interval [a,b]. Let us subdivide the interval by means of points x_1, x_2, x_3, ..., x_7 and construct the upper and lower sums according to the method in the preceding section. In Figure 2-1, the total area of the shaded rectangles represents the difference between the upper and lower sums.

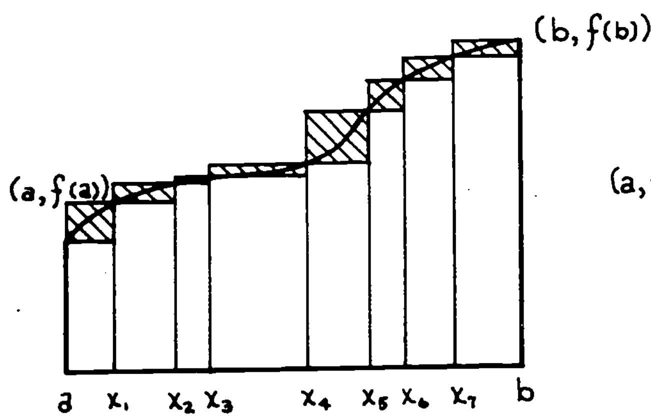


FIGURE 2-1

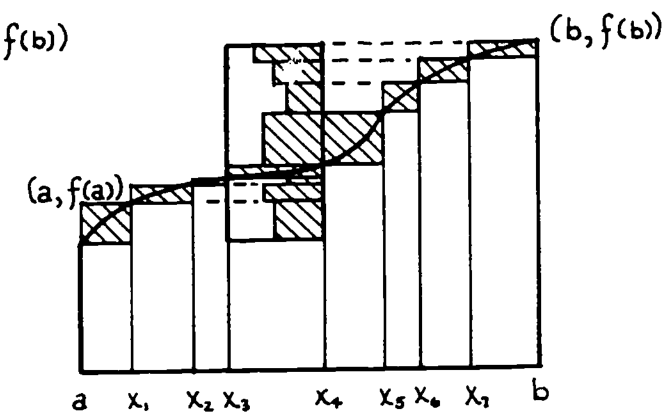


FIGURE 2-2

In Figure 2-2, all these shaded rectangles have been slid horizontally so as to fit, without overlapping, into a rectangle situated above the widest of our subintervals. The area of this rectangle is $|f(b) - f(a)|\cdot(x_4 - x_3)$ and thus our upper sum U and our lower sum L satisfy

$$U - L \leq |f(b) - f(a)|\cdot(x_4 - x_3)$$

In general, this can always be done provided that the function f is monotone, and we will always have

$$U - L \leq |f(b) - f(a)|\cdot\delta$$

where δ is the width of the widest subinterval in our partitioning of the interval [a,b].

Using the average of the upper and lower sums

$$T = \frac{U + L}{2}$$

as an approximation of our area, we have

$$|T - \text{Area}| \leq \frac{U - L}{2} \leq \frac{1}{2}\,|f(b) - f(a)|\cdot\delta$$

Thus, the number $\frac{1}{2}|f(b) - f(a)|\cdot\delta$ is a bound for our error. This bound can be made as small as we like by choosing our partition so as to make δ sufficiently small.

If we take a sequence of partitions of the interval [a,b] where δ_1, δ_2, δ_3, ... converges to zero, then

$$|T_n - \text{Area}| \leq \frac{1}{2}|f(b) - f(a)| \cdot \delta_n$$

so that T_1, T_2, T_3, ... will converge to the Area as a limit.

One simple way of constructing this sequence of partitions is by successively halving the intervals of the preceding partition, as in Figure 2-3. In this way, all the intervals in the n^{th} partitioning of [a,b] have the same length, namely $\frac{b-a}{2^n}$, so that

$$\delta_n = \frac{b-a}{2^n}$$

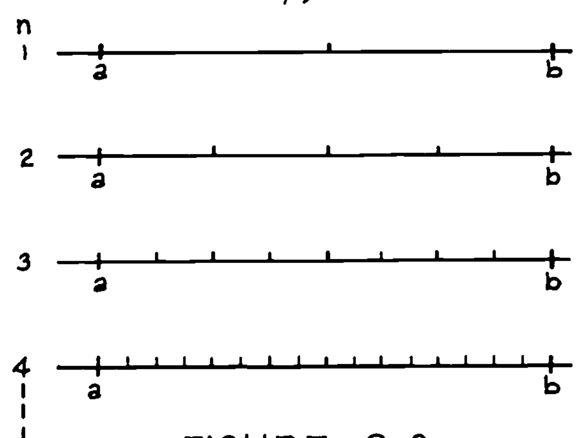


FIGURE 2-3

We can see that in successive computations of the lower sums, each rectangle is replaced by two rectangles with a greater combined area.

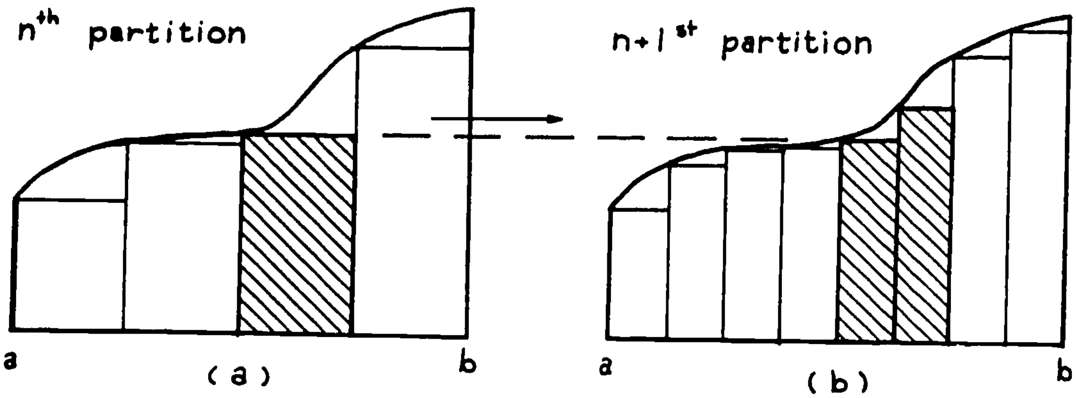


FIGURE 2-4

While in successive computations of the upper sums, each rectangle is replaced by two rectangles with a smaller combined area.

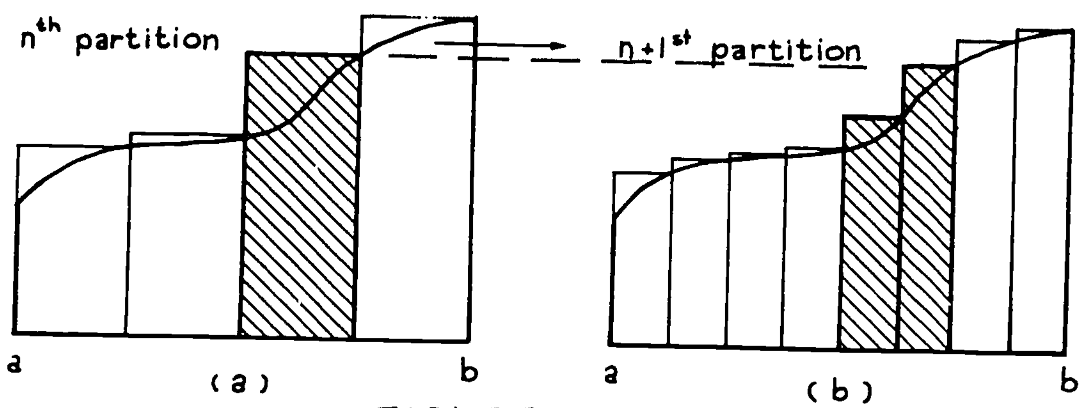


FIGURE 2-5

Now the sequence L_1, L_2, L_3 ... is an increasing sequence while U_1, U_2, U_3, ... is a decreasing sequence. Moreover, $L_n \leq U_n$ for n = 1, 2, ..., and $U_n - L_n$ converges to zero since

$$0 \leq U_n - L_n < \frac{f(b) - f(a)}{2^n}(b - a)$$

Therefore, by our completeness axiom, $\lim_{n\to\infty} U_n = \lim_{n\to\infty} L_n$.

We have obtained this conclusion without using the assumption that the area of the region exists.

The area of the trapezoidal approximation to the area is

$$T_n = \frac{U_n + L_n}{2}$$

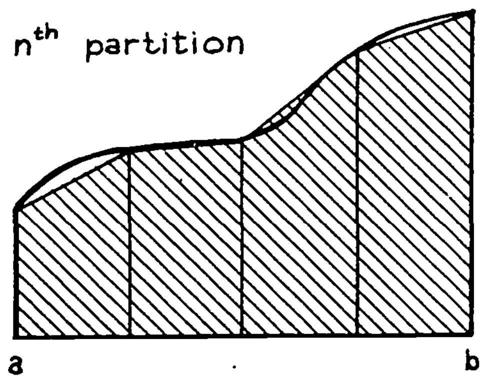


FIGURE 2-6

Since

$$L_n \leq T_n \leq U_n,$$

the squeeze theorem assures us that the sequence T_1, T_2, T_3, ... converges to the common value of $\lim_{n\to\infty} L_n$ and $\lim_{n\to\infty} U_n$.

In computing L_n and U_n, we will let h represent the width of the subintervals. (They all have the same width, namely, $\frac{b-a}{2^n}$) The values of L_n and U_n are then given by

$$L_n = f(a)h + \sum_{k=1}^{2^n-1} f(a + kh)h$$

$$U_n = \sum_{k=1}^{2^n-1} f(a + kh)h \quad + f(b)h$$

and we can make use of the distributive law to write L_n and U_n in the form

$$L_n = h[f(a) + \sum_{k=1}^{2^n-1} f(a + kh)]$$

$$U_n = h[f(b) + \sum_{k=1}^{2^n-1} f(a + k \cdot h)]$$

Introducing a variable SUM to stand for

$$SUM = \sum_{k=1}^{2^n-1} f(a + kh)$$

we have the formulas,

$$L_n = h(f(a) + SUM), \quad U_n = h(f(b) + SUM).$$

The computation of L_n and U_n will thus be accomplished by the process shown in Figure 2-7. The flow chart for the entire process of generating a sequence of upper and lower sums and trapezoidal sums is seen in Figure 2-8.

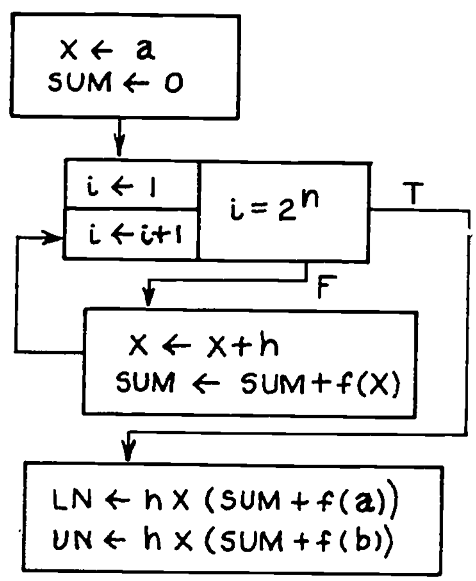


Figure 2-7

Variables NUM ($=2^n$) and FA (=f(a)) and FB (=f(b)) have been introduced to reduce repeated computation. You should especially note how h gets its successive values by being repeatedly halved. The final output value of TN is guaranteed to differ from the true area by less than the input value of ε.

The program works for decreasing functions as well as increasing functions, except that in this case the upper sums will appear as the LN outputs and the lower sums as the UN outputs. You should see why this is the case.

In Figure 2-9, we see a variant of Figure 2-8 which will reduce the computing time by half. The basis

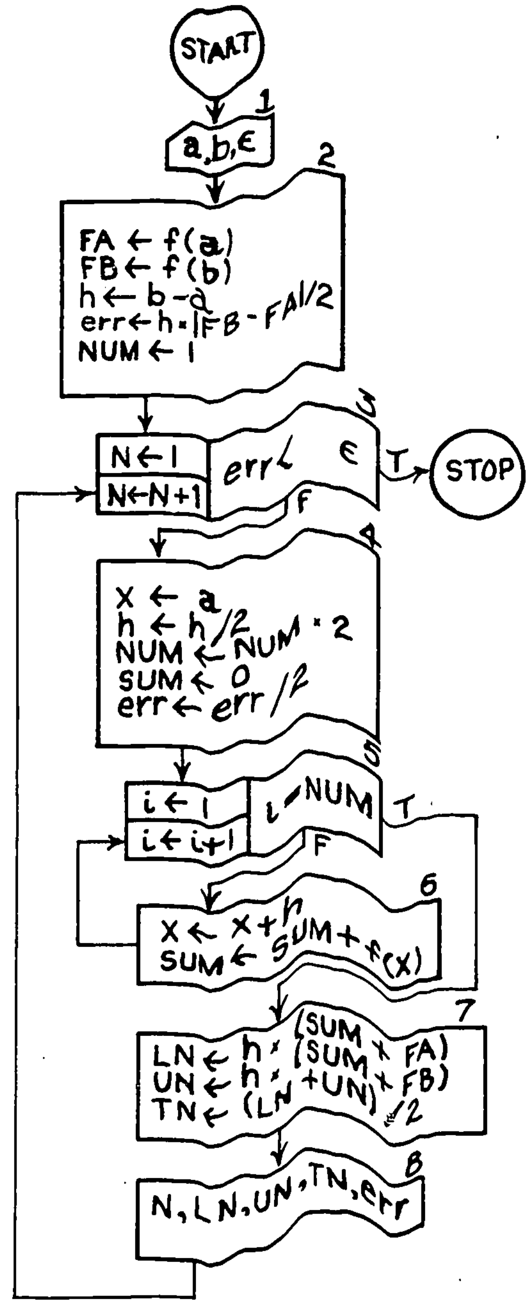


Figure 2-8

for this improvement is as follows. In executing the loops of boxes 5 and 6 of Figure 2-8, the value of f(x) is computed for $2^n - 1$ different values of x. However, $2^{n-1} - 1$ of these values have already been computed and summed in the previous pass through the loop. The variables TOT and d(= 2 x h) are introduced to eliminate this source of inefficiency. How the revised flow chart works is left for you to discover for yourself.

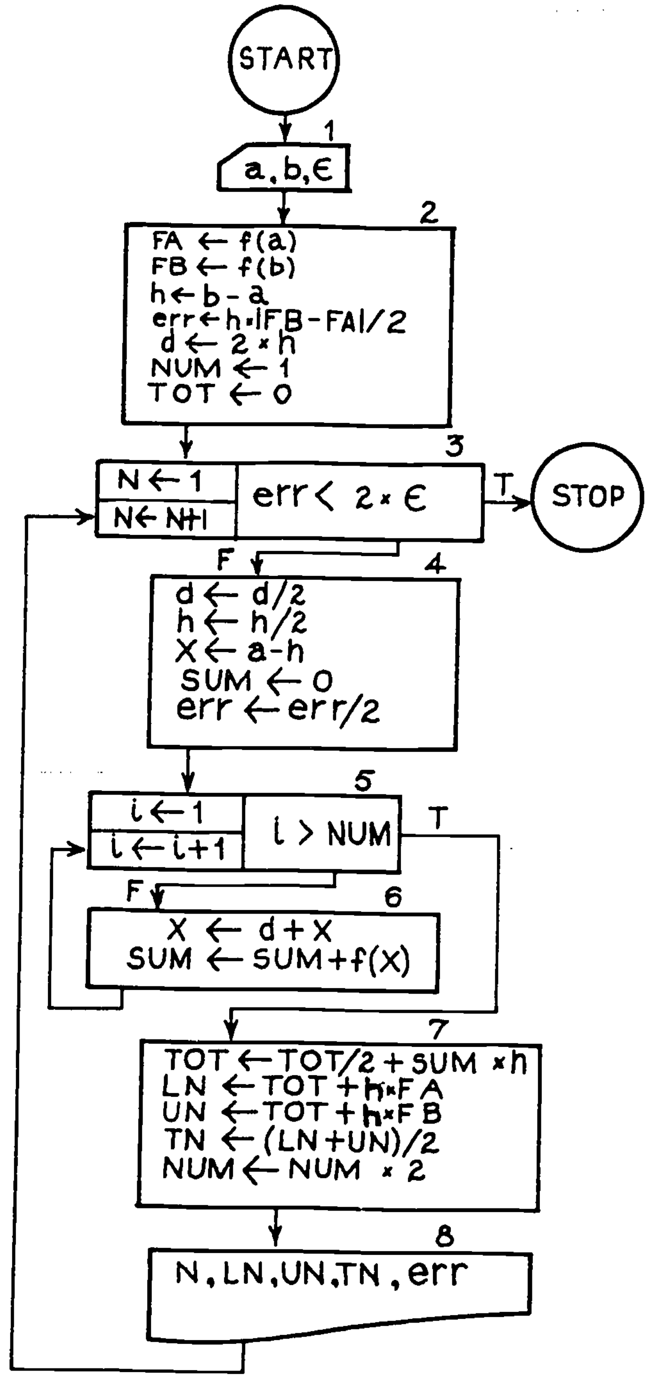


Figure 2-9

PROBLEMS

1. Write a program for the flow chart of Figure 2-8 or Figure 2-9. Run this program with the function in the previous section.

$$f(x) = \sqrt{4 - x^2}, \quad a = 0,\ b = 2$$

Remember that the true value of the area is π. Compare your final value of TN with tabulated values of π.

2. Write a flow chart for computing a trapezoidal approximation T to the area under the curve $y = F(x)$ on the interval [A,B], where F is a monotone function which is nonnegative on [A,B]. First have the program calculate how small the subintervals must be in order to guarantee that T will differ from the true area by no more than ε. Then calculate T using equal subintervals of appropriate width.

3. Write the program flow charted in Problem 2 and use it to approximate the area under each of the following curves. Use $\varepsilon = .001$.

(a) $f(x) = \sqrt{x}$ on [0,1]

(b) $f(x) = \sqrt{x}$ on [1,2]

(c) $f(x) = \sqrt{x + 1}$ on [0,1]

(d) $f(x) = \sin x$ on $[0, \frac{\pi}{2}]$

(e) $f(x) = \sin \frac{x}{2}$ on $[0, \pi]$

(f) $f(x) = x^2$ on $[0, 1]$

(g) $f(x) = \frac{1}{x^2 + 1}$ on $[0, 1]$

(h) $f(x) = \cos x$ on $[0, 1]$

(i) $f(x) = \cos x$ on $[0, \frac{\pi}{2}]$

(j) $f(x) = x^3$ on $[0, 1]$

3. Non-Monotone Functions

The process we have developed suffices to compute the area under the graph of any monotone function. In a certain sense, this will suffice for our needs, because in undergraduate mathematics virtually all the functions we encounter are either monotone or "piecewise monotone". By "piecewise monotone" we mean that the domain can be divided up into a number of intervals of monotonicity as depicted in Figure 3-1. Now the area under the curve can be obtained by computing separately the areas under the monotone pieces and adding as indicated in Figure 3-2.

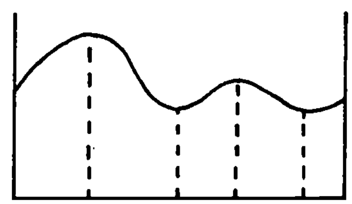


Figure 3-1

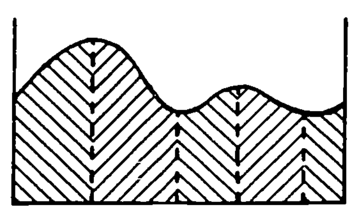


Figure 3-2

In another sense, however, the situation is not quite satisfactory. There are two reasons for this. First, sums and products of piecewise monotone functions are not necessarily piecewise monotone

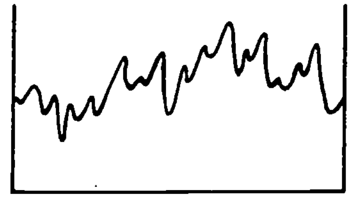


Figure 3-3

(though we will not give an example of such a situation here). This would lead to the necessity of qualifying some of the theorems we wish to prove later on. Secondly, even though a function may be piecewise monotone, it can have a great many maxima and minima (as in Figure 3-3), and the problem of actually locating these points may be a practical impossibility.

There is, however, another means of controlling error in estimating areas which does not require the ability to locate the maxima and minima. For this purpose, we will relax our definitions of upper and lower sums (i.e., make them more general).

Looking at Figure 3-4(a), we see the graph of a function f with the area under the graph shaded. In Figure 3-4(b), we see a rectangular configuration including the area under the graph. We will call the area of such a region an <u>upper sum</u>. In Figure 3-4(c), we see a rectangular configuration entirely contained in the region illustrated in Figure 3-4(a). We will call the area of such a region a <u>lower sum</u>.

If we denote the heights of the rectangles in Figure 3-4(b) as M_1, M_2, ..., M_7 and the heights of those in Figure 3-4(c) as m_1, m_2, ..., m_7, then we see that the upper sum

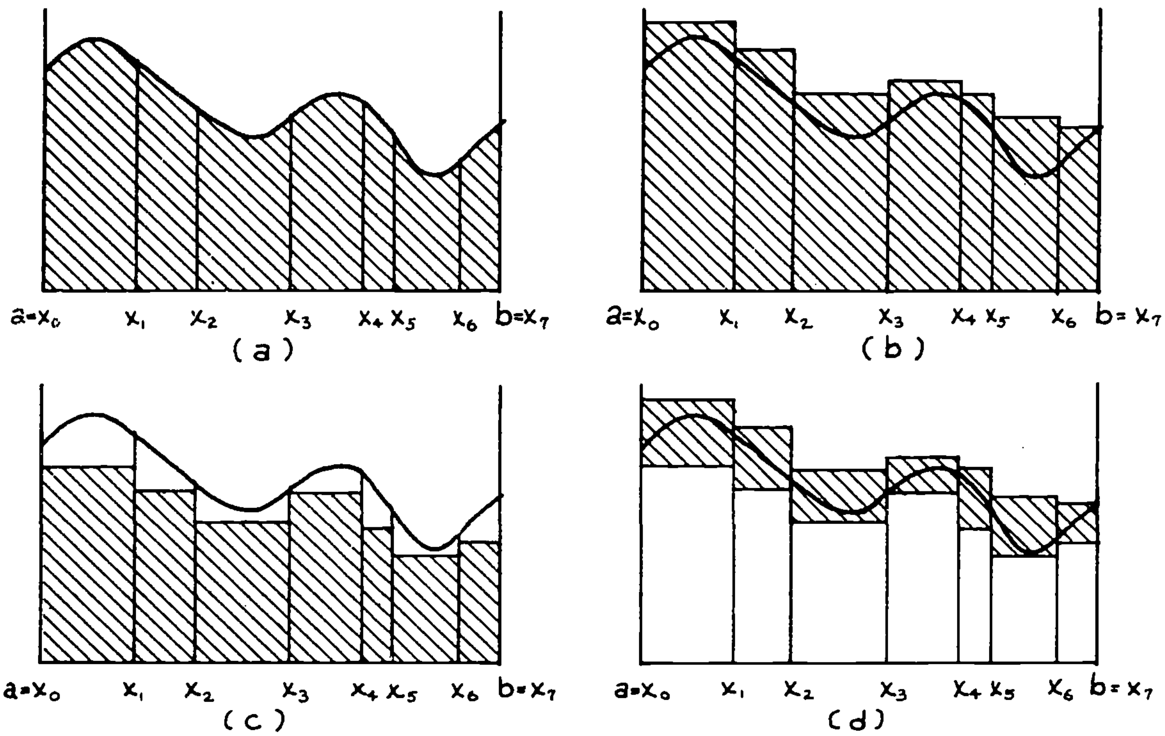


Figure 3-4

U and the lower sum L are given by the formulas

$$\overline{S} = \sum_{k=1}^{7} M_k(x_k - x_{k-1}) \qquad \text{and} \qquad \underline{S} = \sum_{k=1}^{7} m_k(x_k - x_{k-1}).$$

In Figure 3-4(d), the rectangles in Figures 3-4(b) and 3-4(c) have been superimposed and the shaded area represents the difference U - L . The heights of the shaded rectangles in the figure are

$$M_1 - m_1, M_2 - m_2, \ldots, M_7 - m_7 .$$

If, as in Figure 3-4(d), each of these heights is $\leq 2\varepsilon$, then we could as in Figure 3-5(a) "drop each of the shaded rectangles down to the bottom of the elevator shaft", and then we see in Figure 3-5(b) that the whole configuration fits inside a rectangle

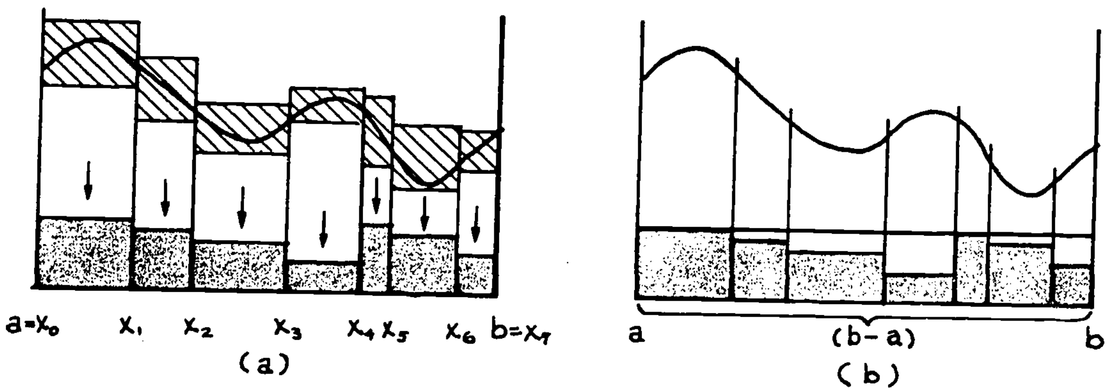
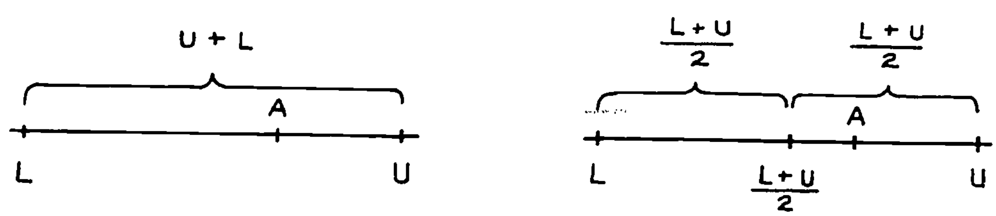


Figure 3-5

of area $2\varepsilon(b - a)$. That is, $U - L \leq 2\varepsilon(b - a)$. And now, since the area under the curve, A, lies between L and U,



we observe that A lies within a distance $(U - L)/2$ of the average $(L + U)/2$ of L and U. That is,

$$\left| A - \frac{L + U}{2} \right| \leq \frac{U - L}{2} \leq \varepsilon(b - a).$$

In this way, we can make the error in estimating the area as small as we like provided that our function has the property that for every $\varepsilon > 0$ we can find a partition $x_0, x_1, x_2, \ldots, x_n$ and upper and lower bounds, M_k and m_k, on each of the subintervals so that for each subinterval we have

$$M_k - m_k \leq 2\varepsilon.$$

But what functions have this property? How are we to find such partitions and such numbers M_k and m_k? We address ourselves to these questions a little later on. What we need for now are the generalized concepts of lower and upper sums.

4. Integrals

The process we have developed for computing areas has many applications. The mathematical name for the limit found by this process is "the integral from a to b of f", written

$$\int_a^b f(x)\,dx$$

(The reason for the "dx" will appear later. At present consider it merely as part of the symbol indicating integration.) We no longer require that the function f be positive over the interval [a,b]. We will be able to find integrals of functions such as that in Figure 4-1.

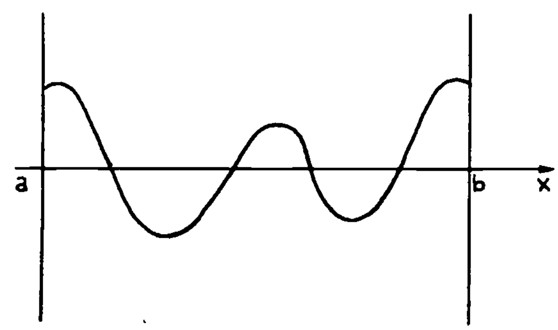


FIGURE 4-1

The area interpretation of such an integral would be the shaded area above the X-axis minus the shaded area below the X-axis, as shown in Figure 4-2.

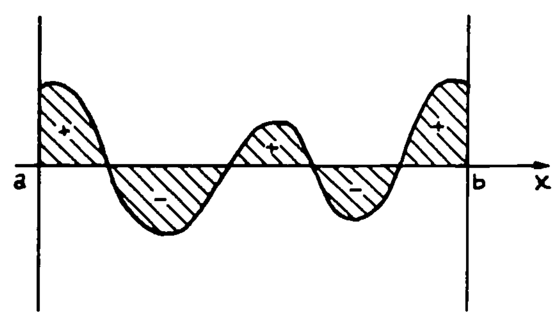


FIGURE 4-2

However, it is not always profitable to think of these integrals in terms of area.

Below we give a formal definition of the integral. You will see that this definition coincides with what we have been doing in finding areas.

Definition: Let f be a function defined on an interval [a,b]. Suppose there is a sequence L_n, n = 1, 2, ..., of lower sums over this interval and a sequence U_n, n = 1, 2, 3, ..., of upper sums with $\lim_{n\to\infty} (U_n - L_n) = 0$. Then the common limit, L, of these two sequences is denoted as

$$\int_a^b f(x)\,dx$$

Up until now we have been using the assumption made in the first section of this Chapter that the regions under consideration have areas. On the basis of this assumption, it is easy to see that if there exist sequences satisfying the conditions in this definition, then they must converge to the area, A, under the curve. For then we have $L_n \leq A \leq U_n$ for all integers n, and since $\lim_{n\to\infty} (U_n - L_n) = 0$, we have also $\lim_{n\to\infty} (U_n - A) = 0$ and $\lim_{n\to\infty} (A - L_n) = 0$.

In order to make our development rigorous, it is necessary to free ourselves from the assumption that regions under the graphs of functions have areas. However, once this assumption is dropped, our definition of the integral is open to serious objections. First, it is not clear that sequences satisfying the conditions in the definition necessarily converge. Second, it is not clear that another pair of sequences L_n' and U_n' satisfying the several conditions, but based on different partitioning of the interval [a,b], will necessarily converge to the same limit. These objections are disposed of in Appendix B to this chapter. The discussion given there is easy-going and informal but somewhat lengthy. We hope that the student will read this discussion now or at least before leaving this chapter in order to appreciate the simple steps necessary to validate the above definition; it makes our theory of integration dependent only on the field, order, Archimedean, and completeness axioms for the real number system.

Although we now allow the possibility that f may assume negative as well as positive values, we will show that we may nevertheless confine our discussion to positive functions. To see that this is so, suppose that f assumes both positive and negative values in the interval [a,b]. Let -K be a lower bound for f in the interval [a,b] so that

$$g(x) = f(x) + K \geq 0 \text{ for } x \text{ in } [a,b].$$

Let

$$L = \sum_{i=1}^{n} m_i(x_i - x_{i-1})$$

be a lower sum, for f in [a,b]. (Here some or all of the m_i may be negative, but let them all be taken $\geq -K$.) Taking $m_i' = m_i + K$, we see that

$$L' = \sum_{i=1}^{n} m_i'(x_i - x_{i-1})$$

is a lower sum for g on [a,b]. Next we calculate

$$L' = \sum_{i=1}^{n} m_i'(x_i - x_{i-1}) = \sum_{i=1}^{n} (m_i + K)(x_i - x_{i-1})$$

$$= \sum_{i=1}^{n} m_i(x_i - x_{i-1}) + \sum_{i=1}^{n} K(x_i - x_{i-1})$$

$$= L + K(b - a)$$

[Conversely, if we had started with L' being given, then taking $m_i = m_i' - K$, we could have computed the lower sum L so that $L = L' - K(b-a)$.] Similar results for upper sums are obtained by replacing the m_i and m_i' by M_i and M_i'. Here we obtain $U' = U + K(b-a)$.

Now if we have sequences L_n and U_n for f, then the sequences L_n' and U_n' constructed according to the above rule satisfy

$$U_n' - L_n' = [U_n + K(b - a)] - [L_n + K(b - a)]$$

$$= U_n - L_n$$

Consequently, if either of the sequences $U_n - L_n$ or $U_n' - L_n'$ converges to zero, then both do since they are the same. Moreover,

$$\int_a^b f(x)dx = \lim_{n\to\infty} L_n = \lim_{n\to\infty} [L_n' - K(b - a)]$$

$$= \lim_{n\to\infty} L_n' - K(b - a) = \int_a^b g(x)dx - K(b - a).$$

Accordingly, we have the following recipe for finding the integral of a function f which assumes negative values:

(1) Find K so that $f(x) + K$ is ≥ 0 throughout $[a,b]$;

(2) Construct g by $g(x) = f(x) + K$;

(3) Calculate the integral $\int_a^b g(x)dx$.

(4) Subtract $K(b-a)$ from this result to find $\int_a^b f(x)dx$.

This is something we never need to do in practice; it is introduced solely for the purpose of justifying the use of methods which apply only to positive functions in proving facts about integrals. We give one example of finding such a number K merely to guarantee that our meaning will be clear.

Example. Let $f(x) = 3x + 5 \sin x$. Find a number K so that $g(x) = f(x) + K$ is ≥ 0 throughout the interval $[-4,6]$.

Solution: Throughout the interval $[-4,6]$, the inequalities

$$x \geq -4 \quad \text{and} \quad \sin x \geq -1$$

hold so that

$$3x \geq -12 \quad \text{and } 5 \sin x \geq -5$$

whence

$$3x + 5 \sin x \geq -17.$$

Therefore,

$$g(x) = f(x) + 17 = 3x + 5 \sin x + 17 \geq 0$$

throughout [-4,6].

Before going on we pause to briefly consider the question: What functions have integrals? That is, for what functions do there exist sequences of lower and upper sums L_n and U_n with $\lim_{n\to\infty} (U_n - L_n) = 0$? We have already seen that such sequences can be found when f is monotone. There is no need in this course for knowing the most general class of functions for which integrals exist, but we will return to this question later in this chapter.

PROBLEMS

1. Repeat Problem 1 of Section 3-1 with the function $y = x^2 - 5$. For part (f), evaluate $\int_0^3 (x^2 - 5)dx$ assuming that $\int_0^3 (x^2 + 5)dx = 24$.

2. Suppose that K is a number and that f is defined on an interval [A,B] by $f(x) = K$. Show that $\int_A^B f(x)dx = K(B - A)$. Interpret this result geometrically in the case $K > 0$.

5. Theory of Integration

We next present several important theorems concerning integrals. In order to streamline the statements of these theorems we adopt the convention that the integrals on the right side of the equal sign are given to exist. The theorem then assures us of the existence of the integral on the left as well as the stated equality. We also assume that a, b, and c belong to an interval in the domain of f.

Theorem 1. If $a < b < c$ then

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx$$

Proof: Since $\int_a^b f(x)\,dx$ and $\int_b^c f(x)\,dx$ are given to exist there are sequences L_n' and U_n' of lower and upper sums over [a,b] and sequences L_n'' and U_n'' of lower and upper sums over [b,c] with

$$\lim_{n\to\infty} (U_n' - L_n') = 0 \qquad \text{and} \qquad \lim_{n\to\infty} (U_n'' - L_n'') = 0 .$$

We define lower and upper sums L_n and U_n for f over [a,c] by

$$L_n = L_n' + L_n'' \qquad \text{and} \qquad U_n = U_n' + U_n''$$

as illustrated in Figure 5-1 where L_n and U_n are respectively represented by the entire shaded areas in Figures 5-1(a) and 5-1(b)

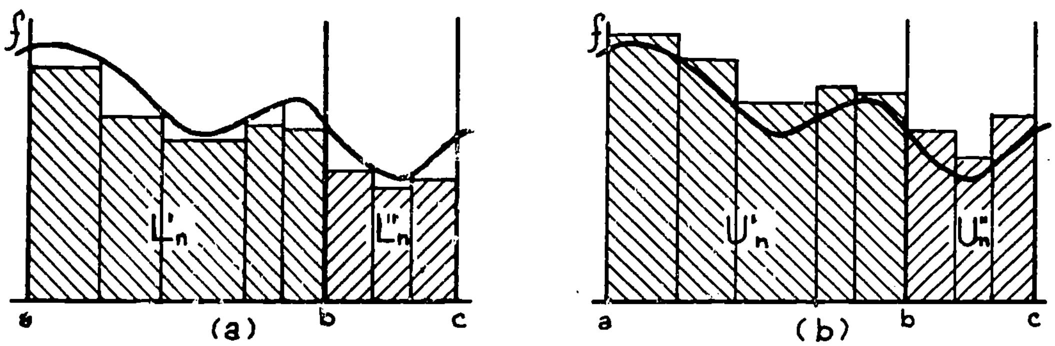


Figure 5-1

Now we see that

$$U_n - L_n = (U_n' + U_n'') - (L_n' + L_n'')$$

$$= (U_n' - L_n') + (U_n'' - L_n'')$$

whence by the theorem on the limit of the sum,

$$\lim_{n\to\infty} (U_n - L_n) = \lim_{n\to\infty} (U_n' - L_n') + \lim_{n\to\infty} (U_n'' - L_n'') = 0 + 0 = 0$$

This shows then that a sequence of the desired form exists so that $\int_a^c f(x)dx$ exists. As for the value of this integral (working with upper sums, although lower sums would do as well)

252

$$\int_a^c f(x)dx = \lim_{n\to\infty} U_n = \lim_{n\to\infty} (U_n' + U_n'')$$

$$= \lim_{n\to\infty} U_n' + \lim_{n\to\infty} U_n''$$

$$= \int_a^b f(x)dx + \int_b^c f(x)dx.$$

<u>Definition 1</u>. $\int_a^a f(x)dx = 0$

<u>Definition 2</u>. If $a > b$, we define

$$\int_a^b f(x)dx = -\int_b^a f(x)dx.$$

<u>Theorem 2</u>. If C is a constant then

$$\int_a^b C\cdot f(x)dx = C\cdot\int_a^b f(x)dx$$

<u>Proof</u>: Consider the case that $C \geq 0$. Let L_n and U_n be lower and upper sums for f on $[a,b]$ with $\lim_{n\to\infty} (U_n - L_n) = 0$. For each value of n, L_n and U_n are expressible in the form

$$L_n = \sum_{k=1}^{n} m_k(x_k - x_{k-1}) \qquad \text{and} \qquad U_n = \sum_{k=1}^{n} M_k(x_k - x_{k-1}).$$

Now we have

$$m_k \leq f(x) \leq M_k \text{ for } x \text{ in } [x_{k-1}, x_k]$$

25?

whence

$$C \cdot m_k \leq Cf(x) \leq C \cdot M_k \text{ for } x \text{ in } [x_{k-1}, x_k].$$

(For negative C these inequalities are reversed.) Hence Cm_k and CM_k are lower and upper bounds for f on $[x_{k-1}, x_k]$, and therefore

$$CL_n = \sum_{k=1}^{n} Cm_k(x_k - x_{k-1}) \quad \text{and} \quad CU_n = \sum_{k=1}^{n} CM_k(x_k - x_{k-1})$$

are lower and upper sums for Cf on [a,b]. Now

$$\lim_{n\to\infty} (CU_n - CL_n) = C \lim_{n\to\infty} (U_n - L_n) = (C)(0) = 0$$

assures us of the existence of $\int_a^b Cf(x)dx$. Finally,

$$\int_a^b Cf(x)dx = \lim_{n\to\infty} CU_n = C \lim_{n\to\infty} U_n = C \int_a^b f(x)dx.$$

Lemma. If L' and U' are lower and upper sums for f over [a,b] and L" and U" are lower and upper sums for g over [a,b], then

$$L = L' + L'' \quad \text{and} \quad U = U' + U''$$

are lower and upper sums for f + g over [a,b].

Proof: We will show this only for lower sums. In Figure 5-2(a) and 5-2(b) we illustrate lower sums for f and g over [a,b]. Figures 5-2(c) and 5-2(d) are the same except that vertical lines have been drawn at all partition points of both partitions. This shows that L' and L" may be regarded as lower sums with respect to the same partition. Let us denote this partition consisting of all the partition points by

$$a = x_0, x_1, x_2, \ldots, x_{n-1}, x_n = b$$

and we let the heights of the rectangles over the subintervals in Figure 5-2(c) be identified as $m_1', m_2', \ldots m_n'$ and those in Figure 5-2(d) as $m_1'', m_2'', \ldots m_n''$.

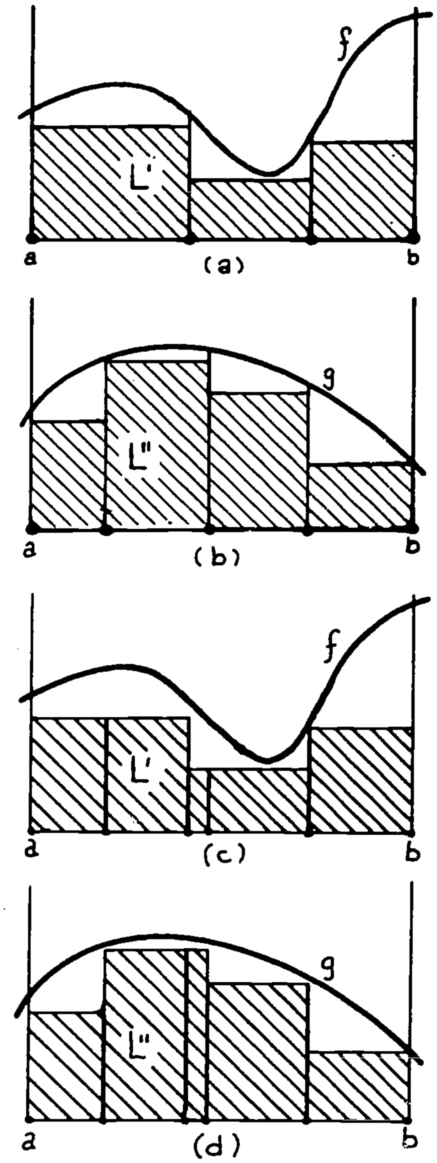


Figure 5-2

Now we see that $m_k' \leq f(x)$ and $m_k'' \leq g(x)$ for x in $[x_{k-1}, x_k]$. Adding, we get

$$m_k' + m_k'' \leq f(x) + g(x) \text{ for } x \text{ in } [x_{k-1}, x_k].$$

Hence $m_k = m_k' + m_k''$ is a lower bound for $f + g$ on $[x_{k-1}, x_k]$ and the sum

$$L = \sum_{k=1}^{n} m_k(x_k - x_{k-1})$$

$$= \sum_{k=1}^{n} (m_k' + m_k'')(x_k - x_{k-1})$$

$$= \sum_{k=1}^{n} m_k'(x_k - x_{k-1}) + \sum_{k=1}^{n} m_k''(x_k - x_{k-1})$$

$$= L' + L''.$$

The result for upper sums is proved in the same way.

Theorem 3. (Integral of the sum)

$$\int_a^b [f(x) + g(x)]dx = \int_a^b f(x)dx + \int_a^b g(x)dx.$$

Proof: Let L_n' and U_n' be sequences of lower and upper sums for f over $[a,b]$, and let L_n'' and U_n'' be sequences of lower and upper sums for g over $[a,b]$ with

$$\lim_{n\to\infty} (U_n' - L_n') = 0 \quad \text{and} \quad \lim_{n\to\infty} (U_n'' - L_n'') = 0.$$

Now the preceding lemma allows us to define a sequence of upper and lower sums for f + g over [a,b] by

$$L_n = L_n' + L_n'' \text{ and } U_n = U_n' + U_n'' \text{ for } n = 1, 2, 3, \ldots$$

Since

$$\begin{aligned} U_n - L_n &= (U_n' + U_n'') - (L_n' + L_n'') \\ &= (U_n' - L_n') + (U_n'' - L_n'') \end{aligned}$$

we can see that

$$\lim_{n\to\infty} (U_n - L_n) = \lim_{n\to\infty} (U_n' - L_n') + \lim_{n\to\infty} (U_n'' - L_n'') = 0 + 0 = 0.$$

This shows that $\int_a^b [f(x) + g(x)]dx$ exists.

Consequently,

$$\begin{aligned} \int_a^b [f(x) + g(x)]dx &= \lim_{n\to\infty} U_n \\ &= \lim_{n\to\infty} (U_n' + U_n'') \\ &= \lim_{n\to\infty} U_n' + \lim_{n\to\infty} U_n'' \\ &= \int_a^b f(x)dx + \int_a^b g(x)dx \end{aligned}$$

It might be thought that Theorems 2 and 3 would be of little practical value from the computational point of view as they say nothing about bounds on the error.

However, if we assume we have shown that $I_1 = \int_0^3 f(x)dx \approx 2.377$ with error $\leq .002$ and that $I_2 = \int_0^3 g(x)dx \approx 1.162$ with error $\leq .005$, then

$$\begin{array}{c} 2.377 - .002 \leq I_1 \leq 2.377 + .002 \\ 1.162 - .005 \leq I_2 \leq 1.162 + .005 \\ \hline 3.539 - .007 \leq I_1 + I_2 \leq 3.539 + .007 \end{array}$$

Then, $\int_0^3 (f(x) + g(x))dx \approx 3.539$ with error $\leq .007$. Thus, in general when approximating the integral of the sum of two functions, we can add the approximations found for the functions separately and add the error bounds.

Similarly, with the conditions as above $I_3 = \int_0^3 4f(x)dx = 4I_1$ satisfies $\quad 4(2.377 - .002) \leq 4I_1 \leq 4(2.377 + .002)$

or $\quad 9.508 - .008 \leq 4I_1 \leq 9.508 + .008.$

Again we see that in approximating the integral of a constant multiple of a function, we multiply the approximation of the original function by the constant and multiply the error bound for the integral of the original function by the absolute value of the constant.

The remaining theorems of this section find several applications in Section 9 of this chapter.

Theorem 4. If $a \leq b$ and $f(x) \geq 0$ for all x in [a,b], and if the indicated integral exists, then

$$\int_a^b f(x)dx \geq 0.$$

Proof: Clearly each of the upper sums U_n in the definition of the integral is greater than or equal to zero and hence

$$\int_a^b f(x)dx = \lim_{n\to\infty} U_n \geq 0$$

If the inequality $f(x) \geq 0$ in Theorem 4 were strengthened to $f(x) > 0$, then we could also replace the inequality in the conclusion by a strict inequality. This is difficult to prove in full generality and the proof will be omitted here.

Theorem 5. If $a < b$ and $f(x) \geq g(x)$ for all x in [a,b], and if the indicated integrals exist, then

$$\int_a^b f(x)dx \geq \int_a^b g(x)dx.$$

Proof: Using the theorem on the integral of the sum, we have

$$\int_a^b f(x)dx = \int_a^b g(x)dx + \int_a^b [f(x) - g(x)]dx$$

and the last integral on the right is ≥ 0 by Theorem 4.

<u>Theorem 6</u>. $$\int_{a+h}^{b+h} f(x - h)dx = \int_a^b f(x)dx$$

<u>Proof</u>: As illustrated in Figure 5-3, the graph of $f(x - h)$ over the interval $[a + h, b + h]$ is merely a shift to the

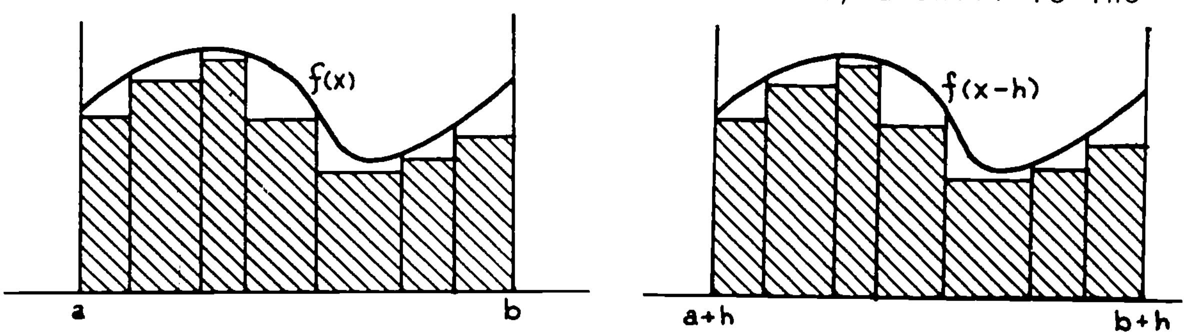


Figure 5-3

right of the graph of f over $[a,b]$. Moreover, from the congruence of the shaded regions in this figure we see that each lower sum for $f(x)$ over $[a,b]$ is also a lower sum for $f(x - h)$ over $[a + h, b + h]$. Similarly, for upper sums. Thus sequences L_n and U_n of lower and upper sums for $f(x)$ over $[a,b]$ with $\lim_{n\to\infty} (U_n - L_n) = 0$ are also sequences of lower and upper sums for $f(x - h)$ over $[a + h, b + h]$. Thus,

$$\int_a^b f(x)dx = \lim_{n\to\infty} U_n = \int_{a+h}^{b+h} f(x - h)dx.$$

PROBLEMS

1. Approximate each of the following integrals and specify an error bound in each case. Use the results of Problem 3 of Chapter 3, Section 2.

(a) $\int_0^1 3\sqrt{x}\, dx$

(b) $\int_0^{\pi/2} -2 \sin x\, dx$

(c) $\int_1^2 \sqrt{2x}\, dx$

(d) $\int_0^1 (x^2 - \sqrt{x})dx$

(e) $\int_0^1 (x^3 + \frac{4}{x^2+1})\, dx$

(f) $\int_0^2 \sqrt{x}\, dx$

(g) $\int_{-1}^1 \sqrt{x + 1}\, dx$

(h) $\int_{-\pi/2}^0 \sin (x + \frac{\pi}{2})dx$

(i) $\int_0^{\pi/2} (\sin x + \cos x)dx$

(j) $\int_0^{\pi/2} (\sin x - 3\cos x)dx$

6. Unicon Functions

For the class of monotone functions we have succeeded in demonstrating the existence of integrals and in approximating these integrals with bounds on the error. In Section 3 we generalized the definition of lower and upper sums so as to pave the way for the consideration of nonmonotone functions. We are ready to introduce another class of functions, the unicon functions, which we will show to have integrals which we can approximate to any desired degree of accuracy.

The idea of a unicon function is quite simple to understand if we think of it in terms of control of error. Suppose we have a function f and we want to compute the values of f(x) for several values of x all in an interval [a,b]. Perhaps these values of x are determined experimentally or perhaps they are subject to computer round off, but anyway suppose they are subject to error. We would like to know that a small error in the value of x will produce a correspondingly small error in the value of f(x).

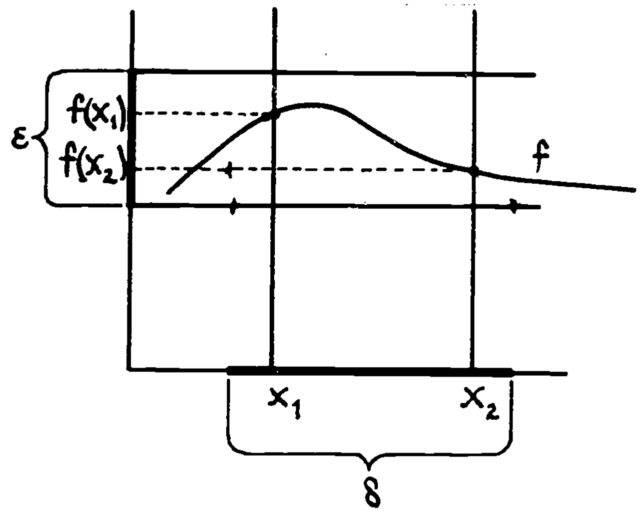


Figure 6-1

Putting it slightly differently, suppose that the maximum error we can permit in our computed values of f(x) is some positive number ε; is there some tolerance, δ, so that when the error in the value of x does not exceed δ then the error in f(x) will not exceed ε? That is, can we find a number δ so that

$$|f(x_1) - f(x_2)| \leq \varepsilon \quad \text{whenever} \quad |x_1 - x_2| \leq \delta \ ?$$

This situation is illustrated in Figure 6-1 for a particular choice of the numbers x_1 and x_2. Here the tolerances ε and δ are represented by the lengths of the heavily drawn intervals.

If the answer to this question is affirmative no matter how small the positive number ε may be, then we say that f is unicon over the interval [a,b]. Putting this as a formal definition, we have:

<u>Definition</u>. A function f is said to be <u>unicon</u> over the interval [a,b] provided that for every positive number ε there can be found a positive number δ so that $|f(x_1) - f(x_2)| \leq \varepsilon$ whenever x_1 and x_2 are in [a,b] and $|x_1 - x_2| \leq \delta$.

Most familiar functions are unicon. We give a few examples. In these examples it will be understood that x_1 and x_2 are always taken to be in the interval [a,b].

<u>Example 1</u>. $f(x) = 3x$ where [a,b] is arbitrary

$$|f(x_1) - f(x_2)| = |3x_1 - 3x_2| = 3|x_1 - x_2|.$$

Thus by choosing $\delta = \varepsilon/3$, we see that if $|x_1 - x_2| \leq \delta$ then

$$|f(x_1) - f(x_2)| = 3|x_1 - x_2| \leq 3\cdot\delta = 3\cdot\varepsilon/3 = \varepsilon.$$

<u>Example 2</u>. $f(x) = x^3$ where [a,b] is arbitrary

$$|f(x_1) - f(x_2)| = |x_1^3 - x_2^3|$$

$$= |(x_1 - x_2)(x_1^2 + x_1x_2 + x_2^2)|$$

$$= |x_1 - x_2|\cdot|x_1^2 + x_1x_2 + x_2^2| \leq |x_1 - x_2|\cdot 3k^2$$

where $k = \max(|a|, |b|)$. Thus letting $\delta = \varepsilon/3k^2$, we see that if $|x_1 - x_2| \leq \delta$ then

$$|f(x_1) - f(x_2)| \leq |x_1 - x_2| \cdot 3k^2 \leq \delta \cdot 3k^2 = (\varepsilon/3k^2) \cdot 3k^2 = \varepsilon.$$

Example 3. $f(x) = \sin x$ where $[a,b]$ is arbitrary.
The result is obvious from the following diagrams of the unit circle

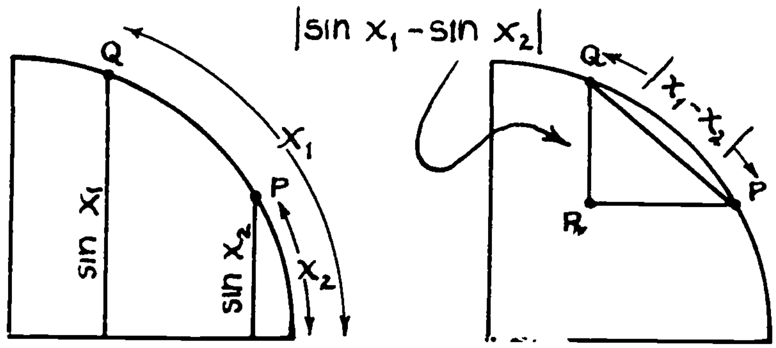


Figure 6-2

We see that $|x_1 - x_2|$ is the length of the arc PQ which is greater than the length of the chord $\overline{PQ}$ since a line segment is the shortest path joining its endpoints. Furthermore the chord $\overline{PQ}$ is longer than the segment $\overline{QR}$ since the hypotenuse of a right triangle is longer than either of its legs. Thus we have

$$|\sin x_1 - \sin x_2| = \text{length of } \overline{QR} \leq \text{length of } \overline{PQ}$$
$$\leq \text{length of PQ} = |x_1 - x_2|$$

Thus, taking $\delta = \varepsilon$ we see that if $|x_1 - x_2| \leq \delta$

$$|\sin x_1 - \sin x_2| \leq |x_1 - x_2| = \delta = \varepsilon.$$

The same method applies when the angles are in different quadrants. Note that we are using radian measure for angles as is always done in calculus for very good reasons which will eventually become apparent. The same argument works equally well for the cosine function.

Example 4. $f(x) = \frac{1}{x}$ where $0 < a < b$.

$$|f(x_1) - f(x_2)| = \left|\frac{1}{x_1} - \frac{1}{x_2}\right| = \left|\frac{x_1 - x_2}{x_1 x_2}\right|$$

$$= |x_1 - x_2| \cdot \frac{1}{|x_1 x_2|} \leq \frac{|x_1 - x_2|}{a^2}$$

Thus, taking $\delta = a^2\varepsilon$ we see that if $|x_1 - x_2| \leq \delta$ then

$$\left|\frac{1}{x_1} - \frac{1}{x_2}\right| \leq \frac{|x_1 - x_2|}{a^2} \leq \frac{\delta}{a^2} = \frac{a^2\varepsilon}{a^2} = \varepsilon.$$

Example 5. $f(x) = \sqrt{x}$ where $0 \leq a \leq b$

(This is a little bit tricky. It can be shown much more easily if a is given to be greater than 0.)

266

Let c and d represent numbers greater than or equal to zero and check the string of inequalities:

$$(c - d)^2 = |c - d| \cdot |c - d| \leq |c - d| \cdot |c + d| = |c^2 - d^2|$$

whence by taking square roots,

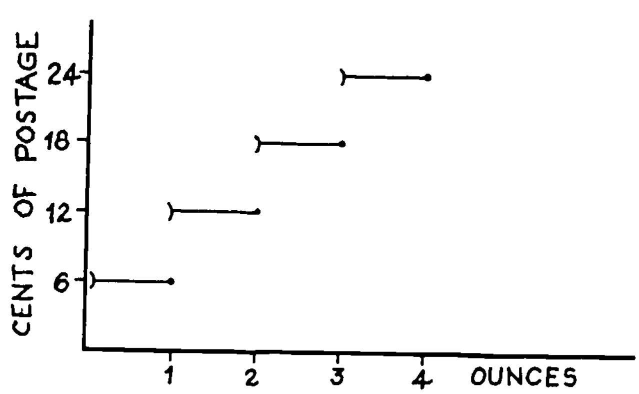
$$|c - d| \leq \sqrt{|c^2 - d^2|}.$$

Now we use this inequality to verify that

$$|f(x_1) - f(x_2)| = |\sqrt{x_1} - \sqrt{x_2}| \leq \sqrt{|x_1 - x_2|}$$

(We took $c = \sqrt{x_1}$ and $d = \sqrt{x_2}$). Hence, taking $\delta = \varepsilon^2$ we see that when $|x_1 - x_2| \leq \delta$ then

$$|f(x_1) - f(x_2)| \leq \sqrt{|x_1 - x_2|} \leq \sqrt{\delta} = \sqrt{\varepsilon^2} = \varepsilon.$$



POSTAGE FUNCTION, P

Figure 6-3

Example 6. As an example of a function which is not unicon, consider the "postage function", P, which gives the number of cents of postage as a function of the weight of the letter according to the formula, "six cents per ounce or fraction thereof." This function is graphed in Figure 6-3. To see that this function is not unicon on the interval [0,4], note that we may choose x_1 and x_2 as close together as we like with $x_1 < 2 < x_2$. For example, take $x_1 = 2 - \frac{1}{n}$ and $x_2 = 2 + \frac{1}{n}$ for some large integer n. Then,

$$P(x_1) = 12 \quad \text{while} \quad P(x_2) = 18.$$

Clearly if now $\varepsilon < 6$, we will not be able to find a $\delta > 0$ so that

$$|P(x_1) - P(x_2)| \leq \varepsilon \quad \text{whenever} \quad |x_1 - x_2| \leq \delta .$$

It is evident that whenever such "jumps" occur in the graph of a function, the function cannot be unicon.

PROBLEMS

1. Gold leaf comes in square sheets in various sizes up to 10" on an edge. What is the maximum tolerance of error in measuring the length of an edge in order that the error in the computed value of the area should never exceed
   (a) 1 square inch?
   (b) .1 square inch?
   (c) .01 square inch?

2. Suppose f is defined on [0,1] by f(x) = 0 or 1 accordingly as x is rational or irrational, respectively. Prove that f is not unicon on [0,1].

3. Suppose $c < d < 0$. If $f(x) = \frac{1}{x}$ on [c,d], prove that f is unicon on [c,d].

4. Prove that if f is unicon on [a,b], and if $g(x) = f(x + c)$, then g is unicon on [a − c, b − c].

5. Use Problem 4 and Example 3 to show that the cosine function is unicon on any closed interval.

6. Find an expression for δ in terms of ε such that $|x_1^2 - x_2^2| < \varepsilon$ whenever $|x_1 - x_2| < \delta$ and x_1 and x_2 are in the interval:

(a) $[0,3]$

(b) $[-10,10]$

(c) $[a,b]$

7. Unicon Functions and Integrals

We are ready to show that unicon functions are integrable and to approximate their integrals with guaranteed error bounds. First we make the following simple observation.

Theorem 1. Suppose f to be defined on [a,b]. Suppose that for each $\varepsilon > 0$ there are lower and upper sums L and U with $U - L < 2\varepsilon\ (b - a)$. Then $\int_a^b f(x)dx$ exists.

Proof: For each positive integer n we let L_n and U_n be upper and lower sums with $U_n - L_n < \frac{2}{n}(b - a)$ so that $\lim_{n\to\infty}(U_n - L_n) = 0$.

Thus, by definition $\int_a^b f(x)dx$ exists.

Accordingly we will now show, for a function f unicon over [a,b], that we can find for any $\varepsilon > 0$ lower and upper sums L and U with $U - L < 2\varepsilon\ (b - a)$.

Suppose that function f, illustrated in Figure 7-1, is unicon in the interval [a,b]. Suppose that a particular value is chosen for ε and a value of δ has been found so that two numbers in [a,b] which differ by no more than δ will have

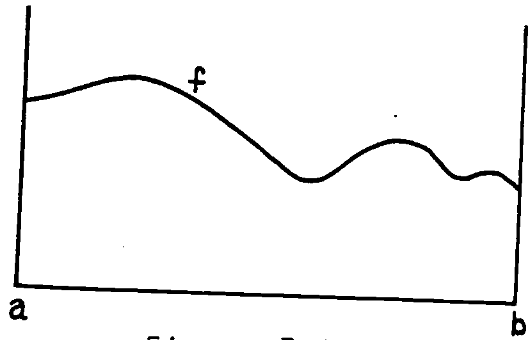


Figure 7-1

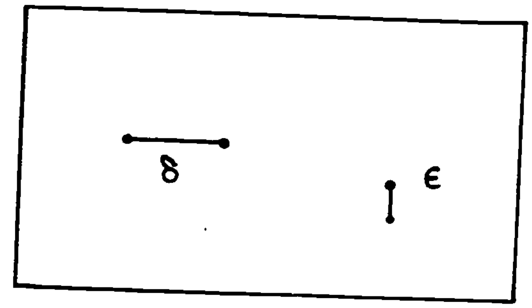


Figure 7-2

functional values which differ by no more than ε. Let these values of ε and δ be represented by the lengths of the segments in Figure 7-2.

Next consider two numbers, c and d, in [a,b] with $c < d$ and $|c - d| < \delta$. Select a number ξ in the interval [c,d] as shown in Figure 7-3(b). Since any number x in [c,d] lies

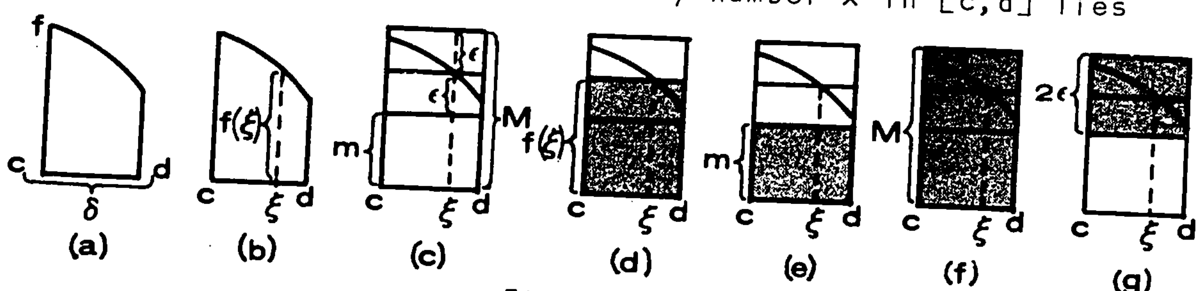


Figure 7-3

within a distance δ of ξ, it follows that f(x) will lie within a distance ε of $f(\xi)$. That is,

$$f(\xi) - \varepsilon \leq f(x) \leq f(\xi) + \varepsilon \qquad \text{for } x \text{ in } [c,d].$$

In other words, the numbers

$$m = f(\xi) - \varepsilon \qquad \text{and} \qquad M = f(\xi) + \varepsilon$$

are lower and upper bounds for function, f, over the interval $[c,d]$. This is exemplified in Figure 7-3(c) by the fact that the graph lies entirely between the horizontal lines. Since $f(\xi)$ is the average of m and M

$$\frac{m + M}{2} = \frac{f(\xi) - \varepsilon + f(\xi) + \varepsilon}{2} = \frac{2f(\xi)}{2} = f(\xi)$$

it is clear that the shaded area in Figure 7-3(d) is the average of those in Figures 7-3(e) and 7-3(f). In Figure 7-3(g) it is seen that the rectangle representing the differenc of the area in 3(f) and 3(e) has height exactly 2ε.

With this preparation, we see that if we partition the interval $[a,b]$ so that each subinterval has length no greater than δ then the above observations hold in each of these subintervals. Thus, in Figure 7-4(a) we have partitioned $[a,b]$ into six subintervals and chosen in Figure 7-4(b) (according to some unspecified rule) numbers ξ_k in each of the subintervals.

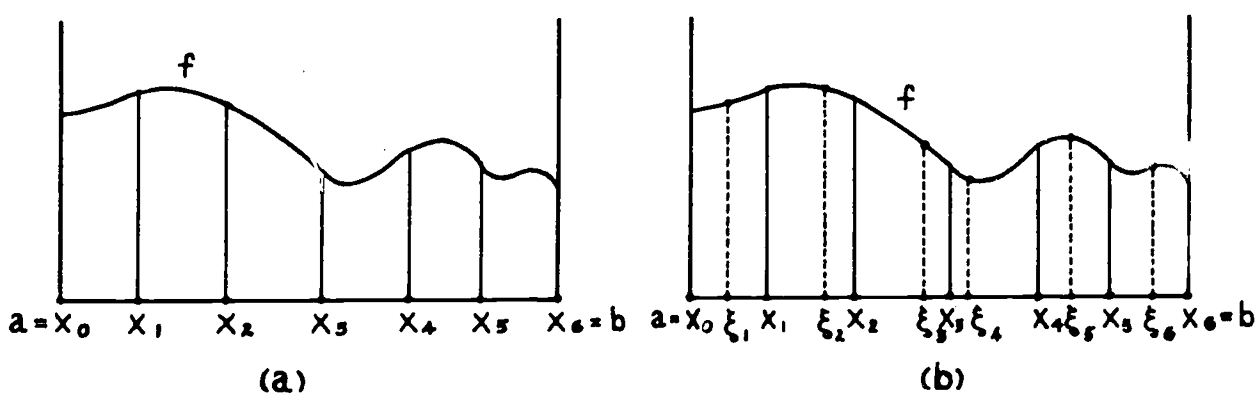


Figure 7-4

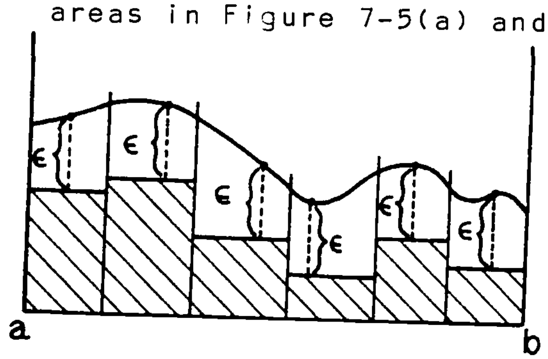
In each subinterval we have lower and upper bounds m_k and M_k given by

$$m_k = f(\xi_k) - \varepsilon, \qquad M_k = f(\xi_k) + \varepsilon \qquad k = 1, 2, \ldots, 6.$$

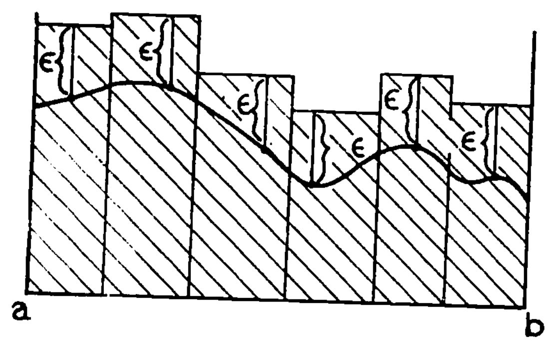
The lower and upper sums

$$L = \sum_{k=1}^{6} m_k(x_k - x_{k-1}) \qquad \text{and} \qquad U = \sum_{k=1}^{6} M_k(x_k - x_{k-1})$$

formed by use of these bounds are represented by the shaded areas in Figure 7-5(a) and 7-5(b).



(a)



(b)

Figure 7-5

The average of these upper and lower sums is

$$S = \sum_{k=1}^{6} f(\xi_k)(x_k - x_{k-1})$$

which is represented by the shaded area in Figure 7-6(a), while the difference of the upper and lower sums represented in Figure 7-6(b) is exactly $2\varepsilon(b-a)$.

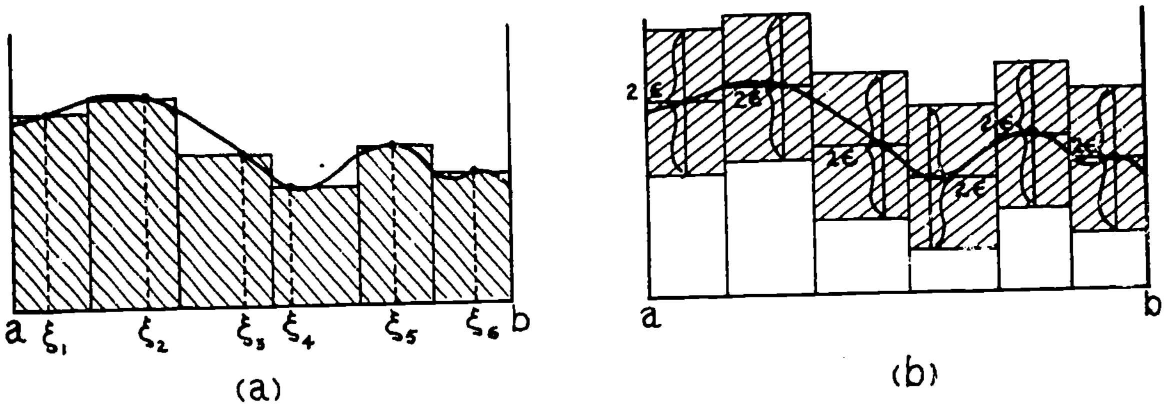
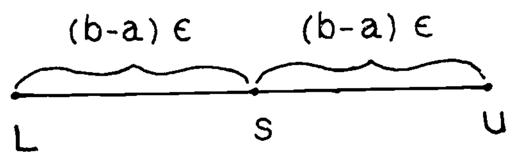


Figure 7-6

According to Theorem I of this section this shows that $I = \int_a^b f(x)dx$ exists. And furthermore, since

$$S = (U + L)/2 \quad \text{and } L \le I \le U$$



we see that the sum

$$S = \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1})$$

approximates the integral $\int_a^b f(x)dx$ with error not exceeding $(b-a)\epsilon$. This error estimate is valid so long as the lengths of subintervals do not exceed δ where δ is so chosen that $|f(x_1) - f(x_2)| \le \epsilon$ for all numbers x_1 and x_2 in $[a,b]$ with $|x_1 - x_2| \le \delta$.

This provides the basis of a computer algorithm for approximating integrals. For this purpose it will usually be convenient to take all the subintervals with the same length. Among the common rules for the choice of the ξ_k are;

i) the midpoint rule, $\xi_k = \frac{x_{k-1} + x_k}{2}$;

ii) the left-end-point rule, $\xi_k = x_{k-1}$;

iii) the right-end-point rule, $\xi_k = x_k$.

Since the approximations obtained by the left-and right-end-point rules both differ from $\int_a^b f(x)dx$ by less than $(b-a)\varepsilon$, so also does their average. This average is the now familiar "trapezoid rule" for approximation of integrals. In the following exercises we will compare estimates made by the two end-point rules, the midpoint rule and the trapezoid rule.

PROBLEMS

1. (a) Draw a flow chart for computing and printing approximations to $\int_A^B F(x)dx$ by the end-point rules, the midpoint rule, and the trapezoidal rule. Each of these four approximations is to be computed and printed for N = 2, 4, 8, 16, ... partitions until $\frac{B - A}{N} <$ DELTA, where DELTA is a positive number to be read.

   (b) Write the program flow charted in (a) and use it to obtain approximations to $\int_0^1 \frac{4}{x^2 + 1} dx$. Use DELTA = .0001.

   (c) With $F(x) = \frac{4}{x^2 + 1}$, show that

   $|F(x_1) - F(x_2)| \leq 8|x_1 - x_2|$ whenever

   $x_1, x_2 \in [0,1]$.

   (d) Show that each of the four approximations obtained for the largest value of N in part (b) is in error by $\leq$ .0008.

   (e) What was the smallest value of N for which the left-end-point approximation obtained in part (b) was in error by $\leq$ .001? the right-end-point approximation? the midpoint approximation? the trapezoidal approximation?

2. Use the program written in problem 1 to approximate each of the following integrals:

(a) $\int_2^4 (x^3 - x^2 + 1)dx$

(b) $\int_1^2 \sqrt{x+1}\, dx$

(c) $\int_2^3 \frac{x^2+1}{x-1}\, dx$

(d) $\int_{-2}^2 (x^4 - 2)dx$

3. Judging from your results in problem 2, which of the four methods is best? Try to compare quantitatively the trapezoid rule and the midpoint rule, by examining their convergence rates in problem 2.

278

8. Formulas for Integrals

So far we have evaluated integrals only by computer methods. With really difficult functions we shall have to rely on computation and approximation of the limiting values, although in many cases representations in forms other than integrals (e.g., infinite series) are available which yield answers easier to evaluate. We shall see some examples of this before too long.

However, there is a small but important class of functions for which the integrals can be found exactly without approximation. This class of functions includes, for example, all polynomials and the functions sin x and cos x and thousands more. However, for such a simple example as $\int_1^2 \frac{\sin x}{x}\,dx$ we can still only approximate the value.

We will show several examples in increasing order of difficulty. All examples will have the same format. We will always integrate from 0 to b. We will always compute the sum S_n obtained by dividing the interval into n equal parts so that each subinterval has length b/n. Moreover,

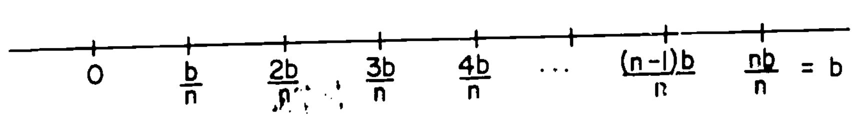


Figure 8-1

the right-hand-end point of the i-th subinterval will always be taken as ξ_i. The sum

$$S_n = \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1})$$

therefore reduces to the form

$$S_n = \sum_{i=1}^{n} f\left(\frac{ib}{n}\right) \frac{b}{n}$$

The factor b/n (which has the same value for all values of i) can be factored out by means of the distributive property giving the even simpler form

$$S_n = \frac{b}{n} \cdot \sum_{i=1}^{n} f\left(\frac{i \cdot b}{n}\right)$$

For several different functions we will find the exact value of this sum and then take the limit as $n \to \infty$.

Example I. $f(x) = 1$

Here $S_n = \frac{b}{n} \sum_{i=1}^{n} 1$. The sum of n one's is simply n so that

$$S_n = \frac{b}{n} \cdot n = b$$

We now have

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} b = b.$$

Of course this result is not very surprising as it merely gives us b as the area of a rectangle of length b and height 1. Still, it is reassuring to see that our method gives correct results in cases where we already know the answer.

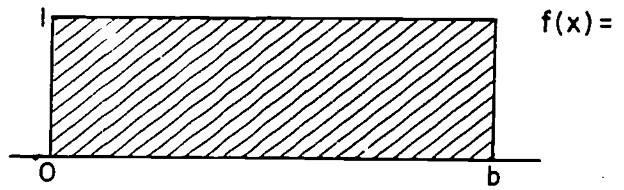


Figure 8-2

<u>Example</u> 2. $f(x) = x$

Here $S_n = \frac{b}{n} \sum_{i=1}^{n} \left(\frac{ib}{n}\right) = \frac{b^2}{n^2} \sum_{i=1}^{n} i$

The sum $\sum_{i=1}^{n} i = 1 + 2 + 3 + 4 + \ldots + (n-1) + n$
is frequently met with in high-school mathematics. It is most easily evaluated by the device:

$$\begin{array}{l} \text{sum} = 1 + 2 + 3 + 4 + \ldots + (n-1) + n \\ \underline{\text{sum} = n + (n-1) + (n-2) + (n-3) + \ldots + 2 + 1} \\ 2(\text{sum}) = (n+1) + (n+1) + (n+1) + (n+1) + \ldots + (n+1) \\ \qquad + (n+1) \end{array}$$

Hence $2(\text{sum}) = n\cdot(n+1)$ whence

$$\text{sum} = \sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

Now

$$S_n = \frac{b^2}{n^2} \cdot \sum_{i=1}^{n} i = \frac{b^2}{n^2} \frac{n(n+1)}{2} = \frac{b^2}{2}(1 + \frac{1}{n})$$

so that

$$\int_0^b x\,dx = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \frac{b^2}{2}(1 + \frac{1}{n}) = \frac{b^2}{2}.$$

Here again the result comes as no surprise as it tells us that the area of an isosceles right triangle with a leg of length b is $\frac{b^2}{2}$.

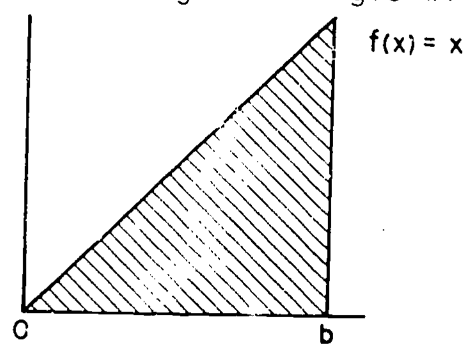


Figure 8-3

Now for our first non-trivial example, a region with a curved boundary.

<u>Example 3</u>. $f(x) = x^2$

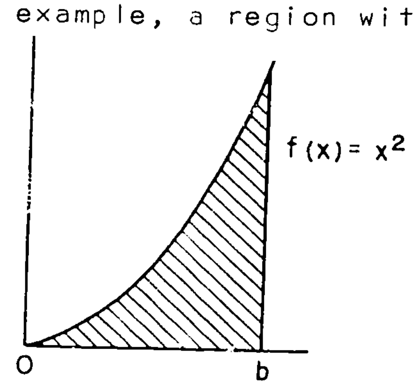


Figure 8-4

$$S_n = \frac{b}{n} \sum_{i=1}^{n} \left(\frac{ib}{n}\right)^2 = \frac{b^3}{n^3} \sum_{i=1}^{n} i^2$$

To find the sum $\sum_{i=1}^{n} i^2 = 1^2 + 2^2 + 3^2 + \ldots + n^2$ we resort to a little trick. We check that

$$i^2 = \frac{1}{3}\left(i + \frac{1}{2}\right)^3 - \frac{1}{3}\left(i - \frac{1}{2}\right)^3 - \frac{1}{12}$$

Substituting 1, 2, 3, 4, ... n in this formula we find that

$$\begin{aligned}
1^2 &= \frac{1}{3}\left(\frac{3}{2}\right)^3 - \frac{1}{3}\left(\frac{1}{2}\right)^3 - \frac{1}{12} \\
2^2 &= \frac{1}{3}\left(\frac{5}{2}\right)^3 - \frac{1}{3}\left(\frac{3}{2}\right)^3 - \frac{1}{12} \\
3^2 &= \frac{1}{3}\left(\frac{7}{2}\right)^3 - \frac{1}{3}\left(\frac{5}{2}\right)^3 - \frac{1}{12} \\
4^2 &= \frac{1}{3}\left(\frac{9}{2}\right)^3 - \frac{1}{3}\left(\frac{7}{2}\right)^3 - \frac{1}{12} \\
&\vdots \\
(n-1)^2 &= \frac{1}{3}\left(n - \frac{1}{2}\right)^3 - \frac{1}{3}\left(n - \frac{3}{2}\right)^3 - \frac{1}{12} \\
n^2 &= \frac{1}{3}\left(n + \frac{1}{2}\right)^3 - \frac{1}{3}\left(n - \frac{1}{2}\right)^3 - \frac{1}{12}
\end{aligned}$$

Now we observe that all the crossed out terms appear once positive and once negative and so cancel out. Adding up what is left, we have

$$\sum_{i=1}^{n} i^2 = \frac{1}{3}(n + \frac{1}{2})^3 - \frac{1}{3}(\frac{1}{2})^3 - n\,\frac{1}{12}\,.$$

Therefore

$$S_n = \frac{b^3}{n^3}\sum_{i=1}^{n} i^2 = b^3\left[\frac{1}{3}\left(\frac{n + \frac{1}{2}}{n}\right)^3 - \frac{1}{3n^3}\,(\frac{1}{2})^3 - \frac{1}{12n^2}\right]$$

so that

$$\int_0^b x^2\,dx = \lim_{n\to\infty} S_n = \lim_{n\to\infty} b^3\left[\frac{1}{3}(1 + \frac{1}{2n})^3 - \frac{1}{24n^3} - \frac{1}{12n^2}\right] = \frac{b^3}{3}.$$

We should pause in wonder and amazement at this point. For the first time in our experience we have derived an exact formula for the area of a region with a curved boundary. In a way it is a sort of miracle that such an involved limiting process should finally boil down to such a simple answer. This is only a harbinger of things to come.

<u>Example 4.</u> $f(x) = x^3$

Here $S_n = \frac{b}{n}\sum_{i=1}^{n}(\frac{ib}{n})^3 = \frac{b^4}{n^4}\sum_{i=1}^{n} i^3$

Use the formula

$$i^3 = \frac{1}{4}\left[\left(i + \frac{1}{2}\right)^4 - \left(i - \frac{1}{2}\right)^4 - i\right]$$

and show as in Example 3 that

$$\sum_{i=1}^{n} i^3 = \frac{1}{4}\left[(n + \frac{1}{2})^4 - (\frac{1}{2})^4 - \frac{n(n-1)}{2}\right]$$

so that $\lim_{n\to\infty} S_n = \frac{1}{4} b^4$. The details are left to an exercise.

Example 5. $f(x) = \cos x$

This is a really difficult example and involves a considerable amount of trigonometry which we develop here. First of all we have

$$S_n = \frac{b}{n} \sum_{i=1}^{n} \cos\left(\frac{ib}{n}\right)$$

Now for the trigonometry. Recalling the familiar formulas

$$\sin(A + B) = \sin A \cos B + \sin B \cos A$$

$$\sin(A - B) = \sin A \cos B - \sin B \cos A$$

and subtracting we have

$$\sin(A + B) - \sin(A - B) = 2 \sin B \cos A$$

Solving for co A, we have

$$\cos A = \frac{1}{2 \sin B}\left[\sin (A + B) - \sin (A - B)\right]$$

Next, substituting $A = \frac{ib}{n}$ and $B = \frac{b}{2n}$, we get

$$\cos\left(\frac{ib}{n}\right) = \frac{1}{2 \sin\left(\frac{b}{2n}\right)}\left[\sin\left(\left(i + \frac{1}{2}\right)\frac{b}{n}\right) - \sin\left(\left(i - \frac{1}{2}\right)\frac{b}{n}\right)\right]$$

Now, substituting $i = 1, 2, 3, \ldots n$, we obtain

$$\cos\left(1 \cdot \frac{b}{n}\right) = \frac{1}{2 \sin\left(\frac{b}{2n}\right)}\left[\sin\left(\frac{3}{2}\frac{b}{n}\right) - \sin\left(\frac{1}{2}\frac{b}{n}\right)\right]$$

$$\cos\left(2 \cdot \frac{b}{n}\right) = \frac{1}{2 \sin\left(\frac{b}{2n}\right)}\left[\sin\left(\frac{5}{2}\frac{b}{n}\right) - \sin\left(\frac{3}{2}\frac{b}{n}\right)\right]$$

$$\cos\left(3 \cdot \frac{b}{n}\right) = \frac{1}{2 \sin\left(\frac{b}{2n}\right)}\left[\sin\left(\frac{7}{2}\frac{b}{n}\right) - \sin\left(\frac{5}{2}\frac{b}{n}\right)\right]$$

$$\vdots$$

$$\cos\left((n-1) \cdot \frac{b}{n}\right) = \frac{1}{2 \sin\left(\frac{b}{2n}\right)}\left[\sin\left(\frac{2n-1}{2}\frac{b}{n}\right) - \sin\left(\frac{2n-3}{2}\frac{b}{n}\right)\right]$$

$$\cos\left(n \cdot \frac{b}{n}\right) = \frac{1}{2 \sin\left(\frac{b}{2n}\right)}\left[\sin\left(\frac{2n+1}{2}\frac{b}{n}\right) - \sin\left(\frac{2n-1}{2}\frac{b}{n}\right)\right]$$

Adding, using the usual "telescoping sum" technique, we have

$$\sum_{i=1}^{n} \cos\left(\frac{ib}{n}\right) = \frac{1}{2 \sin\left(\frac{b}{2n}\right)}\left[\sin\left(\frac{2n+1}{2}\frac{b}{n}\right) - \sin\left(\frac{b}{2n}\right)\right]$$

$$= \frac{\sin\left[\left(n + \frac{1}{2}\right)\frac{b}{n}\right]}{2 \sin\left(\frac{b}{n}\right)} - \frac{1}{2}$$

Therefore

$$S_n = \frac{b}{n} \sum_{i=1}^{n} \cos\left(\frac{ib}{n}\right)$$

$$= \frac{\sin\left[\left(1 + \frac{1}{2n}\right)b\right]}{\frac{\sin(b/2n)}{b/2n}} - \frac{b}{2n}$$

Thus

$$\lim_{n\to\infty} S_n = \frac{\lim\limits_{n\to\infty} \sin\left[\left(1 + \frac{1}{2n}\right) b\right]}{\lim\limits_{n\to\infty} \frac{\sin(b/2n)}{b/2n}} - \lim_{n\to\infty} \frac{b}{2n}$$

Each of the three limits in this formula is easily evaluated.

$$\lim_{n\to\infty} \frac{b}{2n} = 0$$

$$\lim_{n\to\infty} \frac{\sin(b/2n)}{b/2n} = 1.$$

$$\lim_{n\to\infty} \sin\left[\left(1 + \frac{1}{2n}\right) b\right] = \sin\left[\lim_{n\to\infty}\left(1 + \frac{1}{2n}\right) b\right] = \sin b$$

since the sine function is unicon. Therefore

$$\int_0^b \cos x \, dx = \lim_{n\to\infty} S_n = \frac{\sin b}{1} - 0 = \sin b.$$

<u>Example 6.</u> $f(x) = \sin x$

This can be done as in Example 5, or there is the following alter[illegible]:

Since $\sin x = \cos(x - \frac{\pi}{2})$, we see by use of Theorems 6 and [illegible] of Section 5 that

$$\begin{aligned}\int_0^b \sin x \, dx &= \int_0^b \cos(x - \tfrac{\pi}{2})\,dx \\ &= \int_{-\frac{\pi}{2}}^{b-\frac{\pi}{2}} \cos x \, dx \\ &= \int_0^{b-\frac{\pi}{2}} \cos x \, dx - \int_0^{-\frac{\pi}{2}} \cos x \, dx \\ &= \sin(b - \tfrac{\pi}{2}) - \sin(-\tfrac{\pi}{2}) \\ &= -\cos b + 1\end{aligned}$$

Thus we have

$$\int_0^b \sin x \, dx = 1 - \cos b.$$

The formulas derived in these examples show us that we will not need to run to the computer every time we see an integral. In a later chapter an extremely powerful method will be developed by means of which carload lots of formulas of this type can be obtained with very little effort.

PROBLEMS

1. Evaluate each of the following:

(a) $\int_0^5 x\,dx$

(b) $\int_0^5 2x\,dx$

(c) $\int_0^5 (2x + 3)dx$

(d) $\int_0^5 (2x - 3)dx$

(e) $\int_0^5 2(x + 3)\,dx$

(f) $\int_0^{\pi/3} \cos x\,dx$

(g) $\int_{\pi/3}^{2\pi/3} \cos (x - \frac{\pi}{3})\,dx$

2. Suppose f is an even function (that is, f(-x) = f(x) for all x), and a is a positive number.

(a) Explain geometrically why $\int_{-a}^{0} f(x)dx = \int_0^a f(x)dx$.

(b) Explain in terms of upper and lower sums why

$$\int_{-a}^{0} f(x)dx = \int_0^a f(x)dx.$$

(c) Show that each of the following formulas holds for every number c (positive, negative, or zero):

(i) $\int_0^c \cos x\,dx = \sin c$

(ii) $\int_0^c 1\,dx = c$

(iii) $\int_0^c x^2 dx = \frac{c^3}{3}$

3. Suppose f is an odd function (that is, $f(-x) = -f(x)$ for all x), and a is a positive number.

(a) Explain geometrically why $\int_{-a}^0 f(x)dx = -\int_0^a f(x)dx$.

(b) Show that each of the following formulas holds for every number c (positive, negative, or zero):

(i) $\int_0^c x\,dx = \frac{c^2}{2}$

(ii) $\int_0^c x^3 dx = \frac{c^4}{4}$

(iii) $\int_0^c \sin x\,dx = 1 - \cos c$

4. Use the formula $\int_a^b f(x)dx = \int_a^0 f(x)dx + \int_0^b f(x)dx$ to verify each of the following formulas:

(a) $\int_a^b \sin x\,dx = \cos a - \cos b$

(b) $\int_a^b \cos x\,dx = \sin b - \sin a$

(c) $\int_a^b 1\,dx = b - a$

(d) $\int_a^b x\,dx = \frac{1}{2}(b^2 - a^2)$

(e) $\int_a^b x^2 dx = \frac{1}{3}(b^3 - a^3)$

(f) $\int_a^b x^3 dx = \frac{1}{4}(b^4 - a^4)$

5. Evaluate each of the following:

(a) $\int_2^6 1\,dx$

(b) $\int_{-2}^2 x\,dx$

(c) $\int_{-3}^3 x^2 dx$

(d) $\int_0^4 x^3 dx$

(e) $\int_0^{\pi/2} \cos x\,dx$

(f) $\int_{-\pi/2}^0 \sin x\,dx$

(g) $\int_1^3 (x + x^2)\,dx$

(h) $\int_{-\pi/2}^{\pi/2} (\cos x - 2 \sin x + x)dx$

(i) $\int_2^5 (x + 3)^2 dx$

6. Give an example of functions f and g such that

$$\int_0^1 f(x)g(x)dx \neq \left(\int_0^1 f(x)dx\right)\left(\int_0^1 g(x)dx\right)$$

7. Work out the details of Example 4.

9. Tabulating the Sine and Cosine Functions

The formulas derived in the preceding section:

$$\int_0^b 1dx = b, \qquad \int_0^b xdx = \frac{b^2}{2}, \qquad \text{and} \int_0^b x^2dx = \frac{b^3}{3}$$

strongly suggest the general formula,

$$\int_0^b x^n dx = \frac{b^{n+1}}{n+1}$$

for every integer n greater than 0. This is quite correct. We could prove it now, but the proof would be very cumbersome because we have not yet developed the proper techniques for doing it efficiently. In this section we will temporarily "borrow" this formula pending its derivation in a later chapter.

We are going to see how the integration formulas and theorems we have developed so far enable us to calculate the values of sin x and cos x to any desired degree of accuracy. We will see that it is all done by repeating the same process over and over.

From the inequality

$$\cos x \le 1$$

which holds for all x we find by the use of Theorem 5 of Section 5 that

$$\int_0^t \cos x \, dx \le \int_0^t 1 dx$$

for all $t \ge 0$. Evaluating these integrals by the formulas in Section 8, we obtain

$$\sin t \le t.$$

for all $t \ge 0$. Again applying Theorem 5, we get

$$\int_0^x \sin t \, dt \le \int_0^x t \, dt$$

which can be evaluated to give

$$1 - \cos x \le \frac{x^2}{2}$$

This can be rearranged to give the inequality (valid for $x \ge 0$),

$$1 - \frac{x^2}{2} \le \cos x$$

29

Repeating the process over and over we successively find:

$$\int_0^t \left(1 - \frac{x^2}{2}\right) dx \leq \int_0^t \cos x \, dx$$

$$t - \frac{t^3}{2 \cdot 3} \leq \sin t$$

$$\int_0^x \left(t - \frac{t^3}{3!}\right) dt \leq \int_0^x \sin t \, dt$$

$$\frac{x^2}{2} - \frac{x^4}{4!} \leq 1 - \cos x$$

or

$$\cos x \leq 1 - \frac{x^2}{2} + \frac{x^4}{4!}$$

$$\int_0^t \cos x \, dx \leq \int_0^t \left(1 - \frac{x^2}{2} + \frac{x^4}{4!}\right) dx$$

$$\sin t \leq t - \frac{t^3}{3!} + \frac{t^5}{5!}$$

We see that we could continue this process as long as we like.

Collecting what we have learned so far, we see that

$$x - \frac{x^3}{3!} \leq \sin x \leq x - \frac{x^3}{3!} + \frac{x^5}{5!}$$

$$1 - \frac{x^2}{2!} \leq \cos x \leq 1 - \frac{x^2}{2!} + \frac{x^4}{4!}$$

Let us focus our attention on the sine function now. By continuing the process we would obtain a sequence s_1, s_2, s_3, ... where

$$s_n = \sum_{k=1}^{n} \frac{(-1)^{k-1} x^{2k-1}}{(2k-1)!}$$

And sin x lies between each pair of consecutive terms of this sequence as illustrated in Figure 9-1.

(a) n is odd (b) n is even

Figure 9-1

We can see then that

$$|\sin x - s_n| \leq |s_{n+1} - s_n|$$

Now s_{n+1} differs from s_n only in the adjunction of one term, namely

$$(-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

so that

$$|\sin x - s_n| \leq \frac{x^{2n+1}}{(2n+1)!}$$

If x is not too large, say if $0 \le x \le 1$, we see that

$$\sin x - s_n \le \frac{1}{(2n+1)!}$$

which is very small even for relatively small values of n. For example, if n = 5 then

$$\frac{1}{(2n+1)!} = \frac{1}{11!} = \frac{1}{39,916,800}$$

As you well know, it will only be necessary to tabulate values of sin x and cos x for x between 0 and $\pi/2$. In fact it is only necessary to compute the values of sin x and cos x for values of x between 0 and $\frac{\pi}{4}$, owing to the relations

$$\sin x = \cos(\pi/2 - x) \qquad \text{and} \qquad \cos x = \sin(\pi/2 - x)$$

[If $x > \pi/4$ then $\pi/2 - x < \pi/4$.] Therefore, $\pi/4$ is the largest value of x needed for making complete tables. Since $\pi/4 \approx .78$, we see that errors will quickly get very small indeed.

We give in Figure 9-2 a very efficient flow chart for realizing the above process as a computer algorithm. You should check it carefully to see that the successive values of SUM are the terms of the sequence s_1, s_2, s_3, ... described above.

In the exercises which follow you will be asked to modify this flow chart so as to also output approximations for cos x and to further modify it so as to make a table of x, sin x, and cos x for a set of equally spaced values of x.

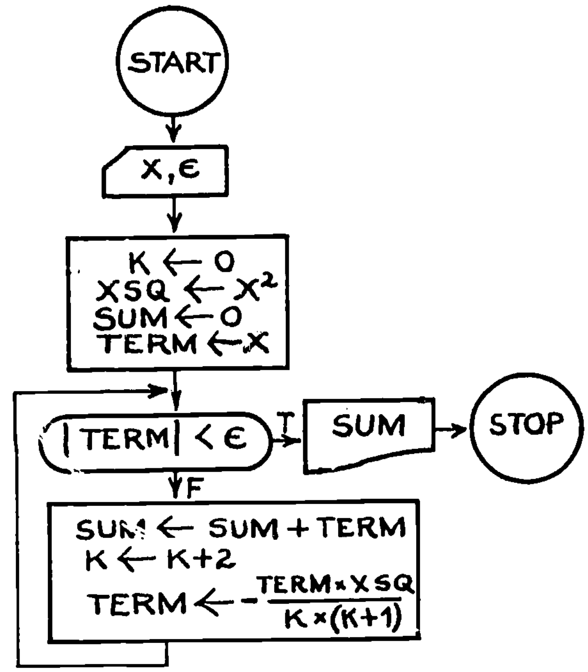


Figure 9-2

PROBLEMS

1. Modify the flow chart in Figure 9-2 so that approximations are obtained for both sin x and cos x at the same time. Try to make the flow chart as efficient as possible.

2. Write a program to prepare a table of sin x and cos x with x in degrees from 0° to 45°. Label the table along the right side from 90° to 45°, so that it will be easy to locate the sine or cosine of any angle between 45° and 90° by using the identities $\sin x = \cos(90^\circ - x)$ and $\cos x = \sin(90^\circ - x)$.

10. The Unicon Modulus

We have seen in connection with controlling the errors in the approximations of integrals that we must be able, given a number $\varepsilon > 0$, to find a number $\delta > 0$ so that

$$|f(x_2) - f(x_1)| \le \varepsilon \qquad \text{whenever} \qquad |x_2 - x_1| \le \delta.$$

The calculations of these values of δ were sometimes rather difficult.

If the tolerance of error in estimating an area were to be reduced, then the value of ε would also be reduced, and it would be necessary to recalculate an appropriate value of δ. It would be a shame to have to repeat these arduous calculations. Clearly, it would be much better to have a formula expressing the value of δ in terms of ε. Fortunately, in the worked out examples of Section 6 such formulas were derived. In Figure 10-1 we see for some of these examples the relations between δ and ε given by means of formulas and also graphically.

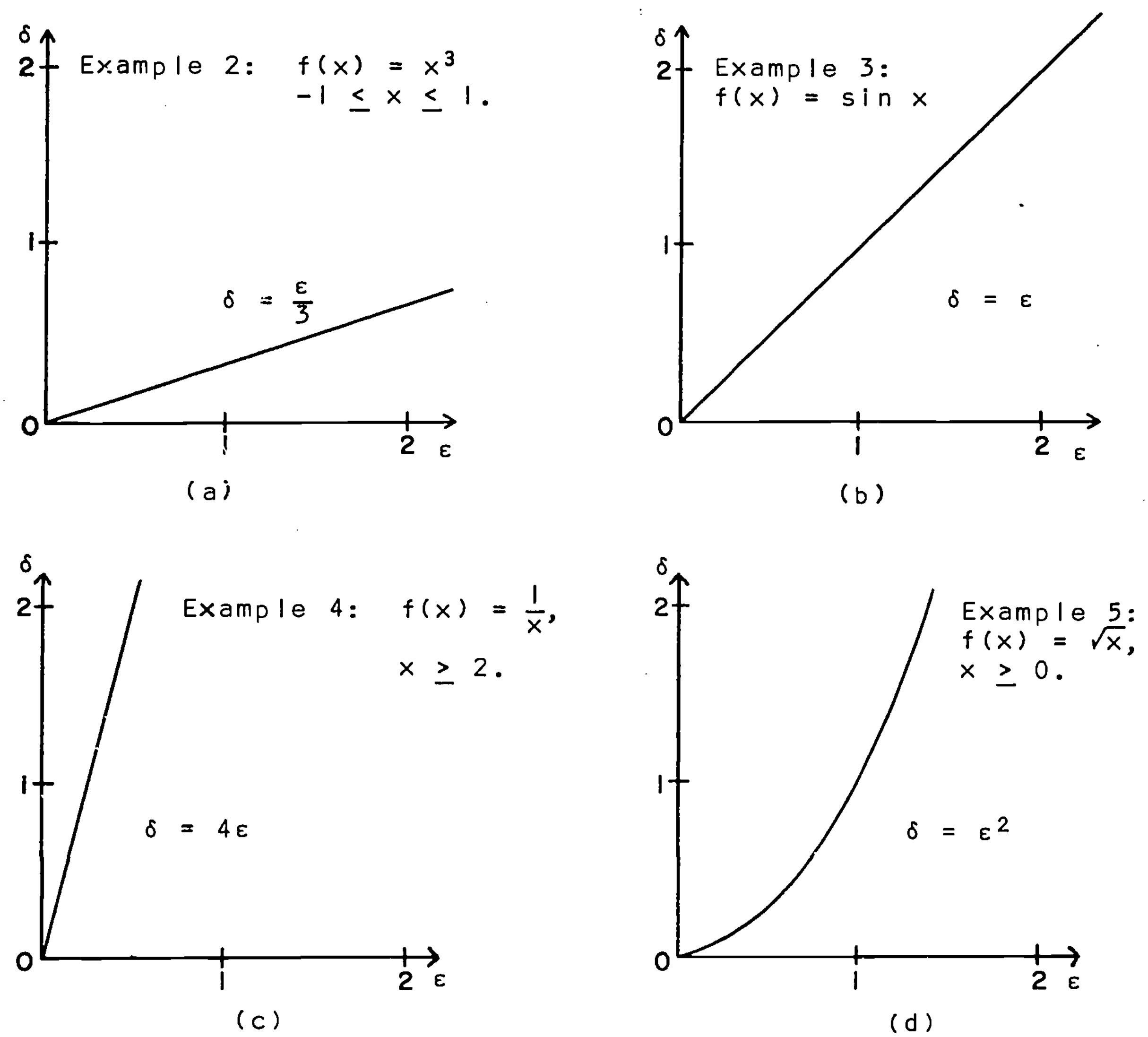


Figure 10-1

We can see in these figures that the value of ε determines the value of δ so that these graphs are graphs of functions. At the risk of a slight confusion we will use the letter "δ" to designate these functions. Thus we have:

in Figure 10-1(a), $\quad \delta(\varepsilon) = \frac{\varepsilon}{3}$;

in Figure 10-1(b), $\quad \delta(\varepsilon) = \varepsilon$;

in Figure 10-1(c), $\delta(\varepsilon) = 4\varepsilon$;

in Figure 10-1(d), $\delta(\varepsilon) = \varepsilon^2$.

This slight change in our attitude will bring a correspondingly slight change in our formulation of the definition of unicon functions.

Definition (Alternative): We say that f is unicon on an interval [a,b] provided that there exists a function δ on the positive numbers to the positive numbers so that if x_1, x_2 are in [a,b] and $|x_1 - x_2| \leq \delta(\varepsilon)$ then $|f(x_1) - f(x_2)| \leq \varepsilon$.

This is entirely equivalent to the previous definition. Any such function, δ, is called <u>a unicon modulus</u> for f over the interval [a,b]. Remember that $\delta(\varepsilon)$ tells us how close together to take x_1 and x_2 in order to guarantee that their functional values $f(x_1)$ and $f(x_2)$ differ by no more than ε. This function is not uniquely determined. The "best" $\delta(\varepsilon)$ would be the largest one as this would allow us the most leeway in the difference between x_1 and x_2. In general finding the best possible δ is very difficult and of very little practical value. [It so happens that in the four examples in Figure 10-1 the δ's are best possible, but that is only because the functions f are so simple.]

Lipschitzian Functions

We observe that in three cases out of four in Figure 10-1, the graph of the function δ is a straight line through the origin. That is, in each of these cases

$$\delta(\varepsilon) = \frac{\varepsilon}{K}.$$

In such cases the modulus is generally relatively easy to find and to work with. A function f having a modulus of this form over an interval [a,b] is said to be "Lipschitzian" or to "satisfy a Lipschitz condition" over the interval. An alternative and more convenient definition of Lipschitzian is given below.

> Definition: A function f is Lipschitzian over the interval [a,b] provided that there is a number K such that
>
> $$|f(x_2) - f(x_1)| \leq K|x_2 - x_1|$$
>
> for all x_1, x_2 in [a,b]. The number K is called the Lipschitz coefficient.

We have seen above that the sine function is Lipschitzian over the whole line with Lipschitz coefficient, 1, i.e.,

$$|\sin x_2 - \sin x_1| \leq 1 \cdot |x_2 - x_1| \text{ for all } x_1, x_2.$$

This is also true for the absolute value function, i.e.,

$$||x_2| - |x_1|| \leq 1 \cdot |x_2 - x_1| \text{ for all } x_1, x_2,$$

which statement is essentially equivalent to the triangle inequality for the absolute value function.

It will turn out that most functions we consider are Lipschitzian over finite intervals. The geometrical meaning of a function, f, being Lipschitzian is that there is an upper bound, K, for the slopes of all chords drawn on the graph of f. That is,

$$\left|\frac{f(x_2) - f(x_1)}{x_2 - x_1}\right| \leq K$$

which is clearly equivalent to

$$|f(x_2) - f(x_1)| \leq K|x_2 - x_1|.$$

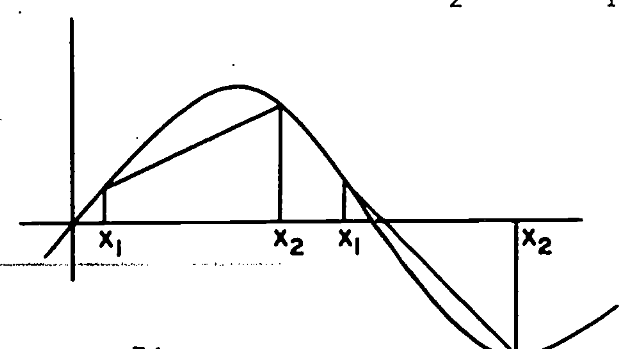


Figure 10-2

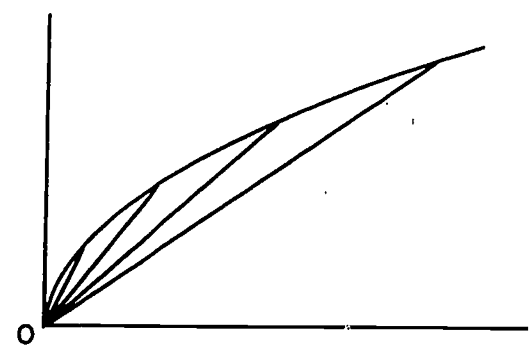


Figure 10-3

The reason that the square root function, $f(x) = \sqrt{x}$, $0 \leq x \leq 1$ is not Lipschitzian is that chords drawn with one end-point at 0 can have arbitrarily large slopes. Over an interval such as $[\frac{1}{4},2]$ the square root function is Lipschitzian.

We have tabulated below the unicon moduli for several functions, some obtained from examples in the text, others obtained from the following problem set.

Unicon Moduli for Various Functions over an interval $[a,b]$. Where applicable, K denotes max $[|a|, |b|]$.

Function	Conditions on Interval	Modulus
$f(x) = \sin x$	none	$\delta(\varepsilon) = \varepsilon$
$f(x) = \frac{1}{x}$	$0 < a < b$	$\delta(\varepsilon) = a^2\varepsilon$
$f(x) = \sqrt{x}$	$0 \leq a \leq b$	$\delta(\varepsilon) = \varepsilon^2$
$f(x) = \sqrt{x}$	$0 < a < b$	$\delta(\varepsilon) = 2\sqrt{a}\ \varepsilon$
$f(x) = \|x\|$	none	$\delta(\varepsilon) = \varepsilon$
$f(x) = x$	none	$\delta(\varepsilon) = \varepsilon$
$f(x) = x^2$	none	$\delta(\varepsilon) = \frac{\varepsilon}{2K}$
$f(x) = x^3$	none	$\delta(\varepsilon) = \frac{\varepsilon}{3K^2}$
$f(x) = x^n$	none	$\delta(\varepsilon) = \frac{\varepsilon}{nK^{n-1}}$

We are now in a position to show that sums, differences, products, quotients, and compositions of unicon functions are again unicon. Since the proofs of these facts are rather difficult they have been relegated to Appendix A at the end of this chapter. Every student should read this material although complete mastery of the techniques is not required.

The fact that sums, differences, products, quotients, and compositions of unicon functions are again unicon is one reason that unicon functions are important. Recall our earlier remark that piecewise monotone functions are not so well behaved.

PROBLEMS

1. Verify the entries in the Unicon Moduli Table that were not established in Section 3-6.

2. In each of the following examples use the Unicon Moduli Table to find a δ small enough to guarantee that an end-point approximation will yield an error no larger than specified.

(a) $\int_0^{\pi/2} \sin x \, dx$, error $\leq .001$

(b) $\int_0^{\pi} \sin x \, dx$, error $\leq .001$

(c) $\int_2^4 \frac{1}{x} \, dx$, error $\leq .0001$

(d) $\int_2^3 \sqrt{x} \, dx$, error $\leq .001$

(e) $\int_0^3 \sqrt{x} \, dx$, error $\leq .001$

(f) $\int_{-2}^2 |x| \, dx$, error $\leq .00001$

(g) $\int_{-2}^2 5|x| \, dx$, error $\leq .00005$

APPENDIX A

Moduli for Combinations of Unicon Functions

It was announced in Section 10 that sums, products, quotients and compositions of unicon functions are also unicon. We will establish these facts in this Section. This will be done by constructing moduli for these various combinations out of the moduli of the component parts. These constructions involve some techniques unfamiliar to you. Consequently many students find these methods rather discouraging. Take heart in the fact you can get a good working knowledge of calculus without entirely mastering these derivations; millions of students have done it. On the other hand, the techniques introduced here form the basis of advanced work in mathematical analysis so that there is a handsome pay-off for learning them as well as you can. In any event you should have no qualms about using the results derived here regardless of whether you have mastered the techniques.

We will suppose throughout this Section that f and g are unicon in an interval [a,b] with moduli ϕ and γ respectively.

[This supposition will be somewhat modified when we come to the last (and easiest) derivation on composition.]

Modulus of the Sum

Let $s = f + g$, that is, $s(x) = f(x) + g(x)$ for x in $[a,b]$. We must find a function, δ, so that for x_1 and x_2 in $[a,b]$ with $|x_1 - x_2| \leq \delta(\varepsilon)$, we have $|s(x_1) - s(x_2)| \leq \varepsilon$. Accordingly we express $s(x_1) - s(x_2)$ in terms of f and g and rearrange some terms

$$\begin{aligned} s(x_1) - s(x_2) &= f(x_1) + g(x_1) - f(x_2) - g(x_2) \\ &= [f(x_1) - f(x_2)] + [g(x_1) - g(x_2)]. \end{aligned}$$

Hence, by the triangle inequality

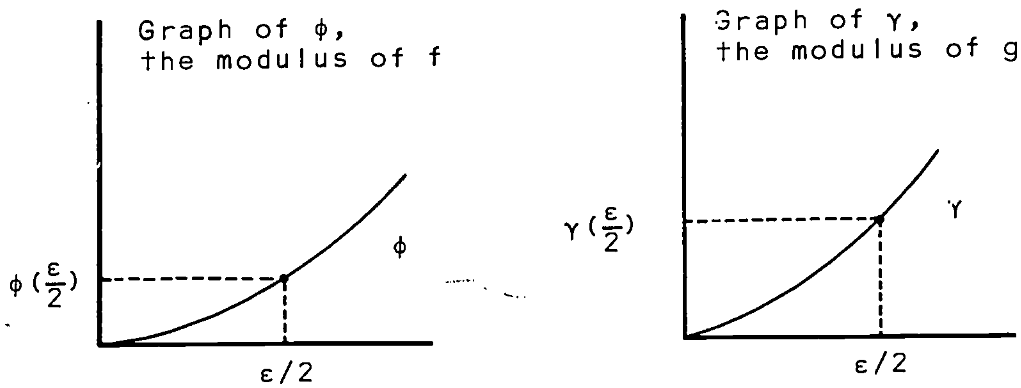
$$(1) \quad |s(x_1) - s(x_2)| \leq |f(x_1) - f(x_2)| + |g(x_1) - g(x_2)|$$

Thus, in order to have $|s(x_1) - s(x_2)| \leq \varepsilon$, it will suffice to force each of the quantities

$$|f(x_1) - f(x_2)| \qquad \text{and} \qquad |g(x_1) - g(x_2)|$$

to be less than or equal to $\frac{\varepsilon}{2}$.

Now look at the graphs of the moduli ϕ and γ of f and g.



If x_1 and x_2 are chosen so that $|x_1 - x_2| \leq \phi(\frac{\varepsilon}{2})$, then $|f(x_1) - f(x_2)| \leq \frac{\varepsilon}{2}$

If x_1 and x_2 are chosen so that $|x_1 - x_2| \leq \gamma(\frac{\varepsilon}{2})$, then $|g(x_1) - g(x_2)| \leq \frac{\varepsilon}{2}$

Figure A-1

From this figure and the accompanying remarks we can see that if x_1 and x_2 are chosen to be less than both $\phi(\frac{\varepsilon}{2})$ and $\gamma(\frac{\varepsilon}{2})$, then both the consequences

$$|f(x_1) - f(x_2)| \leq \frac{\varepsilon}{2} \quad \text{and} \quad |g(x_1) - g(x_2)| \leq \frac{\varepsilon}{2}$$

will hold true which will in turn, as seen in Formula (1), yield

$$|s(x_1) - s(x_2)| \leq \varepsilon.$$

Therefore we choose our modulus for s by

$$\delta(\varepsilon) = \text{minimum of } \phi\left(\frac{\varepsilon}{2}\right) \text{ and } \gamma\left(\frac{\varepsilon}{2}\right) = \min\left[\phi\left(\frac{\varepsilon}{2}\right), \gamma\left(\frac{\varepsilon}{2}\right)\right]$$

Going over this once more; with $\delta(\varepsilon)$ so chosen, then we see that

if $|x_1 - x_2| \leq \delta(\varepsilon)$

then

$|x_1 - x_2| \leq \phi\left(\frac{\varepsilon}{2}\right)$

so that

$|f(x_1) - f(x_2)| \leq \frac{\varepsilon}{2}$

and

$|x_1 - x_2| \leq \gamma\left(\frac{\varepsilon}{2}\right)$

so that

$|g(x_1) - g(x_2)| \leq \frac{\varepsilon}{2}$

whence

$$|s(x_1) - s(x_2)| \leq |f(x_1) - f(x_2)| + |g(x_1) - g(x_2)| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

In the particular case illustrated, it is obvious from the graph that $\phi\left(\frac{\varepsilon}{2}\right) < \gamma\left(\frac{\varepsilon}{2}\right)$ so that in this case $\min\left[\phi\left(\frac{\varepsilon}{2}\right), \gamma\left(\frac{\varepsilon}{2}\right)\right]$ is $\phi\left(\frac{\varepsilon}{2}\right)$, but in general either one could be the minimum.

Example 1: Let $f(x) = x^3$, $g(x) = \sin x$, $[a,b] = [-1,1]$. Moduli for these functions are $\phi(\varepsilon) = \frac{\varepsilon}{3}$, $\gamma(\varepsilon) = \varepsilon$ as seen in Figures 1(e) and 1(b) of Section 10. Thus

$$\phi(\tfrac{\varepsilon}{2}) = \frac{\varepsilon}{6} \quad \text{and} \quad \gamma(\tfrac{\varepsilon}{2}) = \frac{\varepsilon}{2}$$

so that a modulus for $(f + g)(x) = x^3 + \sin x$ over the interval $[-1,1]$ is given by

$$\phi(\tfrac{\varepsilon}{2}) = \min\,[\phi(\tfrac{\varepsilon}{2}),\ \gamma(\tfrac{\varepsilon}{2})] = \min\,(\tfrac{\varepsilon}{6},\ \tfrac{\varepsilon}{2}) = \frac{\varepsilon}{6}\ .$$

Modulus of Constant Multiple of a Function.

Let $m = c \cdot f$, that is, $m(x) = c \cdot f(x)$ for x in $[a,b]$, where c is a non-zero constant. We wish to make $|m(x_1) - m(x_2)| \leq \varepsilon$. The calculations are very simple.

$$m(x_1) - m(x_2) = cf(x_1) - cf(x_2) = c[f(x_1) - f(x_2)]$$

so that

$$|m(x_1) - m(x_2)| = |c| \cdot |f(x_1) - f(x_2)|.$$

Therefore it will suffice to make $|f(x_1) - f(x_2)|$ less than or equal to $\varepsilon/|c|$. Accordingly we choose $\delta(\varepsilon) = \phi(\varepsilon/|c|)$

so that if $|x_1 - x_2| \le \delta(\varepsilon)$ [which means $|x_1 - x_2| \le \phi(\varepsilon/|c|)$], we have $|f(x_1) - f(x_2)| \le \varepsilon/|c|$ so that

$$|m(x_1) - m(x_2)| = |c|\cdot|f(x_1) - f(x_2)| \le |c|\cdot\frac{\varepsilon}{|c|} = \varepsilon$$

Bounds for Unicon Functions.

We will show here that a function f which is unicon on an interval [a,b] is bounded. That is to say, there is a number F so that $|f(x)| \le F$ for all x in [a,b]. We will show how to find such a number. In doing this we use only one value of ε, namely $\varepsilon = 1$.

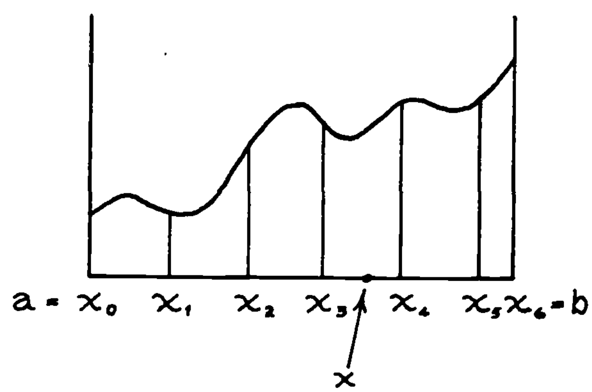


Figure A-2

Consider a partition of the interval [a,b] with each of the subintervals having length $\phi(1)$ except for the last one which is taken to have whatever length is left over. In the illustration in Figure 2 there are six subintervals, but in general the number of subintervals N will be the smallest integer greater than or equal to $\frac{b - a}{\phi(1)}$.

Now consider a number x in [a,b] as illustrated in Figure A-2. Adding and subtracting some terms, we obtain

$$f(x)-f(a)=f(x)-f(x_3)+f(x_3)-f(x_2)+f(x_2)-f(x_1)+f(x_1)-f(a)$$

Now using the triangle inequality, we get

$$|f(x)-f(a)|\leq|f(x)-f(x_3)|+|f(x_3)-f(x_2)|+|f(x_2)-f(x_1)|+|f(x_1)-f(a)|.$$

According to the way the partition was chosen, each of the terms on the right is less than or equal to one so that

$$|f(x) - f(a)| \leq 1 + 1 + 1 + 1 = 4$$

In general the absolute difference $|f(x) - f(a)|$ cannot exceed the number of subintervals passed through in travelling from a to x. Thus, for any choice of x in [a,b] we have

$$|f(x) - f(a)| \leq N$$

so that

$$|f(x)| \leq |f(a)| + N \qquad \text{for all } x \text{ in } [a,b],$$

which gives the required bound.

Obtaining a modulus for the product of two functions is the hardest of our derivations. There is a "trick" used in this process which is so frequently used in connection

with "differences of products" that it is worth a little special attention.

Consider one rectangle with dimensions L and W and a second rectangle with slightly smaller dimensions L' and W'. In Figure A-3 we see these rectangles drawn separately and then superimposed with the differences of their areas shaded.

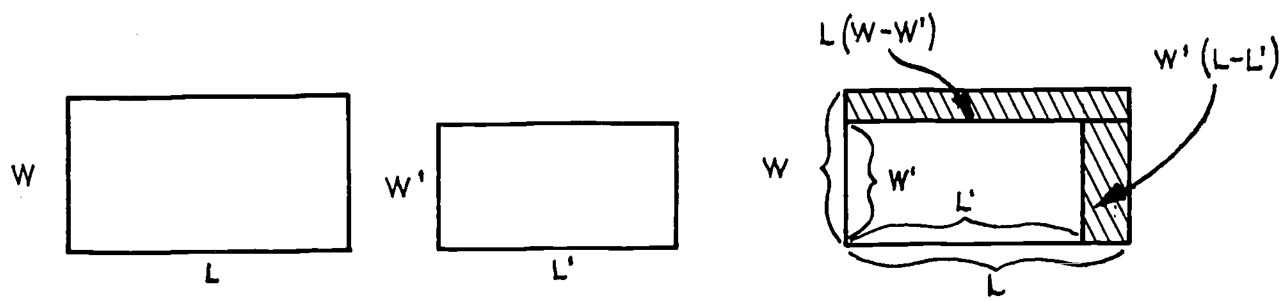


Figure A-3

The difference of these areas can of course be represented as $LW - L'W'$, but we also see that this difference is represented by the two shaded rectangles with areas $L(W - W')$ and $W'(L - L')$ and so we have

$$LW - L'W' = L(W - W') + W'(L - L').$$

The expression on the right has the advantage of indicating how the difference in the areas depends on the difference in the dimensions, $L - L'$ and $W - W'$. In applications we will not refer to areas but will obtain this result by the following "adding-and-subtracting trick."

$$LW - L'W' = LW - LW' + LW' - L'W'$$
$$= L(W - W') + W'(L - L').$$

Done in this way it makes no difference whether $L - L'$ and $W - W'$ are positive or negative. The difference in the products LW and $L'W'$ has been expressed in terms of the differences $L - L'$ and $W - W'$. You will see this trick used in the following derivation.

Modulus of the Product.

Let F and G represent bounds for f and g on [a,b], that is, $|f(x)| \leq F$ and $|g(x)| \leq G$ for x in [a,b]. Also let $p = f \cdot g$, that is $p(x) = f(x) \cdot g(x)$ for x in [a,b]. Our objective is to make $|p(x_1) - p(x_2)| \leq \varepsilon$. Accordingly we express the difference $p(x_1) - p(x_2)$ in terms of f and g and use our "adding-and-subtracting trick" to obtain

$$p(x_1) - p(x_2) = f(x_1)g(x_1) - f(x_2)g(x_2)$$
$$= f(x_1)g(x_1) - f(x_1)g(x_2) + f(x_1)g(x_2) - f(x_2)g(x_2)$$
$$= f(x_1)[g(x_1) - g(x_2)] + g(x_2)[f(x_1) - f(x_2)]$$

Now using the triangle inequality and the bounds on f and g,

$$|p(x_1) - p(x_2)| \leq |f(x_1)[g(x_1) - g(x_2)]| + |g(x_2)[f(x_1) - f(x_2)]|$$

$$= |f(x_1)|\cdot|g(x_1) - g(x_2)| + |g(x_2)|\cdot|f(x_1) - f(x_2)|$$

$$\leq \underbrace{F\cdot|g(x_1) - g(x_2)|}_{\text{to be made } \leq\, \varepsilon/2} + \underbrace{G\cdot|f(x_1) - f(x_2)|}_{\text{to be made } \leq\, \varepsilon/2}.$$

Here we see that we will have $|p(x_1) - p(x_2)| \leq \varepsilon$ as desired if each of the two terms in this last expression is $\leq \varepsilon/2$.

This in turn will hold true if

$$|g(x_1) - g(x_2)| \leq \frac{\varepsilon}{2F} \qquad \text{and} \qquad |f(x_1) - f(x_2)| \leq \frac{\varepsilon}{2G}$$

The first of these inequalities will hold if $|x_1 - x_2| \leq \gamma(\frac{\varepsilon}{2F})$ and the second will hold true if $|x_1 - x_2| \leq \phi(\frac{\varepsilon}{2G})$ where γ, you will recall, denotes the unicon modulus of g and ϕ that of f. Accordingly we choose

$$\delta(\varepsilon) = \min\left[\phi(\frac{\varepsilon}{2G}),\ \gamma(\frac{\varepsilon}{2F})\right]$$

With $\delta(\varepsilon)$ so chosen we see that if

$$|x_1 - x_2| < \delta(\varepsilon), \qquad \text{then} \qquad |f(x_1) - f(x_2)| \leq \frac{\varepsilon}{2G}$$

$$\text{and} \qquad |g(x_1) - g(x_2)| \le \frac{\varepsilon}{2F}$$

so that

$$|p(x_1) - p(x_2)| \le F \cdot |g(x_1) - g(x_2)| + G \cdot |f(x_1) - f(x_2)|$$

$$\le F \cdot \frac{\varepsilon}{2F} + G \cdot \frac{\varepsilon}{2G} = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Example 2. We saw in Figures 1(d) and 1(b) of Section 10 that the functions $f(x) = \sqrt{x}$ and $g(x) = \sin x$ have moduli, respectively, $\phi(\varepsilon) = \varepsilon^2$ and $\gamma(\varepsilon) = \varepsilon$ over $[0,4]$. Bounds for these functions on this interval are $F = 2$, $G = 1$. Thus, a modulus for the product

$$(f \cdot g)(x) = \sqrt{x} \sin x$$

is given by

$$\delta(\varepsilon) = \min\left[\phi\left(\frac{\varepsilon}{2G}\right), \gamma\left(\frac{\varepsilon}{2F}\right)\right] = \min\left[\left(\frac{\varepsilon}{2 \cdot 1}\right)^2, \frac{\varepsilon}{4}\right]$$

$$= \min\left[\frac{\varepsilon^2}{4}, \frac{\varepsilon}{4}\right]$$

so that

$$\delta(\varepsilon) = \begin{cases} \frac{\varepsilon^2}{4} & \text{for } \varepsilon \le 1 \\ \frac{\varepsilon}{4} & \text{for } \varepsilon > 1 \end{cases}$$

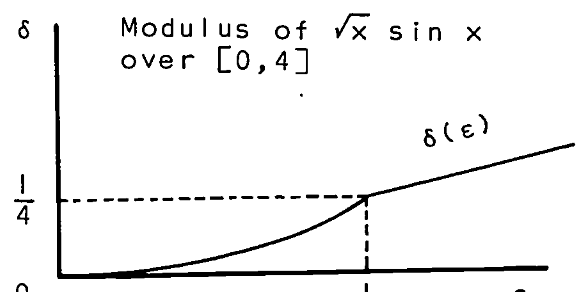


Figure A-4

This function is graphed in Figure A-4.

Modulus of the Reciprocal of a Function.

Here we assume that $g(x)$ is positive for x in $[a,b]$ and that there is a strictly positive number Γ so that

$$g(x) > \Gamma \qquad \text{for all } x \text{ in } [a,b].$$

Now we let $r(x) = \frac{1}{g(x)}$ and calculate:

$$r(x_1) - r(x_2) = \frac{1}{g(x_1)} - \frac{1}{g(x_2)} = \frac{g(x_2) - g(x_1)}{g(x_1)\, g(x_2)}$$

Taking absolute values and noting that $1/|g(x_1)| \leq 1/\Gamma$ and $1/|g(x_2)| \leq 1/\Gamma$, we see that

$$|r(x_1) - r(x_2)| = \frac{|g(x_2) - g(x_1)|}{|g(x_1)|\cdot|g(x_2)|} \leq \frac{|g(x_2) - g(x_1)|}{\Gamma^2}$$

The last expression is made less than or equal to ε by making $|g(x_2) - g(x_1)| \leq \Gamma^2\varepsilon$, which is accomplished by taking

$$|x_2 - x_1| \leq \gamma(\Gamma^2\varepsilon).$$

Accordingly we choose

$$\delta(\varepsilon) = \gamma(\Gamma^2\varepsilon).$$

(This derivation also works when $g(x)$ is negative throughout $[a,b]$ with $\Gamma > 0$ such that $|g(x)| \geq \Gamma$ for x in $[a,b]$.)

Modulus of the Quotient.

This formula is derived by using the product and reciprocal formulas. As in the derivation of the product formula, let F be a bound for f on $[a,b]$ so that $|f(x)| \leq F$ for x in $[a,b]$. Moreover, if $0 < \Gamma < |g(x)|$ for every x in $[a,b]$, then $|\frac{1}{g(x)}| < \frac{1}{\Gamma}$ so that $G = \frac{1}{\Gamma}$ is a bound for $\frac{1}{g(x)}$ in $[a,b]$. Now let

$$q(x) = \frac{f(x)}{g(x)} = f(x) \cdot \frac{1}{g(x)}$$

Denoting by $\mu(\varepsilon)$ the modulus for $\frac{1}{g(x)}$ derived above, we see that

$$\mu(\varepsilon) = \gamma(\Gamma^2\varepsilon)$$

The modulus of $f(x) \cdot \frac{1}{g(x)}$ is, by the modulus of the product,

$$\delta(\varepsilon) = \min\left[\phi\left(\frac{\varepsilon}{2G}\right), \mu\left(\frac{\varepsilon}{2F}\right)\right]$$

$$= \min\left[\phi\left(\frac{\Gamma\varepsilon}{2}\right), \gamma\left(\frac{\Gamma^2\varepsilon}{2F}\right)\right]$$

Example 3: Modulus of $q(x) = \frac{\sin x}{x}$ over an interval $[a,b]$ with $0 < a < 1$. Here $f(x) = \sin x$ and $g(x) = x$ so that we may take $\phi(\varepsilon) = \varepsilon, \gamma(\varepsilon) = \varepsilon$, $F = 1$, $\Gamma = a$. Therefore

$$\delta(\varepsilon) = \min\ [\phi(\frac{a\varepsilon}{2}),\ \gamma(\frac{a^2\varepsilon}{2F})]$$

$$= \min\ [\frac{a\varepsilon}{2},\ \frac{a^2\varepsilon}{2}]$$

$$= \frac{a^2\varepsilon}{2}$$

Modulus of the Composition.

This is the easiest to derive of our modulus formulas, but our assumptions are slightly changed. Here we assume that g is unicon with modulus γ on the interval $[a,b]$ which is mapped by g into the interval $[A,B]$ on which f is unicon with modulus ϕ.

We let $c(x) = f(g(x))$. And now

$$|c(x_1) - c(x_2)| = |f(g(x_1)) - f(g(x_2))|$$

which will be $\leq \varepsilon$ provided that

$$|g(x_1) - g(x_2)| \leq \phi(\varepsilon)$$

which will in turn hold true provided that

$$|x_1 - x_2| \leq \gamma(\phi(\varepsilon)).$$

Thus we choose

$$\delta(\varepsilon) = \gamma(\phi(\varepsilon)).$$

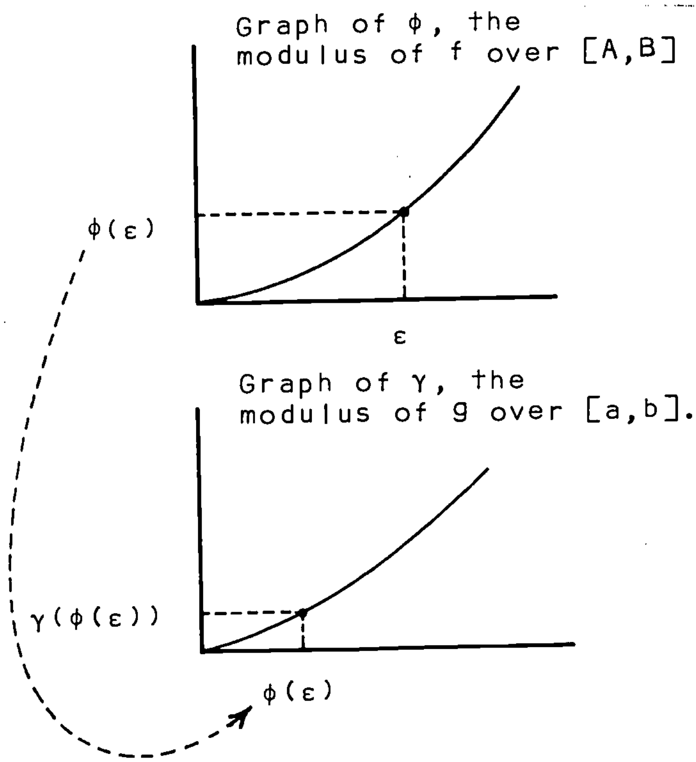
Hence for x_1 and x_2 in [a,b] with

$$|x_1 - x_2| \leq \delta(\varepsilon) = \gamma(\phi(\varepsilon))$$

we have $g(x_1)$ and $g(x_2)$ in [A,B] with

$$|g(x_1) - g(x_2)| \leq \phi(\varepsilon)$$

so that



Modulus of composition of f and g.

Figure A-5

$$|f(g(x_1)) - f(g(x_2))| \leq \varepsilon.$$

Hence we see that the modulus of the composition of two functions is the composition of their moduli in the <u>opposite order</u>. This is illustrated in Figure A-5.

Modulus of Function Plus Constant.

It is left to the reader to establish that $f(x) + k$ has the same modulus over $[a,b]$ as does $f(x)$.

Example: Modulus of $(\sqrt{x} + 1)^3$ over $[0,4]$.

Letting $f(y) = y^3$ and $g(x) = \sqrt{x} + 1$ we see that

$$f(g(x)) = (\sqrt{x} + 1)^3$$

The modulus of $g(x)$ is the same as that of $\sqrt{x}$ which has been seen to be $\gamma(\varepsilon) = \varepsilon^2$. The interval $[0,4]$ is mapped by g into the interval $[1,3]$ on which f has the modulus $\phi(\varepsilon) = \frac{\varepsilon}{27}$. Thus the modulus δ of $f(g(x))$ is given by

$$\delta(\varepsilon) = \gamma(\phi(\varepsilon)) = \left(\frac{\varepsilon}{27}\right)^2 = \frac{\varepsilon^2}{729}$$

It should be mentioned that this is an extremely bad result as far as actual computing is concerned. If we wished to know the area under the curve $y = (\sqrt{x} + 1)^3$ between $x = 0$ and $x = 4$ with error $\leq 10^{-4}$ then we would have to choose ε so that $(b - a)\cdot\varepsilon \leq 10^{-4}$ or $\varepsilon \leq \frac{10^{-4}}{4}$ whence $\delta(\varepsilon)$, the spacing of our subintervals, would have to be $\leq (10^{-4}/4)^2/729 \approx 8.6\cdot 10^{-13}$. The number of calculations of the functional values required for this problem would be about $2.3\cdot 10^{12}$.

Even with the high speed of computers such a number of calculations is far beyond our scope. Before long we shall develop methods for avoiding such difficulties.

In the table below we have collected the results of this section.

TABLE I

LEGEND

ϕ is a modulus of f on [a,b].
γ is a modulus of g on [a,b].
δ is a modulus of indicated combinations on [a,b]

Where applicable:

F is an upper bound for f on [a,b].
G is an upper bound for g on [a,b].
Γ is a positive lower bound for g(x) on [a,b].

Name	Combination	Modulus of Combination
Sum	$s(x) = f(x) + g(x)$	$\delta(\varepsilon) = \min[\phi(\frac{\varepsilon}{2}), \gamma(\frac{\varepsilon}{2})]$
Constant Multiple	$m(x) = K f(x)$	$\delta(\varepsilon) = \phi(\frac{\varepsilon}{\lvert K\rvert})$
Product	$p(x) = f(x) g(x)$	$\delta(\varepsilon) = \min[\phi(\frac{\varepsilon}{2G}), \gamma(\frac{\varepsilon}{2F})]$
Reciprocal	$r(x) = \frac{1}{g(x)}$	$\delta(\varepsilon) = \gamma(\Gamma^2\varepsilon)$
Quotient	$q(x) = f(x)/g(x)$	$\delta(\varepsilon) = \min[\phi(\frac{\Gamma\varepsilon}{2}), \gamma(\frac{\Gamma^2\varepsilon}{2F})]$
Additive Constant	$k(x) = f(x) + K$	$\delta(\varepsilon) = \phi(\varepsilon)$
Composition	$c(x) = f(g(x))$	$\delta(\varepsilon) = \gamma(\phi(\varepsilon))$ (Here g maps [a,b] into [A,B] on which interval f has modulus ϕ.)

<u>Example</u>: To compute a modulus for

$$\frac{x^3 + x + 1}{(x^2 + 2)^3} \quad \text{over} \quad 0 \le x \le 1.$$

Here moduli of x, x^2 and x^3 over [0,1] are ε, $\frac{\varepsilon}{2}$, $\frac{\varepsilon}{4}$, respectively. The function $g(x) = x^2 + 2$ maps [0,1] into [2,3] over which interval $f(y) = y^3$ has modulus $\frac{\varepsilon}{27}$. Now from our table:

modulus of $(x^3 + x)$	is $\min(\varepsilon/8, \varepsilon/2) = \varepsilon/8$
modulus of $(x^3 + x + 1)$	is $\varepsilon/8$
modulus of $x^2 + 2$	is $\varepsilon/2$
modulus of $(x^2 + 2)^3$	is $(\varepsilon/27)/2 = \varepsilon/54$

Now the maximum of $x^3 + x + 1$ on [0,1] is 3 while minimum of $(x^2 + 2)^3$ on [0,1] is 8 so that

$$\delta(\varepsilon) = \text{modulus of } \frac{x^3 + x + 1}{(x^2 + 2)^3} \quad \text{is } \min\left[\frac{8\varepsilon}{2}/3,\ \frac{8^2 \cdot \varepsilon}{2 \cdot 3}/54\right]$$

$$= \min\left[\frac{4}{3}\varepsilon,\ \frac{16}{81}\varepsilon\right]$$

$$= \frac{16}{81}\varepsilon.$$

PROBLEMS

1. Prove that if c is any constant, then the unicon modulus of $f(x) + c$ on an interval [A,B] is the same as the unicon modulus of $f(x)$ on [A,B].

2. Cite previous problems and results (but do not use epsilons and deltas) to prove that each of the following functions is unicon on the interval specified.

   (a) $f(x) = x + 4$ on $[3,8]$

   (b) $f(x) = 2x - 3$ on $[-1,1]$

   (c) $f(x) = x^2$ on $[-5,2]$

   (d) $f(x) = \frac{1}{x} + \sin x$ on $[2,4]$

   (e) $f(x) = \frac{1}{x^3} + 2 + 4x^2$ on $[1,2]$

   (f) $f(x) = x\sqrt{\sin x}$ on $[\frac{\pi}{2},\pi]$

   (g) $f(x) = \sqrt{\sin(x^2)}$ on $[-\pi,\pi]$

   (h) $f(x) = \frac{1}{x \sin(x)}$ on $[\frac{\pi}{3},\frac{\pi}{2}]$

3. Prove by induction that any polynomial function $f(x) = a_0 + a_1x + \ldots + a_nx^n$ is unicon on any closed interval [c,d].

APPENDIX B

Dispensing with Area

The purpose of this section is to free our definition of the integral from any preliminary assumption that regions under curves have areas. In fact, we will take a reverse point of view. When f is positive over [a,b] and $\int_a^b f(x)\,dx$ exists, we then *define* the area under the graph of f from a to b to be equal to this integral.

In spite of our claim to be making our development independent of existence of areas, you will see many references in this section to areas of regions built up of rectangles such as seen in Figure I. Such areas enjoy a different status than areas of regions with curved boundaries. For example, the shaded region in Figure B-I can be regarded as merely a geometrical method of representing the sum

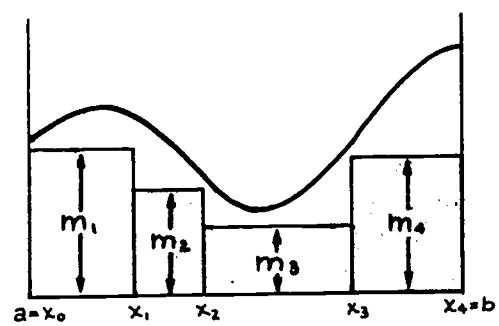


Figure B-I

$$\sum_{k=1}^{4} m_k(x_k - x_{k-1})$$

We assume throughout this section that all lower and upper sums are for the same function f over the same interval [a,b]. There are three things to be established:

(i) every lower sum is less than every upper sum;

(ii) if L_n and U_n, n = 1, 2, 3, ..., are sequences of lower and upper sums with $\lim_{n\to\infty} (U_n - L_n) = 0$, then these two sequences converge to the same limit, I;

(iii) the value of this limit I is independent of the choice of the sequences U_n and L_n enjoying the property $\lim_{n\to\infty} (U_n - L_n) = 0$.

We consider first the question: Can we see that every lower sum is less than every upper sum without making appeal to the area under the curve and reasoning

$$L \leq A \quad \text{and} \quad A \leq U \quad \text{so that} \quad L \leq U \text{ ?}$$

The answer is certainly yes in the case that the upper and lower sums are based on the same partition. For then we see that

$$m_k \leq M_k \qquad k = 1, 2, 3, \ldots, n$$

and since the differences $x_k - x_{k-1}$ are positive,

$$m_k(x_k - x_{k-1}) \leq M_k(x_k - x_{k-1}) \qquad k = 1, 2, 3, \ldots, n$$

and hence

$$L = \sum_{k=1}^{n} m_k(x_k - x_{k-1}) \leq \sum_{k=1}^{n} M_k(x_k - x_{k-1}) = U$$

To see that the relation $L \leq U$ holds true even when based on different partitions we consider Figure B-2.

In Figure B-2(a) and B-2(b) we see shaded areas representing lower and upper sums for a function f over an interval [a,b] but based on different partitions. In Figure B-2(c) the two Figures B-2(a) and B-2(b) have been superimposed thereby creating a new partition consisting of all the partition points of both earlier figures. The areas L and U are unchanged and we see that they can be considered as lower and upper sums based on the new partition. Since we have already seen that when based on the same partition, lower sums are less than or equal to upper sums we now find that $L \leq U$.

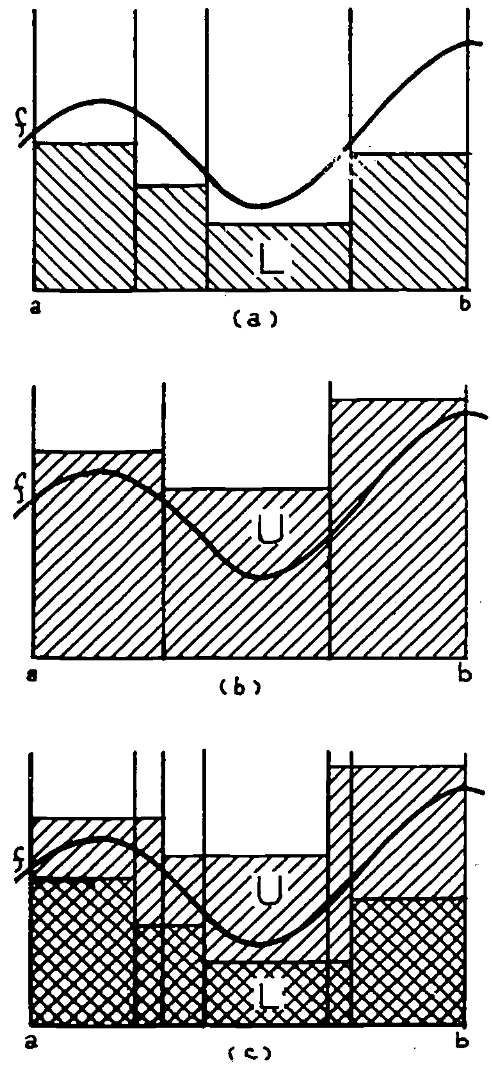


Figure B-2

Next we will show, given two lower sums, how to obtain a third one greater than or equal to each of the two given ones. At the risk of a little confusion we will use the letters U and L to denote both the rectangular regions and their areas, letting the context carry the information as to whether regions or areas are being considered.

In Figures B-3(a) and B-3(b) we see two lower sums L_1 and L_2 for f over the interval [a,b]. In Figure B-3(c) the entire shaded area L which is the union of L_1 and L_2 ($L_1 \cup L_2$). Clearly since $L_1 \subset L$ and $L_2 \subset L$ we have for the areas, $L_1 \leq L$ and $L_2 \leq L$. Figure B-3(d) is a repeat of Figure B-3(c) with some of the unnecessary lines omitted.

In Figure B-4 we do a similar thing for upper sums. Here, given two upper sums U_1 and U_2, we show how to obtain a third one less than or equal to the two given ones.

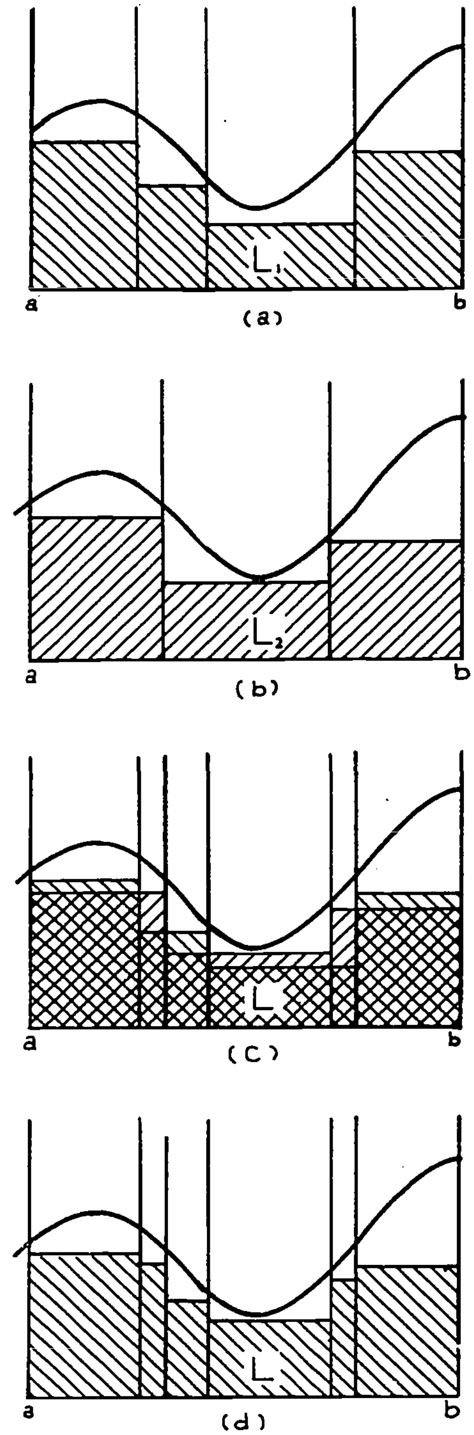


Figure B-3

In Figure B-4 we see Figures B-4(a)

and B-4(b) superimposed and observe that U, the intersection of U_1 and U_2 ($U_1 \cap U_2$), the doubly cross-hatched region in Figure B-4(a), again represents an upper sum. This intersection, $U = U_1 \cap U_2$ is shown again in Figure B-4(d) with extraneous detail omitted. Since $U \subset U_1$ and $U \subset U_2$ we see that for the areas we have $U \leq U_1$ and $U \leq U_2$.

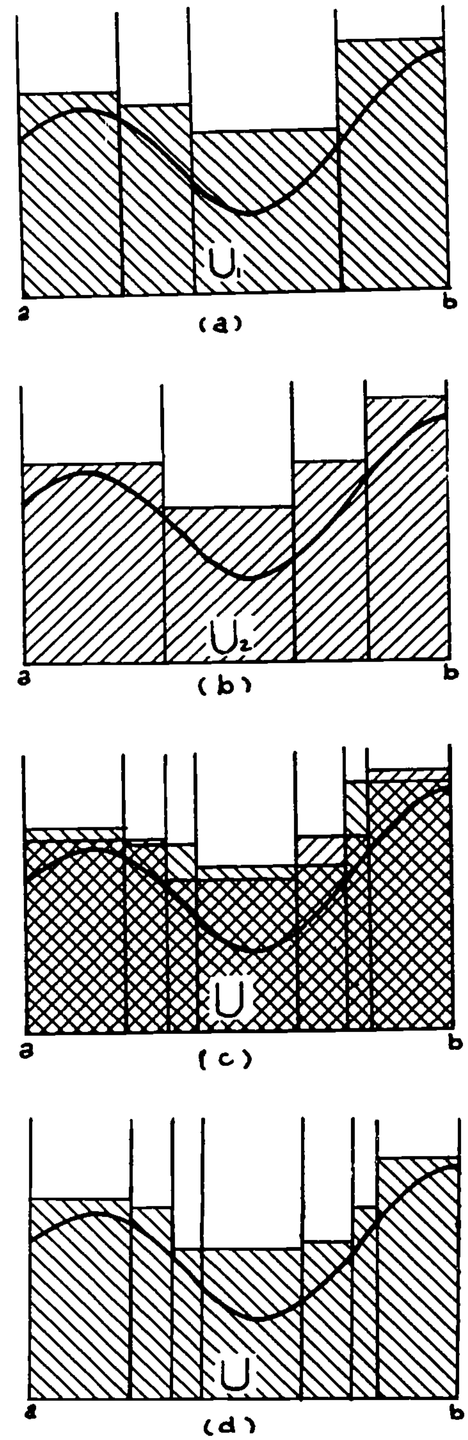


Figure B-4

Now we are in a position to dispose of the first objection to our definition of the integral. We will show that sequences L_n, $n = 1, 2, \ldots$ and U_n, $n = 1, 2, \ldots$ of lower and upper sums satisfying $\lim_{n\to\infty} (U_n - L_n) = 0$ must converge to a common limit.

Considering that we are given sequences L_n and U_n with $\lim_{n\to\infty} (U_n - L_n) = 0$ we construct new sequences of lower and upper sums L'_n and U'_n according to the following rule:

$$L_1' = L_1, \text{ and for } n > 1 \quad L_n' = L_n \cup L_{n-1}';$$

$$U_1' = U_1, \text{ and for } n > 1 \quad U_n' = U_n \cap U_{n-1}'.$$

Study this rule of formation carefully to see exactly what it says, particularly noting where primes do and do not occur and the use of intersection and union.

Previous observations now guarantee that:

(i) L_n' and U_n' are lower and upper sums for $n = 1, 2, \ldots;$

(ii) $L_n' \geq L_{n-1}'$ so that the sequence $L_1', L_2', L_3', \ldots$ is increasing;

(iii) $U_n' \leq U_{n-1}'$ so that the sequences $U_1', U_2', U_3', \ldots$ is decreasing;

(iv) $L_n \leq L_n'$ and $U_n' \leq U_n$ for $n = 1, 2, \ldots;$

(v) $L_n' \leq U_n'$ since any lower sum is less than any upper sum.

The last two remarks show that for each value of n we have the ordering indicated in Figure B-5.

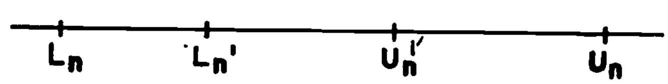


Figure B-5

Therefore $0 \leq U_n' - L_n' \leq U_n - L_n$ and since $\lim_{n\to\infty} (U_n - L_n) = 0$

the squeeze theorem shows that $\lim_{n\to\infty} (U'_n - L'_n) = 0$ also.

Therefore, by remarks (ii) and (iii) above we see that L'_n and U'_n are respectively increasing and decreasing sequences whose differences tend to zero so that these sequences converge to the same limit by the completeness axiom.

Again using the squeeze with Figure B-5 we see that

$$0 \leq L'_n - L_n \leq U_n - L_n \quad \text{and} \quad 0 \leq U_n - U'_n \leq U_n - L_n$$

whence

$$\lim_{n\to\infty} (L'_n - L_n) = 0 \quad \text{and} \quad \lim_{n\to\infty} (U_n - U'_n) = 0$$

which assures us that the limits

$$\lim_{n\to\infty} L_n \quad \text{and} \quad \lim_{n\to\infty} U_n$$

actually exist (because $\lim_{n\to\infty} L'_n$ and $\lim_{n\to\infty} U'_n$ do)

and

$$\lim_{n\to\infty} L_n = \lim_{n\to\infty} L'_n = \lim_{n\to\infty} U'_n = \lim_{n\to\infty} U_n.$$

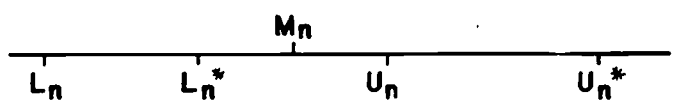
The second objection to the definition of the integral is next disposed of by similar methods.

Suppose that L_n and L^*_n are sequences of lower sums and that U_n and U^*_n are sequences of upper sums. And further suppose that

$$\lim_{n\to\infty} (U_n - L_n) = 0 \quad \text{and} \quad \lim_{n\to\infty} (U_n^* - L_n^*) = 0.$$

As already seen, L_n and U_n will converge to a common limit, $\mathcal{I}$, and L_n^* and U_n^* will converge to a common limit $\mathcal{I}^*$. We wish to show that $\mathcal{I}^* = \mathcal{I}$.

Since every lower sum is less than every upper sum, we see that for each integer n each of the lower sums L_n and L_n^* is less than each of the upper sums U_n and U_n^*. Thus if we select M_n halfway between the larger lower sum and the smaller upper sum,



we have both the inequalities

$$L_n \le M_n \le U_n \quad \text{and} \quad L_n^* \le M_n \le U_n^*.$$

Since

$$\lim_{n\to\infty} L_n = \mathcal{I} = \lim_{n\to\infty} U_n \quad \text{and} \quad \lim_{n\to\infty} L_n^* = \mathcal{I}^* = \lim_{n\to\infty} U_n^*$$

the squeeze theorem guarantees that

$$\lim_{n\to\infty} M_n = \mathcal{I} \quad \text{and} \quad \lim_{n\to\infty} M_n = \mathcal{I}^*.$$

The uniqueness of the limit of a sequence therefore assures us that $\mathcal{I}^* = \mathcal{I}$.

CHAPTER 4

APPLICATIONS OF INTEGRALS

1. Calculating Areas

We have seen that an area like that of the total cross-hatched region in Figure 1-1 is expressible as

$$\int_a^b f(x)dx.$$

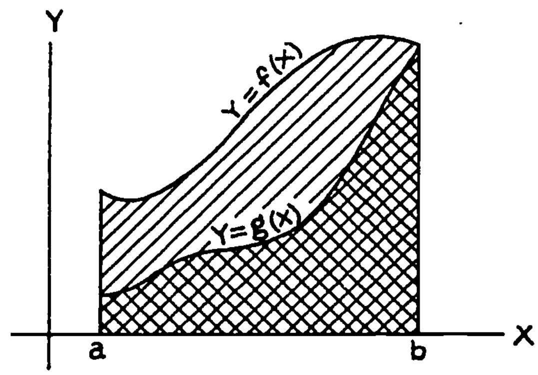


Figure 1-1

Similarly the area of the doubly cross-hatched region is

$$\int_a^b g(x)dx.$$

It follows that the area of the singly cross-hatched region lying between the two curves is the difference of these two integrals, and by a property of integrals (Corollary 1 of Section 3-5) this is

$$\int_a^b [f(x)-g(x)]dx.$$

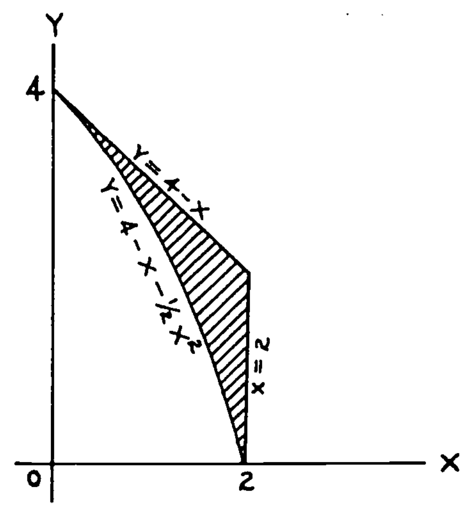


Figure 1-2

As noted on page 210 many areas can be found in this way.

Example 1. Find the area of the region bounded by the curve $y = 4 - x - \frac{1}{2}x^2$ and the lines $y = 4 - x$ and $x = 2$. In a problem like this it is important to draw a careful figure and see just what the region looks like. Here we obtain Figure 1-2. The horizontal limits of the region are 0 and 2, and so the area is

$$\int_0^2 [(4-x) - (4 - x - \tfrac{1}{2}x^2)]dx$$

$$= \int_0^2 \tfrac{1}{2}x^2 dx = \tfrac{1}{2} \cdot \tfrac{1}{3} 8 = \tfrac{4}{3} ,$$

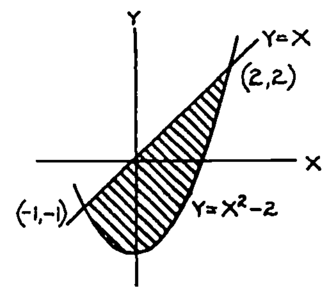


Figure 1-3

using the result of Section 3-8, Example 3.

Example 2. Find the area of the region bounded by the curve $y = x^2 - 2$ and the line $y = x$. By solving the two equations simultaneously we find that the curve and the line intersect at $(-1, -1)$ and $(2,2)$, giving Figure 1-3. Here the problem is more complicated since part of the region lies below the x-axis and we have seen in Section 3-4 that this gives a negative contribution to the integral. However, we can get back

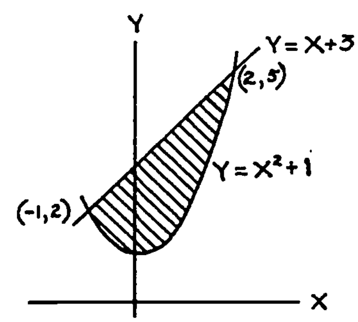


Figure 1-4

to the case of Example 1 by using the trick introduced at the end of Section 5. That is, we consider the curves $y = x^2 - 2 + k$ and $y = x + k$, where k is chosen so that the region bounded by the two curves lies entirely above the x-axis, as in Figure 1-4, in which k = 3. This amounts to lifting the region by an amount k without changing its shape or size. Hence the area is, as before,

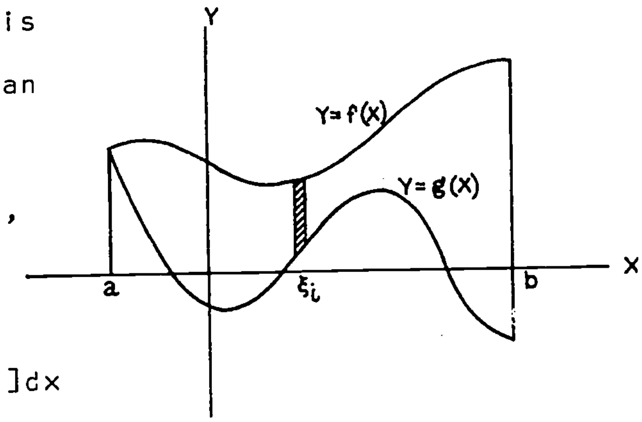


Figure 1-5

$$\int_{-1}^{2} [(x + k) - (x^2 - 2 + k)]dx$$

$$= \int_{-1}^{2} [x - (x^2 - 2)]dx$$

$$= \int_{-1}^{2} x\, dx - \int_{-1}^{2} x^2 dx + \int_{-1}^{2} 2\, dx$$

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

using the result of Problem 4, Section 3-8.

Example 2 shows that the general expression

(1) $A = \int_a^b [f(x) - g(x)]dx$

gives the area bounded on top by $y = f(x)$, on the bottom by

$y = g(x)$, on the left by $x = a$ and on the right by $x = b$, regardless of how this region is located with respect to the coordinate axes. One way of looking at this is to think of this region as cut into many vertical strips, one of which is shown cross-hatched in Figure 1-5. The height of such a strip, at position ξ_i is approximately $f(\xi_i) - g(\xi_i)$, and the sum of the areas of all the strips is similar to the sum (1) on page 265. In fact, the whole theory of Section 3-7 can easily be extended so as to give a proof of formula (1) above. We shall not do this, however, since in Chapter 12 we shall present an approach to the integral that covers not only this case but others much more complicated. For the present we can accept the evidence of Example 2, which can obviously be made completely general, that formula (1) is correct.

Like so many textbook problems Examples 1 and 2 were made numerically easy to facilitate concentration on the principles involved. A more realistic example is the following.

<ins>Example 3</ins>. Find the area common to the ellipse

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

and the circle

$$(x - 2)^2 + (y - 1)^2 = 4.$$

(Figure 1-6)

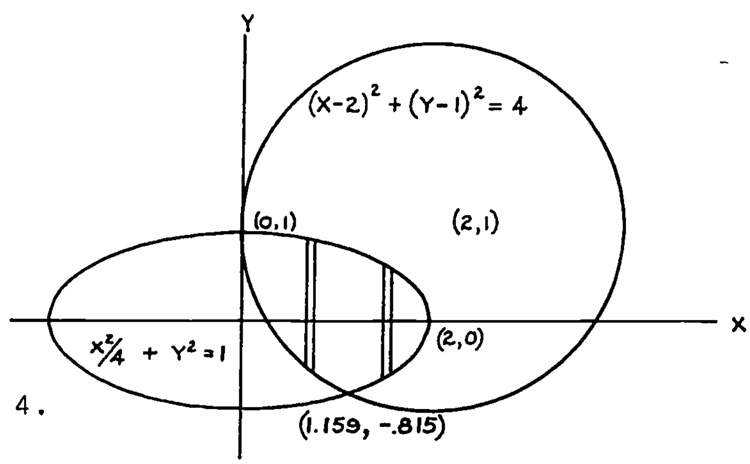


Figure 1-6

It is evident from the figure that one point of intersection of the two curves is (0,1). The other point is much more difficult to determine and is left as a problem (Problem 1). To three decimal places it is (1.159, -0.815).

The desired area is, as before,

$$(2) \qquad A = \int_0^2 [f(x) - g(x)]dx,$$

where y = f(x) is the equation of the curve at the top of a vertical strip and y = g(x) is the equation of the curve at the bottom. Evidently

$$f(x) = \sqrt{1 - \frac{x^2}{4}} \quad ,$$

since the top of the strip lies on the ellipse. But the bottom of the strip is sometimes on the ellipse and sometimes on the circle. Precisely:

$$g(x) = \begin{cases} 1 - \sqrt{4 - (x-2)^2} & \text{if } 0 \le x \le 1.159, \\ -\sqrt{1 - \frac{x^2}{4}} & \text{if } 0 \le 1.159 \le 2. \end{cases}$$

Thus $A = \int_0^2 h(x)dx$, where

$$(3) \qquad h(x) = \begin{cases} \sqrt{1 - \frac{x^2}{4}} - 1 + \sqrt{4 - (x-2)^2} & \text{if } 0 \leq x \leq 1.159, \\ 2\sqrt{1 - \frac{x^2}{4}} & \text{if } 1.159 \leq x \leq 2. \end{cases}$$

This integral can be evaluated by one of the programs of Section 3-2, to give A = 2.403 to three decimal places.

Another way of expressing this integral, avoiding the 2-part expression for g(x), is to use Theorem 1 of Section 5 to get

$$A = \int_0^{1.159} \left(\sqrt{1 - \frac{x^2}{4}} - 1 + \sqrt{4 + (x-2)^2}\right) dx + \int_{1.159}^{2} 2\sqrt{1 - \frac{x^2}{4}}\, dx.$$

There is no advantage in this if machine computation is to be used but later methods will enable us to evaluate this without the use of a machine.

Many area problems are given to us not in terms of functions but in terms of equations of curves. Indeed, Example 3 is of this type. In treating it we could equally well regard each curve as defining x as a function of y. Then we take horizontal strips, as in Figure 1-7, and the length of such a strip is

$G(\eta_i) - H(\eta_i)$,

where

$G(y) = 2\sqrt{1 - x^2}$,

$H(y) = 2 - \sqrt{4 - (y-1)^2}$.

The area is therefore

(4) $$\int_{-.815}^{1} (2\sqrt{1-x^2} - 2 + \sqrt{4 - (y-1)^2})dy.$$

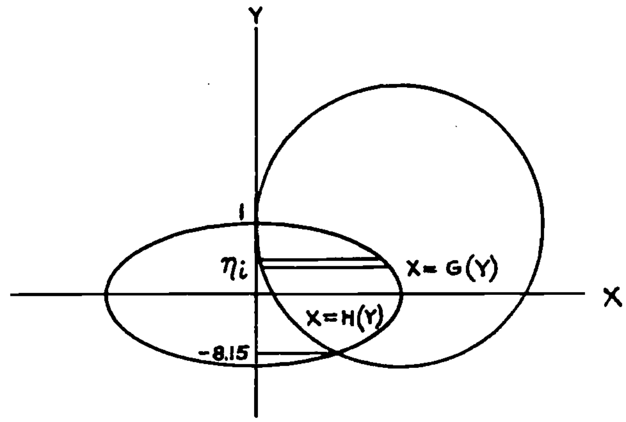


Figure 1-7

This of course must also evaluate to 2.403. Evidently this is a somewhat neater way to do the problem.

Here is an example that emphasizes the differences in the two methods.

<u>Example 4</u>. Find the area of the region bounded by the parabola $y^2 = 4x$ and the line $2x - y = 4$. (Figure 1-8) Solving the second equation for x and substituting in the first gives $y^2 + 2y-8 = 0$,

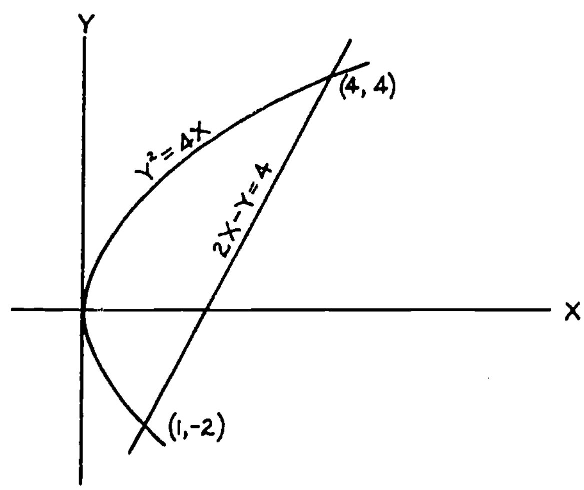


Figure 1-8

which has roots $y = 4$ and $y = -2$. The points of intersection are thus (4,4) and (1,-2). The use of vertical strips is awkward, giving

$$\int_0^1 4\sqrt{x}\,dx + \int_1^4 [2\sqrt{x} - (2x - 4)]dx.$$

With horizontal strips we get the much nicer form

$$\int_{-2}^4 [\tfrac{1}{2}(4 + y) - \tfrac{1}{4} y^2]dy$$

which by the methods of Section 3-8 evaluates to 31/3.

Whether, in a given problem, horizontal strips are better than vertical strips is something that can only be learned by experience. In most cases it isn't of great importance since, fundamentally, all problems can be done either way.

PROBLEMS

1. To find the points of intersection of $x^2 + 4y^2 - 4 = 0$ and $x^2 + y^2 - 4x - 2y + 1 = 0$:

   (a) Solve one equation for y^2 and substitute in the other;

   (b) Solve the resulting equation for y in terms of x and x^2;

   (c) Substitute in the first equation to get an equation in x only;

   [Ans. $9x^4 - 96x^3 + 320x^2 - 256x = 0$.]

   (d) This equation must have one root $x = 0$. From Figure 1-6 there must be another root between 0 and 2. Use your program from Section 2-2 to compute it to at least 5 decimal places; [Ans. 1.15947275]

   (e) Use the result of (b) to get the corresponding value of y. [Ans. -.814804107]

2. Find the areas of these regions. The integrals, if properly set up, can all be evaluated by the formulas of Section 8.

   (a) The region bounded by the curves $4y = x^2$ and $y = 5 - x^2$. Ans. $\frac{10}{3}$.

   (b) The region in the first quadrant bounded by the y-axis and the curves $y = \sin x$ and $y = \cos x$. Ans. $\sqrt{2} - 1$.

(c) The entire region bounded by the curve

$y = x^3 - 4x^2 + 4x + 1$ and the line $y = x + 1$.

(Be careful!) Ans. $\frac{37}{12}$.

(d) The region in the first quadrant bounded by the curve

$y^2 = 4x$, the x-axis, and the line $x = 4$. Ans. $\frac{32}{3}$.

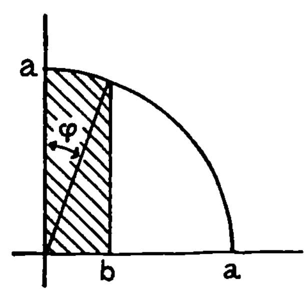
3. The interior of the ellipse $x^2 + 9y^2 = 9$ is divided into two regions by the line $x + y = 2$. Set up, but do not evaluate, one or more integrals expressing the area of the smaller region, using (a) vertical strips, (b) horizontal strips.

4. (a) Why is $\int_0^a \sqrt{a^2 - x^2}\, dx = \frac{1}{4}\pi a^2$?

(b) Show that the area of the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

is πab.



5. (a) Use the adjacent figure to derive the integration formula

$$\int_0^b \sqrt{a^2 - x^2}\, dx = \frac{1}{2}(b\sqrt{a^2 - b^2} + a^2\phi),$$

where $0 \leq b \leq a$ and $\sin \phi = b/a$, ϕ being measured in radians.

(b) Use this formula to evaluate

$$\int_0^1 \sqrt{4 - x^2}\, dx.$$

to 3 decimal places without using the computer.

[Ans. 1.913]

(c) Use the result of Problem 2(c) to derive

$$\int_{-c}^0 \sqrt{a^2 - x^2}\, dx = \int_0^c \sqrt{a^2 - x^2}\, dx, \; 0 \leq c \leq a.$$

(d) Evaluate

$$\int_{-1}^2 \sqrt{9 - x^2}\, dx$$

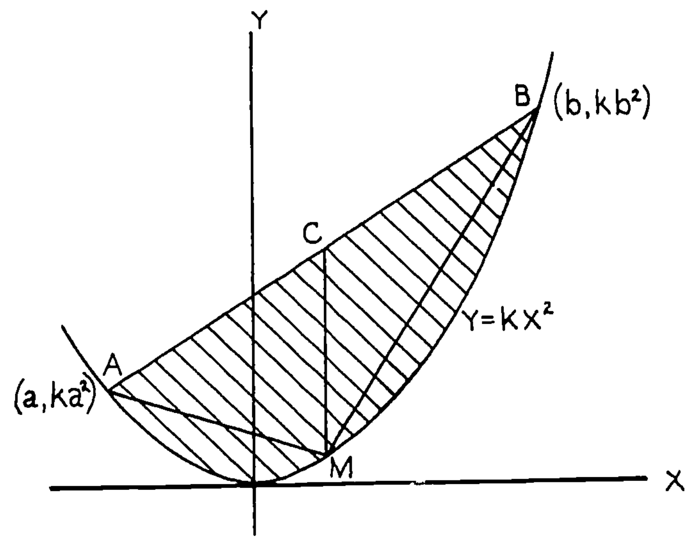
to 2 decimal places without using the computer.

[Ans. 3.46]

(e) Find the area in Problem 3 to 2 decimal places without using the computer. [Ans. 1.19]

6. (a) Find the area of the cross-hatched region.

(b) If c is the midpoint of AB show that the area in (a) is $\frac{4}{3}$ the area of the triangle AMB. This property of the parabola was first proved by Archimedes in about 250 B.C.



2. Calculating Volume

Associated with the concept of volume there are properties closely analogous to those of area, as follows:

(i) Corresponding to every three dimensional region R, there is a number $V(R) \geq 0$, called the volume of R.

(ii) If a region R is contained in a region S, then $V(R) \leq V(S)$. [Thus, in Figure 2-1 the volume of the sphere will be $\leq$ the volume of the cone.]

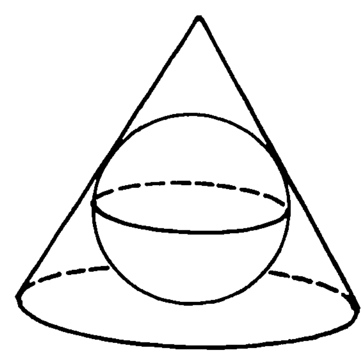
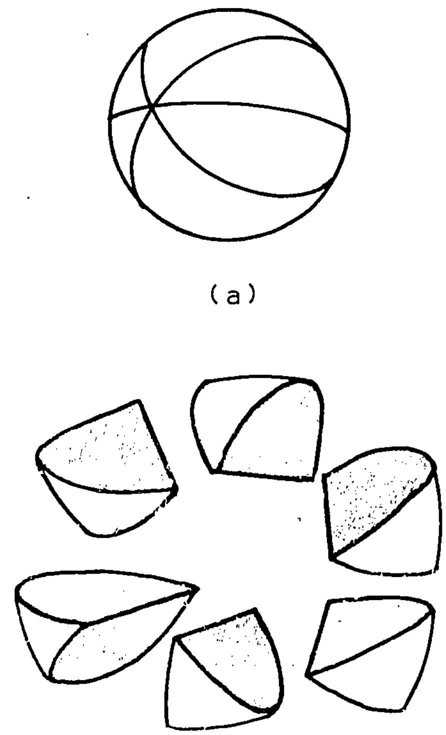


Figure 2-1

(iii) Congruent regions have the same volume.

(iv) If a region R is decomposed into two or several nonoverlapping parts then the volume of R is equal to the sum of the volumes of the individual parts. [Thus, for example, the volume of the "orange" in Figure I-2(a) is the sum of the volumes of the "segments" shown in the "exploded" diagram of Figure I-2(b).]



(a)

(b)

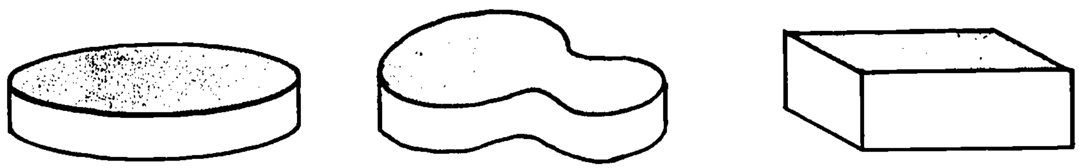
Figure I-2

(v) The volume of a "right cylinder" is the product of the height and the area of the base.

In connection with (v) we denote as right cylinders a more general type of region than the familiar circular cylinders. In general by a right cylinder we denote the region bounded by two planes and by a "wall" composed of line segments perpendicular to the two planes.

The plane figures cut off on the two parallel planes are congruent and are called the "bases" of the cylinder.

When we refer to the "area of the base" we refer to the area of one of these bases.



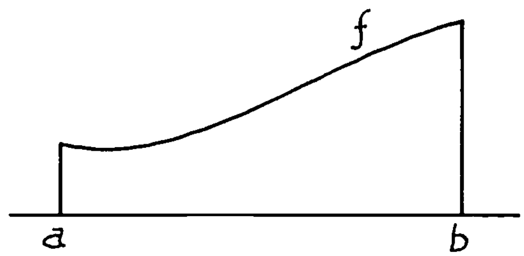
(a) Circular (b) Guitar shaped (c) Box shaped

Figure 1-3 Right Cylinders

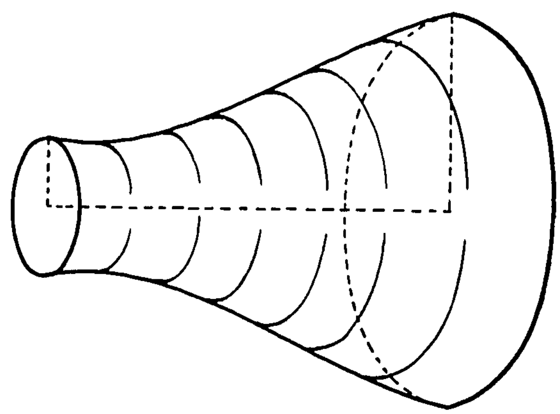
In analogy with area, we will see in the next section how the volumes of certain solids, which are not cylinders, can be represented as integrals.

2. Solids of Revolution

If the region under the graph of a function, as seen in Figure 2-1(a), is rotated about the X-axis, the space-capsule-shaped solid swept out [Figure 2-1(b)] is called a <u>solid of revolution</u>. Many familiar geometrical solids have this form such as the cone and the sphere in Figure 2-2.



(a)



(b)

Figure 2-1

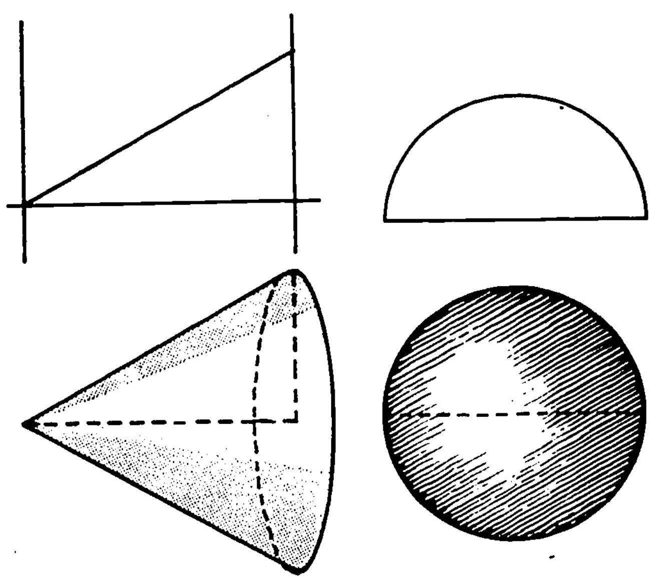


Figure 2-2

Now we are ready to see how the properties of volume in the preceding section enable us to determine the volumes of solids of revolution.

In Figures 2-3(a) and 2-3(c) we have returned to the picture in Figure 2-1(a), partitioned the interval [a,b]. and formed upper and lower rectangular configurations. Then in Figures 2-3(b) and 2-3(d) we rotated these configurations about the X-axis obtaining wedding-cake-shaped cylindrical configurations respectively contained in and containing the space-capsule-shaped solid shown in Figure 2-1(b).

According to the second of the properties of volume in the preceding section, the volumes of these "wedding cakes" yield respectively lower and upper approximations of the volume of the space capsule.

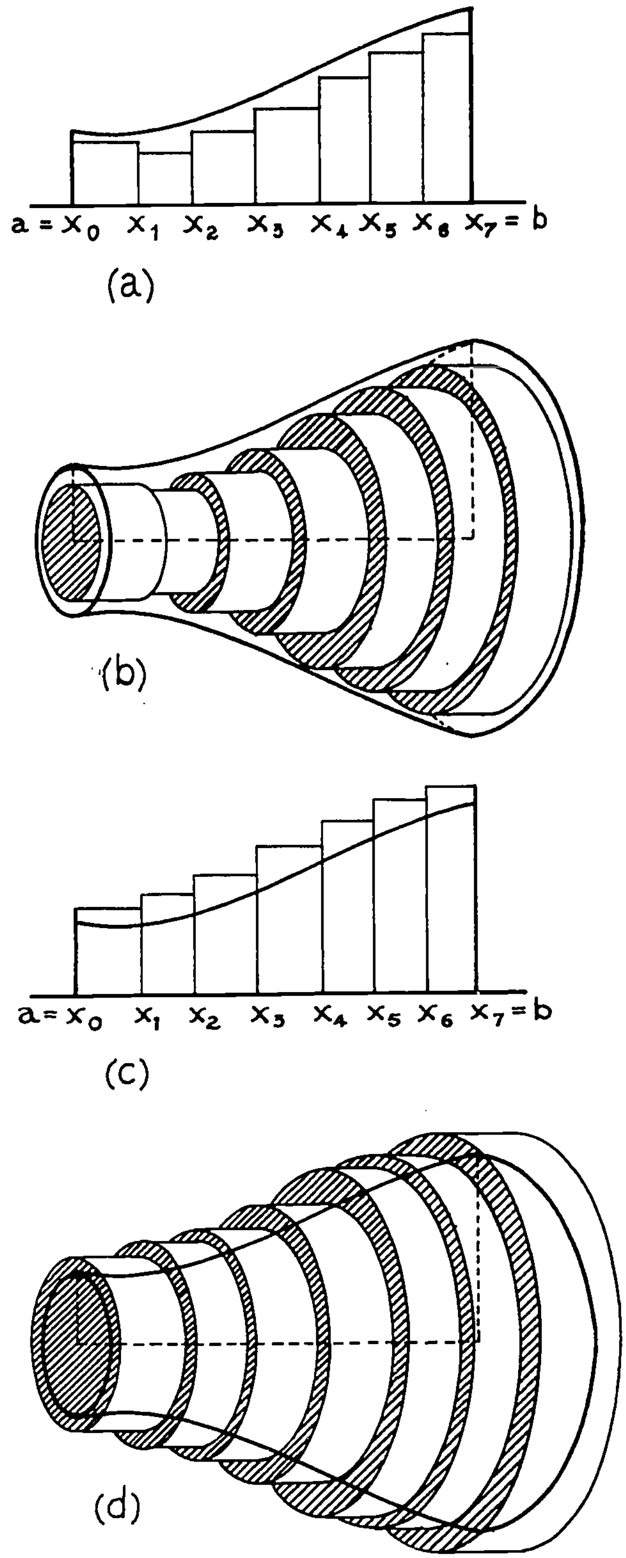


Figure 2-3

According to the fourth of the properties of volume in the preceding section, the volume of such a wedding cake is the sum of the volumes of the individual layers. To find the volume of such a layer, we have in Figure 2-4(a) singled out one of the rectangles in Figure 2-3(c) and shown the cylinder swept out by this rectangle in Figure 2-4(b). According to the last property of volume in the preceding section, the volume of this cylinder is the area of the base times the height or

$$\pi M_k^2(x_k - x_{k-1}).$$

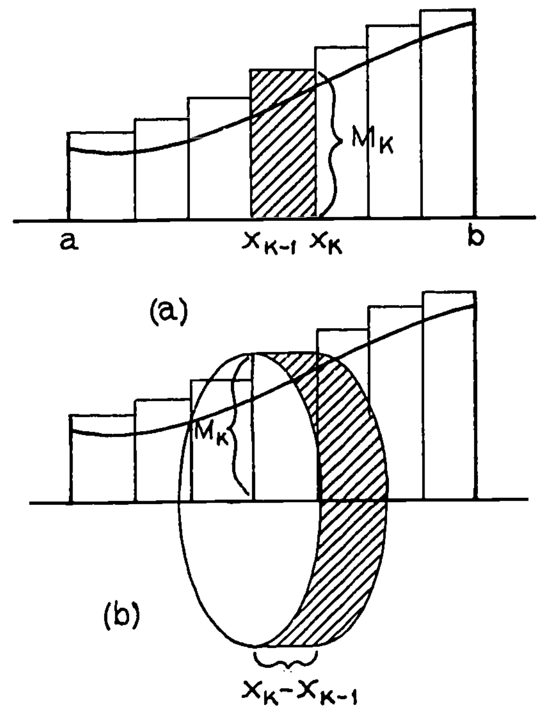


Figure 2-4

Adding up the volumes found in this way for the individual layers, we find for the volume of the outer wedding cake in Figure 2-3(d)

$$U = \sum_{k=1}^{n} \pi M_k^2(x_k - x_{k-1})$$

where n denotes the number of subintervals in the partition of [a,b] and the M_k denote the heights of the rectangles in Figure 2-3(c). Similarly the volume of the inner wedding cake in Figure 2-3(b) is given by

$$L = \sum_{k=1}^{n} \pi m_k^2(x_k - x_{k-1})$$

where the m_k denotes the heights of the rectangles in Figure 2-3(a). Moreover we have the inequalities

$$L \leq V \leq U$$

where V represents the volume of the space capsule.

From the inequalities

$$m_k \leq f(x) \leq M_k \quad \text{for } x_{k-1} \leq x \leq x_k$$

it follows that

$$\pi m_k^2 \leq \pi[f(x)]^2 \leq \pi M_k^2 \quad \text{for } x_{k-1} \leq x \leq x_k$$

or in other words

$$\pi m_k^2 \quad \text{and} \quad \pi M_k^2$$

are lower and upper bounds for the function $F(x) = \pi[f(x)]^2$ on the interval $[x_{k-1}, x_k]$. Consequently, L and U are lower and upper sums for this function F on the interval [a,b].

Assuming that f is integrable over [a,b] so that $F(x) = \pi[f(x)]^2$ is also integrable over this interval, we can construct sequences L_1, L_2, ... and U_1, U_2, ... of such lower and upper sums so that $U_n - L_n$ converges to 0 as $n \to \infty$. From the theory in Chapter 3,

$$\lim_{n\to\infty} L_n = \lim_{n\to\infty} U_n = \int_a^b F(x)\,dx = \int_a^b \pi[f(x)]^2 dx$$

However, since

$$L_n \le V \le U_n,$$

it is also clear that

$$\lim_{n\to\infty} L_n = V = \lim_{n\to\infty} U_n.$$

Thus we see that

$$V = \lim_{n\to\infty} L_n = \int_a^b \pi[f(x)]^2 dx$$

This is typical of the reasoning involved in representing geometrical and physical quantities by means of integrals, but in future applications we will greatly condense the explanation.

Example 1: To find the volume of a sphere of radius a.

Solution: First we will find the volume of the hemisphere swept out by rotating about the X-axis the region under the graph of

$$f(x) = \sqrt{a^2 - x^2}$$

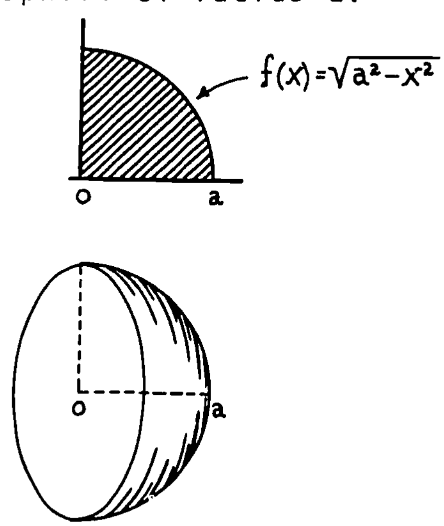


Figure 2-5

between $x = 0$ and $x = a$.

According to the above discussion the volume of this hemisphere is given by

$$\int_0^a \pi[f(x)]^2 dx = \int_0^a \pi(\sqrt{a^2 - x^2})^2 dx$$

$$= \int_0^a \pi(a^2 - x^2) dx$$

$$= \pi(a^2 x - \frac{x^3}{3}) \Big|_0^a = \pi(a^3 - \frac{a^3}{3}) = \frac{2}{3}\pi a^3$$

(The notation $F(x)\Big|_a^b$ as used above is a convenient shorthand for $F(b)-F(a)$.) Thus the volume of the entire sphere is $\frac{4}{3}\pi a^3$.

The method presented here can be extended to finding the volumes of more complicated regions such as that obtained by rotating about the X-axis the region indicated in Figure 3-6(a).

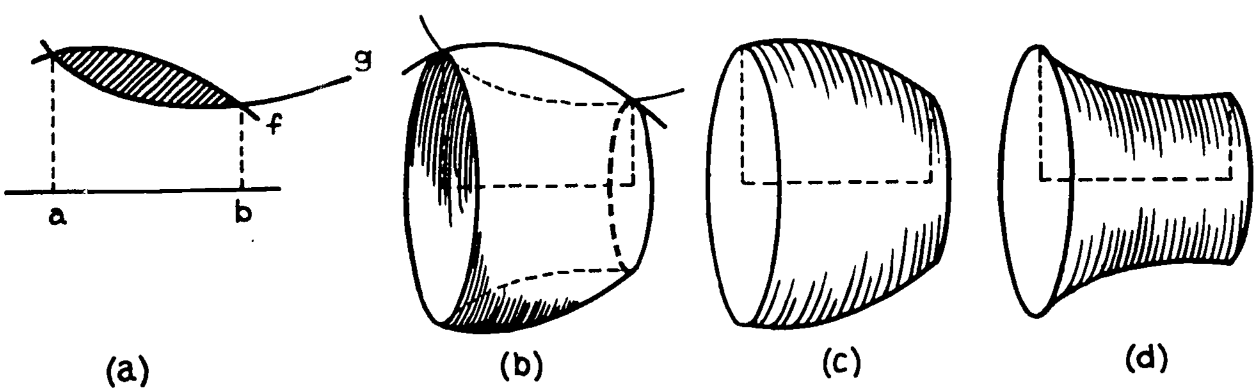


Figure 3-6

The volume of this solid depicted in Figure 2-6(a) is clearly the difference of the volumes of the solids in Figures 2-6(c) and 2-6(d) obtained by rotating the regions under the graph of f and that under the graph of g. The volume in Figure 2-6(b) is therefore given by

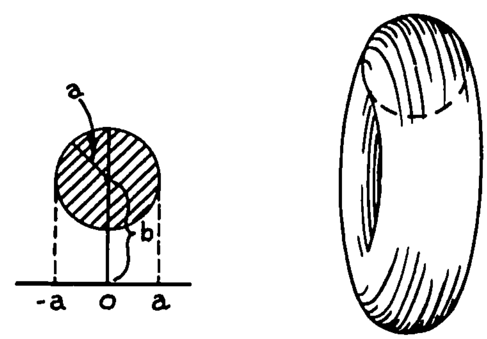
$$V = \int_a^b \pi[f(x)]^2 dx - \int_a^b \pi[g(x)]^2 dx$$

$$= \pi \int_a^b ([f(x)]^2 - [g(x)]^2) dx.$$

Example 2: To find the volume of the "doughnut" obtained by rotating about the X-axis the region inside the circle

$$x^2 + (y - b)^2 = a^2$$

where $a \leq b$.



(a) (b)

Figure 2-7

Solution: Solving the above equation for y, we obtain

$$y = b \pm \sqrt{a^2 - x^2}$$

so that the upper and lower bounding curves for the region in Figure 2-7(a) are given by

$f(x) = b + \sqrt{a^2 - x^2}$ and $g(x) = b - \sqrt{a^2 - x^2}$ where $-a \leq x \leq a$.

Now $[f(x)]^2 - [g(x)]^2$ is calculated to be

$$[f(x)]^2 - [g(x)]^2 = (b + \sqrt{a^2 - x^2})^2 - (b - \sqrt{a^2 - x^2})^2$$
$$= 4b\sqrt{a^2 - x^2} .$$

Using the technique of integrating from 0 to a and doubling, we have

$$V = 2\pi \int_0^a 4b\sqrt{a^2 - x^2}dx = 8\pi b\int_0^a \sqrt{a^2 - x^2}dx$$

At first glance we seem to be at an impasse in finding a formula for the volume of the doughnut for we as yet have no technique for evaluating the integral

$$\int_0^a \sqrt{a^2 - x^2}dx .$$

However, we quickly recall that this integral also represents the area of a quarter circle of radius a (as dealt with at considerable length in the early sections of Chapter 3).

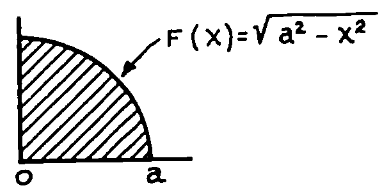


Figure 3-8

Therefore,

$$\int_0^a \sqrt{a^2 - x^2}dx = \frac{\pi a^2}{4}$$

and the volume of our doughnut is now seen to be

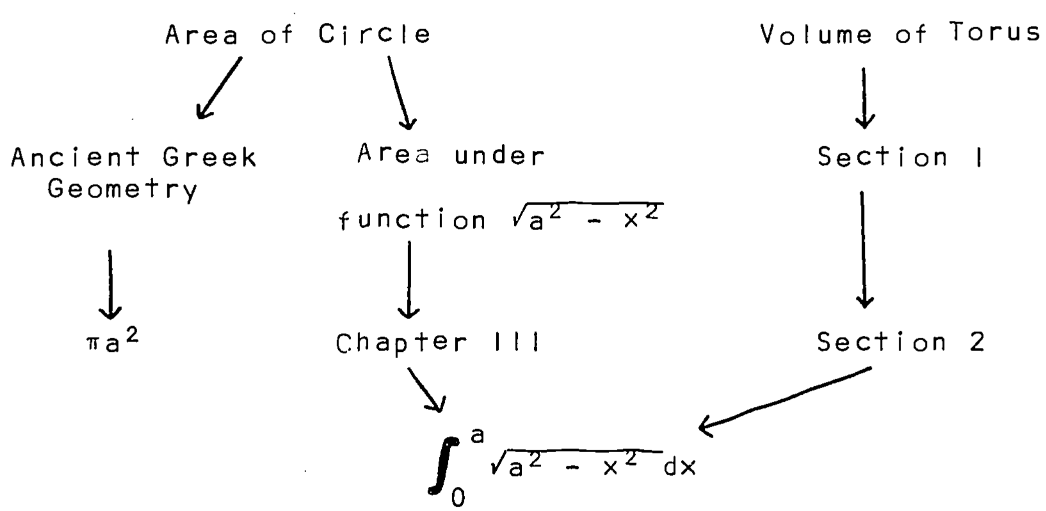
$$V = 8\pi b \int_0^a \sqrt{a^2 - x^2}dx = 8\pi b \frac{\pi a^2}{4} = 2\pi^2 a^2 b.$$

The trick that we used here is an example of an important technique known as the Method of Analogy. Suppose that we have two problems that can be reduced, at some stage, to the same mathematical model; in our case this model was the integral

$$\int_0^a \sqrt{a^2 - x^2}\, dx.$$

If we know the solution of this model by means of some other approach to the first problem, then we can apply this knowledge to the second problem.

We can illustrate our two problems with the following rough diagram:



Example 3. Suppose that a gold ring is made by cutting a cylinder of radius r out of a sphere of radius R.

(a) What is the ring's width?

(b) What is the volume of the ring? Hint: We may think of the ring as being obtained by rotating about the X-axis the region between the graphs of $y = \sqrt{R^2 - x^2}$ and $y = r$. (See Figure 3-9).

(c) Write a formula for the volume V of the ring in terms of its width W.

Solution: The two graphs intersect when $r = \sqrt{R^2 - x^2}$, that is, $x = \pm\sqrt{R^2 - r^2}$. Thus the ring extends from $x = -\sqrt{R^2 - r^2}$ to $x = \sqrt{R^2 - r^2}$, and its width is $2\sqrt{R^2 - r^2}$. Its volume is

$$2\int_0^{\sqrt{R^2 - r^2}} (\pi(\sqrt{R^2 - x^2})^2 - \pi r^2)dx$$

$$= 2\pi\int_0^{\sqrt{R^2 - r^2}} (R^2 - x^2 - r^2)dx$$

$$= 2\pi(R^2\sqrt{R^2 - r^2} - \frac{(R^2 - r^2)^{3/2}}{3} - r^2\sqrt{R^2 - r^2})$$

$$= \frac{4}{3}\pi(R^2 - r^2)^{3/2}$$

We have $W = 2\sqrt{R^2 - r^2}$ and $V = \frac{4}{3}\pi(R^2 - r^2)^{3/2}$. Therefore, $V = \frac{\pi}{6} W^3$.

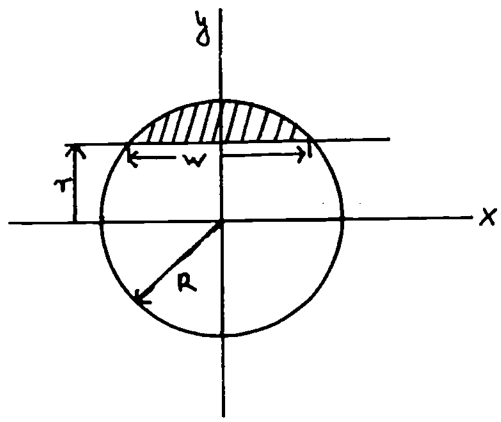


Figure 2-9

PROBLEMS

1. Compute the volume of the solid obtained by rotating each of the following figures around the X-axis.

   (a) The rectangle with vertices at (0,0), (5,0), (5,3), (0,3).

   (b) The triangle with vertices at (-1,0), (1,0), (0,3).

   (c) The region bounded by the curves $y = x^2$, $y = 0$, $x = -1$, $x = 1$.

   (d) The region bounded by $x = 4y^2$ and $x = 7$.

   (e) The region bounded by $y = x$ and $x = y^2$.

   (f) The region bounded by $y = x^2$ and $x = y^2$.

   (g) The square with vertices at (7,2), (7,4), (9,2), (9,4).

   (h) The triangle with vertices at (-1,2), (1,2), (0,5).

   (i) The region bounded by $y = 0$, $y = \sqrt{\sin x}$, $x = 0$, $x = \frac{\pi}{2}$.

   (j) The region bounded by $y = 1$, $y = \sqrt{\sin x}$, $x = 0$, $x = \frac{\pi}{2}$.

2. If A and B are congruent regions, and if V_1 and V_2 are volumes of the solids obtained by rotating A and B, respectively, about the X-axis, is it necessarily true that $V_1 = V_2$?

3. Suppose A is a region and B is a region obtained by translating the region A to the right or left. Is the volume obtained by rotating A about the X-axis necessarily equal to the volume obtained by rotating B about the X-axis?

3. Another Characterization of the Integral

When electrical users are billed, they are charged for the number of kilowatt-hours they have consumed during the month. The units require some explanation. The "watt" is a unit of electrical power; the kilowatt is a somewhat more convenient unit equal to 1000 watts. The kilowatt-hour is a unit of electrical energy consumed in drawing one kilowatt of power for one hour. For example, ten 100 watt light bulbs draw one kilowatt of power; thus ten 100 watt bulbs switched on for one hour consume one kilowatt-hour of electrical energy. If power is used at a constant rate (i.e., not varying with time), then the formula for the energy consumed is

$$\text{ENERGY} = \text{POWER} \times \text{TIME}.$$

Thus when a five kilowatt air conditioner is operated for three hours, the amount of energy consumed is

$$5 \text{ kilowatts} \times 3 \text{ hours} = 15 \text{ kilowatt-hours}.$$

(In the future we will suppress the units in such calculations.)

If we look at the power drawn by an entire town during a 24 hour day, the power will certainly not be constant but will vary with time. In Figure 3-1 there is a graph of such a power function, f. There are dips in the graph when factories shut down at lunch and at the end of the working day, and there is a hump when lights and televisions come on in the evening. In the middle of the night the main power consumption is due to street lighting and refrigerators. Now we ask the question, if the power function is given, how can we determine the number of kilowatt-hours of energy used over a given time interval?

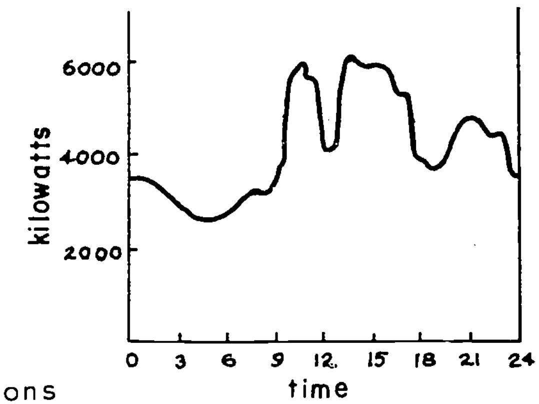


Figure 3-1

First we introduce a notation to represent this quantity. We will use

$$E_{t_1}^{t_2}(f)$$

to denote the energy consumption between the times t_1 and t_2 using a power function, f.

Next we take note of three relatively obvious properties of energy consumption.

(i) The number of kilowatt-hours of energy consumed between 3 o'clock and 6 o'clock is the sum of the amount consumed between the hours of 3 and 5 and that consumed between 5 and 6. In general, if $t_1 \leq t_2 \leq t_3$ then

$$E_{t_1}^{t_3}(f) = E_{t_1}^{t_2}(f) + E_{t_2}^{t_3}(f).$$

(ii) If user A at all times draws less power than user B, then user A will use less energy than user B. We express this principle in the form: If $f(t) \leq g(t)$ for $t_1 \leq t \leq t_2$ (as seen in Figure 3-2) then $E_{t_1}^{t_2}(f) \leq E_{t_1}^{t_2}(g)$.

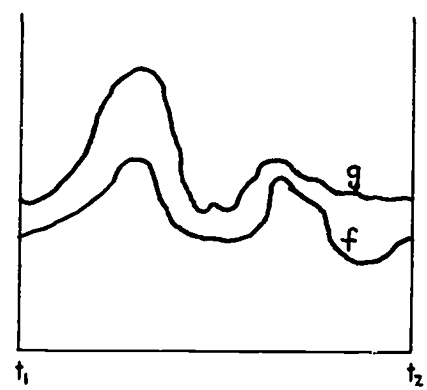


Figure 3-2

(iii) As already observed above, if power is drawn at a constant rate, then the energy consumption is the product of the power and the time. That is, if $f(t) = k$ for $t_1 \leq t \leq t_2$, then

36

$$E_{t_1}^{t_2}(f) = k(t_2 - t_1).$$

These three properties suffice for us to prove that for any power function f and any time interval [a,b],

$$E_a^b(f) = \int_a^b f(t)\,dt$$

provided of course that f is integrable over [a,b].

We will see this by showing that the inequality

$$L \le E_a^b(f) \le U$$

holds whenever L and U are lower and upper sums for f over [a,b]. We show this for upper sums in three steps using one of the above properties in each step.

Step 1. Consider an upper sum

$$U = \sum_{k=1}^{n} M_k(t_k - t_{k-1})$$

for f over [a,b] and focus attention momentarily on the subinterval $[t_0, t_1]$. We let

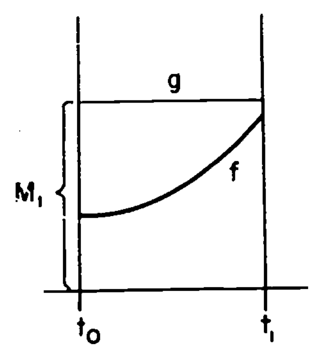


Figure 3-3

$g(t) = M_1$ for $t_0 \le t \le t_1$. Since $f(t) \le M_1 = g(t)$ for $t_0 \le t \le t_1$, we see by property (ii) that

$$E_{t_0}^{t_1}(f) \le E_{t_0}^{t_1}(g).$$

Step II. Since g is constant in $[t_0, t_1]$, property (iii) yields $E_{t_0}^{t_1}(g) = M_1(t_1 - t_0)$ so that

$$E_{t_0}^{t_1}(f) \le M_1(t_1 - t_0).$$

By the same reasoning applied to each subinterval $[t_{k-1}, t_k]$, we obtain

$$E_{t_{k-1}}^{t_k}(f) \le M_k(t_k - t_{k-1}) \qquad \text{for} \qquad k = 1, 2, \ldots, n.$$

Step III. Adding these inequalities for $k = 1, 2, \ldots, n$, we find

$$\sum_{k=1}^{n} E_{t_{k-1}}^{t_k}(f) \le \sum_{k=1}^{n} M_k(t_k - t_{k-1}) = U$$

Finally, repeated application of property (i) above yields

$$E_a^b(f) = \sum_{k=1}^{n} E_{t_{k-1}}^{t_k}(f) \le U$$

Conclusion. We have shown that $E_a^b(f) \leq U$ for every upper sum U. It may be shown in a similar manner that $L \leq E_a^b(f)$ for every lower sum L. It now follows that

$$E_a^b(f) = \int_a^b f(x)dx.$$

Example. Suppose that a certain pogo stick company consumes energy according to the power function $p(t) = 20000t - 2500t^2$ during the time interval [0,8]. How much energy is consumed?

Solution.
$$\int_0^8 p(t)dt = \int_0^8 (20000t - 2500t^2)dt$$
$$= 20000 \int_0^8 t dt - 2500 \int_0^8 t^2 dt$$
$$= 20000 \left(\frac{8^2}{2}\right) - 2500 \frac{8^3}{3}$$
$$= 640000 - 426667 = 213333$$

If we extract the mathematical content from the foregoing discussion and forget about power and energy, we have the following principle.

We assume that for each function f and each interval [a,b], a number $E_a^b(f)$ is determined and that the following properties are satisfied:

(i) $E_a^b(f) + E_b^c(f) = E_a^c(f)$ whenever $a \leq b \leq c$;

(ii) If $f(x) \leq g(x)$ for x in $[a,b]$, then $E_a^b(f) \leq E_a^b(g)$;

(iii) If $f(x) = k$ (constant) over $[a,b]$, then
$E_a^b(f) = k(b - a)$.

Under these conditions, we may conclude that

$$E_a^b(f) = \int_a^b f(x)dx$$

Many applications of the integral (but by no means all) fall into this category. Let us see one more.

Let

$$D_{t_1}^{t_2}(f)$$

represent the distance travelled by a moving object where the function, f, denotes its speed, i.e.,

$$f(t) = \text{speed of object at time } t.$$

We shall show that

$$D_{t_1}^{t_2}(f) = \int_{t_1}^{t_2} f(t)dt$$

by verifying the three properties specified in the principle above.

(i) $D_{t_1}^{t_2}(f) + D_{t_2}^{t_3}(f) = D_{t_1}^{t_3}(f)$

(This says that the distance travelled between times t_1 and t_2 plus the distance travelled between times t_2 and t_3 equals the total distance travelled between times t_1 and t_3.)

(ii) If $f(t) \leq g(t)$ for $t_1 \leq t \leq t_2$, then

$$D_{t_1}^{t_2}(f) \leq D_{t_1}^{t_2}(g).$$

(This says that the faster moving object covers the greater distance in a given time.)

(iii) If $f(t) = k$ for $t_1 \leq t \leq t_2$, then

$$D_{t_1}^{t_2}(f) = k(t_2 - t_1).$$

(This i the well-known formula: distance = speed x time.)

Having verified that these properties hold, we can now conclude

$$D_{t_1}^{t_2}(f) = \int_{t_1}^{t_2} f(t)\,dt$$

Example: Suppose that in a drag race a car speeds up in such a way that its velocity (or speed) is given by

$$v(t) = 10t$$

where the time, t, is measured in seconds and the velocity is measured in feet/second. How far will the car travel in 16 seconds after leaving the starting line?

Solution. According to our formula and results from Chapter 3,

$$D_0^{16}(v) = \int_0^{16} 10t\,dt = \frac{10}{2}\cdot 16^2 = 1280$$

Further illustrations of these principles will be found in the exercises.

Let us now consider another situation in which the three principles apply.

If an object is displaced a distance d by a constant force F acting in the direction of the displacement (as in lifting a weight) the work done (according to a definition in physics) is F · d. In numerous situations we deal with forces

which are not constant but depend on the position of the object. (For example the force required to displace an object attached to a spring becomes greater as the spring is stretched.) In general, then, the force required to displace an object may be represented as a function, $f(x)$, of the position, x. We let $W_a^b(f)$ denote the work done in moving an object from $x = a$ to $x = b$ with a force $f(x)$ and we accept from physics the intuitively reasonable properties:

(i) $W_a^b(f) + W_b^c(f) = W_a^c(f)$, if $a \leq b \leq c$;

(ii) $W_a^b(f) \leq W_a^b(g)$ if $f(x) \leq g(x)$ for x in $[a,b]$;

(iii) if $f(x) = k$ (constant) over $[a,b]$, then $W_a^b(f) = k(b - a)$. (The third property is merely a reformulation of the condition in the first sentence of this problem.)

It follows that $W_a^b(f) = \int_a^b f(x)\,dx$.

<u>Example</u>: Suppose a tow truck is pulling a car which is becoming increasingly difficult to pull because its tires are going flat. Indeed suppose that the force required is $2000 + 40t^2$ for time t between 0 and 10. How much work is done over the indicated period of time?

Solution. $W_0^{10}(2000 + 40t^2) = \int_0^{10} (2000 + 40t^2)\,dt$

$= (2000)(10) + \frac{40}{3}(10^3) = 20000 + 13333 = 33333.$

372

PROBLEMS

1. Suppose that the Jinglehammer Hair Curler factory measures its power requirements every half hour during an eight-hour work day and obtains the following results, where f is the power function (in kilowatts):

t	8	8.5	9	9.5	10	10.5	11	11.5	12	12.5	13	13.5	14	14.5	15	15.5	16	16.5	17
$f(t)$	0	50	100	150	140	145	150	60	0	0	0	50	100	140	150	150	150	50	0

(a) Estimate the day's energy consumption by using the trapezoid rule to approximate $\int_8^{17} f(t)dt$.

(b) If the same energy is consumed every day of a twenty-day work month, how much energy is consumed during the month?

(c) What is the electric bill for the month if the electric company charges 4¢ per kilowatt-hour for the first 2000 kilowatt-hours and 2¢ per kilowatt-hour for each additional kilowatt-hour?

2. For each of the following power functions f, compute the energy consumed over the indicated time interval.

(a) $f(t) = 200$, $[0,10]$

(b) $f(t) = 3t^2$, $[2,5]$

(c) $f(t) = \sin t$, $[0,\pi]$

(d) $f(t) = t - t^2$, $[0,1]$.

3. Suppose a certain power function is given by $f(t) = 200t$ for nonnegative t. Find the positive number B such that the energy expended over the time interval $[0,B]$ will be 1000 kilowatt-hours.

4. (a) Let f be a nonnegative integrable function defined on an interval $[a,b]$. If the graph of f is rotated around the X-axis, a solid of revolution is generated. For $a \leq c \leq d \leq b$ let $V_c^d(f)$ represent the volume of the part of this solid bounded by the planes $x = c$ and $x = d$. Then:

(i) $V_c^d(f) = V_c^x(f) + V_x^d(f)$ for $c \leq x \leq d$;

(ii) $V_c^d(f) \leq V_c^d(g)$ if $f(x) \leq g(x)$ for all x in $[c,d]$;

but (iii) if $f(x) = k$ (constant), then $V_c^d(f) = \pi k^2(d - c)$.

Show that $V_c^d f \neq \int_c^d f(x)dx$ (except for special cases).

(b) For each x in $[a,b]$, $A(x) = \pi(f(x))^2$ represents the area of the cross-section of the solid of revolution in the plane perpendicular to the X-axis at x. For $a \leq c \leq d \leq b$, let $V_c^d(A)$ represent the volume of the part of the solid bounded by the planes $x = c$

and $x = d$. Which of the following properties hold?

(i) $V_c^d(A) = V_c^x(A) + V_x^d(A)$ for $c \leq x \leq d$;

(ii) $V_c^d(A) \leq V_c^d(B)$ if $A(x) \leq B(x)$ for all x in $[c,d]$.

(iii) if $A(x) = k$ (constant, then $V_c^d(A) = k(d - c)$.

(c) Does $V_c^d(A) = \int_c^d A(x)dx = \int_c^d \pi(f(x))^2 dx$?

5. Work problems.

(a) A 40 pound bucket of coal is being pulled up a 100 foot chute by a chain weighing 2 pounds per foot. Thus when the bucket has been raised x feet, the force with which the bucket must be pulled is $40 + 2(100 - x)$, that is, the weight of the bucke plus the weight of the part of the chain not yet drawn in. How much work is done in raising the bucket to the top of the chute?

(b) A load of lime is being pulled up the side of a 500 foot cliff in a rainstorm. The load initially weighed 5000 pounds, but it is being washed away at a rate of 20 pounds per minute. If the load is raised 50 feet per minute, how much work is done altogether? (Ignore the weight of the cables.)

(c) According to Hooke's Law, as long as a spring is not stretched beyond its elastic limit, the force on the spring is equal to kx, where k is the spring constant and x is the distance that the spring is extended beyond its natural length. Thus if a spring with natural length L and spring constant k is stretched to a length M, the force required is given by $f(x) = kx$, for $0 \leq x \leq M - L$, and the work required is $\int_0^{M-L} f(x)dx = \int_0^{M-L} kxdx$.

How much work is required to double the length of a 2 meter spring with spring constant k = 5 newtons per meter?

(d) How much work is done in extending a 1 meter spring an additional centimeter if the spring constant is 3 newtons per meter?

(e) Suppose that a rubber strap 50 centimeters long has a "spring constant" of 2 newtons per meter. How much work is done in stretching the strap to a length of 3 meters?

(f) The good witch Zarc is flying from her hut to the cave of the giant Carx, 2000 meters away. Carx is blowing a wind toward her with a force $f(x) = 3x^2$ (newtons), where x is the distance from Zarc's hut. How much work will be done by Zarc over the course of the flight?

CHAPTER 5

LOCAL PROPERTIES OF FUNCTIONS

1. Speed and Limits

Since our elementary school days we have all been familiar with the formula

$$\text{speed} = \frac{\text{distance}}{\text{time}}.$$

A typical problem might be: In a drag race a car travelled a quarter of a mile in 12 seconds. What was its speed? We get the answer in feet per second by

$$\text{speed} = \frac{1320 \text{ ft.}}{12 \text{ sec.}} = 110 \text{ ft./sec.}$$

In miles per hour the calculation would be

$$\text{speed} = \frac{\frac{1}{4} \text{ mile}}{\frac{12}{3600} \text{ hrs.}} = 75 \text{ mph.}$$

We all know the speed referred to in this example is "average speed." At the start of the race the speed was zero and as the car crossed the finish line its speed was well over 100 mph.

We have made a sort of apology for talking about average speed, but what other kind of speed is there? What do we mean by saying that the speed of the car at the instant it crossed the finish line was well over 100 mph? Presumably everyone has a feeling for the concept of "instantaneous velocity" but a definition of it is no trivial matter. We can probably agree to the validity of the following method of approximating the instantaneous velocity at the finish line and that will lead us to a definition.

Suppose in the example under discussion we were somehow able to determine the position of the car $\frac{1}{100}$ of a second after it crossed the finish line; let's say that it was $1\frac{3}{4}$ feet past the finish line. Then during this $\frac{1}{100}$ second the average velocity was

$$\frac{1\frac{3}{4}\text{ feet}}{\frac{1}{100}\text{ sec.}} = 175 \text{ feet/sec.} \qquad \text{or} \qquad 119.3 \text{ mph.}$$

Our reasoning here is that the speed could not have changed much over such a short time interval so that the instantaneous velocity will not be much different from the average velocity. The shorter the time interval the less change in speed and the better the average velocity approximates the instantaneous velocity. Thus if we can imagine that we measured that the car travelled $2\frac{3}{32}$ inches in $\frac{1}{1000}$ second after crossing the finish line we would calculate the average velocity over this time interval as

$$\frac{2\frac{3}{32}\text{ inches}}{\frac{1}{1000}\text{ seconds}} = \frac{\frac{67}{384}\text{ feet}}{\frac{1}{1000}\text{ sec}} = 174.5\text{ ft./sec. or }119.0\text{ mph.}$$

We can begin to see that the computation of the exact instantaneous velocity involves a limiting process. We would define the instantaneous velocity as the limit of the average velocity over shorter and shorter time intervals. Such a definition is entirely impractical as the measurements of the distances and times involved would be subject to great inaccuracies, and it is obviously impossible to make infinitely many such measurements in order to take the limit.

However, in mathematical discussions of velocity we deal not with actual measurements but with mathematical models in which the positions of the moving objects are expressed as functions of the time by means of formulas. Here the definition of instantaneous velocity as a limit is quite satisfactory.

Example: A disc one foot in radius is attached to a motor which turns at a uniform rate. [By this we mean that for some constant k the angle θ through which the shaft turns in a time t is given by the formula $\theta = k \cdot t$. For simplicity of calculation we will take $k = 1$ so that the angle θ (in radians) is numerically equal to the time (in seconds).] There is a small ball on the rim of the disc. The apparatus is depicted in Figure 1-1.

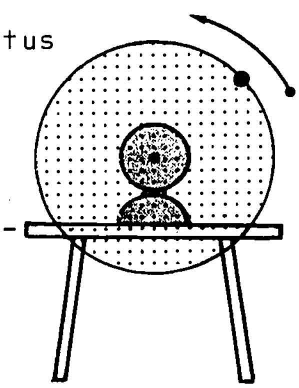


FIGURE 1-1

Light is coming in a window horizontally to the disc so as to project on a wall a shadow of the disc and ball as seen in Figure 1-2 where the motor and table are not shown. We assume that at the starting time the ball is level with the center and to right of it. The problem is to find a formula for the instantaneous velocity of the shadow of the ball at time t.

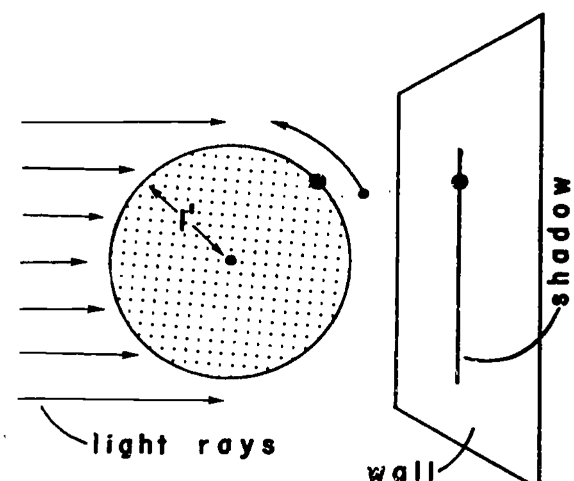


FIGURE 1-2

Discussion: We see that as the disc rotates, the ball moves in a circle while the shadow moves up and down on a line. Though the ball traverses equal distances in equal times this is not so of its shadow. For when the ball is near the top or bottom of the disc it is travelling nearly horizontally so that the shadow hardly moves at all, while when the ball is near the level of the center it is travelling nearly vertically so that the shadow moves almost at the same speed as the ball.

Solution. We introduce coordinates with the origin at the center of the disc. The elevation, $y(t)$, of the shadow above the X-axis is the same as that of the ball. Since the angle (measured in radians) turned by the disc from its starting position is numerically equal to the time, we have

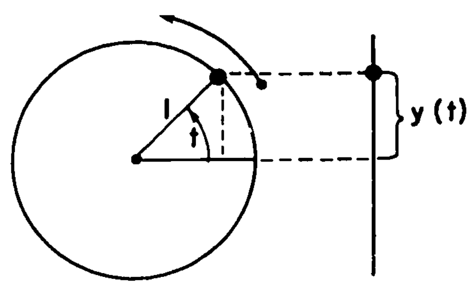


FIGURE 1—3

$$y = \sin t$$

The average velocity of the shadow between times t_0 and t_1 is

$$\frac{\text{distance}}{\text{time}} = \frac{y(t_1) - y(t_0)}{t_1 - t_0}.$$

If we consider a sequence of times t_1, t_2, t_3, ... with $t_n \to t_0$ as $n \to \infty$ (with none of the t_n actually equal to t_0) then the instantaneous velocity at time t_0 is given by

$$v(t_0) = \lim_{n\to\infty} \frac{y(t_n) - y(t_0)}{t_n - t_0} = \lim_{n\to\infty} \frac{\sin(t_n) - \sin(t_0)}{t_n - t_0}.$$

It turns out that we will be able to evaluate this limit.

For our purposes it will simplify matters to introduce the notation

$$h_n = t_n - t_0 \qquad \text{so that} \qquad t_n = t_0 + h_n$$

and note that $\lim_{n\to\infty} h_n = 0$. With this notation

$$v(t_0) = \lim_{n\to\infty} \frac{\sin(t_0 + h_n) - \sin(t_0)}{h_n}$$

a few trigonometric and algebraic manipulations bring this limit into a form where the value is easily recognizable. Thus

$$\frac{\sin(t_0 + h_n) - \sin(t_0)}{h_n} = \frac{\sin(t_0)\cos(h_n) + \sin(h_n)\cos(t_0) - \sin(t_0)}{h_n}$$

$$= \cos(t_0)\,\frac{\sin(h_n)}{h_n} - \sin(t_0)\,\frac{1 - \cos(h_n)}{h_n} .$$

The theorems on limits of sums and products from Chapter 2 assure us that

$$v(\quad = \cos(t_0) \lim_{n\to\infty} \frac{\sin(h_n)}{h_n} - \sin(t_0) \lim_{n\to\infty} \frac{1 - \cos(h_n)}{h_n} .$$

The two limits $\lim_{n\to\infty} \frac{\sin(h_n)}{h_n}$ and $\lim_{n\to\infty} \frac{1 - \cos(h_n)}{h_n}$, where $h_n \to 0$ as $n \to \infty$, were evaluated in section 2-8 respectively as 1 and 0.

Thus

$$\begin{aligned} v(t_0) &= (\cos t_0) \cdot 1 - (\sin t_0) \cdot 0 \\ &= \cos t_0 . \end{aligned}$$

This important example has shown that in cases where the position of a moving object is given by formula we can find the instantaneous velocity by culation, whereas any attempt to do so by measurement is clearly impossible.

We have shown in this example that

$$\lim_{n\to\infty} \frac{\sin(t_n) - \sin(t_0)}{t_n - t_0} = \cos t_0$$

provided that $\lim_{n\to\infty} t_n = t_0$ and that $t_n \neq t_0$ for $n = 1, 2, \ldots$. We stress that this holds true for any sequence having these properties. Put slightly differently, if we let

$$F(t) = \frac{\sin(t) - \sin(t_0)}{t - t_0}$$

then

$$\lim_{n\to\infty} F(t_n) = \cos t_0$$

for any sequence $t_1, t_2, \ldots$ satisfying the above conditions. This holds true because of a property of the function F which can be expressed quite independently of sequences.

Limit of a Function at a Point

For simplicity let us now take $t_0 = 0$ whence

$$F(t) = \frac{\sin(t) - \sin(0)}{t - 0} = \frac{\sin(t)}{t}$$

and

$$\lim_{n\to\infty} F(t_n) = \cos(0) = 1 .$$

The graph of F is seen in 1-4.

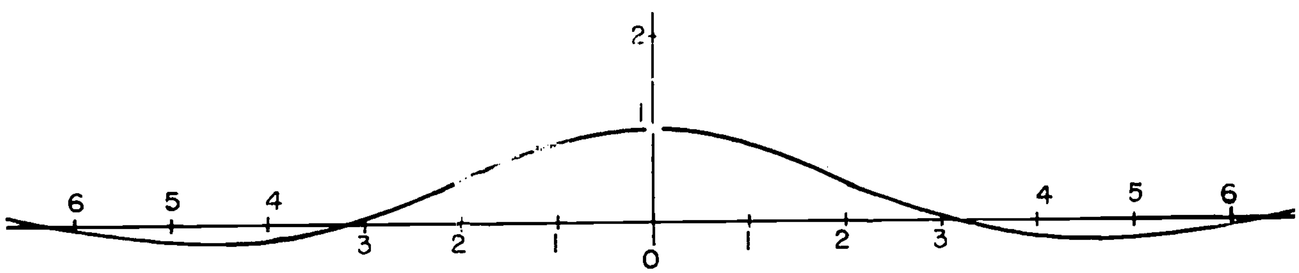


FIGURE 1-4

We see a small gap on this graph as it crosses the Y-axis to suggest the fact that F(t) is not defined for t = 0 since substitution in the formula for F(t) would yield

$$F(0) = \frac{\sin(0)}{0} = \frac{0}{0}$$

which is of course meaningless.

However, we get the impression that the value of F(0) "ought to be" 1. If we filled in the point (0,1) on the graph in Figure 1-4 we would have a nice smooth unbroken

curve. In other words, for values of t close to 0 the values of $F(t)$ approximate the number 1 very closely. We can pin down this vague idea by use of some of the inequalities developed in Chapter 3.

In that chapter we learned that

$$t - \frac{t^3}{3!} \leq \sin(t) \leq t$$

for $t \geq 0$ with the inequalities reversed for negative values of t. Hence dividing through by t we obtain

$$1 - \frac{t^2}{3!} \leq \frac{\sin(t)}{t} \leq 1$$

which holds for all values of t different from zero, positive or negative. Subtracting 1 from all members of this inequality yields

$$\frac{-t^2}{3!} \leq \frac{\sin(t)}{t} - 1 \leq 0$$

or

$$0 \leq 1 - \frac{\sin(t)}{t} \leq \frac{t^2}{6} \qquad \text{for } t \neq 0 .$$

Now we can see what it means to say that $\frac{\sin(t)}{t}$ is a close approximation of 1 for values of t close to but not equal to zero. For example if t is different from zero but differs from zero by less than $\frac{1}{100}$ (i.e., $0 < |t - 0| < \frac{1}{100}$) then $\frac{\sin(t)}{t}$ differs from 1 by less than $\frac{1}{60,000}$ (i.e.,

$\left|1 - \frac{\sin(t)}{t}\right| \le \frac{t^2}{6} = \frac{|t - 0|^2}{6} < \frac{1}{60,000}$). Or we could ask how close to zero t must be in order to guarantee that $\frac{\sin(t)}{t}$ differs from 1 by less than $\frac{1}{1,000,000}$. The answer is obtained by solving

$$\frac{t^2}{6} < \frac{1}{1,000,000}$$

or $|t - 0| < \frac{\sqrt{6}}{1000}$.

And in general, given a tolerance of error, ε, then $\frac{\sin(t)}{t}$ will differ from 1 by less than ε provided that $0 < \frac{t^2}{6} < \varepsilon$, i.e., provided that $t \neq 0$ but differs from 0 by less than $\sqrt{6\varepsilon}$. This analysis provides the basis for a new kind of limit — the limit of a function at a point. We will say that

$$\lim_{t \to 0} F(t) = 1 .$$

The general definition, motivated by this example is:

<u>Definition</u>: We say that

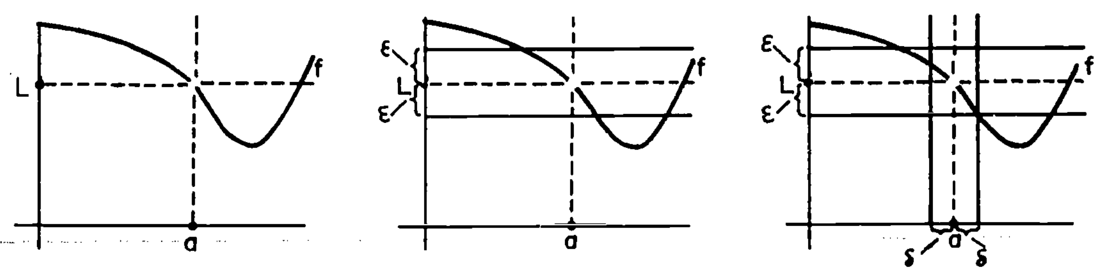
$$\lim_{x \to a} f(x) = L$$

provided that for each positive number ε there is a corresponding positive number δ so that $|f(x) - L| \le \varepsilon$ whenever $0 < |x - a| \le \delta$.

In a more descriptive wording the last line of the definition could be stated:

"f(x) differs from L by no more than ε provided that x is not equal to a but differs from a by less than δ."

The idea may also be described pictorially as seen in Figure 1-5.



(a) (b) (c)

FIGURE 1-5

In Figure 1-5(a) we see the graph of a function f in the vicinity of a. In Figure 1-5(b) we have drawn in horizontal lines ε units above and below L. In Figure 1-5(c) we have drawn in vertical lines δ units to the right and left of a.. An we see that all points of the graph lying between the two vertical lines, (excluding the point, if any, lying directly above a) also lie between the two horizontal lines. [This is the geometrical equivalent of the statement $|f(x) - L| \leq \varepsilon$ whenever $0 < |x - a| \leq \delta$.] If we are able to find such numbers δ for all ε no matter how small, then we say that $\lim_{x \to a} f(x) = L$.

In Figure 5(c) the value of δ has been chosen as large as possible. If it had been chosen any larger, the part of the graph in the vertical strip would lie outside the horizontal strip. Nothing in the definition requires us to choose δ as large as possible. Any smaller (positive) choice of δ would have served as well. In practice it is often convenient to restrict the choice of δ to be less than one (or some other fixed positive number). If this is done then we see that the questions as to whether $\lim_{x \to a} f(x) = L$ depend only on values of x taken from the interval [a - 1, a + 1]. These remarks are comprised in the following theorem which is essentially an alternative definition of the limit.

<u>Theorem 1.</u> Suppose that $c < a < d$. Further suppose that for each $\varepsilon > 0$ there is a $\delta > 0$ such that

$$|f(x) - L| \leq \varepsilon$$

for every number x in [c, d] which satisfies the inequality $0 < |x - a| \leq \delta$. Then $\lim_{x \to a} f(x) = L$.

To clarify the definition of limit we note that the inequality

$$0 < |x - a|$$

occurring in that definition has the effect of omitting the case $x = a$ from consideration. Hence we remain non-committal concerning the validity of the inequality

$$|f(x) - L| \leq \varepsilon$$

in case x were given the value a. For most familiar functions it turns out that $\lim_{x \to a} f(x)$ is simply equal to $f(a)$. In this case the inequality

$$|f(x) - f(a)| \leq \varepsilon$$

certainly holds true for $x = a$ whatever positive value ε may have. However, in most important applications of $\lim_{x \to a} f(x)$, it turns out (as with $\lim_{t \to 0} \frac{\sin(t)}{t}$) that $f(x)$ is undefined for $x = a$.

In the following section we discuss numerous properties of limits. We present just one theorem here to tie back into the ideas that motivated this presentation.

<u>Theorem 2</u>: If $\lim_{x \to a} f(x) = L$ and $x_1, x_2, \ldots$ is a sequence with $x_n \neq a$ but with $\lim_{n\to\infty} x_n = a$, then $\lim_{n\to\infty} f(x_n) = L$.

In order to prove this theorem we must show that for every positive ε there is a number N such that

$$|f(x_n) - L| \leq \varepsilon \quad \text{whenever } n \geq N.$$

Let $\varepsilon > 0$. Since $\lim_{x\to a} f(x) = L$ there is a positive number δ so that

$$|f(x) - L| \leq \varepsilon \quad \text{whenever} \quad 0 < |x - a| \leq \delta.$$

Since $\lim_{n\to\infty} x_n = a$ we can find N so that $|x_n - a| \leq \delta$ whenever $n \geq N$. Since $x_n \neq a$ it follows also that $0 < |x_n - a|$. Hence for $n \geq N$ we have

$$0 < |x_n - a| \leq \delta \qquad \text{from which } |f(x_n) - L| \leq \varepsilon.$$

390

PROBLEMS

1. Recalling from Chapter 3 the inequality

$$1 - \frac{x^2}{2!} \leq \cos(x) \leq 1 - \frac{x^2}{2!} + \frac{x^4}{4!}$$

(a) Find the value L so that

$$\lim_{x \to 0} \frac{1 - \cos(x)}{x^2} = L$$

(b) Find a formula giving δ in terms of ε so that

$$\left| \frac{1 - \cos(x)}{x} - L \right| \leq \varepsilon \quad \text{whenever } |x| \leq \delta$$

(c) If $|x| \leq \frac{1}{100}$ find a bound on

$$\left| \frac{1 - \cos(x)}{x^2} - L \right|$$

(d) How close to 0 must x be taken in order to guarantee that

$$\left| \frac{1 - \cos(x)}{x^2} - L \right| \leq \frac{1}{6{,}000{,}000} \ ?$$

2. Write a program to compute and print the difference quotient $\frac{f(x) - f(a)}{x - a}$ for a given function f, a given number a, and a given set of values $x_1, x_2, \ldots, x_n$ for x. Have the program read values for a and n, but let the values $x_1, \ldots, x_n$ be given by a function SEQ(J).

3. Suppose that the temperature of a certain furnace is given by the formula $T(t) = 20t^2 + 40t$ (degrees Centigrade) for t between 0 and 10 (minutes). Use the program written in Problem 2 to calculate the difference quotient at $x = 3 + \frac{7}{k^2}$ for $1 \leq k \leq 50$. Estimate the instantaneous rate of change of the temperature at t = 3.

4. A man jumps from a balloon carrying an altimeter. He notices that his altitude is given by $A(t) = 2280 + 2t - 16t^2$ (feet), where t is measured in seconds from the time that he jumps.

   (a) How high is he when he jumps?

   (b) Estimate his initial velocity by using the program written in Problem 2 to calculate several terms of the difference quotient at $t = \frac{1}{n}$.

   (c) Find the time t_1 when he hits the ocean below.

(d) Estimate his velocity at impact by calculating several terms of the difference quotient at $t = t_1 - \frac{1}{n}$.

5. Use the program written in Problem 2 to calculate several terms of the sequence $\frac{\sqrt{x_n} - \sqrt{3}}{x_n - 3}$ for some sequence x_n converging to 3. Try to pick a sequence x_n that no one else in the class is likely to come up with. Do all sequences obtained by the class seem to converge to the same limit?

6. Let P be the postage function of Example 6, Section 3-6. Find two sequences x_n and y_n that converge to 2 such that $\lim_{n\to\infty} P(x_n) \neq \lim_{n\to\infty} P(y_n)$. How does this relate to Theorem 2 of this section?

7. Prove that $\lim_{x\to 0} x \sin(\frac{\pi}{x}) = 0$.

2. Limit Theorems and Continuity

Let us compare the meanings of the two kinds of limits we have considered, the sequential limit and the functional limit.

We see that

$$\lim_{n\to\infty} a_n = L$$

tells us that a_n closely approximates L provided that n is sufficiently large, while

$$\lim_{x\to c} f(x) = L$$

tells us that f(x) closely approximates L provided that x is sufficiently close to c (but not equal to c). Naturally we would suppose that if f(x) closely approximates L and g(x) closely approximates M, then f(x) + g(x) closely approximates L + M and f(x)g(x) closely approximates LM, etc. Consequently, we suppose that the following theorems hold true.

39

Theorem 1. If $\lim_{x \to c} f(x) = L$ and $\lim_{x \to c} g(x) = M$, then

$$\lim_{x \to c} [f(x) + g(x)] = L + M.$$

Theorem 2. If $\lim_{x \to c} f(x) = L$ and k is a constant, then

$$\lim_{x \to c} [kf(x)] = kL.$$

Theorem 3. If $\lim_{x \to c} f(x) = L$ and $\lim_{x \to c} g(x) = M$, then

$$\lim_{x \to c} [f(x)g(x)] = LM.$$

Theorem 4. If $\lim_{x \to c} f(x) = L$ and $\lim_{x \to c} g(x) = M$, then

$$\lim_{x \to c} \frac{f(x)}{g(x)} = \frac{L}{M} \text{ provided } M \neq 0.$$

Not only are these theorems true, but their proofs are mere paraphrasing of the proofs of the corresponding theorems for sequential limits. We therefore omit these proofs. The conscientious reader who would like to refresh his memory of the techniques involved is invited to refer to Chapter 2 and write out the paraphrasing of these proofs.

The only illustrations of functional limits examined so far have been ones in which the function under consideration is undefined at the limiting point. For example, in

$$\lim_{t \to 0} \frac{\sin t}{t}$$

we see that $\frac{\sin t}{t}$ is not defined for $t = 0$. With most ordinary garden variety functions this will not be the case. It will often turn out that $\lim_{x \to c} f(x)$ is just what you would expect it to be, namely, $f(c)$. Such a function is said to be continuous at c.

Definition. f is <u>continuous at c</u> provided that

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

Example 1: To show that $\lim_{x \to 2} x^2 = 4$.

Solution. Let us first confine our attention to the interval $1 \leq x \leq 3$ as is permissible by Theorem 1-1. Now our problem is to show that for any positive number ε we can find a positive number δ so that

$$\text{if } 1 \leq x \leq 3 \text{ and } 0 \neq |x - 2| \leq \delta, \text{ then } |x^2 - 4| \leq \varepsilon.$$

The pattern in such demonstrations is to defer the determination of the value of δ until the desired value becomes apparent. We write

$$\text{if } 1 \leq x \leq 3 \text{ and } 0 \neq |x - 2| \leq \delta$$

$$\text{then } |x^2 - 4| = |x + 2||x - 2| \leq |x + 2|\delta \leq 5\delta.$$

(The last inequality in this string follows from $1 \le x \le 3$ so that $|x + 2| \le 5$.) Now it is apparent how δ should be chosen to obtain the desired inequality, $|x^2 - 4| \le \varepsilon$; namely, we choose $\delta = \frac{\varepsilon}{5}$.

The final write-up of such a demonstration is generally presented in the condensed form: Let $\varepsilon > 0$. Choose $\delta = \frac{\varepsilon}{5}$. Now if $1 \le x \le 3$ and $0 \ne |x - 2| \le \delta$, then $|x^2 - 4| = |x + 2||x - 2| \le |x + 2| \delta \le 5\delta = 5 \cdot \frac{\varepsilon}{5} = \varepsilon$.

This leaves everyone wondering, "How did you have the foresight to choose $\delta = \frac{\varepsilon}{5}$ at that early stage of the proceedings?" The answer of course is that you didn't.

An ε, δ definition of continuity at c is obtained by replacing the L in the limit definition by f(c) obtaining: "f is continuous at c provided that for every $\varepsilon > 0$ there is a $\delta > 0$ so that

$$|f(x) - f(c)| \le \varepsilon \text{ whenever } 0 \ne |x - c| \le \delta."$$

However the restriction "$0 \ne |x - c|$" is not necessary in this case since

$$|f(x) - f(c)|$$

is zero when $x = c$, and thus less than any positive number ε. Hence we can give the definition of continuity in the following (often useful) form.

Definition (Alternative): f is continuous at c provided that for every $\varepsilon > 0$ there is a $\delta > 0$ so that

$$|f(x) - f(c)| \leq \varepsilon \text{ whenever } |x - c| \leq \delta.$$

A number of facts concerning continuity are now entirely trivial.

Theorem 5. If f and g are continuous at c and k is a constant, then:

(i) $f + g$ is continuous at c;

(ii) kf is continuous at c;

(iii) fg is continuous at c; and

(iv) f/g is continuous at c provided that $g(c) \neq 0$.

We give only the proof of (i), as the others all go in standard patterns.

Proof of (i). Let $h(x) = f(x) + g(x)$. By Theorem 1,

$\lim_{x \to c} h(x) = \lim_{x \to c} f(x) + \lim_{x \to c} g(x) = f(c) + g(c) = h(c)$, so that h is continuous at c.

Theorem 5 tells us that once we know some functions continuous at a point, then all the functions constructed from them by the ordinary arithmetic processes are also continuous at this point. A particularly important class of continuous functions results from applying this principle to the identity function and constant functions.

Theorem 6. The functions $f(x) = x$ and $g(x) = k$ (a constant) are continuous at all points.

Proof. (i) [for $f(x)$] Let $\varepsilon > 0$. Take $\delta = \varepsilon$. Now if $|x - c| \leq \delta$ then $|f(x) - f(c)| = |x - c| \leq \delta = \varepsilon$.

(ii) [for $g(x)$] Let $\varepsilon > 0$. Let δ be an arbitrary positive number. Now if $|x - c| \leq \delta$ then $|g(x) - g(c)| = |k - k| = 0 \leq \varepsilon$.

Recall from Chapter 0 (page 44), that all polynomials such as

$$p(x) = 5x^7 - 11x^5 + \frac{3}{4}x^4 - 2x^2 + 6$$

are built up by repeated additions and multiplications of constant functions and the identity function $f(x) = x$. Hence, all polynomials are continuous at all points. When division is also permitted, we also get rational functions such as

$$\frac{3x^4 - x^3 + 7}{x^2 - 3x + 2}.$$

Rational functions are therefore continuous at all points except where the denominators are zero. We see that we now have a large class of functions for which the evaluation of limits is merely a matter of substitution.

We have defined what is meant by a function being continuous at a point. If a function is continuous at every point of its domain, we simply call it continuous. And at times, we shall want to restrict our attention to functions "continuous on an interval," by which is meant continuous at each point of the interval.

Another important class of continuous functions is the class of unicon functions introduced in Chapter 3. We recall the definition: f is unicon on [a,b] provided that for every $\varepsilon > 0$ there is a $\delta > 0$ so that $|f(x_1) - f(x_2)| \leq \varepsilon$ whenever x_1 and x_2 are in [a,b] with $|x_1 - x_2| \leq \delta$.

Theorem 7. If f is unicon on [a,b] and $a < c < b$, then f is continuous at c.

Proof. Let $\varepsilon > 0$. Choose $\delta > 0$ so that $|f(x_1) - f(x_2)| \leq \varepsilon$ whenever x_1 and x_2 are in $[a,b]$ with $|x_1 - x_2| \leq \delta$. Now we see that if x is in $[a,b]$ with $0 \neq |x - c| \leq \delta$ then $|f(x) - f(c)| \leq \varepsilon$. Thus (by Theorem 1, Section 1) $\lim_{x \to c} f(x) = f(c)$.

To motivate the next very useful theorem we return to the subject of speed.

Example 2: Suppose that an object travels in a straight line in such a way that the distance travelled (in feet) t seconds after the start is given by

$$d(t) = 16t^2.$$

What is the velocity 3 seconds after the start?

Solution. The average velocity between time 3 and time t is the distance travelled divided by the time or

$$\frac{d(t) - d(3)}{t - 3} = \frac{16t^2 - 144}{t - 3}.$$

The instantaneous velocity at $t = 3$ is the limit of this average velocity as $t \to 3$ or

$$v(3) = \lim_{t \to 3} \frac{16t^2 - 144}{t - 3}.$$

The function $F(t) = \frac{16t^2 - 144}{t - 3}$ is a rational function, but the limit cannot be evaluated by substituting 3 for t since the denominator would then be zero. However for $t \neq 3$,

$$\frac{16t^2 - 144}{t - 3} = \frac{16(t^2 - 9)}{t - 3} = \frac{16(t - 3)(t + 3)}{t - 3} = 16(t + 3).$$

We might have some momentary qualms about concluding that

$$(1) \qquad \lim_{t \to 3} \frac{16t^2 - 144}{t - 3} = \lim_{t \to 3} 16(t + 3)$$

since the function on the right is defined for $t = 3$ while the one on the left is not. However, we recall that in the definition of $\lim_{x \to c} f(x)$, the stipulation $0 \neq |x - c|$ eliminates from consideration what happens to $f(x)$ when $x = c$. Thus, $\lim_{x \to c} f(x)$ depends on the values of $f(x)$ for other values of x but not for $x = c$. These remarks lead us to conclude that the statement in formula (1) is valid, and now the limit on the right can be evaluated by substitution since the function $g(x) = 16(t + 3)$ is continuous at 3. Thus the limit is $g(3) = 16(3 + 3) = 96$, so that the speed is 96 ft/sec.

The technique employed in this example may be somewhat exasperating. We first admonished you about the impossibility of substituting 3 for t in the expression $\frac{16t^2 - 144}{t - 3}$; then we made a simplification valid only for t different from 3; and finally, we substituted 3 for t in the result. We have some sympathy for the student who said, "In high school, they tell you that you can't divide by zero. In college they show you how to do it." We hope that the following theorem will dispel any confusion about the above technique.

Theorem 8. If $f(x) = g(x)$ for $x \neq c$ and g is continuous at c, then

$$\lim_{x \to c} f(x) = g(c).$$

Proof. Let $\varepsilon > 0$. Choose $\delta > 0$ so that $|g(x) - g(c)| \leq \varepsilon$ whenever $0 \neq |x - c| \leq \delta$. Now if $0 \neq |x - c| \leq \delta$ then $|f(x) - g(c)| = |g(x) - g(c)| \leq \varepsilon$.

The crucial step in this proof was the ability to substitute $g(x)$ for $f(x)$ under the condition that $0 \neq |x - c|$.

The following three theorems are stated without proof since these proofs are entirely analogous to those of the corresponding theorems for sequences.

Theorem 9. If for some number $\delta > 0$ we have $f(x) \geq K$ whenever $0 \neq |x - a| \leq \delta$, then $\lim_{x \to a} f(x) \geq K$ provided that this limit exists.

Theorem 10. (Squeeze Theorem) If for some number $\delta > 0$, $f(x)$ lies between $g(x)$ and $h(x)$ whenever $0 \neq |x - a| \leq \delta$ and $\lim_{x \to a} g(x) = L = \lim_{x \to a} h(x)$, then $\lim_{x \to a} f(x) = L$.

Theorem 11. (Convex Combination Theorem) If for some $\delta > 0$, $f(x)$ is a convex combination (or weighted average) of $g(x)$ and $h(x)$ whenever $0 \neq |x - a| \leq \delta$, and if $\lim_{x \to a} g(x) = L = \lim_{x \to a} h(x)$, then $\lim_{x \to a} f(x) = L$. [Recall that $f(x)$ is a convex combination of $g(x)$ and $h(x)$ means that $f(x) = r(x)g(x) + s(x)h(x)$ where $r(x)$ and $s(x)$ are non-negative with $r(x) + s(x) = 1$.]

PROBLEMS

1. Find each of the following limits if they exist.

(a) $\lim_{x \to 0} x \csc x$

(b) $\lim_{x \to 2} \frac{x - 3}{x - 4}$

(c) $\lim_{x \to 2} \frac{x^2 - 7x + 12}{x - 4}$

(d) $\lim_{x \to 4} \frac{x^2 - 7x + 12}{x - 4}$

(e) $\lim_{x \to 4} \frac{x - 3}{x - 4}$

(f) $\lim_{x \to \frac{\pi}{4}} \tan x$

(g) $\lim_{x \to 0} \frac{4x}{\sin x}$

(h) $\lim_{x \to 0} \sin \left(\frac{1}{x}\right)$

(i) $\lim_{x \to 2} [(x - 2) \sin (x^5 - 4x^2 + \frac{3}{x})]$

(j) $\lim_{x \to 3} \frac{\sqrt{x} - \sqrt{3}}{x - 3}$ (Hint: For $x \neq 3$,

$$\frac{\sqrt{x} - \sqrt{3}}{x - 3} = \frac{(\sqrt{x} - \sqrt{3})(\sqrt{x} + \sqrt{3})}{(x - 3)(\sqrt{x} + \sqrt{3})}$$

$$= \frac{x - 3}{(x - 3)(\sqrt{x} + \sqrt{3})}$$

$$= \frac{1}{\sqrt{x} + \sqrt{3}}$$

. Use Theorem 8.

(k) $\lim_{x \to 5} \frac{\sqrt{x} - \sqrt{5}}{x - 5}$

(l) $\lim_{x \to 8} \frac{\sqrt{x + 1} - 3}{x - 8}$

(m) $\lim_{x \to -1} \frac{x + 1}{x^3 + 1}$

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

(o) $\lim_{x \to 2} \frac{x - tx - 2 + 2t}{x - 2}$

2. To which parts of Problem 1 did Theorem 8 apply?

3. For each function f and number a, find $\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$.

(a) $f(x) = x$. Do for each of the following values of a: $2, 4, -3, 0, \pi, t$.

(b) $f(x) = x^2$. $a = 3, 17, t, \pi^2 + 1$.

(c) $f(x) = x^3$. $a = 4, -1, t$.

(d) $f(x) = |x|$. $a = 2, -2, t, 0$.

(e) $f(x) = \sqrt{x}$. $a = 3, 5$.

(f) $f(x) = \cos x$. $a = \frac{\pi}{4}, \frac{\pi}{3}, t$.

4. For several of the functions f and numbers a of Problem 3, find $\lim_{h \to 0} \frac{f(a + h) - f(a)}{h}$.

5. (a) Show for $2 \le x \le 4$ that $19 \le x^2 + 3x + 9 \le 37$, and thus $|x^2 + 3x + 9| \le 37$.

(b) Show by definition, as in Example 1, that

$$\lim_{x \to 3} x^3 = 27.$$

(c) Show again that $\lim_{x \to 3} x^3 = 27$, this time using Theorem 7.

6. Determine where each of the following functions is not continuous.

(a) $f(x) = \frac{x^2}{x^2 - 4}$ (c) $f(x) = \frac{x}{\sin x}$

(b) $f(x) = \tan x$

7. Give a $\delta - \varepsilon$ proof that $f(x) = |x|$ is continuous.

8. Prove that $f(x) = \sqrt{x}$ is continuous at a for every $a > 0$.

9. Be a "conscientious reader." Write out the proofs of Theorems 1, 2, 3, and 4.

10. Prove parts ii, iii, and iv of Theorem 5.

3. Composition and Intermediate Value Thecrems.

What can we say about such limits as

$$\lim_{x\to a} f(g(x))?$$

Suppose that $\lim_{x\to a} g(x) = b$ and that $\lim_{y\to b} f(y) = c$. Reasoning intuitively we might say: "If x is close to a, then g(x) is close to b so that f(g(x)) ought to be close to c. Ergo $\lim_{x\to a} f(g(x)) = c$."

Alas, this reasoning is not quite valid. But what could be wrong with it? The answer is that the following unpleasantness could take place:

(i) for some values of x close to (but different from) a we have g(x) = b; and

(ii) f(b) is undefined or different from the limit, c.

If <u>both</u> (i) and (ii) occur then we will have values of x near a for which f(g(x)) = f(b) is not close to c. Since <u>both</u> (i) and (ii) must occur to invalidate the conclusion that $\lim_{x \to a} f(g(x)) = c$, we will have this conclusion provided that <u>either</u> of the possibilities (i) or (ii) can be ruled out.

The possibility of (ii) occurring is ruled out if f is continuous at b. The possibility of (i) occurring is ruled out if for some positive number λ, we have $g(x) \neq b$ for any x with $0 \neq |x - a| \leq \lambda$. (In this case we say that <u>g excludes the value b in some deleted neighborhood of a</u>.)

Thus we have the following theorem.

<u>Theorem 1</u> (Composition Theorem for Limits)

If $\lim_{x \to a} g(x) = b$ and $\lim_{y \to b} f(y) = c$ then $\lim_{x \to a} f(g(x)) = c$ provided that <u>either</u> of the following additional hypotheses holds:

(i) g excludes the value b in some deleted neighborhood of a; or,

(ii) f is continuous at b.

There are several guises in which this theorem will be encountered frequently. Note first that the conclusion of the theorem could have been written in the form

$$\lim_{y \to b} f(y) = \lim_{x \to a} f(g(x)).$$

Expressed in this way we can think of the theorem as a "change of variable" theorem. Thus we can replace the variable y by g(x) provided we also replace $y \to b$ by $x \to a$. Of course, the hypotheses of the theorem must be satisfied.

In the case that f is continuous at b, the conclusion of the theorem could be expressed in the form

$$\lim_{x \to a} f(g(x)) = f(\lim_{x \to a} g(x)).$$

Expressed in this way we can think of the theorem as a "commutativity" or "change of order" theorem, i.e., the order of taking the limit and applying the function f can be interchanged.

In the case that f is continuous at b <u>and</u> g is continuous at a, the theorem tells us that

$$\lim_{x \to a} f(g(x)) = f(\lim_{x \to a} g(x)) = f(b) = f(g(a))$$

which tells us that the composite function f(g(x)) is continuous at a. This fact we reformulate as follows.

<u>Corollary</u>: The composition of continuous functions is continuous.

The final theorem of this section applies exclusively to continuous functions. We tend to think of continuous functions as having graphs which are unbroken curves. This being the case, if we have a function, f, continuous on an interval [a,b] and if K is a number between f(a) and f(b) then there should be a number c in the interval [a,b] for which f(c) = K. That is to say there should be a point (or perhaps several points as in Figure 3-1) where the graph of f crosses the horizontal line y = K.

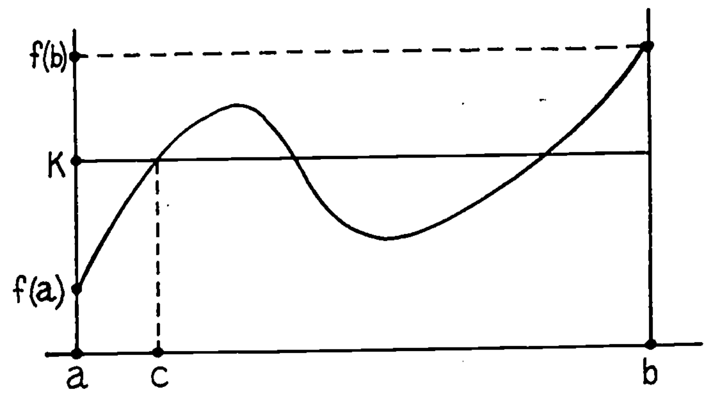


Figure 3-1

As a matter of fact, back in Section 2-2 we presented an algorithm for finding such points to any desired degree of accuracy. But we were cheating slightly back there as we made no mention of hypothesis of continuity necessary to make the conclusion valid. We were depending on the natural tendency of students to think of graphs of functions as being unbroken curves. Now the time has come to nail down this theorem and give a correct proof.

For simplicity lets just consider the case that f(a) < K and f(b) > K. We go through the old bisection algorithm for finding a root of $f(x) = K$ which is flow-charted in Figure 3-2.

In this algorithm, unless we hit a root exactly and terminate in box 9, we produce two infinite sequences $L_1, L_2, L_3, \ldots$ and $R_1, R_2, R_3, \ldots$ respectively increasing and decreasing and such that

$$R_n - L_n = \frac{b-a}{2^n}$$

so that $R_n - L_n \to 0$ as $n \to \infty$. By our completeness axiom both sequences converge to a common limit which we call c.

From the continuity of f at c we know

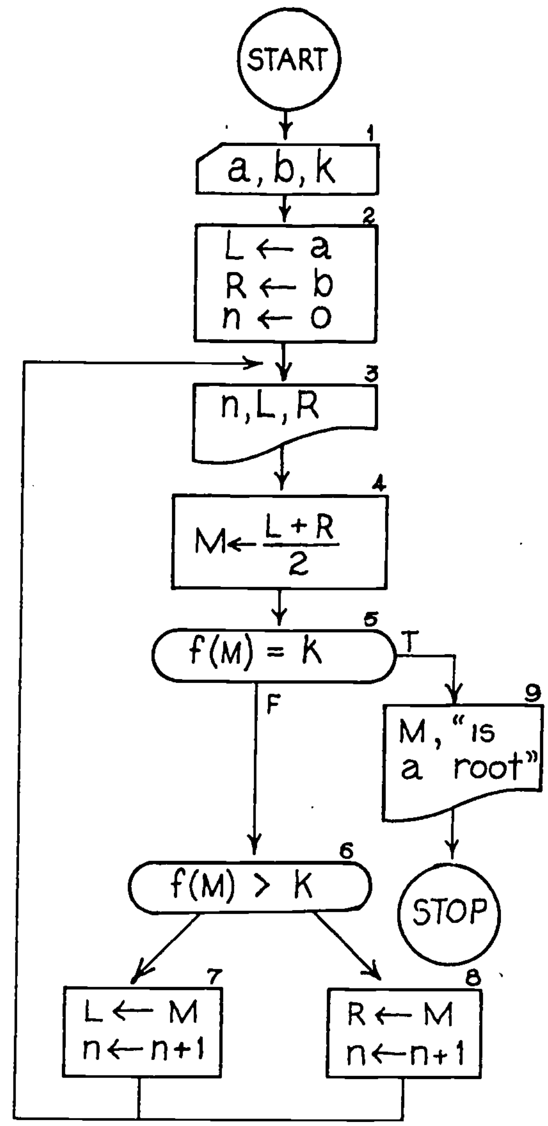


Figure 3-2

that $\lim_{n \to c} f(x) = f(c)$ and since $L_n \to c$ as $n \to \infty$ and $R_n \to c$ as $n \to \infty$ we note by Theorem 2 of Section 5-1 that

$$\lim_{n \to \infty} f(L_n) = f(c) \quad \text{and} \quad \lim_{n \to \infty} f(R_n) = f(c).$$

But now we also note that

$$f(L_n) \leq K \quad \text{and} \quad f(R_n) \geq K$$

so that

$$\lim_{n \to \infty} f(L_n) \leq K \quad \text{and} \quad \lim_{n \to \infty} f(R_n) \geq K$$

(This obvious conclusion follows from Problem 5(a) of Section 2-4.) Both of these limits have been shown to be equal to $f(c)$, hence we may write

$$f(c) \leq K \quad \text{and} \quad f(c) \geq K$$

so that

$$f(c) = K.$$

Thus we have proved the desired result which we formulate below.

Theorem 2. (Intermediate Value Theorem)

If f is continuous on the interval $[a,b]$, and if K is a number between $f(a)$ and $f(b)$, then there is a number c in $[a,b]$ for which $f(c) = K$.

This theorem is often verbalized as "a continuous function takes on all values between any two values it assumes." It is an extremely useful theorem. A common special case is that in which $K = 0$. We state this case as a corollary.

Corollary. If f is continuous in $[a,b]$ and if $f(a)$ and $f(b)$ have opposite signs then the equation $f(x) = 0$ has at least one root in $[a,b]$.

This removes the intuitive aspect from Section 2-2 where we spoke of "unbroken curves" which we can now interpret to mean "graph of a continuous function."

PROBLEMS

1. Evaluate $\lim_{x \to a} f(g(x))$, where

(a) $a = -3$, $f(y) = y + 1$ and $g(x) = x^2 + 1$

(b) $a = 2$, $f(y) = y^{3/2}$ and $g(x) = -x^3 + 3x^2$

(c) $a = 4$, $f(y) = y^2$ and $g(x) = \frac{2(x^2 - 7x + 12)}{(x - 4)}$

(d) $a = 0$, $f(y) = \sqrt{y + 8}$ and $g(x) = \frac{\sin x}{x}$

(e) $a = 2\pi$, $f(y) = \tan y$ and $g(x) = \sin \frac{x}{2}$

(f) $a = 2$, $f(y) = \sqrt{y}$ and $g(x) = x^2 + x + 1$

(g) $a = 1$, $f(y) = \frac{1}{y}$ and $g(x) = x^2 + 3$.

2. State the values for which the given function is discontinuous.

(a) $f(x) = \frac{4}{x^2}$

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

(c) $f(x) = \frac{1}{x^3}$

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

3. Let $f(y) = \frac{\sin y}{y}$, and let $g(x) = x \sin \frac{\pi}{x}$. Recall from Section 1 that $\lim_{y \to 0} f(y) = 1$ and from Problem 7 of Section 1 that $\lim_{x \to 0} x \sin \frac{1}{x} = 0$. Is it true that $\lim_{x \to 0} f(g(x)) = \lim_{y \to 0} f(y)$? Let $a_n = \frac{1}{n}$, then

(a) find $g(a_n)$

(b) find $f(g(a_n))$

(c) use your results from above to conclude that

$\lim_{x \to 0} f(g(x))$ does not exist.

4. Derivatives

The limit

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a},$$

if it exists, is called the derivative of f at a, written $f'(a)$.

We have already had some experience with derivatives. In the first section of this chapter we considered a moving object where position at time t was $s(t)$. Then the instantaneous velocity at time t_0 was given by

$$v(t_0) = \lim_{t \to t_0} \frac{s(t) - s(t_0)}{t - t_0}.$$

According to the above definition, the instantaneous velocity at time t_0 is just the derivative $s'(t_0)$.

The derivative also has an important geometrical representation connected with tangent lines. From your high school geometry you will recall that a line tangent

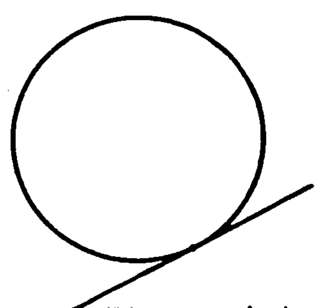


Figure 4-1

to a circle touches the circle at one point, and that the circle lies entirely on one side of the line. We will take this criterion for tangency as a starting point though we will presently come up with a more general notion of tangency.

Let us try to apply this criterion to find the equation of the line tangent to the parabolic graph of the function $f(x) = x^2$ at the point $P(1,1)$ as depicted in Figure 4-2. Here we see that we know a point (1,1) on the desired line so that it is only necessary to find the slope, m, in order to write the equation

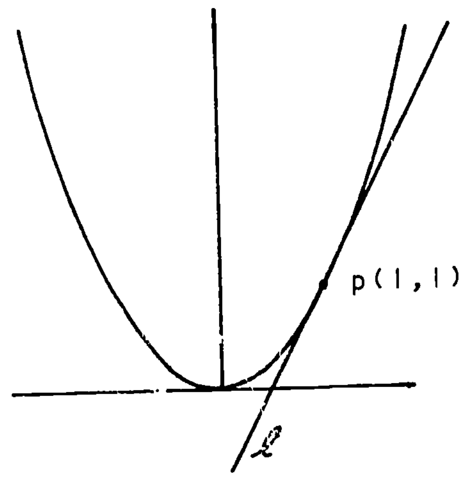


Figure 4-2

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

Consider for the moment a line, ℓ, passing through the point (1,1) not tangent to the curve but intersecting it in a point $Q(x,x^2)$ as illustrated in Figure 4-3(a). The slope of this line is

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

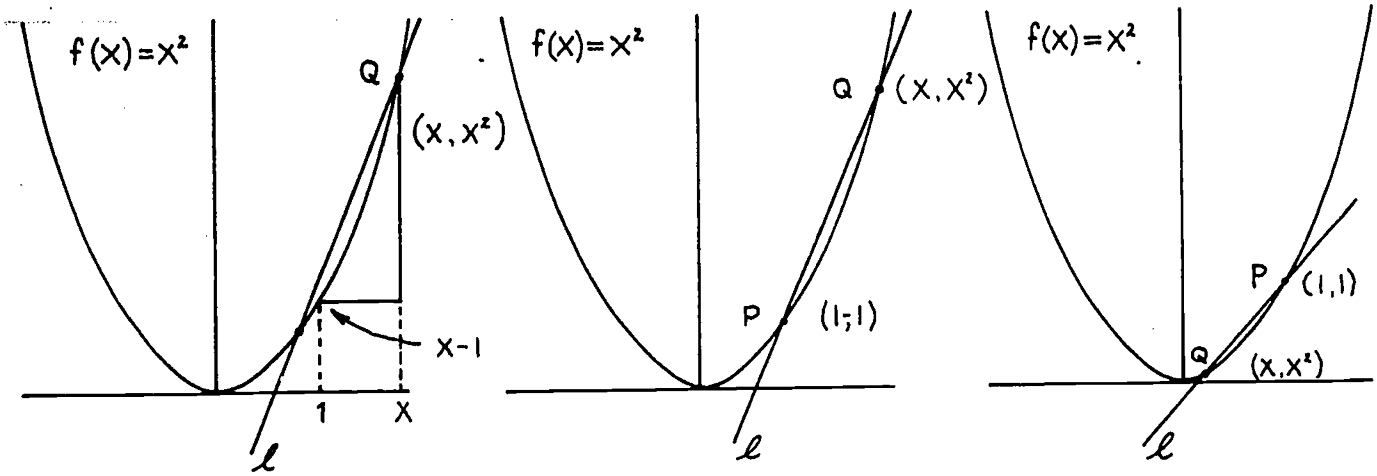


Figure 4-3

Now consider that the line ℓ rotates clockwise about the point (1,1). The point Q moves to the left as seen in Figure 4-3(b). After a sufficient rotation, the point Q will appear to the left of P as seen in Figure 4-3(c). Somewhere in between, the point Q must have coincided with P. In that position, the line ℓ was tangent to the graph as indicated in Figure 4-2.

But what was the slope of the line when this position occurred? We cannot obtain this slope by substituting 1 for x in the expression

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

for then we would have our old nemesis $\frac{0}{0}$. The answer is found by taking the limit as x tends to 1,

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

This limit is easily evaluated by techniques discussed in previous sections,

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

Finally, then, the tangent line to the graph of $f(x) = x^2$ at the point (1,1) is the line through (1,1) with slope 2. The equation is

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

In general, we define the slope of the tangent line to the graph of a function f at a to be $f'(a)$. As seen again in Figure 4-4, this slope $f'(a)$ is the limiting value as $x \to a$ of the slope $\frac{f(x) - f(a)}{x - a}$ of a secant line intersecting the graph of f at (a, f(a)) and (x, f(x)).

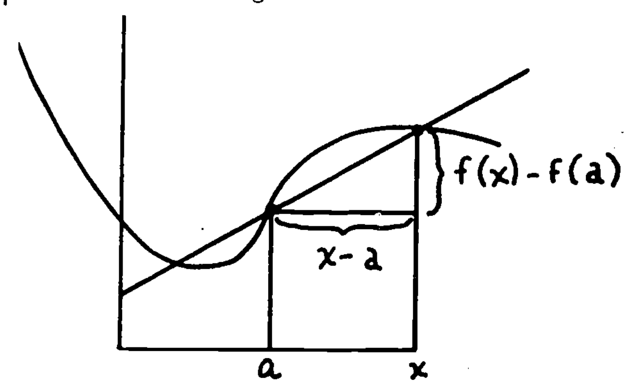


Figure 4-4

This definition is more general than the criterion for tangency discussed earlier in the section. It often happens that the curve does not lie entirely on one side of the tangent line.

Example 1. Find the line tangent to the graph of $y = x^3$ at (2,8) and find another point where this line intersects the graph.

Solution: Let $f(x) = x^3$. The slope of the tangent line at (2,8) is

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

$$= \lim_{x\to 2} \frac{x^3 - 8}{x - 2}$$

$$= \lim_{x\to 2} \frac{(x - 2)(x^2 + 2x + 4)}{x - 2}$$

$$= \lim_{x\to 2} (x^2 + 2x + 4) = 12 .$$

The equation of the tangent line is therefore

$$y - 8 = 12(x - 2) \qquad \text{or} \qquad y = 12x - 16.$$

The points of intersection of this line with the graph of $y = x^3$ are found by solving

$$x^3 = 12x - 16 \qquad \text{or} \qquad x^3 - 12x + 16 = 0 .$$

Aided by the knowledge that $x = 2$ (the abscissa of the point of tangency) must be a solution, we obtain the factorization

$$x^3 - 12x + 16 = (x - 2)^2(x + 4).$$

The other point of contact is therefore (-4,-64).

In the next example we see the "rationalizing" trick. (In high school mathematics you often had to rationalize denominators. In calculus it is usually the numerators which need rationalizing.) We also need the continuity of the square root function in this example. You will recall that in Chapter 3 it was shown that the square root function is unicon and it is therefore continuous.

Example 2. Let $f(x) = \sqrt{x}$ and find $f'(2)$.

Solution: $f'(2) = \lim_{x\to 2} \frac{\sqrt{x} - \sqrt{2}}{x - 2}$

$$= \lim_{x\to 2} \frac{\sqrt{x} - \sqrt{2}}{x - 2} \cdot \frac{\sqrt{x} + \sqrt{2}}{\sqrt{x} + \sqrt{2}}$$

$$= \lim_{x\to 2} \frac{x - 2}{(x - 2)(\sqrt{x} + \sqrt{2})}$$

$$= \lim_{x\to 2} \frac{1}{\sqrt{x} + \sqrt{2}} = \frac{1}{\sqrt{2} + \sqrt{2}} = \frac{1}{2\sqrt{2}} .$$

Another notation used in connection with derivatives is

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

This notation is justified by use of the composition theorem for limits. Letting $F(x) = \frac{f(x) - f(a)}{x - a}$ and $g(h) = a + h$ we see that $\lim_{x\to a} F(x) = f'(a)$ and that $\lim_{h\to 0} g(h) = a$. Moreover g excludes the value a in a deleted neighborhood of 0 so that by the composition theorem

$$\lim_{x\to a} F(x) = \lim_{h\to 0} F(g(h)) = \lim_{h\to 0} \frac{f(g(h)) - f(a)}{g(h) - a}$$

$$= \lim_{h\to 0} \frac{f(a + h) - f(a)}{a + h - a} = \lim_{h\to 0} \frac{f(a + h) - f(a)}{h}$$

The validity of this "change of variable" can also be seen geometrically in Figure 4-5 where we again see that

$$\lim_{h\to 0} \frac{f(a + h) - f(a)}{h}$$

gives the slope of the tangent line to the graph of f at (a,f(a)). Next we see an example of the use of this notation.

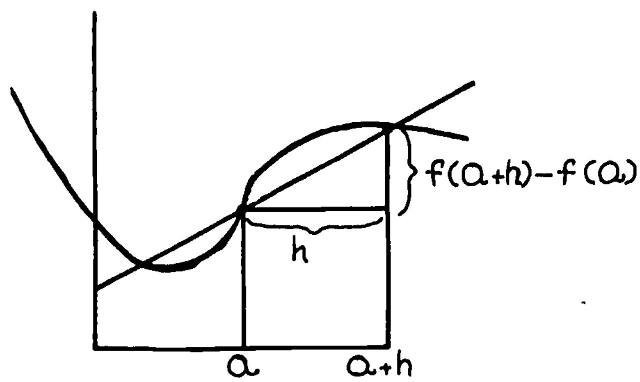


Figure 4-5

<u>Example 3</u>. Find $f'(3)$ where $f(x) = x^4$.

<u>Solution</u>. $f'(3) = \lim_{h\to 0} \frac{f(3 + h) - f(3)}{h}$

$$= \lim_{h\to 0} \frac{(3 + h)^4 - 3^4}{h}$$

$$= \lim_{h\to 0} \frac{3^4 + 4\cdot 3^3 h + 6\cdot 3^2 h^2 + 4\cdot 3h^3 + h^4 - 3^4}{h}$$

$$= \lim_{h\to 0} \frac{4\cdot 3^3 h + 6\cdot 3^2 h^2 + 4\cdot 3h^3 + h^4}{h}$$

$$= \lim_{h\to 0} (4\cdot 3^3 + 6\cdot 3^2\cdot h + 4\cdot 3h^2 + h^3)$$

$$= 4\cdot 3^3 = 108.$$

PROBLEMS

1. For each function f and number a, find $f'(a)$.

(a) $f(x) = x^2$. $a = 2, -1, 3, \pi$, and 5.

(b) $f(x) = \frac{1}{x+1}$. $a = 1, 0, 2, t$.

2. For $f(x) = \frac{1}{x+1}$, find $f'(5)$ by substituting 5 for t in the expression obtained for $f'(t)$ in Problem 1(b).

3. (a) Find the line tangent to the graph of $y = x^3$ at the point (0,0).

(b) Does that tangent line intersect the curve at any other point?

(c) Graph the function $y = x^3$ and the tangent line at (0,0). Note that the tangent line does not lie entirely on one side of the curve.

4. Find the slope of the tangent line to the graph of $y = x^3 + c$ at the point (0,c); at the point $(c, c^3 + c)$.

5. For each function f and point (a,b), not on the graph of $y = f(x)$, find the tangent line(s) to the graph of

y = f(x) passing through the point (a,b) not on the graph of y = f(x). Hint: Find the equation of the tangent line to y = f(x) at the point (c, f(c)); then determine for what values of c the line will pass through the point (a,b).

(a) $f(x) = 3x^2$; (2,0)

(b) $f(x) = x^2 - 2x + 1$; (8,1)

(c) $f(x) = x^2$; (2,5)

5. The Derived Function

One of the advantages of the "h" notation introduced at the end of the preceding section is that it enables us to talk about $f'(x)$. Just as we write

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

so we may write

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

In this way, we can think of a derivative function (or derived function) f', as is certainly suggested by the notation $f'(x)$. In most simple applications of derivatives, the function f is expressed by some simple formula, and it is as easy to calculate a general formula for $f'(x)$ as to calculate $f'(2)$ or $f'(3)$ or $f'(1)$. We then have the advantage that particular derivative values can be found by mere substitution. The derived function f' then has the property that for each number x, the value of $f'(x)$ is the slope of the tangent line to the graph of f at the point $(x, f(x))$.

Example 1. Find the point where the lines tangent to the graph of $y = x^4$ at $(1,1)$ and at $(2,16)$ intersect.

Solution: Let $f(x) = x^4$ and compare the following with Example 3 of Section 5-4.

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{(x+h)^4 - x^4}{h}$$

$$= \lim_{h\to 0} \frac{x^4 + 4x^3h + 6x^2h^2 + 4xh^3 + h^4 - x^4}{h}$$

$$= \lim_{h\to 0} \frac{4x^3h + 6x^2h^2 + 4xh^3 + h^4}{h}$$

$$= \lim_{h\to 0} (4x^3 + 6x^2h + 4xh^2 + h^3)$$

$$= 4x^3.$$

The slope of the tangent lines to the graph of f at (1,1) and (2,16) are respectively $f'(1) = 4\cdot 1^3 = 4$ and $f'(2) = 4\cdot 2^3 = 32$. The equations of these tangent lines are

$$y - 1 = 4(x - 1) \quad \text{and} \quad y - 16 = 32(x - 2).$$

Solving simultaneously, we have

$$4x - 3 = 32x - 48$$

or

$$x = \frac{45}{28}, \; y = \frac{24}{7}.$$

It is important in using the notation

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

that we think of x as a fixed number whose value is unspecified or perhaps just unavailable to us. The results we get are valid regardless of the value of x or, in other words, these results are valid for all x for which the derivative exists. Thus, the domain of f' is taken to be the set of all numbers x for which f'(x) exists. For example, the domain of $f(x) = \sqrt{x}$ is $[0,\infty)$ while the domain of its derived function $f'(x) = \frac{1}{2\sqrt{x}}$ is $(0,\infty)$.

The derivative of the sine function has in actuality been marked out in previous sections. We repeat it here, however, for good measure.

Example 2. Given f(x) = sin x; calculate f'(x).

Solution:

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{\sin(x+h) - \sin x}{h} \\ &= \lim_{h\to 0} \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \\ &= \lim_{h\to 0} \cos x \cdot \frac{\sin h}{h} - \sin x \cdot \frac{1 - \cos h}{h} \\ &= \cos x \cdot \lim_{h\to 0} \frac{\sin h}{h} - \sin x \cdot \lim_{h\to 0} \frac{1-\cos h}{h} \\ &= (\cos x)\cdot 1 - (\sin x)\cdot 0 \\ &= \cos x \end{aligned}$$

The derivative of the cosine function can be calculated by use of the same techniques. This is left as an exercise.

A useful concept in physics is that of acceleration which is defined as the instantaneous rate of change of velocity. Acceleration is what you feel when you "step on the gas" or apply the brakes. If an object is moving along a straight line with velocity $v(t)$ at time t then the acceleration $a(t)$ is defined by

$$a(t) = \lim_{h \to 0} \frac{v(t+h) - v(t)}{h} = v'(t)$$

We already know that

$$v(t) = s'(t)$$

where $s(t)$ is the position of the object at time t. Thus the acceleration function is the derivative of the derivative of the position function s. Accordingly we say that $a(t)$ is the second derivative of $s(t)$ and write

$$a(t) = s''(t)$$

where the double prime denotes the second derivative.

<u>Example 3</u>: An object moves along a line so that its distance $s(t)$ from the starting point is given by

$$s(t) = 5t^3$$

Find the acceleration when $t = 3$ and when $t = 4$.

Solution: $s'(t) = \lim_{h \to 0} \frac{5(t + h)^3 - 5t^3}{h}$

$$= 5 \lim_{h \to 0} \frac{(t^3 + 3t^2h + 3th^2 + h^3) - t^3}{h}$$

$$= 5 \lim_{h \to 0} (3t^2 + 3th + h^2)$$

$$= 15t^2$$

And now

$$s''(t) = \lim_{h \to 0} \frac{15(t + h)^2 - 15t^2}{h}$$

$$= 15 \lim_{h \to 0} \frac{t^2 + 2th + h^2 - t^2}{h}$$

$$= 15 \lim_{h \to 0} (2t + h)$$

$$= 30t.$$

Therefore $s''(3) = 90$ and $s''(4) = 120$.

Second derivatives have geometrical as well as physical significance as will be seen a little later on. It is also clear that we may define the third derivative as the derivative of the second derivative and so on. Applications of these higher derivatives will be seen much later.

PROBLEMS

1. For each function f, find the derived function f'.

   (a) $f(x) = x$

   (b) $f(x) = x^2$

   (c) $f(x) = x^3$

   (d) $f(x) = \cos x$

   (e) $f(x) = \sqrt{x}$

2. Find the point where the lines tangent to the graph of $y = x^2 + 2x + 1$ at $(1,4)$ and $(2,9)$ intersect.

3. Show that for every number c, the tangent lines to the graph of $y = x^2$ at (c,c^2) and $(c + 2,\ c^2 + 4c + 4)$ intersect at a point on the graph of $y = x^2 - 1$.

4. Let L and M be tangent lines to the graph of $y = \frac{1}{x}$. Then one triangle is determined by the X-axis, the Y-axis, and the line L, and another triangle is determined by the X-axis, the Y-axis and the line M. Show that these two triangles have the same area.

5. Neglecting air resistance, the height s(t) (in feet) of a freely falling object near the Earth's surface is

approximately $-16t^2 + ct + d$ for some constants c and d, where t is measured in seconds.

(a) What is the "initial" height of the object, that is, the height when $t = 0$?

(b) What is the "initial" velocity of the object?

(c) Find the acceleration function $a(t) = s''(t)$.

(d) Suppose a man throws a hammer upward from a 200 foot tower with (initial) velocity of 20 ft./sec. Find the height function $s(t)$ of the hammer.

6. A stone is thrown upward from the top of a building 128 feet high, with an initial velocity of 64 ft/sec. Find the maximum height of the stone. Hint: When the stone reaches its maximum height, it stops rising and starts falling so its instantaneous velocity is zero.

7. A brick falls from a tower 144 ft. high. How much time does it take to reach the ground and what is the velocity of the brick when it hits the ground?

8. A cannon ball is shot 400 ft/sec. at an angle ϕ relative to the Earth's surface. Thus its height is given by $h(t) = -16t^2 + (400 \sin \phi)t$, and its horizontal distance is $(400 \cos \phi)t$.

(a) How long is the cannon ball in the air?

(b) How far away does it land?

(c) For what value of ϕ does it go the farthest?

6. Derivative Theorems

We have seen the calculations of the derivatives of a number of functions in the preceding section. We will see more a little later. Right now we will present theorems for differentiation of combinations of functions which enable us to write out the derivatives for a great many functions without going through the limit process.

As a preliminary we formulate a theorem connecting derivatives and continuity which is needed in the proof of the product theorem below.

Theorem I. If $f'(a)$ exists, then f is continuous at a.

Proof: We wish to prove that $\lim_{x\to a} f(x) = f(a)$. It will be equivalent to prove that $\lim_{x\to a} [f(x) - f(a)] = 0$. Now

$$\lim_{x\to a} [f(x) - f(a)] = \lim_{x\to a} \frac{f(x) - f(a)}{x - a} (x - a)$$

$$= \lim_{x\to a} \frac{f(x) - f(a)}{x - a} \cdot \lim_{x\to a} (x - a)$$

$$= f'(a) \cdot 0 = 0.$$

The statement of this theorem could be reformulated as: if for some number x, the derivative f'(x) exists, then $\lim_{h \to 0} f(x + h) = f(x)$. This is the form to be used in the proof of the product theorem.

It is brought out in the exercises that the converse of Theorem I, above, does not hold. That is, a function f may be continuous at a number a without the necessity that f'(a) exists.

In order to simplify the statements of these theorems we introduce another notation for the derivative - one which has the advantage of permitting substitutions. We write

$$D_x\ f(x) \text{ to stand for } f'(x).$$

With this notation we can write

$$D_x\ x^2 = 2x$$

instead of: letting $f(x) = x^2$, then $f'(x) = 2x$. Or, in the following theorem we may write

$$D_x\ [f(x) + g(x)] = f'(x) + g'(x)$$

instead of: letting $s(x) = f(x) + g(x)$ then $s'(x) = f'(x) + g'(x)$.

The following theorems are existence theorems as well as giving the values of the derivatives. They tell us that the derivatives on the left exist for all values of x for which all the derivatives on the right exist. The proofs are by now quite routine.

<u>Theorem 2</u>. $D_x\,[f(x) + g(x)] = f'(x) + g'(x)$.

<u>Theorem 3</u>. $D_x\,[k\cdot f(x)] = k\cdot f'(x)$.

<u>Theorem 4</u>. $D_x\,[f(x)\cdot g(x)] = f'(x)\cdot g(x) + f(x)\cdot g'(x)$.

<u>Theorem 5</u>. $D_x\,[1/g(x)] = -\dfrac{g'(x)}{[g(x)]^2}$.

<u>Theorem 6</u>. $D_x\,[f(x)/g(x)] = \dfrac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}$.

<u>Theorem 7</u>. $D_x[k] = 0$.

We give the proofs of these theorems omitting those of the trivial theorems 3 and 7.

<u>Proof of Sum Theorem 2</u>: Let $s(x) = f(x) + g(x)$.

$$\text{Then}\quad s'(x) = \lim_{h\to 0} \frac{s(x+h) - s(x)}{h}$$

$$= \lim_{h\to 0} \frac{[f(x+h) + g(x+h)] - [f(x) + g(x)]}{h}$$

$$= \lim_{h \to 0} \left(\frac{f(x+h) - f(x)}{h} + \frac{g(x+h) - g(x)}{h} \right)$$

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

$$= f'(x) + g'(x).$$

<u>Proof of Product Theorem 4</u>: Let $p(x) = f(x) \cdot g(x)$.

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

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

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

$$= \lim_{h \to 0} \left(\frac{f(x+h) - f(x)}{h} g(x+h) + f(x) \frac{g(x+h) - g(x)}{h} \right)$$

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

$$= f'(x)g(x) + f(x)g'(x).$$

<u>Proof of Reciprocal Theorem 5</u>. Let $r(x) = \frac{1}{g(x)}$.

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

$$= \lim_{h \to 0} \frac{\frac{1}{g(x+h)} - \frac{1}{g(x)}}{h}$$

$$= \lim_{h \to 0} \frac{g(x) - g(x+h)}{h \cdot g(x)g(x+h)}$$

$$= \lim_{h \to 0} - \frac{1}{g(x)} \cdot \frac{g(x+h) - g(x)}{h} \cdot \frac{1}{g(x+h)}$$

$$= - \frac{1}{g(x)} \cdot \lim_{h \to 0} \frac{g(x+h) - g(x)}{h} \cdot \lim_{h \to 0} \frac{1}{g(x+h)}$$

$$= - \frac{1}{g(x)} \cdot g'(x) \cdot \frac{1}{g(x)} = - \frac{g'(x)}{[g(x)]^2} .$$

The proof of the quotient theorem 6 now becomes a corollary of the product and reciprocal theorems and can be proved without direct use of limits.

Proof of Quotient Theorem 6: Using the product and reciprocal theorems we see that

$$\begin{aligned} D_x \frac{f(x)}{g(x)} &= D_x \, f(x)\cdot\frac{1}{g(x)} \\ &= D_x[f(x)]\cdot\frac{1}{g(x)} + f(x)\cdot D_x \frac{1}{g(x)} \\ &= f'(x)\cdot\frac{1}{g(x)} + f(x)\cdot - \frac{g'(x)}{[g(x)]^2} \\ &= \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2} . \end{aligned}$$

As a first application of these theorems, we will show that

$$D_x \, x^n = nx^{n-1}$$

for every positive integer n. This formula has already been demonstrated for n = 1, 2, 3, 4 in the preceding section. We establish the truth of this formula by mathematical induction. That is, we show that it is true for n = 1 (as has already been done) and then show that if the formula holds for some integer value of n, then it holds for the next integer value of n as well. Put differently, we show that if true for n it is also true for n + 1.

To this end suppose that n is an integer for which $D_x\, x^n = nx^{n-1}$. Then by use of the product theorem

$$D_x\, x^{n+1} = D_x\, [x^n \cdot x]$$

$$= D_x x^n \cdot x + x^n \cdot 1 = nx^{n-1} \cdot x + x^n \cdot 1$$

$$= nx^n + x^n = (n + 1)x^n.$$

Armed with this result we can now differentiate polynomials term by term by repeated use of the sum theorem. For example:

$$D_x\, (x^5 - 3x^4 + 7x^2 - 5x + 8)$$

$$= 5x^4 - 3 \cdot 4x^3 + 7 \cdot 2x - 5 \cdot 1 + 0$$

$$= 5x^4 - 12x^3 + 14x - 5$$

We can also see immediately how to differentiate the other basic trigonometric functions.

Example: To find $D_x \tan x$.

Solution. $D_x \tan x = D_x \dfrac{\sin x}{\cos x}$

$$= \frac{D_x[\sin x]\cos x - \sin x\, D_x[\cos x]}{\cos^2 x}$$

$$= \frac{\cos x \cdot \cos x - \sin(x)(-\sin x)}{\cos^2 x}$$

$$= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = \sec^2 x.$$

The next theorem and its two corollaries deal with the behavior of a function in the immediate vicinity of a point where the derivative exists. This theorem can be expressed in geometric terms as follows: Suppose that $f'(a)$ exists and denote by ℓ the line tangent to the graph of f at the point $(a,f(a))$ [See Figure 6-1(a)]. Now let ℓ' and ℓ'' be lines through $(a,f(a))$ on either side of ℓ as in Figure 6-1(b). Then we

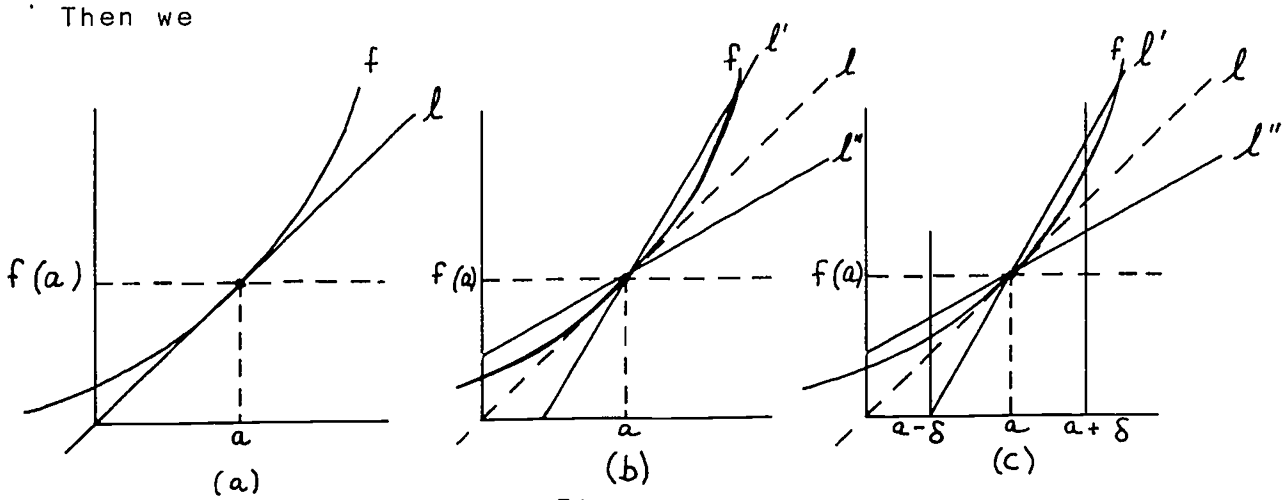


Figure 6-1

can find an interval centered at a so that over this interval the graph of f lies between ℓ' and ℓ''. An analytical formulation of this theorem is:

<u>Theorem 8.</u> If $f'(a)$ exists and if K and L are numbers with $K < f'(a) < L$ then there is a number $\delta > 0$ so that

$$K \leq \frac{f(x) - f(a)}{x - a} \leq L \quad \text{for } 0 \neq |x - a| \leq \delta. \tag{1}$$

<u>Proof:</u> $f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$ by definition. In terms of

ε and δ, this becomes: for every $\varepsilon > 0$ there is a $\delta > 0$ so that

$$f'(a) - \varepsilon \leq \frac{f(x) - f(a)}{x - a} \leq f'(a) + \varepsilon \quad \text{for } 0 \neq |x - a| \leq \delta .$$

Choosing ε so that both $f'(a) + \varepsilon \leq L$ and $f'(a) - \varepsilon \geq K$ we have the desired result.

Corollary 1: If $f'(a)$ exists, then there are positive numbers M and δ so that

$$|f(x) - f(a)| \leq M|x - a| \text{ whenever } |x - a| \leq \delta.$$

[Note that we do not exclude the case that $x = a$.] To see that this is so we take M to be the larger of $|K|$ and $|L|$ in the theorem above. Then we have $-M \leq K$ and $L \leq M$ so that from (1) we have

$$-M \leq \frac{f(x) - f(a)}{x - a} \leq M \quad \text{for } 0 \neq |x - a| \leq \delta$$

which can be rewritten as

$$\left|\frac{f(x) - f(a)}{x - a}\right| \leq M \text{ or } |f(x) - f(a)| \leq M|x - a| \text{ for } 0 \neq |x - a| \leq \delta.$$

We identify the property described in this corollary by saying that <u>f is locally Lipschitzian at a</u>. [We recall from Chapter 2 that a function was said to be Lipschitzian on an interval provided that for some number M we had

$|f(x_1) - f(x_2)| \leq M|x_1 - x_2|$ for all x_1 and x_2 in the interval. In our local version of the property just such an inequality holds in a sufficiently small interval about a provided that x_2 is fixed with the value a.] The geometrical meaning of the corollary is seen in Figure 6-2, where, over the interval $[a - \delta, a + \delta]$, the graph of f is confined to the shaded region. For an example of a function which fails to be locally Lipschitzian at a point, see in Figure 6-3 the graph of $f(x) = \sqrt[3]{x}$ in the vicinity of $x = 0$.

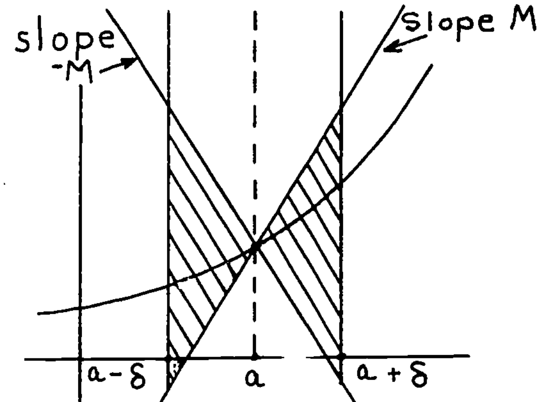


Figure 6-2

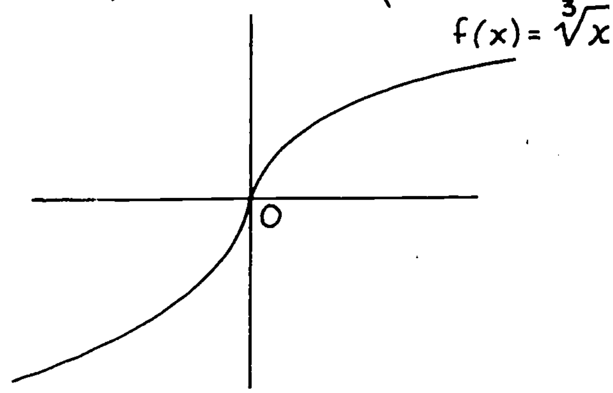


Figure 6-3

Corollary 2: If f'(a) exists and is different from 0, then f excludes the value f(a) in some deleted neighborhood of a.

To see that this is so, refer back to Figure 6-1(c). Here we see that in case $f'(a) > 0$ then ℓ' and ℓ'' can be chosen with positive slopes. Since the graph of f over the interval $[a - \delta, a + \delta]$ lies between ℓ' and ℓ'' then the graph of f cannot intersect the horizontal line $y = f(a)$ in this interval except at $x = a$. In the case that $f'(a) < 0$ we choose ℓ' and ℓ'' both with negative slopes and use a similar argument.

Corollary 3: If $f'(a)$ exists and is different from 0, then in any neighborhood of a, $f(x)$ has values $> f(a)$ and values $< f(a)$.

Using the same argument as for Corollary 2, we see that if $f'(a) > 0$ then for $a < x < a + \delta$ we must have $f(x) > f(a)$, and for $a - \delta < x < a$ we must have $f(x) < f(a)$.

PROBLEMS

1. (a) Find $D_x x|x|$.

   (b) Find $D_x |x|$.

   (c) Is the function $f(x) = |x|$ continuous at $x = 0$? Is it differentiable there?

   (d) What can now be said about the converse of Theorem 1?

2. In each case find $D_x y$. Specify the domain of the derived function whenever it is different from the domain of the given function.

   (a) $y = x^5 + 12x^4 - 3x^3 + 2x - 36$

   (b) $y = 7x^6 + 4x^3 - 2$

   (c) $y = 8$

   (d) $y = \frac{\sqrt{x}}{x + 2}$

   (e) $y = \cot x$

   (f) $y = (x^2 + 4)(x^3 + 2x + 1)$

   (g) $y = x \tan x$

   (h) $y = (x + 3)(x + 4)(x - 5)$

   (i) $y = \frac{x - 2}{x + 4}$

   (j) $y = \sin x \tan x$

   (k) $y = \frac{\cos x}{\tan x}$

(l) $y = \dfrac{\sqrt{x}}{\sqrt{x} - 2}$

(m) $y = \dfrac{x^2 + 4}{x^2 + 1}$

(n) $y = 6(x^3 + 2)$

(o) $y = \cos x - \sin x$

(p) $y = \tan x - x^3$

(q) $y = \dfrac{1}{(x^3 - 4x^2 + 1)}$

(r) $y = \dfrac{x^4 - 16}{x^4 + 16}$

(s) $y = (1 + \frac{2}{x})(2 + \frac{1}{x})$

(t) $y = (3x^2 + 1)^2$

(u) $y = x^2 + \dfrac{1}{x^2}$

(v) $y = \dfrac{3}{x} + \dfrac{7}{x} - x^{-3}$

(w) $y = \dfrac{(x - 2)(x - 3)}{x^3}$

(x) $y = \dfrac{1}{\sqrt{x}} + \dfrac{\sqrt{x}}{9}$

3. For $y = 2x^3 + 21x^2 + 72x + 24$ find all points where the tangent line is horizontal.

4. Find all points where the tangent line to the graph of $y = \sin x$ has slope 1.

5. (a) Recall that f is locally Lipschitzian at a if there are positive numbers M and δ so that

$$|f(x) - f(a)| \le M|x - a| \quad \text{whenever} \quad |x - a| \le \delta.$$

Now show that the function $y = |x|$ is locally Lipschitzian at 0.

(b) True or false: If a function is locally Lipschitzian at $x = a$ then it is differentiable at $x = a$.

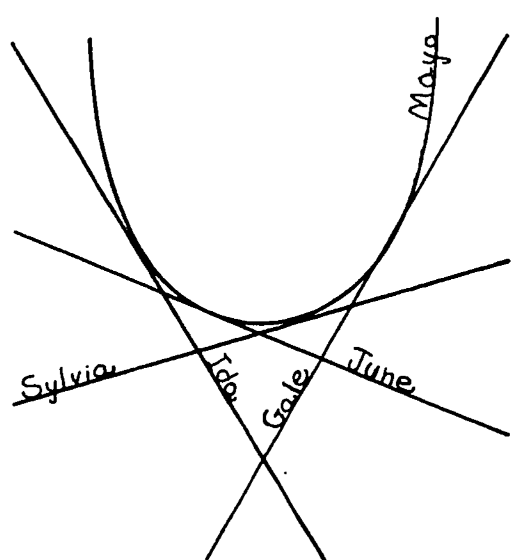
6. Let $f(x) = x^2 - 1$.

(a) Show that $1 < f'(2) < 5$.

(b) Find a number $\delta > 0$ such that $1 \le \dfrac{f(x) - f(2)}{x - 2} \le 5$ whenever $0 \ne |x - 2| \le \delta$.

7. Suppose that f is a function for which $f'(4) = 6$. Show that there is no interval (a,b) containing the point 4 such that f is constant on the interval (a,b).

8. In Finlaysonville, the avenues Ida, June, Sylvia, and Gale are straight avenues tangent to the parabolic street Mayo. Suppose that the formula for Mayo Street is $y = x^2$



and that Ida Avenue and Gale Avenue intersect Mayo Street at $x = -2$ and $x = 2$, respectively. June Avenue and Sylvia Avenue each intersect Mayo Street between the Ida-Mayo and Gale-Mayo intersections. Show that the distance from the Ida-June intersection to the Ida-Gale intersection plus the distance from the Ida-Gale intersection to the June-Gale intersection is equal to the sum of the distances from Ida-Sylvia to Ida-Gale and from Ida-Gale to Sylvia-Gale.

9. For each function f, find the higher derivatives f', f'', f''', and $f^{(4)}$ (the fourth derivative).

(a) $f(x) = x$
(b) $f(x) = x^2$
(c) $f(x) = x^3$
(d) $f(x) = x^4$
(e) $f(x) = x^5$
(f) $f(x) = 4x^5$
(g) $f(x) = \sin x$
(h) $f(x) = \cos x$
(i) $f(x) = 2 \sin x - 3 \cos x$

10. Suppose that a witch is suspended by a spring above a vat of boiling oil and that she bobs in and out of the oil with height $w(t) = 3 \sin t$ (feet).

(a) To what depth does she descend?
(b) What is her fastest velocity?

(c) How long is she under the surface each time?

(d) What is her acceleration at maximum depth?

(e) What is her acceleration at the surface?

(f) How loud does she scream?

11. Use the method of finding $D_x x^n$ to do the following.

(a) Derive a formula for $D_x(ax + b)^n$, where a and b are any constants.

(b) Derive the formula $D_x(x^m + a)^n = mnx^{m-1}(x^m + a)^{n-1}$.

(c) Prove that if f(x) is a differentiable function then $D_x[f(x)]^n = nf(x)^{n-1}f'(x)$.

Chapter 6

MAXIMA. THE MEAN VALUE THEOREM

1. An Example.

Henri Dupré, the great automobile manufacturer, decides to market a new luxury car, the Interstellar, otherwise known as the Supré-Dupré. His engineers tell him that to set up the production line will cost \$175,000,000, after which the cars can be turned out at a cost of \$3877 apiece. His economists predict that the number of sales at any given selling price can be represented approximately by the graph in Figure 1-1. (The break in the graph is due to the exhaustion of the mass market. Only the hard-core wealthy will go above \$6000 and their reluctance to buy increases slowly with increasing price.) What should be the selling price to give M. Dupré the most profit?

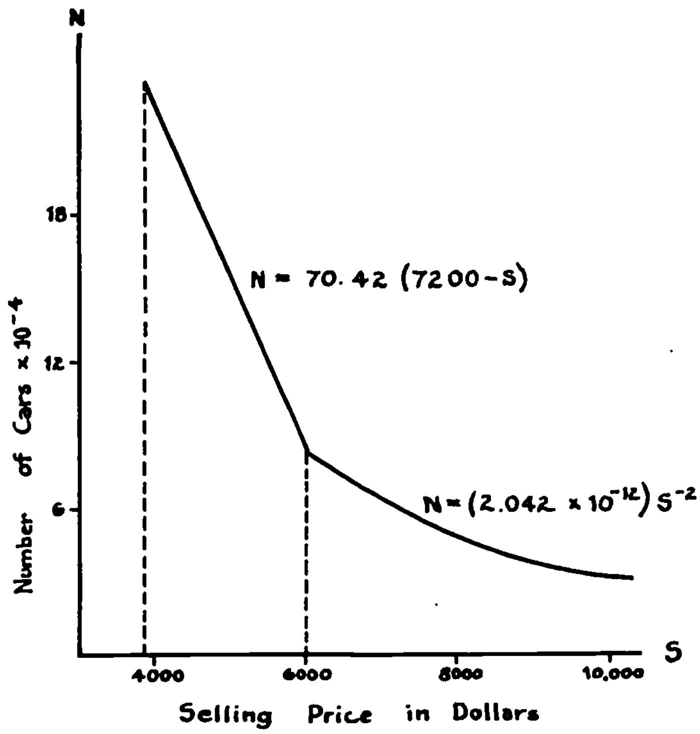


Figure 1-1

If we designate by N the function in Figure 1-1 then at selling price S the profit is

$$P(S) = (S - 3877)N(S) - 175 \times 10^6$$

$$= \begin{cases} 70.42(7200 - S)(S - 3877) - 175 \times 10^6 & \text{if } S \leq 6000, \\ (3.042 \times 10^{12})(S - 3877)S^{-2} - 175 \times 10^6 & \text{if } S \geq 6000. \end{cases}$$

The adjacent table gives the values of P(S) for various values of S. It looks as if P(S) is a maximum near S = 8000, but the function is changing rapidly near S = 5500 and more computation would be needed to make a clear-cu[illegible]se. Actually, the function N(S) as a model of the market is too inaccurate to justify saying more than that the prices $5500 and $8000 are equally good. M. Dupré, scornful of the mass market, unhesitatingly sets the price at $8000.

S	$P(S) \times 10^{-6}$
4000	-152.29
4500	-56.55
5000	-1.02
5250	13.54
5500	19.30
5750	16.25
6000	4.40
6500	13.86
7000	18.88
7500	20.93
8000	20.97
8500	19.65
9000	17.40
9500	14.53
10,000	11.26

This example illustrates a type of problem of frequent occurrence. In these problems, the mathematical model eventually reduces to the question: at what point on a given interval does a certain function assume its maximum value? Methods for finding such points, and/or the corresponding maximum value of the function, are among the most useful of the applications of the calculus. In this chapter we shall consider this problem and some related theory and applications.

In certain cases one wants to get the minimum of a function rather than the maximum. One way to do this is simply to replace the function by its negative; this interchanges maxima and minima, and anything proved about maxima becomes, with suitable modification, applicable to minima. Most of our theorems and definitions will be stated for maxima, leaving to the student the statements and proofs for minima.

The word "extremum" is often useful to cover both "maximum" and "minimum".

2. The Maximum Theorem.

We state here, without proof, the basic theorem regarding maxima. A proof and discussion of this theorem will be given in Section 7.

Theorem I. If f is continuous on [a,b] then there is at least one point m in [a,b] such that

$$f(m) \geq f(x) \quad \text{for any } x \text{ in } [a,b].$$

We shall call m a maximum point of f, and f(m) the maximum value, or simply the maximum, of f.

That both the closure of the interval and continuity are needed is seen by the following examples.

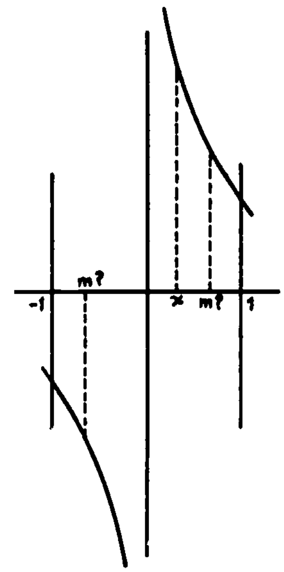


Figure 2-1

Example 1. f is defined on $[-1,1]$ by $f(0) = 0$, $f(x) = 1/x$ if $x \neq 0$. Obviously there is no maximum, for given any m we can always find an x with $f(x) > f(m)$ simply by taking $0 < x < m$ if $m > 0$, or any $x > 0$ if $m < 0$. Of course f is not continuous at $x = 0$.

Example 2. f is defined by $f(x) = x$ on $(-1,1)$. For any m satisfying $-1 < m < 1$ we can find an x, say $x = (1 + m)/2$, that satisfies the same inequalities and such that $x > m$, i.e. $f(x) > f(m)$. However, on the closed interval we could take $m = 1$ and this would give the maximum.

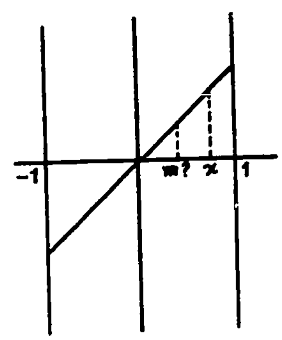


Figure 2-2

The continuous function illustrated in Figure 2-3 has a maximum at m and a minimum at b (also at b_1 - the minimum, or maximum, may be attained at more than one point).

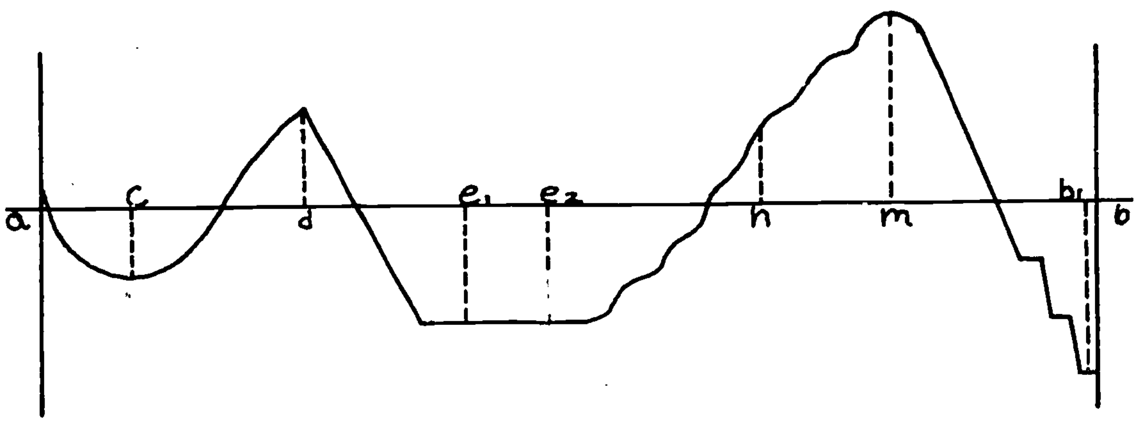


Figure 2-3

But points like c, d, e_1, e_2 are also of interest. They could be maxima or minima if the domain of the variable were sufficiently restricted, as contrasted with points like n which could not. These "local" extrema are important enough to be given a specific definition.

<u>Definition</u>: m is a local maximum point of f if there is a $\delta > 0$ such that $f(m) \geq f(x)$ for any x in $[m-\delta, m+\delta]$.

Evidently an "absolute" or "global" maximum is either a local maximum or an end-point of the interval, so from now on we concentrate on the local maxima. The key theorem is the following.

Theorem 2. If m is a local maximum point of f, and if $f'(m)$ exists, then $f'(m) = 0$.

This is an immediate consequence of Corollary 3, Section 5-6, which says, in effect, that if $f'(m) \neq 0$ then $f(m)$ is neither a local maximum nor a local minimum.

We thus arrive at the following classification of possible values for the maximum point:

a. End-points of the interval,

b. Points of non-differentiability of the function,

c. Roots of $f'(x) = 0$.

The last type are called "critical" points; i.e., m is critical point of f if $f'(m) = 0$.

The most general technique in finding maximum points is to determine all these points, find the value of f at each such point, and pick out the largest. Most problems have various short-cuts, however, some of which are indicated in the following examples.

Example 3. Little Johnny has a board 1" thick and 24" wide with which he wishes to make a low shelf for his clothes closet. He decides to cut it lengthwise and nail it together to give the cross-

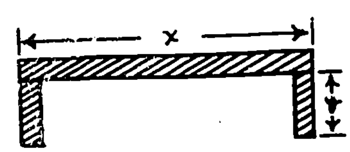


Figure 2-4

section shown in Figure 2-4. Since he wants to use the space underneath the shelf as a hidey-hole he wants to make this space as large as possible. What should be the widths, x and y, of the boards?

We neglect the width of the saw cuts and assume in our model that $x + 2y = 24$. This is certainly good enough for Johnny. We wish to maximize the area $A = y(x-2)$. To apply our theory we must first of all express A as the value of a function of one variable. Since the two variables at present in A are connected by a simple equation it is easy enough to do this by solving this equation for one of the variables in terms of the other and substituting in the expression for A, thus:

$$x = 24 - 2y,$$

$$A = y(24 - 2y - 2) = 22y - 2y^2.$$

Now our problem is to maximize (i.e., find a maximum point for) the function $A(y) = 22y - 2y^2$.

First we note that the conditions of our problem require that we restrict ourselves to the interval [0,11]; certainly we cannot have $y < 0$, and for $y > 11$ we would have $x < 0$. (One might question whether $y = 0$ or $y = 11$ make any sense in the physical set-up. Regardless of whether they do or not we include them so as to get the closed interval needed for the application of our theory.) On this interval A(y) is continuous and differentiable, and so the maximum occurs either at an end-point or at a critical point. But $A(y) = 0$ at both end-points (this is physically obvious, but of course it follows from the form of A(y) if you want to keep this part of the argument mathematically pure), so the maximum comes at a critical point. $A'(y) = 22 - 4y$, so the only critical point is $y = 5.5$. Hence this is the maximum point. It follows at once that $x = 13$.

This last argument is not uncommon. If we have only one candidate for a local maximum, either a critical point or a point of non-differentiability, and if the possibility of an end-point maximum is ruled out by physical considerations, then the candidate must be the global maximum. We can also adapt the argument to serve as a test to tell whether a possible local maximum really is one; this is useful information in some cases.

Local Maximum Test I. Let f be continuous on [c,d] and differentiable except possibly at the point m in [c,d]. Let m be either a critical point or a point of non-differentiability, and be the only point of either kind in [c,d]. If f(c) and f(d) are each less than f(m) then m is a local maximum point, and conversely.

Example 4. $f(x) = 1/(1-x^2)$ gives $f'(x) = 2x/(1-x^2)^2$, and $x = 0$ is the only critical point. Since f(x) is not defined at -1 and 1, to apply Test I we must choose c and d in (-1,1) and on opposite sides of 0. Choosing $c = -1/2$ and $d = 1/2$ we get $f(c) = f(d) = 4/3$, $f(0) = 1$. Hence 0 is a local minimum point.

Example 5(a). $f(x) = x^2 + 1$. Obviously, $x = 0$ is a minimum since $f(x)$ increases with $|x|$.

(b). $f(x) = \dfrac{1}{(x^2 + 1)^{5/2}}$. Hence, $x = 0$ is a maximum and $f(x)$ decreases steadily as $|x|$ increases.

This example shows that calculus is not always needed to determine extrema.

Example 6. Let us take another look at the Supré-Dupré of Section 1. The function $P(S)$ is continuous for all S and differentiable everywhere but at $S = 6000$, but we are obviously interested only in the range $S \geq 3877$.

For $S < 6000$ we have $P'(S) = 70.42(11077 - 2S)$, giving $S = 5538.5$ as a critical point. For $S > 6000$, $P'(S) = (3.042 \times 10^{12})(-S^{-2} + 7754S^{-3})$, giving $S = 7754$ as another. At these two points the values of $P(S)$ are respectively 19.4002×10^6 and 21.1568×10^6. Since these are both greater than $P(6000)$ we see that 7754 is actually the maximum point for P.

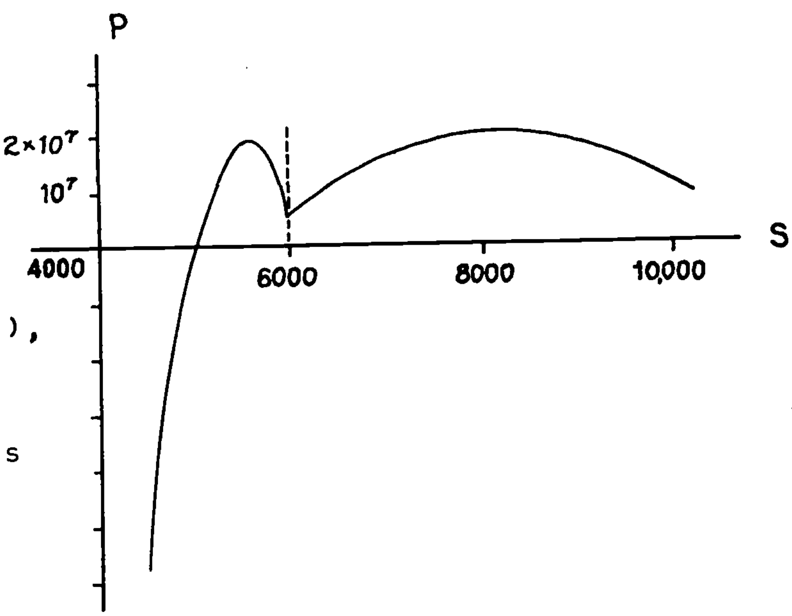


Figure 2-5

Example 7. What are the dimensions of the right circular cylinder of largest lateral area that can be cut from a sphere of radius R?

In Figure 2-6 the lateral area of the cylinder is $A = 4\pi rh$, and r and h are related by $r^2 + h^2 = R^2$. If we eliminate h from these two equations we get

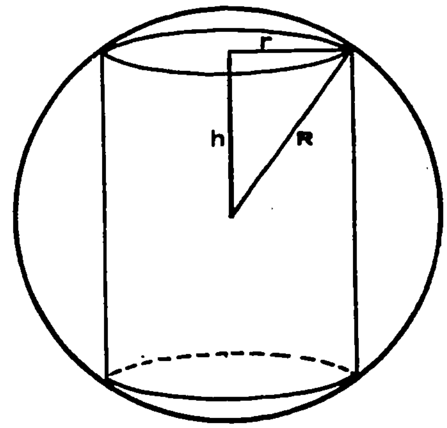


Figure 2-6

$$A(r) = 4\pi r\sqrt{R^2 - r^2}$$

as our function to be maximized. Now we cannot (as yet) differentiate this function, but if we let $B(r) = [A(r)]^2$, then obviously A will be a maximum if and only if B is a maximum, and

$$B(r) = 16\pi^2(R^2r^2 - r^4)$$

can easily be handled by the methods of this Section. Thus:

$$B'(r) = 16\pi^2(2R^2r - 4r^3)$$

equals zero when $r = 0$ or $\pm R/\sqrt{2}$. r is obviously restricted to the interval $[0,R]$, so $-R/\sqrt{2}$ need not be considered. $B(r)$ certainly assumes some positive values, and since $B(0) = B(R) = 0$, $r = R/\sqrt{2}$ must be the maximum point. For this value of r we find $h = R/\sqrt{2}$ also.

PROBLEMS

Formulas for volumes and areas of simple geometrical figures are given here for reference.

r = radius $\qquad$ h = altitude

Circle: Circumference = $2\pi r$,

Area = πr^2

Circular sector: $A = \frac{1}{2}r^2\theta$, where θ is the central angle measured in radians.

Sphere: Volume = $\frac{4}{3}\pi r^3$,

Area = $4\pi r^2$.

Right circular cylinder: Volume = $\pi r^2 h$,

Lateral area = $2\pi rh$.

Right circular cone: Volume = $\frac{1}{3}\pi r^2 h$,

Lateral area = πrs, where s = slant height = $\sqrt{r^2 + h^2}$.

1. Discuss each of the following functions defined in the given interval with regard to local maxima and minima.

a) $f(x) = 4 - x^2$ $\qquad$ x in $[-3,3]$

b) $f(x) = x^2 + \frac{1}{x^2}$ $\qquad$ x in $[-2,2]$ and $x \neq 0$

c) $f(x) = \frac{x - 1}{x + 1}$ $\qquad$ $x \neq -1$

d) $f(x) = 13 - 15x + 9x^2 - x^3$ $\qquad$ x in $[-2,6]$

e) $f(x) = x + \frac{1}{x}$ $\qquad$ $x \neq 0$

f) $f(x) = 10x/(1 + 3x^2)$ $\qquad$ x in $[-2,2]$

g) $f(x) = \frac{1}{3}x^3 - x^2 - 3x + 2$ $\quad$ x in $[-1,4]$

h) $f(x) = x - 2 \sin x$ $\quad$ x in $[\frac{3\pi}{2}, 2\pi]$

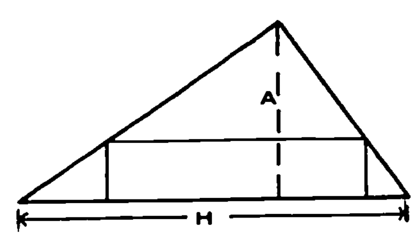
i) $f(x) = \sin x + 2 \cos x - 1$ $\quad$ x in $[-2\pi, 2\pi]$

j) $f(x) = \sin x + \cos x - x$ $\quad$ x in $[-2\pi, 2\pi]$

2. A man wants to build a rectangular pen adjacent to his house. There is to be fencing on three sides since the side on the house needs no fencing. If the man has 100ft of fencing, what should be the dimensions of the pen in order that it will have a maximum area?

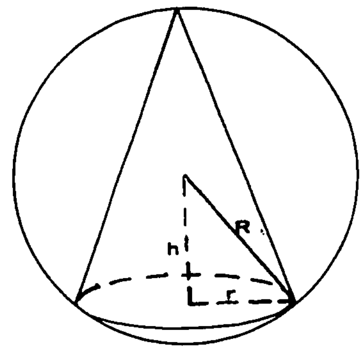
3. A piece of wire 90in. long is bent in the shape of a rectangle. Find the length and width that give the maximum area.

4. Find the dimension of a rectangle of largest area which has one side along the hypotenuse of a right triangle and the ends of the opposite side on the legs of the right triangle.



Assume that the hypotenuse is of length H and that the altitude to it is of length A.

5. Find the dimension of the right circular cone of maximum volume which can be inscribed in a sphere of radius R.



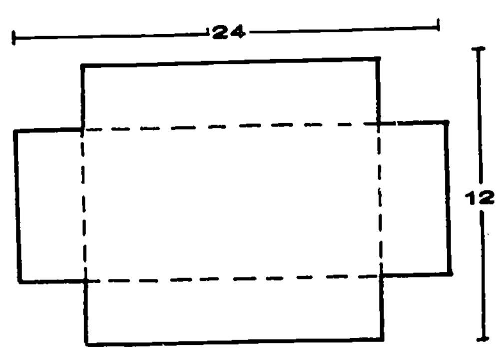
6. Find the dimension of the rectangle of maximum area which can be inscribed in the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

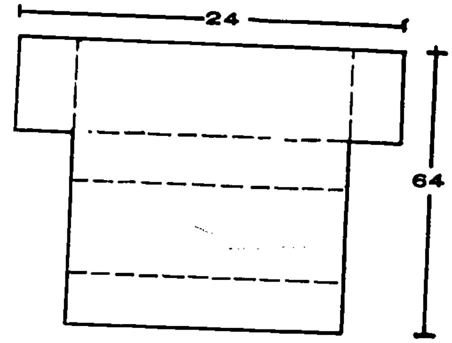
7. Find two positive numbers whose sum is 16 and the sum of whose cubes is a minimum.

8. Prove that of all rectangles inscribed in a fixed circle the square has the largest area. Can you do this without using calculus?

9. An open box is to be made by cutting out squares from the corners of a rectangular piece of cardboard and then turning up the sides. If the piece of cardboard is 12 in. by 24 in., what are the dimensions of the box of largest volume made in this way?

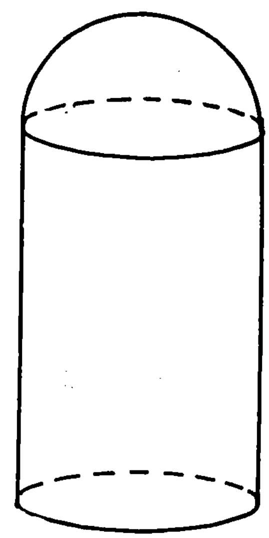


10. A closed box is to be made from a 24 x 64 piece of cardboard by cutting and folding as indicated in the figure. What are the dimensions of the box of maximum volume?



11. Consider a right circular cone with a given volume, and with altitude h and radius of base r. What relationship between these quantities is needed to obtain a minimum lateral area?

12. A silo of given volume is to be built in the form of a right cylinder surmounted by a hemisphere. If the cost per square foot of the material is the same for floor, walls and top, find the most economical proportions.

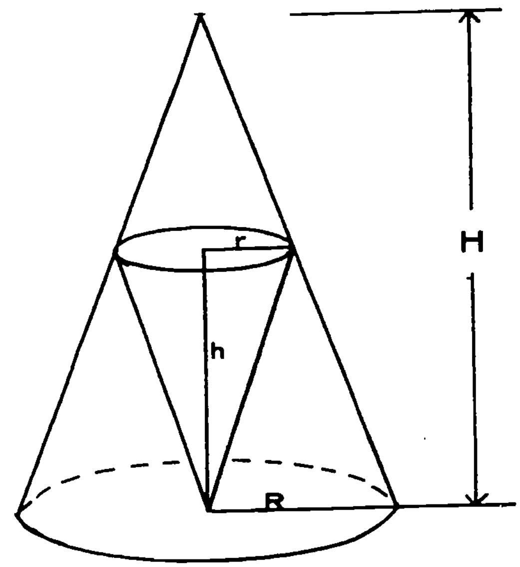


13. One right circular cone is contained within another, as shown in the figure. R and H are given, and we wish to determine r so that the lateral area of the inner cone is a maximum. Consider the three cases:

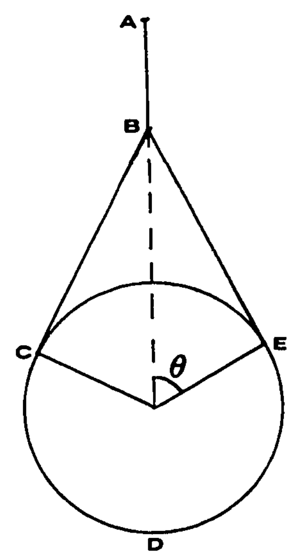
(a) H = 10, R = 2;

(b) H = 10, R = 3;

(c) H = 10, R = 4.

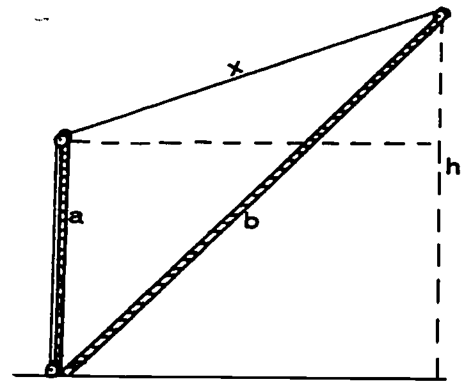


14. A cylinder hangs, like a yo-yo, in a loop of string tied at B and fastened at A. If the total length of string ABCDEB is constant what should be the angle θ so that the cylinder hangs as low as possible?



15. Discuss the local maxima and minima of f(x) = x sin x. Find two positive critical points to 3 decimal place accuracy and determine whether they are local maxima or minima.

16. On the adjacent derrick the boom, of length b, is held by a cable going over the top of the mast, of height a. Let x be the length of the cable between the tops of the mast and the boom, and h the height of the top of the boom.



a. Show that

$$h = f(x) = \frac{1}{2a}(a^2 + b^2 - x^2).$$

b. Since $f'(x) = -x/a$, $x = 0$ is the only local extremum, $f(0) = \frac{1}{2a}(a^2 + b^2)$, and for $x > 0$ or $x < 0$, $f(x)$ is obviously less than $f(0)$. This would seem to show that $x = 0$ is the maximum point and $\frac{1}{2a}(a^2 + b^2)$ the maximum value of f. Do you think this is the case? If not, analyze the situation carefully, considering both the maximum and minimum of $f(x)$, and explain why the standard method seemed to break down.

3. Numerical Methods

The methods of Section 2 are fine if we can find f' and solve f'(x) = 0, but suppose one or the other of these tasks is impossible? In Section 2-2, a method of approximating the roots of any continuous function was developed, so if we can get an analytic expression for f'(x) we can use this method to solve f'(x) = 0 to as great an accuracy as we wish (or as our machine will handle). On the other hand, even though Chapter 7 will introduce a tremendously powerful method of differentiation, one may encounter functions whose derivatives either unobtainable in any reasonable form or are so complicated that even the evaluation of f'(x) for one value of x is a major task. In such cases it is often best to abandon the sophisticated methods of Section 2 and go back to the crude hunting method used at the end of the Supré-Dupré problem.

Even for the simplest case, the type of Example 3, in which we know that f has exactly one local maximum and no local minimum in the interval, a general hunting program is rather complicated. One trouble is that you cannot compare just two values of f at each step but you need three, the idea being to keep the middle value larger than the others.

One flow chart for this case is given at the end of this section; you may well be able to devise others.

If there is more than one extremum, the problem becomes much more difficult. One approach is to take a fairly large number of points, say a hundred, equally spaced over the interval, and essentially examine each consecutive triple, x_{i-1}, x_i, x_{i+1}, for a possible extremum; e.g., if $f(x_i) > f(x_{i-1}), f(x_{i+1})$ then there is a maximum in (x_{i-1}, x_{i+1}). But if the function has a very tiny bump in its graph the bump might come between x_i and x_{i+1} and then it will be missed. In fact, there is no guarantee in any numerical method that we may not miss such bumps. If we do find an extremum between x_{i-1} and x_{i+1} for some i then we can divide this small interval into a hundred parts and repeat the process.

From the purely mathematical point of view the computer calculation of a maximum is a trivial problem. Any given computer can work with only a fixed finite set of numbers. If, for example, the computer word has ten digits and sign there are exactly 2×10^{10} numbers that can be used. (This statement must be modified if "multiple precision" is used.) Hence, within any interval there are only a finite number, say N, of values of x that we can use, and all we need to do is to compute f(x) for these N values and pick out the largest. The catch is that N is too large to allow this to be done in any reasonable time.

In Section I we remarked that we sometimes want to find the maximum point m and sometimes the maximum value M of the function. A glance at a graph will show that in the case where m is a critical point it is much easier to get M to a given accuracy $\pm\varepsilon$ than it is to get m to the same accuracy. In Figure 3-1 any estimate of m between the dotted lines will give an error in M of less than ε. (If the curve bends very sharply at the maximum ε will have to be taken very small before this phenomenon shows up.) Hence, if M is what you want, it is wasteful to require too much accuracy of m. In a computer program it is best to make the stopping condition depend on consecutive values of f(x) as they approach M, rather than on the values of x as they approach m.

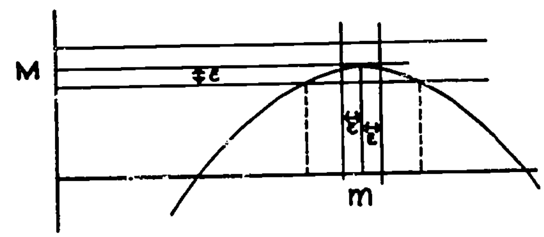


Figure 3-1

Example I. Flow chart for finding the maximum point of a function on a closed interval.

Given:

terval, [a,b];

nction, f; having exactly one maximum point, either local or at an endpoint, and no other local extrema;

Measure of accuracy, ε; i.e. if computed maximum point is m, true maximum point lies in (m-ε, m+ε).

General method:

Let c, m, d divide [a,b] into four equal subintervals of length h_1. At the i-th step, given c_i, m_i, d_i, with $h_i = m_i - c_i = d_i - m_i$, pick m_{i+1} as that one of c_i, m_i, d_i that gives the largest value of f. Let $h_{i+1} = h_i/2$, $c_{i+1} = m_{i+1} - h_{i+1}$, $d_{i+1} = m_{i+1} + h_{i+1}$. Continue until $h_i < \varepsilon$.

Comment: This flow chart does not take account of roundoff error. That roundoff can cause trouble is shown in Table I, which results from using this flow chart to find the maximum point of

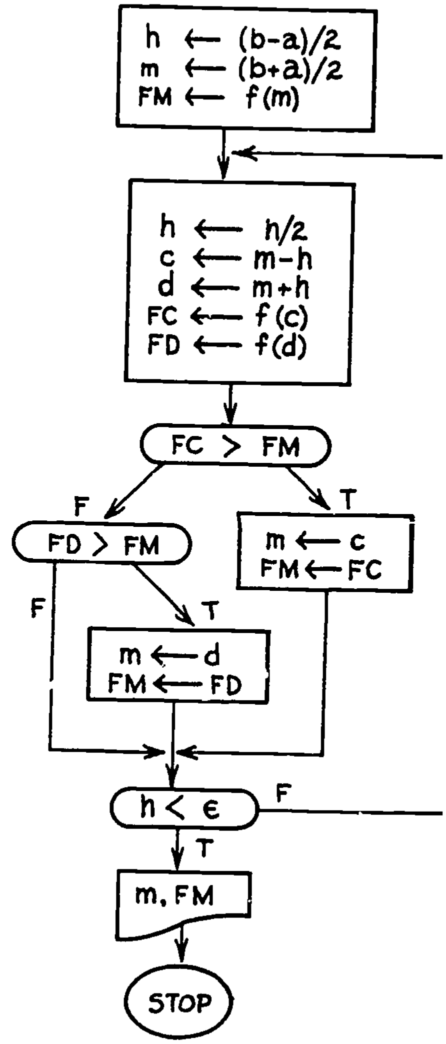


Figure 3-2

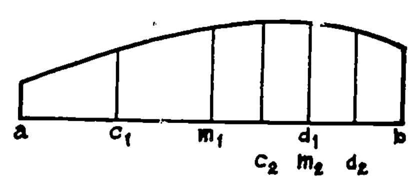


Figure 3-3

$f(x) = 2.25x - x^3$ on the interval [0,1], to 8D, (i.e. 8 decimal places, corresponding to $\varepsilon = 5 \times 10^{-9}$). The exact maximum value is

is

$$3\sqrt{3}/4 \approx 1.29903810567666$$

at

$$x = \sqrt{3}/2 = 0.86602540378444.$$

Note that f(m) is correct to 10D while h is still 3×10^{-5}, whereas m never gets more accurate than 7D. What has happened is that the errors in f(m) became so small that they were obscured by the roundoff in the machine computation of f(m). At this point the branch conditions, FC > FM and FD > FM became random choices dependent on roundoff, and the rest of the program is meaningless.

MAXIMUM OF F(X) = X*(2.25 - X*X) ON [0,1]

VALUE OF H	VALUE OF M	VALUE OF F(M)
.2500000000	.5000000000	1.0000000000
.1250000000	.7500000000	1.2656250000
.0625000000	.8750000000	1.2988281250
.0312500000	.8750000000	1.2988281250
.0156250000	.8750000000	1.2988281250
.0078125000	.8593750000	1.2989234924
.0039062500	.8671875000	1.2990345955
.0019531250	.8671875000	1.2990345955
.0009765625	.8652343750	1.2990364805
.0004882813	.8662109375	1.2990380162
.0002441406	.8662109375	1.2990380162
.0001220703	.8659667969	1.2990380968
.0000610352	.8659667969	1.2990380968
.0000305176	.8660278320	1.2990381057
.0000152588	.8660278320	1.2990381057
.0000076294	.8660278320	1.2990381057
.0000038147	.8660278320	1.2990381057
.0000019073	.8660240173	1.2990381057
.0000009537	.8660259247	1.2990381057
.0000004768	.8660249710	1.2990381057
.0000002384	.8660254478	1.2990381057
.0000001192	.8660254478	1.2990381057
.0000000596	.8660254478	1.2990381057
.0000000298	.8660254478	1.2990381057
.0000000149	.8660254180	1.2990381057
.0000000075	.8660254180	1.2990381057
.0000000037	.8660254180	1.2990381057

These results emphasize the need for caution in accepting the outcome of a machine calculation. If the given flow chart had not been modified so as to output M at each step we would not have detected this situation and would probably have assumed that the final value was correct to 8D. What is needed in the flow chart is something to stop the process when $|FC - FM|$ and $|FD - FM|$ are each less than ε_{min}, where ε_{min} is the minimum accuracy, introduced in Section 2-1, beyond which the combination of machine and programming language cannot go. We leave it to the reader to make the necessary modifications, including an adjustment of the output to tell us what has happened and what accuracy we are actually getting.

PROBLEMS

1. Write a flow chart for the method suggested in the [illegible] paragraph of this section.

2. Program either your flow chart of Problem 1 or the one at the end of the section.

3. Modify your flow chart and program from Problem 2 to compute the maximum value of the function to a given accuracy.

4. Given the function

$$f(x) = \frac{4x^7}{2x^{16} + x + 1}, \qquad 0 \leq x \leq 1 .$$

(a) Find the maximum point correct to 6 decimal places.

(b) Find the maximum value correct to 6 decimal places.

Answers: .979278, 1.013118.

4. The Mean Value Theorem.

The following fact seems obvious from a picture: If a curve joining P and Q has a tangent at every point then at some point the tangent is parallel to the line PQ. (Figure 4-1(a)) Expressed in terms of functions this observation leads to a simple equation that is extremely useful.

Mean Value Theorem. If f is continuous on [a,b] and differentiable on (a,b) then there is a point ξ in (a,b) such that

$$f(b) - f(a) = (b - a)f'(\xi).$$

That is, the tangent to the curve y = f(x) at the point $(\xi, f(\xi))$ has the same slope as the line L joining the endpoints; i.e., $f'(\xi) = \dfrac{f(b) - f(a)}{b - a}$. (Figure 4-1 (b))

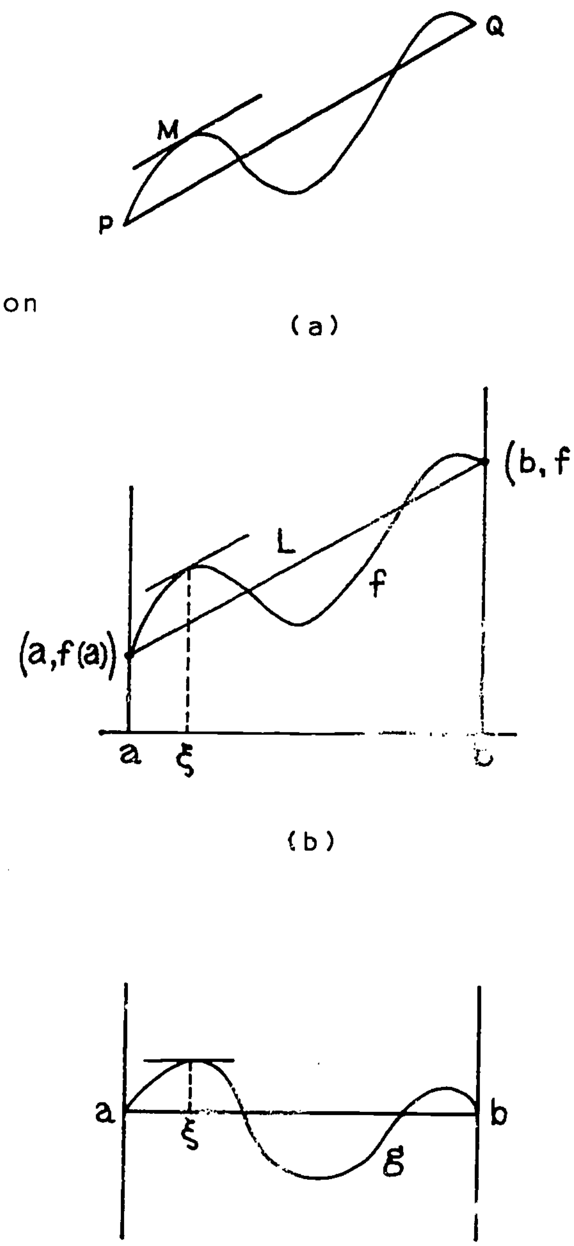


Figure 4-1

Proof. We simplify the picture by subtracting the linear function L from f to give

$$g(x) = f(x) - L(x)$$
$$= f(x) - [f(a) + \frac{f(b) - f(a)}{b - a}(x - a)],$$

(Figure 4-1 (c)).

We have $g(a) = g(b) = 0$. If $g(x)$ is ever positive then g must have a local maximum; if $g(x)$ is ever negative then g has a local minimum; the only remaining case is for $g(x)$ to be zero for all x, and in that case every point of (a,b) is both a local maximum and a local minimum. Therefore g always has at least one local extremum ξ in (a,b) and so

$$0 = g'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a},$$

which proves the theorem.

The special case of the Mean Value Theorem when applied to functions like g is known as Rolle's Theorem:

If f is continuous on [a,b] and differentiable on (a,b) and if $f(a) = f(b) = 0$ then there is a point ξ in (a,b) for which $f'(\xi) = 0$.

If our function f has a second derivative on (a,b) we can get an equation similar to the MVT (Mean Value Theorem) but involving $f''(\xi)$ instead of $f'(\xi)$. To see the analogy more clearly write the MVT in the form

$$f(b) = f(a) + (b - a)f'(\xi).$$

The extended MVT is:

If f' is continuous on $[a,b]$ and if f'' exists on (a,b) then there is a point ξ in (a,b) such that

$$f(b) = f(a) + (b - a)f'(a) + \frac{1}{2}(b - a)^2 f''(\xi).$$

<u>Proof</u>. This time we consider a function of the form

$$(1) \qquad g(x) = f(x) - [f(a) + (x - a)f'(a) + (x - a)^2 C]$$

where C is a constant chosen so that $g(b) = 0$; i.e.

$$C = [f(b) - f(a) - (b - a)f'(a)]/(b - a)^2.$$

It is easy to see that $g(a) = 0$, so, applying Rolle's Theorem to $g(x)$, there is a ξ_1 in (a,b) such that $g'(\xi_1) = 0$. Differentiation of (1) gives

$$g'(x) = f'(x) - [f'(a) + 2(x - a)C],$$

from which follows $g'(a) = 0$. We can therefore apply Rolle's Theorem to the function $g'(x)$ on the interval $[a,\xi_1]$, to get $g''(\xi) = 0$ for some ξ in (a,ξ_1) and hence in (a,b). Since further differentiation gives $g''(x) = f''(x) - 2C$ we conclude that

$$\frac{1}{2}f''(\xi) = [f(b) - f(a) - (b - a)f'(a)]/(b - a)^2,$$

from which the desired result follows on solving for f(b).

It is evident that this process can be extended to higher derivatives, and we state without proof a general theorem of this sort.

First a little notation: f' is the derivative, f'' the second derivative, f''' the third derivative, but we can't continue this indefinitely. For higher derivatives we usually write $f^{(4)}$, $f^{(5)}$, ...; this has the advantage that $f^{(n)}$ is the n-th derivative, etc. For consistency of the notation it is often convenient to define $f^{(0)} = f$.

<u>Taylor's Theorem</u>. If $f, f', \ldots, f^{(n)}$ are defined and continuous on $[a,b]$ and if $f^{(n+1)}$ is defined on (a,b) then there is a point ξ in (a,b) such that

$$f(b) = f(a) + (b - a)f'(a) + \frac{1}{2!}(b - a)^2 f''(a) + \frac{1}{3!}(b - a)^3 f'''(a)$$

$$+ \ldots + \frac{1}{n!}(b - a)^n f^{(n)}(a) + \frac{1}{(n+1)!}(b - a)^{n+1} f^{(n+1)}(\xi).$$

We shall give a more direct proof of this theorem later on, from quite a different point of view.

As an example of how the extended theorem can be used, let us examine more carefully the statement in Section

3 that if m is a local maximum point and if $f'(m) = 0$ then it is easier to get accuracy in $f(m)$ than in m. In the Extended MVT replace a by m and b by x to get

$$f(x) = f(m) + (x - m)f'(m) + \frac{1}{2}(x - m)^2 f''(\xi), \quad m < \xi < x.$$

Now in deriving the EMVT we made no use of the fact that $b > a$; the same process could be applied to give the formula with a and b interchanged. If in <u>this</u> formula we replace a by x and b by m we get exactly the same equation but with the inequality $x < \xi < m$. Combining these two cases, we can state the EMVT in the form: If f' is continuous on [a,b] and f" exists on (a,b) then for any two points m and x in [a,b] we have

$$f(x) = f(m) + (x - m)f'(m) + \frac{1}{2}(x - m)^2 f''(\xi)$$

where ξ is between m and x. The EMVT and Taylor's Theorem are usually stated in this form.

In our case we have $f'(m) = 0$, so

$$f(x) = f(m) + \frac{1}{2}(x - m)^2 f''(\xi)$$

with ξ between m and x. Now assume that f" is bounded in some neighborhood of m; that is, that there is a $\delta > 0$ and a $B > 0$ such that $|f''(x)| < B$ for all x satisfying $|x - m| < \delta$. Then for $|x - m| < \delta$ we have

$$|f(x) - f(m)| < \frac{1}{2}B\,|x - m|^2 < \frac{1}{2}B\,\delta^2.$$

Thus f(x) approaches f(m) like δ^2, much more rapidly than x approaches m.

In the particular case of Example 1 of the last section we have

$$f(x) = 2.25x - x^3, \quad f''(x) = -6x.$$

Since $m = \sqrt{3}/2 \approx .87$, for $|x - m| < .1$ we will certainly have $|f''(x)| < 1$.

Hence

$$|f(x) - f(m)| < .5\,|x - m|^2.$$

Now if x is a 7D approximation to m then $|x - m| \leq 5 \times 10^{-8}$, and so

$$|f(x) - f(m)| < 2 \times 10^{-15}.$$

This is less than the accuracy of the machine on which Example 1 was run, so it is not surprising that we were unable to get m to more than 7D accuracy.

PROBLEMS

1. In geometric terms, the Mean Value Theorem states that there is a ξ between a and b such that the tangent at $(\xi, f(\xi))$ is parallel to the line joining the points $[a, f(a)]$ and $[b, f(b)]$. In each of the problems (a) through (e) a curve and the end points of the interval are given. Find a value of ξ satisfying the requirements of the Mean Value Theorem.

(a)	$f(x) = x^2$	$a = 2$	$b = 3$
(b)	$f(x) = \sqrt{x}$	$a = 25$	$b = 36$
(c)	$f(x) = x^3 - 9x + 1$	$a = -3$	$b = 4$
(d)	$f(x) = x^2 - 2x - 3$	$a = -1$	$b = 3$
(e)	$f(x) = x^3 - 2x^2 + 3x - 2$	$a = 0$	$b = 2$

2. Suppose that $f'x) = 2x^{2/3}$ and let $a = -1$ and $b = +1$. Show that there is no number ξ between a and b which satisfies $f'(\xi) = \frac{f(b) - f(a)}{b - a}$, given that $f'(x) = \frac{4}{3}x^{-1/3}$.

   Explain, and sketch the graph.

3. Use the Mean Value Theorem to show that $\sqrt{x+1} - \sqrt{x}$ approaches zero as x increases without bound. Can you find a way to show this without using calculus?

4. Prove Taylor's Theorem for the case n = 2. [Use the results of Problem 11, Section 5-6, to differentiate the powers of $(x - a)$.]

5. Approximations.

The Mean Value Theorem and its extensions are basic tools in the approximate computation of values of functions. Their application is best shown by an example.

Example 1. What is the approximate value of $\sin 61^\circ$?

This is a pretty vague question as it stands. 0 is an approximation, and so is 100, but the errors are of the same or greater orders of magnitude than the value itself. $\sin 60^\circ = .86603$ is a much better approximation, but this is so obvious that presumably we want a still better one. To get this we apply the MVT to the function $f(x) = \sin x$; since we are going to get derivatives we will express all angles in radians. For convenience let $a = \pi/3$ ($=60^\circ$), $b = \pi/3 + \pi/180$. Then the MVT gives

$$f(b) = f(a) + (b - a)f'(\xi), \qquad a < \xi < b,$$

or

$$\sin b = .86603 + \frac{\pi}{180} \cos \xi, \qquad \frac{\pi}{3} < \xi < \frac{\pi}{3} + \frac{\pi}{180}.$$

Evidently $\cos \xi$ is close to $\cos(\pi/3) = .5$, so

$$\sin b = .86603 + .5\frac{\pi}{180} = .87475.$$

This is a much better estimate than simply .86603. [Various symbols are used for "approximately equal to":

$\doteq$, $\sim$, $\approx$, $\simeq$, and possibly others. We are using $\simeq$ since it seems to be increasing in popularity, possibly because it is the one most apt to appear on a typewriter.]

This type of computation is common in the rough analysis of certain types of errors. The word "error" in this connection does not mean "mistake" but refers to the difference between the exact value of a quantity and a computed value. Essentially all computations involve errors, and error analysis is one of the most important parts of numerical mathematics.

From this point of view we can replace the question of Example 1 by the following.

<u>Example 2</u>. A pole 100 ft. long leans against a vertical wall. The base angle x is measured and found to be 60°, giving 86.6 ft. as the height H at which the pole touches the wall. However there is a possible error of about 1° in

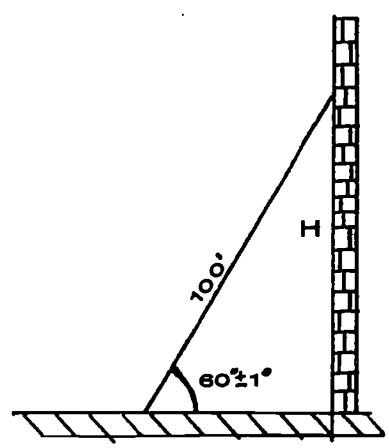


Figure 5-1

the measured value of the angle x. What is the possible error in the computed height?

We wish to find the change in H corresponding to a change in x of $1^\circ = \pi/180$. The phrase "increment of x" is usually used in these situations and designated by the symbol Δx. Similarly we have ΔH, the increment of H. What we do is to apply the method of Example 1, using x, $x + \Delta x$, H, and $H + \Delta H$ respectively for a, b, f(a) and f(b). The MVT gives directly

$$\Delta H = f'(\xi)\Delta x,$$

which, approximating $f'(\xi)$ by $f'(x)$, gives

$$\begin{aligned} \Delta H &\approx f'(x)\Delta x \\ &= (100 \cos x)\Delta x \\ &= 50\ (\pi/180) \\ &\approx .9 \text{ ft.} \end{aligned}$$

Thus, H can vary from 85.7 to 87.5, and a reasonable statement about the height is that it is about $86\frac{1}{2}$ ft. with a possible error of about a foot.

The general approximation formula,

$$\Delta f(x) \approx f'(x)\Delta x,$$

is by far the most useful one in approximation problems.

Now let us return to Example 1 and ask the further question: "Just how accurate is the value .87475?" Of course we cannot hope to answer this question exactly, for if we could we would merely correct the computed value and get the exact value of sin 61°. What we want is an "error bound," a number B (presumably small) for which we can prove that

$$|\sin 61^\circ - .87475| < B.$$

Such an error bound could be obtained by a more careful examination of the quantity $\cos\xi$, but it is easier to apply the extended MVT. This gives us

$$\sin b = \sin a + (b - a)\cos a + \frac{1}{2}(b - a)^2(-\sin \xi),$$

or

$$\sin 61^\circ = .86603 + \frac{\pi}{180}(.5) - \frac{1}{2}\left(\frac{\pi}{180}\right)^2 \sin \xi,$$

from which we get

$$|\sin 61^\circ - .87475| < \frac{1}{2}\left(\frac{\pi}{180}\right)^2 < 1.6 \times 10^{-4},$$

since certainly $|\sin \xi| < 1$.

Do we want more accuracy? Simply use the next case of Taylor's Theorem:

$$\sin b = \sin a + (b - a) \cos a + \frac{1}{2}(b - a)^2(-\sin a) + \frac{1}{6}(b - a)^3(-\cos \xi)$$

$$= .86603 + \frac{\pi}{180}(.5) + \frac{1}{2}\left(\frac{\pi}{180}\right)^2(-.86603) + \frac{1}{6}\left(\frac{\pi}{180}\right)^3(-\cos \xi)$$

$$= .86596 + E$$

where $|E| < \frac{1}{6}\left(\frac{\pi}{180}\right)^3 < 3 \times 10^{-6}$. For still higher accuracy we merely use more terms in Taylor's Theorem.

Figure 5-2 gives a graphic interpretation of these results. We replace 61° by x and consider the two approximations

$$y_1 = \sin a + (x - a) \cos a$$

and

$$y_2 = \sin a + (x - a) \cos a + \frac{1}{2}(x - a)^2(-\sin a).$$

The first is the linear approximation, given by the tangent line. This is the "best" linear approximation, in the sense that if $y_0 = p + qx$ is any other first degree approximation then $\lim_{x \to a} \frac{\sin x - y_1}{\sin x - y_0} = 0$; that is, the error $\sin x - y_1$, is an order of magnitude smaller than the error $\sin x - y_0$. In the same sense y_2 shown by the dotted line in the figure, is the best quadratic approximation in the neighborhood of a.

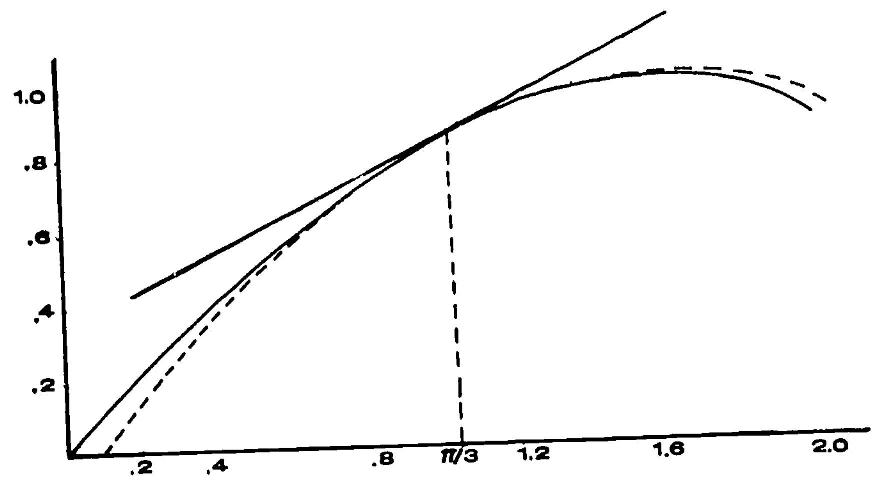


Figure 5-2

The replacement of a function by its linear approximation is a very common device in applied mathematics. Quite frequently when a model of a physical problem is first set up it is too complicated for mathematical analysis. It must then be further simplified, and one way to do this is to "linearize" some or all of the functions that appear in it. For example, when one sets up the equation for the motion of a simple pendulum swinging through a small angle, the function $\sin \theta$ comes into the equation. This equation is impossible to solve in elementary terms, so one linearizes it by replacing $\sin \theta$ by

$$\sin 0 + (\theta - 0)(\cos 0) = \theta,$$

and with this modification the equation is readily solved to give simple harmonic motion.

Of course linearization introduces an error, which depends on the size of the neglected term $\frac{1}{2}(x - a)^2 f''(\xi)$ in the EMVT, in addition to whatever errors were involved in setting up the model to begin with. It is up to the model maker, in cooperation with the mathematician, to decide whether the extra errors can be tolerated.

It is worth mentioning that one nice thing about computers is that they are much less dependent on linearity than classical mathematical analysis. A computer will solve the pendulum equation just about as easily with $\sin\theta$ in it as with $\sin\theta$ replaced by θ. On the other hand, in neither case will it tell you anything about the relation of the solution to simple harmonic motion.

PROBLEMS

1. A wooden cube, supposed to be 3" on a side, was found to be 3.03". It was brought down to size by sandpapering off the excess. About how much material was removed?

2. Using linear approximations, find approximate values, with error bounds, for each of the following:

   (a) $\sqrt{4.12}$

   (b) $\tan 47^\circ$

   (c) $\cos 1$

3. Compute $\sqrt{5}$ to 3 decimals by using $\sqrt{5} = \frac{1}{4}\sqrt{80}$ and the fact that $\sqrt{80} = 9$.

4. Since the attraction of gravity falls off inversely as the square of the distance from the center of the earth, the weight of an object h feet high is given by

$$w(h) = w(0)\,\frac{R^2}{(R + h)^2},$$

where R is the radius of the earth and w(0) the weight of the object at the ground.

(a) We ordinarily use the approximation $w(h) \approx w(0)$. Show that for an object 100 miles high this is in error by about 5% [use R = 4000 miles].

(b) Show that for the same object the linear approximation is in error by less than 0.2%.

(c) Show that the quadratic approximation is in error by less than 0.01%.

(d) How high could one go before getting a 1% error in the linear approximation? In the quadratic approximation?

5. (a) Make a flow chart for a program to approximate sin x by a fifth degree polynomial in x, for $0 \le x \le .5$, and to give an error bound. [Use Taylor's Theorem].

(b) Write the program from (a) and use it to compute sin x for x = .2 and .5.

6. Monotone Functions

Between an adjacent local maximum and local minimum a curve steadily rises or steadily sinks - that is, the function is monotone. It seems evident that a tangent to a rising curve has a positive slope, and conversely that if a curve has a positively sloping tangent at every point then it is rising. We shall prove the correctness of these observations and relate them to the behavior of a function at its extremum and critical points.

We shall state our theorems for increasing functions and positive slopes, leaving to the student the statements and proofs of the corresponding theorems for decreasing functions and negative slopes.

If f is increasing on an interval [c,d], then each chord drawn on the graph of f between points (x, f(x)) and (x + h, f(x + h)) must have its slope,

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

greater than or equal to zero, regardless of the sign of h.

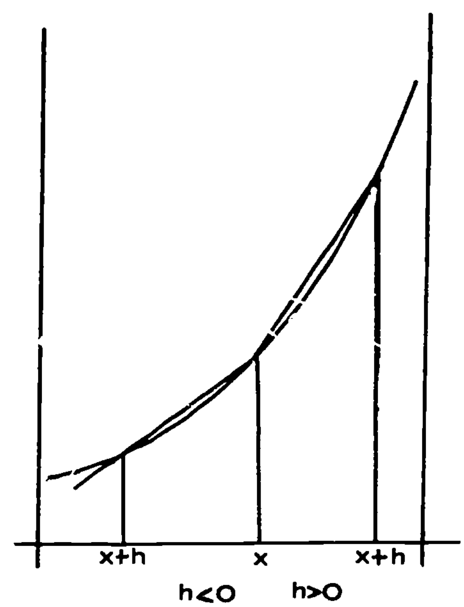


Figure 6-1

Hence

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

whenever this limit exists.

We state this result as

Theorem 1. If f is increasing on [c,d] and if x is any point of (c,d) at which f'(x) exists then f'(x) $\geq$ 0.

Knowledge of monotonicity thus tells us something about the derivative. Can we somehow reverse our argument and use the derivative to tell us about possible monotonicity? Example 1 of Section 2 shows the need for caution, for this function has a negative derivative, $-x^{-2}$, at all points except x = 0, and it certainly is not decreasing over the interval. We obviously need to say something about all points of the interval.

Theorem 2. If f is continuous on [c,d] and if f'(x) $\geq$ 0 for all x in (c,d) then f is increasing in [c,d].

Proof. If p and q are any two points in [c,d], with q > p we have by the MVT,

$$f(q) - f(p) = (q - p)f'(\xi).$$

By assumption f'(ξ) $\geq$ 0, and so f(q) $\geq$ f(p).

Theorem 2, but not Theorem 1, is true if $f'(x) \geq 0$ and "increasing" are replaced by $f'(x) > 0$ and "strictly increasing." We leave to the student the verification of this statement.

<u>Example 1.</u> Determine the behavior of the function

$$f(x) = \frac{x^3}{x^2 - 1}$$

with regard to domain of definition, local extrema, critical points, and domains of monotonicity. Use this information to graph the function. The discussion of this example is divided into several parts, each devoted to one aspect of the function's behavior. Be sure you understand each part before going on to the next one.

(a) First of all it is evident that the function is defined and differentiable for all values of x except -1 and 1. Hence the three intervals $(-\infty,-1)$, $(-1, 1)$, $(1,\infty)$ must be considered separately.

(b) The critical points are the solution of $f'(x) = 0$. Since

$$f'(x) = \frac{(x^2 - 1)\ D_x x^3 - x^3\ D_x\ (x^2 - 1)}{(x^2 - 1)^2}$$

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

$$= \frac{x^2(x^2 - 3)}{(x^2 - 1)^2} ,$$

the critical points are $-\sqrt{3}$, 0, $\sqrt{3}$.

(c) f is monotone in each of the intervals $(-\infty, -\sqrt{3})$, $(-\sqrt{3}, -1)$, $(-1, 0)$, $(0, 1)$, $(1, \sqrt{3})$, $(\sqrt{3}, \infty)$. To determine where f is increasing and where decreasing we need check $f'(x)$ at only one point in each interval, for $f'(x)$ cannot change sign in one of these intervals. We easily see, without bothering to compute exact values of $f'(x)$, that for

x =	-2,	-1.5,	-.5,	.5,	1.5,	2,
f'(x) is	+	-,	-,	-,	-,	+.

Hence f is increasing on $(-\infty, -\sqrt{3})$ and $(\sqrt{3}, \infty)$ and decreasing on $(-\sqrt{3}, -1)$, $(-1, 1)$, and $(1, \sqrt{3})$.

(d) Coming now to the extrema, note that with no further computation we can tell that $-\sqrt{3}$ is a local maximum point, $\sqrt{3}$ is a local minimum point, and the remaining critical point, 0, is not an extremum point at all. However, to graph the function we do wish to compute its values for all significant values of x, and we find

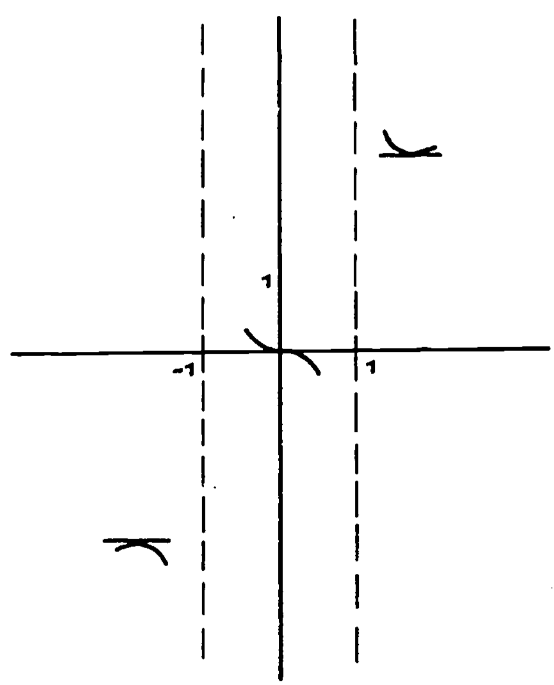


Figure 6-2(a)

x	$-\sqrt{3}$	0	$\sqrt{3}$
$f(x)$	$-\sqrt{3}/2$	0	$3\sqrt{3}/2$

Note that the local maximum is less than the local minimum. (Figure 6-2(a)).

(e) To draw a good picture of the curve we need two more bits of information. First, how does the function behave in the neighborhood of the points where it is not defined? To see what happens near $x = 1$ put $x = 1 + h$ and express y as a function of h. We get

$$y = \frac{(1 + h)^3}{h(2 + h)} .$$

Now for very small values of h, $(1 + h)^3/(2 + h)$ is approximately $\frac{1}{2}$, and so y behaves much like $\frac{1}{2h}$. (Figure 6-2(b)).

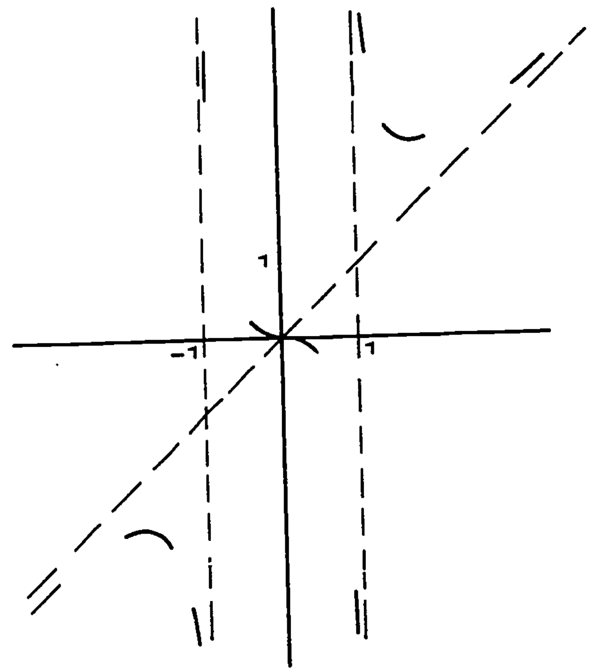


Figure 6-2(b)

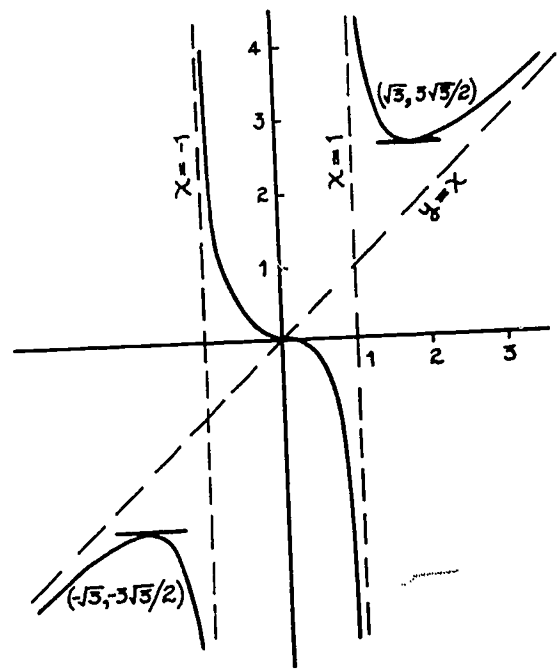


Figure 6-2(c)

(f) Finally, what happens as $x \to \infty$; i.e. as x gets larger and larger without bound? To see this divide numerator and denominator of f(x) by x^k, where k is the smaller of the degrees of the two polynomials. In our case these degrees are 3 and 2, and dividing by x^2 gives

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

Now as $x \to \infty$, $\frac{1}{x^2} \to 0$, and so for large values of x, f(x) is approximately x, (Figure 6-2(b)).

(We can do better if we wish,

$$f(x) - x = \frac{1}{x} \frac{1}{1 - \frac{1}{x^2}} \approx \frac{1}{x} ,$$

which is positive for $x > 0$ and negative for $x < 0$. Hence our graph lies above the line $y = x$ on the right and below it on the left.)

We now have so much information that the graph practically draws itself. The values at $x = \pm 3/4$, i.e. $y = \pm 1$, give a little more firmness to the middle branch. The carefully drawn and labeled graph is shown in Figure 6-2(c).

Comments on the example.

1. Lines like $y = x$, $x = 1$, and $x = -1$ in this example are called <u>asymptotes</u> of the curve (or to the curve). A

formal definition of this term is complicated and is not needed for our description of graphs so we merely say that an asymptote is a line that "gets arbitrarily close" to the curve as we move out along the line. The two most common ways of determining asymptotes are illustrated in (e) and (f); i.e. finding the behavior of $f(x)$: as x approaches a value where a denominator of $f(x)$ becomes zero; and as $x \to \infty$ or $x \to -\infty$.

2. Part (d) illustrates the second important method of testing for a local maximum.

<u>Local Maximum Test 2</u>. Let f satisfy the same conditions as in Test 1. If $f'(c) > 0$ and $f'(d) < 0$ then m is a local maximum, and conversely.

An advantage of Test 2 over Test 1 is that no functional values need be computed. We need only determine whether $f'(c)$ and $f'(d)$ are positive or negative. As in Test 1, c and d may be chosen wherever convenient. We can even avoid using any particular points; for instance, in texting $x = \sqrt{3}$ we see from the form of $f'(x)$ that its sign depends only on the factor $x^2 - 3$, since x^2 and $(x^2 - 1)^2$ are never negative, and that $x^2 - 3$ is negative if x is slightly less than $\sqrt{3}$ and positive if x is slightly more. Hence $\sqrt{3}$ is a local minimum point.

3. If f has a second derivative at m there is still another test that tells (in most cases) whether m gives a maximum.

<u>Local Maximum Test 3</u>. Let m be a critical point of f at which $f''(m)$ exists. If $f''(m) < 0$ then m is a local maximum point (but not necessarily conversely).

<u>Proof</u>. Since $f''(m) = \lim_{x \to m} \frac{f'(x) - f'(m)}{x - m} < 0$, for x sufficiently near to m we must have

$$\frac{f'(x) - f'(m)}{x - m} = \frac{f'(x)}{x - m} < 0.$$

Taking x slightly larger than m gives $x - m > 0$ and so $f'(x) < 0$. Taking x slightly smaller than m gives $f'(x) > 0$. By Test 2, m then a local maximum point.

In our example we can find $f''(x)$ by writing $f'(x)$ in the form

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

Then

$$f''(x) = \frac{(x^4 - 2x^2 + 1)(4x^3 - 6x) - (x^4 - 3x^2)(4x^3 - 4x)}{(x^4 - 2x^2 + 1)^2}$$

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

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

Indicating the signs only, we have

at $x = -\sqrt{3}$, $f''(x) = \frac{-\,(+)}{+} = -$, therefore a maximum;

at $x = \sqrt{3}$, $f''(x) = \frac{+\,(+)}{+} = +$, therefore a minimum;

at $x = 0$, $f''(x) = \frac{0\,(+)}{-} = 0$, therefore no conclusion can be drawn.

This example illustrates the two disadvantages of Test 3; the relative difficulty (in some cases) of getting $f''(x)$, and the fact that we get no information at all if $f''(x) = 0$ at the tested point. In the latter case the function may have a maximum, a minimum, or neither, as is shown by the cases $-x^4$, x^4, x^3. Nevertheless, if $f''(x)$ is easy to find Test 3 is usually tried first, and then one of the other two in case of failure.

4. To graph a function it is rarely necessary to get all the information that we did in this example. On the other hand, there are cases where additional information may be needed. How much, and what kind of information is useful depends on the particular function, but also on our

reason for drawing the graph. A graph, after all, is simply a method of describing a function - highly inaccurate in a quantitative sense but very good qualitatively - and it should tell us those things about the function that we want to know. For example, before evaluating $\int_{-1}^{5} \frac{x}{x^2 + 1} dx$ we should know something about the behavior of the function. Where is it negative; is it monotone; does it have extrema? However, the precise location of a local extremum is of little importance. All in all, considerable practice with a wide variety of functions is needed before one acquires the knack of quickly sketching a given function with a degree of roughness appropriate to the situation.

PROBLEMS

1. Discuss each of the following functions with regard to local extrema, critical points, and domains of monotonicity. Sketch the graph of each function.

a) $f(x) = x - x^2$ $[-1, 2]$

b) $f(x) = \frac{x - 1}{x^2 + 3}$ $(-\infty, +\infty)$

c) $f(x) = \sqrt{9 - x^2}$

d) $f(x) = (x - 3)\sqrt{x}$

e) $f(x) = \dfrac{3x + 1}{x^2 + x + 3}$ [-4, -1]

f) $f(x) = \dfrac{x^2}{1 + x^2}$

g) $f(x) = x^3 + 6x^2$ [-4, 0]

h) $f(x) = 2 \cos x + \cos 2x$ [0, 2π]

2. Find all the critical points of $f(x) = x + \sin x$. Prove that this curve is always increasing and hence has neither a local maximum or minimum. Sketch the curve.

3. Find constants a, b, and c, so that the curve $y = ax^2 + bx + c$ goes through the point (3,0) and has a local extremum at (1,2).

4. Show that (8,8) is the point of the curve $x^2 - 8y = 0$ closest to the point (2,11).

5. Graph $y = \sqrt{x + 1} - \sqrt{x}$ (See Problem 3 Section 4).

6. The cost of manufacturing x articles is $c(x) = a + bx$. And they can be sold at a price $p(x)$, given by $p(x) = m - nx$. Show that the profit is a maximum when $x = (m - b)/2n$.

*7. Existence of a Maximum.

Theorem. On a closed interval a continuous function has a maximum.

Let f be continuous on [a,b]. We wish to show the existence of a point m in [a,b] such that $f(m) \geq f(x)$ for any x in [a,b]. Our procedure will be to construct an algorithm to determine such a point m if there is one, and then to prove that the determined point does have the required property. The continuity of f will be used only in the latter step - the algorithm converges whether or not f is continuous.

The algorithm, a flow chart for which is given in Figure 7-1, is a variation of the familiar bisection process, the critical step being the branch condition which tells us whether we choose the right or the left half of the bisected segment. This condition is the following:

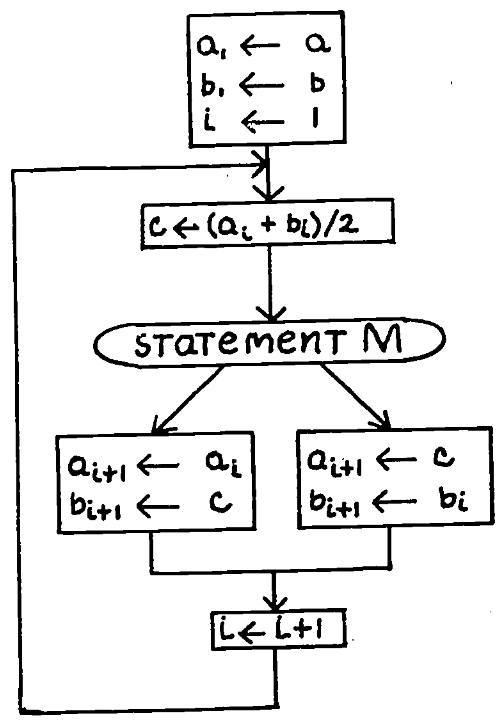


Figure 7-1

Statement M : There is an x in $[a_i, c]$ such that $f(x) \geq f(y)$ for any y in $[c, b_i]$.
If there is such an x we choose the half-segment containing it, that is, the left half; if not we choose the right half.

Statement M is rather complicated and deserves some discussion. But for the present we make only two comments: first, at each step statement M is either true or false, that is, either there is such a point x or there isn't; second, if f has a maximum on the interval Statement M will select the half-interval containing the maximum (it will select the left half-interval if both contain a maximum).

The proof that the sequences $a_1, a_2, \ldots$ and $b_1, b_2, \ldots$ satisfy the conditions of our Axiom of Continuity and so converge to a point m is the same as in earlier applications of the bisection process and we shall not repeat it.

We have now to prove that $f(m) \geq f(x)$ for any x in $[a,b]$. Let $x = x_1$ be any such point. If $x_1 = m$ the inequality is certainly true, so suppose $x_1 \neq m$. There must be a step in the algorithm at which x_1 and m are separated, i.e., one in a left half-interval and one in a right. Now at this point the algorithm chose the half-interval containing m, not x_1, and so there must be an x_2 in the same half-interval as m with $f(x_2) \geq f(x_1)$.

Now we proceed with x_2 just as we did with x_1. If $x_2 = m$ then we have $f(m) \geq f(x_1)$ as desired; if not there is some further step at which x_2 and m are separated and there is an x_3 in a smaller sub-interval with m such that $f(x_3) \geq f(x_2)$. Continuing the process gives us

$$f(x_1) \leq f(x_2) \leq f(x_3) \leq \cdots$$

If the process stops with $x_k = m$ we have $f(x_1) \leq f(m)$ as desired. If not we have a sequence x_1, x_2, ... which converges to m (details of proof left to the student). Since f is a continuous function $f(x_1)$, $f(x_2)$, ... converges to f(m). Since each $f(x_n) \geq f(x_1)$ it follows (Section 2-4, problem 5(a)) that the limit $f(m) \geq f(x_1)$, as was to be proved.

Note that continuity of the function comes in only at the very end. It is a useful exercise to carry through the steps of the proof up to this point for Example 2-1 and for Example 2-2 extended to the closed interval by $f(-1) = f(1) = 0$.

We have said that Statement M is either true or false, but the question arises "How do you tell?" In some special cases, like Examples 2-1 and 2-2, it is easy enough, but is there a general method that will work for any function, or at least any continuous function? The answer is "no," the basic trouble being that we would have to consider an infinite number of x's and an infinite number of y's and this can't be done in a finite time. Because of this situation this proof of the existence of a maximum point m is said to be "non-constructive." It has been proved that this state of affairs cannot be avoided - <u>any</u> proof of the existence of a maximum is necessarily non-constructive. This, of course, does not prevent us from using our knowledge of the <u>existence</u> of a maximum to help us in finding ways to locate the maximum - which is exactly what we did in Sections 2 to 5.

Chapter 7

THE CHAIN RULE

1. Differentiation of Composite Functions

In Chapter 5 we developed formulas for computing derivatives of certain basic functions, notably x^n, sin x, and cos x. Theorems 2 to 6 of Sections 5-6 enable us to differentiate complicated functions formed from these by the four arithmetical operations. But these will not even enable us to differentiate the simple function $f(x) = \sqrt{4 - x^2}$ which arose in Chapter 2 in consideration of a circle. Is this the best we can ever hope to do? No, we can certainly derive other basic formulas as is done in the last problem in the Chapter 5, by clever tricks; or we can go back to the definition of a derivative and try to derive other formulas, as we did for sin x. But such methods turn out to be unnecessary unless we encounter completely new types of functions.

The secret lies in the one derivative theorem we have not yet developed, the formula for the derivative of a composite function. Take the function $f(x) = \sqrt{4 - x^2}$. Introducing the functions

$$g(x) = 4 - x^2, \quad h(y) = \sqrt{y} ,$$

enables us to write $f(x) = h(g(x))$. If we can express f' in terms of g' and h' we will have what we want, for g' and h' can easily be calculated (for h' see Example 2, Section 5-4). The Chain Rule tells us how to do this.

How, in general, would we go about getting the derivative of $f(x) = h(g(x))$ at the point a? By definition

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a} \frac{h(g(x)) - h(g(a))}{x - a}$$

Let $y = g(x)$ and $b = g(a)$. Then we are tempted to say; since certainly $y \to b$ as $x \to a$;

$$f'(a) = \lim_{x \to a} \frac{h(y) - h(b)}{x - a}$$

$$= \lim_{x \to a} \frac{h(y) - h(b)}{y - b} \quad \frac{g(x) - g(a)}{x - a}$$

$$= \lim_{x \to a} \frac{h(y) - h(b)}{y - b} \quad \lim_{x \to a} \frac{g(x) - g(a)}{x - a}$$

$$= \lim_{y \to b} \frac{h(y) - h(b)}{y - b} \quad \lim_{x \to a} \frac{g(x) - g(a)}{x - a}$$

$$= h'(b)g'(a) = h'(g(a))g'(a).$$

Unfortunately this "proof" resembles the one in aerodynamics about which it was remarked that all the steps were wrong except the last one. The trouble is that as $x \to a$, <u>but always with</u> $x \neq a$, we do have $y \to b$, <u>but</u> not necessarily with $y \neq b$. (If you want an example consider the function $g(x) = x \sin(1/x)$.) This makes it meaningless to write

$$\lim_{x \to a} \frac{h(y) - h(b)}{y - b}$$

since the denominator may be zero and the function not defined for values of x arbitrarily close to a. For the same reason we cannot replace $\lim_{x \to a}$ by $\lim_{y \to b}$.

In spite of its inaccuracy this "proof" indicates why the chain rule has the form it does. (The aerodynamics "proof" has the same virtue.) Our problem is now to fix up our "proof" by finding a way to eliminate the division by y-b. We do this by a simple trick that gives a formula that will be useful later on.

Since, by definition,

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = f'(a);$$

if we define

$$z(x) = \frac{f(x) - f(a)}{x - a} - f'(a)$$

then $\lim_{x \to a} z(x) = 0$. $z(x)$ is not defined at $x = a$, so, in consideration of the last equation, we define $z(a) = 0$.

Then z(x) is continuous at $x = a$. This result is important enough to be stated formally.

Lemma. If f is differentiable at a, then there is a function z, defined and continuous wherever f is, and with $z(a) = 0$, such that

$$f(x) - f(a) = (x - a)[f'(a) + z(x)].$$

We can now easily prove the Chain Rule.

Theorem 1. Let $f(x) = h(g(x))$, where g is differentiable at a and h is differentiable at g(a). Then f is differentiable at a and $f'(a) = h'(g(a))g'(a)$.

Proof. Apply the Lemma to the function h, using $b = g(a)$ instead of a and $y = g(x)$ instead of x. We get

$$h(y) - h(b) = (y - b)[h'(b) + z(y)],$$

with $z(b) = 0$. Then

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

$$= \frac{y - b}{x - a}[h'(b) + z(y)]$$

$$= \frac{g(x) - g(a)}{x - a}[h'(b) + z(y)].$$

Now

$$\lim_{x \to a} z(y) = z(\lim_{x \to a} y) = z(b) = 0,$$

since z(y) is continuous (see page 388). And

$$\lim_{x \to a} \frac{g(x) - g(a)}{x - a} = g'(a)$$

by definition of $g'(a)$. Hence

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = h'(g(a))g'(a),$$

which proves the theorem.

If one wishes to think in terms of functions rather than values at a point, the following formulation is convenient.

Chain Rule. If $f(x) = h(g(x))$ then $f'(x) = h'(g(x))g'(x)$ for all values of x for which the indicated derivatives exist.

Example 1. Find the equation of the line tangent to the circle at the point (3,4).

This problem can be done by analytic geometry, since the tangent to a circle has a a special property. The slope of the radius to the point (3,4) is 4/3; hence the slope of the tangent, perpendicular to this radius, is -3/4, and the equation of the tangent is

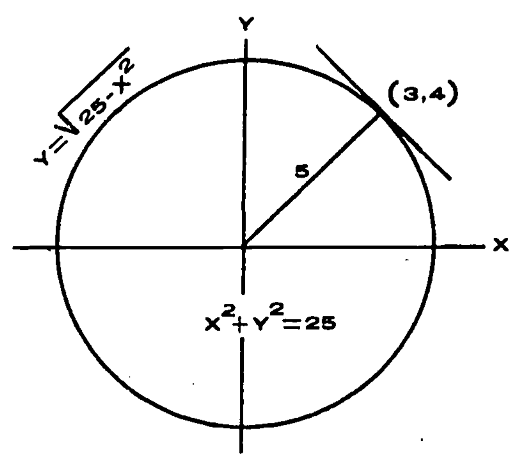


Figure 1-1

$$y - 4 = -\frac{3}{4}(x - 3), \quad \text{or} \quad y = -\frac{3}{4}x + \frac{25}{4}.$$

To avoid using this special property of the circle we get the slope of the tangent as a derivative. The point (3,4) lies on the upper semicircle, which, as we saw in Chapter 3, is the graph of the function

$$f(x) = \sqrt{25 - x^2}, \quad x \text{ in } [-5,5].$$

Now f(x) can be thought of as h(g(x)), where $g(x) = 25 - x^2$ and $h(y) = \sqrt{y}$. Then $g'(x) = -2x$ and $h'(y) = \frac{1}{2}y^{-1/2}$. Applying the chain rule gives

$$f'(x) = h'(g(x))g'(x)$$

$$= \frac{1}{2}(25 - x^2)^{-1/2}(-2x) = -x/\sqrt{25 - x^2}.$$

For x = 3 this gives for the slope of the tangent line $f'(3) = -\frac{3}{4}$, as before.

Example 2. To differentiate $\sin \frac{1 - x}{1 + x}$ we consider this function as

$$f(x) = h(y) = \sin y, \quad y = g(x) = \frac{1 - x}{1 + x}.$$

It is convenient to use y' to represent g'(x). Then we can write—

$$f'(x) = (\cos y)y' = (\cos y)\frac{(1 + x)(-1) - (1 - x)(1)}{(1 + x)^2}$$

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

<u>Example 3</u>. Differentiate $\sin^2 \sqrt{x^2+1}$. Note first that $f(x) = y^2$, $y = \sin\sqrt{x^2+1}$, so that $f'(x) = 2yy'$. Now $y = g(x) = \sin z$, $z = \sqrt{x^2+1}$, and so $y' = (\cos z)z'$. Next $z = \sqrt{w}$, $w = x^2 + 1$, so $z' = \frac{1}{2}w^{-1/2}w'$, and, finally, $w' = 2x$. Putting all this together gives

$$
\begin{aligned}
f'(x) &= 2y(\cos z)(\tfrac{1}{2}w^{-1/2})(2x) \\
&= 2(\sin\sqrt{x^2+1})(\cos\sqrt{x^2+1})\frac{1}{2\sqrt{x^2+1}}(2x) \\
&= \frac{x}{\sqrt{x^2+1}}\,2(\sin\sqrt{x^2+1})(\cos\sqrt{x^2+1}) \\
&= \frac{x}{\sqrt{x^2+1}}\sin(2\sqrt{x^2+1}).
\end{aligned}
$$

The last two steps of this example are algebraic simplification. It is often important to get the derivative, which may at first look extremely complicated, into a reasonably compact form - if, indeed, this can be done at all. Simplification of the answer is therefore an important part of any differentiation problem. It may be the hardest part, since there are almost no rules for guidance and one must rely on his own experience and ingenuity. With the chain rule and the theorems of Chapter 5 formal differentiation becomes a straightforward process that can be programmed for a computer without much difficulty. But so far nobody has come up with an effective program for simplifying the results.

One general rule that is often useful is the following: if an expression appears as a factor of two or more terms, to possibly different powers, factor it out of these terms to the smallest power to which it appears.

__Example 4__. Find the critical points of $y = x^3\sqrt{1-x^2}$ and test them for local extrema.

We start by treating this as a product of two functions, giving us

$$f'(x) = 3x^2\sqrt{1-x^2} + x^3 D_x\sqrt{1-x^2}.$$

As in Example 1, $D_x\sqrt{1-x^2} = -x(1-x^2)^{-1/2}$, so

$$f'(x) = 3x^2(1-x^2)^{1/2} - x^4(1-x^2)^{-1/2}$$

Our rule says to factor out x^2 and $(1-x^2)^{-1/2}$; this gives

$$f'(x) = x^2(1-x^2)^{-1/2}[3(1-x^2) - x^2]$$

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

$f'(x) = 0$ for $x=0$ and $\pm\sqrt{3}/2$. We certainly do not want to find the second derivative, so Test 2 seems the best. x^2 and $\sqrt{1-x^2}$ are always positive, so $f'(x)$ can change sign only at $-\sqrt{3}/2$ and $+\sqrt{3}/2$. Evidently f' is positive if $|x| < \sqrt{3}/2$ and negative if $|x| > \sqrt{3}/2$, so $-\sqrt{3}/2$ is a local minimum point and $\sqrt{3}/2$ a local maximum point. $x=0$ is neither, since $f(x)$ increases on both sides of it. A rough sketch of the curve is given in Figure 1-2.

Let us return for a moment to the statement of the Lemma. Writing this equation in terms of increments (Chapter 6, Section 5) we have

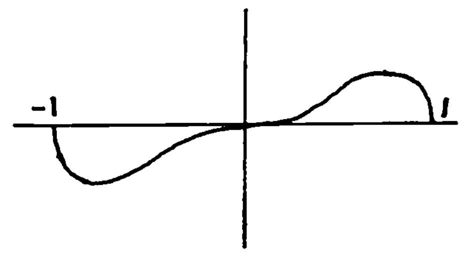


Figure 1-2

(1) $\Delta f(a) = f'(a)\Delta a + z(x)\Delta a.$

Compared with the approximation

(2) $\Delta f(a) \approx f'(a)\Delta a,$

we see that the error in the approximation is just $z(x)\Delta a$. Since $z(x) \to 0$ as $x \to a$ this justifies our use of (2) as an approximation independently of the complicated machinery of Chapter 6. What Chapter 6 does is to give us expressions for $z(x)$ that can be estimated quantitatively, either

$$z(x) = f'(\xi) - f'(a),$$

by the MVT, or

$$z(x) = \frac{1}{2}f''(\xi)\Delta a,$$

by the EMVT.

PROBLEMS

1. Differentiate each of the following functions and simplify if possible

(a) $(1 - x)^{10}$

(b) $(2 - x^{-3})^{-1}$

(c) $(x^3 - 4)^5$

(d) $(x + 1)^{5/2}$

(e) $[1 + (1 + x^2)^{5/2}]^{-3/2}$

(f) $(3 - 2x)^5(3x^2 + 4)^3$

(g) $\left(\frac{x^3 + 2x + 1}{x^2 + 1}\right)^3$

(h) $\left(\frac{x}{1 - x}\right)^{1/2}$

(i) $\frac{x}{\sqrt{1 - x^2}}$

(j) $x\sqrt{x}$

(k) $\sqrt{\sqrt{x} + 1}$

(l) $\sqrt{\frac{1}{x} + 4}$

(m) $\cos 3x - 2 \sin x$

(n) $(2 - x)\cos x^2 + 2x^2 \sin x^4$

(o) $\sin(\cos^2 x)\cos(\sin^2 x)$

(p) $\sin[\sin(\sin x)]$

(q) $(x + 3)^2(x^2 + 2x + 1)^3(x^2 + 4)^4$

(r) $\frac{(x + 3)^2(x - 4)^3}{(x - 4)^4}$

(s) $\frac{x^2 + 2}{\sqrt{x^3 + 4}}$

(t) $\frac{(x^2 + 2x - 1)(x^3 + 3x - 4)^2}{(2x + 6)^2}$

(u) $(x^4 + 5x - 6x^{-1})^3$

(v) 14

(w) $x \sin x$

(x) $\sqrt{\sin x}$

(y) $\sin \sqrt{x}$

(z) $\cot^2 x$

2. Find first and second derivative of each of the following functions.

(a) $\sqrt{x^2 + 1}$

(b) $\frac{x}{(x^2 + 1)^2}$

(c) $(\sin x + \cos x)^2$

(d) $\tan 3x$

(e) $x^2 \sin 2x$

(f) $\sqrt{\tan x + \cot x}$

3. Do Problem 11, Section 5-6, by using the chain rule.

4. If f is an even function (See Problems 2 and 3, Section 3-8) prove that f' is an odd function, and vice versa.

5. Prove that if f is differentiable at 0, then 0 is a critical point of the function $g(x) = f(x^3)$.

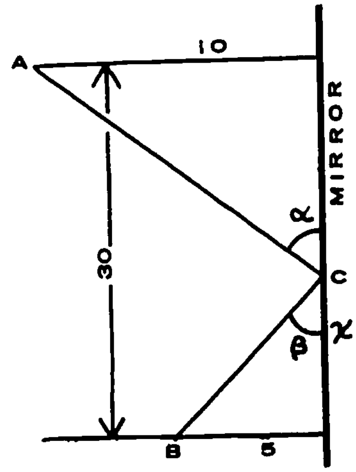
6. Find the local extrema and draw the graphs of each of the following equations.

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

(b) $y = \sqrt[3]{\dfrac{3x^2}{x^3 + 4}}$

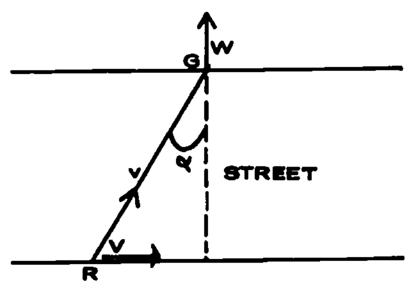
7. (a) A ray of light, passing from A to B by reflection in a mirror at some point C, takes the path that makes

the total length of path, AC + CB, as short as possible. If the relative positions of A, B and the mirror are as indicated in the figure show that x = 10, and that $\alpha = \beta$.



(b) Replacing the specific numbers 10, 5 and 30 by a, b and c, show that in all cases we have $\alpha = \beta$. [It is not necessary to solve for x to show this.]

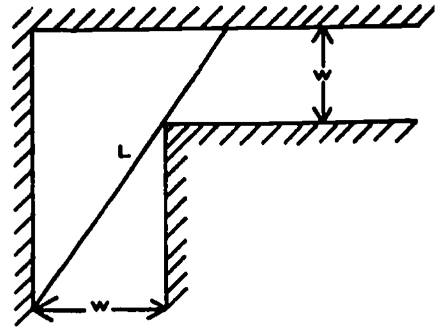
8. (Conceived while trying to cross a street). Being late for class, I am in a tremendous hurry to proceed along the path W, across the street, and because of a fence I must go through the gate G. On this side there is no such restriction. Traffic is heavy, and I see that the first break will come when the right rear corner R of a certain car moves past. I therefore station myself slightly "up-stream," ready to dash to G as soon as R passes me. If my velocity is v and the

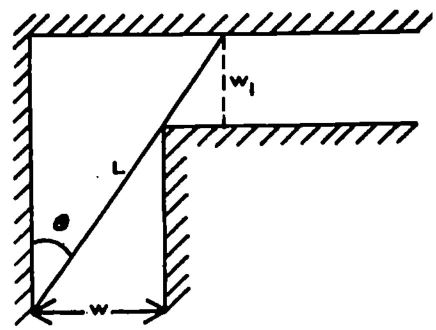


car's is V, with V > v, where should I stand?

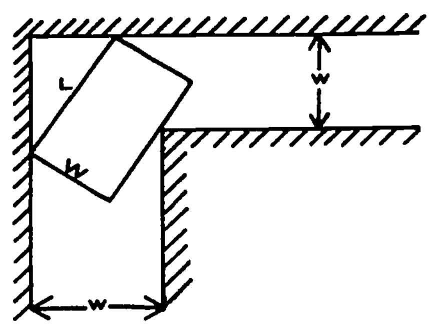
Ans. So that $\sin \alpha = v/V$.

9. What is the length L of the longest straight horizontal rod that can be moved around a corner in a passage-way of width w? [Hint 1. Find, instead, the width w_1, necessary to contain a rod of fixed length L moving around the corner. Then determine L so that $w_1 = w$. Hint 2. You may find it convenient to use the angle θ as the independent variable]





10. If you got the rod around the corner you might try getting a desk of size L × W around the same corner. Formulate a reasonable question and answer it.



2. Implicit Functions

There is an easier way of doing Example I above. Instead of actually solving for y as a function of x we merely imagine that y has been solved for, writing $y = f(x)$ without further specification. We say that we are treating y as an implicit function of x, in contrast to the explicit function $\sqrt{25 - x^2}$ that was used in Example I.

On the upper semicircle, or, more particularly, in the neighborhood of $x=3$, we now have

(I) $$x^2 + [f(x)]^2 = 25$$

identically in x, with $f(3) = 4$.

To say that (I) is true "identically" means that for every value of x the left hand side of the equation has the same value as the right hand side. Or, in other words, the function g defined by $g(x) = x^2 + [f(x)]^2$ is the same function as the one defined by $h(x) = 25$, i.e., $g=h$. Hence $g'=h'$; that is, we can differentiate both sides of (I) and obtain another identity. This is the essence of the method of differentiating implicit functions.

Differentiating (I), using the chain rule to handle $[f(x)]^2$, gives the identity

$$2x + 2f(x)f'(x) = 0.$$

Now putting $x=3$ gives, since $f(3) = 4$, our familiar result $f'(3) = -3/4$.

We usually abbreviate this, as in Example 2 above, by using y' instead of $f'(x)$. Thus from $x^2 + y^2 = 25$ we get at once $2x + 2yy' = 0$, and so in general $y' = -x/y$. Note that this form applies as well to the lower semicircle as to the upper; for example, the slope at $(2,-\sqrt{21})$ is $-2/(-\sqrt{21}) = 2/\sqrt{21}$.

<u>Example 1.</u> The graph of $x^3 + y^3 = 3xy$ is shown in Figure 2-1. The point $P:(1,.348)$ was found by solving $1 + y^3 = 3y$ by an approximation method like that of Chapter 3, Section 2. We wish to find the slope of the curve at P.

The figure strongly suggests that there is a differentiable function $f(x)$ such that $f(1) = .348$ and such that $x^3 + [f(x)]^3 = 3xf(x)$ for all x in, say $[.7,1.3]$; namely, the heavy part of the curve. If we let y designate

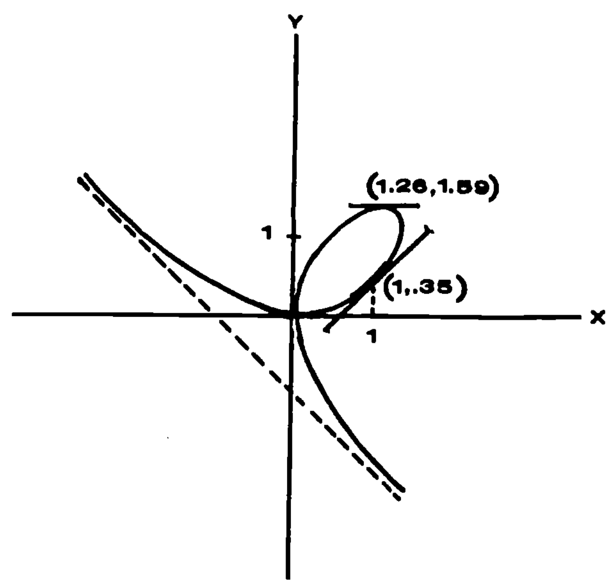


Figure 2-1

this function and y' its derivative then by differentiating $x^3 + y^3 = 3xy$ we get

(2) $$3x^2 + 3y^2y' = 3(xy' + y).$$

The expression in parentheses comes from the derivative of the product

$$D_x(xy) = x(D_xy) + (D_xx)y = xy' + 1y,$$

or more precisely, avoiding abbreviations, from

$$D_x(xf(x)) = xf'(x) + f(x).$$

Solving (2) for y' then gives us

(3) $$y' = \frac{y - x^2}{y^2 - x} = .742 \text{ at } (1, .348).$$

Having turned the crank to get the answer we must now be critical and ask whether the process is actually justified. Is there really such a function as f(x), or are we being misled by a picture? How do we know that as x changes from the value 1 the equation $y^3 - 3xy + x^3 = 0$ has a root in y that changes continuously and differentiably with x? The answer is that there is such a function, although it cannot be expressed in an explicit form (involving real numbers only). The proof of the existence of such a function is a special case of an Implicit Function Theorem of great generality, a topic for Advanced Calculus. In this book we shall adopt the point of view that each use of implicit functions is tacitly preceded by a remark of

the sort: "From graphical or other evidence it looks as if there is an implicit function of the required kind. I shall proceed on the assumption that the function exists, knowing that my results are dependent on the truth of this assumption."

Example 2. Find the local maxima of the curve of Example 1.

It is fairly obvious how we want to define a local maximum of a curve given by an equation in x and y; namely, a local maximum of any implicit function of x defined by the curve. If f(x) is any such function then, by (3), $f'(x) = 0$ only if $y = x^2$. Since x and y must also satisfy $x^3 + y^3 = 3xy$, we get the possibilities by solving these two equations simultaneously. Eliminating y gives $x^3(x^3 - 2) = 0$, or $x=0$, $\sqrt[3]{2}$, and correspondingly $y=0$, $\sqrt[3]{4}$.

($\sqrt[3]{2}$, $\sqrt[3]{4}$) looks like a reasonable extremum - we shall test for a maximum later - but (0,0) certainly does not. For one thing, although the numerator of $f'(x)$ is zero at this point so is the denominator, so $f'(0)$ is not defined by (3). You might maintain that there is a smooth piece of the curve that goes through the origin tangent to the x-axis and that this defines an implicit function having x=0 as a local minimum point. This, in fact, is the point of view of the branch of mathematics known as Algebraic Geometry. On the other hand the Implicit Function Theorem mentioned above does not apply to a point like

(0,0) on this curve. We will follow this track and exclude such points from consideration.

What test do we apply to $(\sqrt[3]{2}, \sqrt[3]{4})$ to see if it is a maximum? Because of the difficulty of computing values of y, Test 1 is not attractive. Test 2 can be used by showing that a small increase in x will make $y'<0$ and a small decrease will make $y'>0$. (We leave it to the student to fill in the details.) Can we use Test 3? To get to the heart of the matter, can we find second - and higher - derivatives of implicit functions? The answer is yes, of course; we simply differentiate equation (2) or (3). The latter is usually the most convenient, since y' is not mixed up with x and y. Remembering always that y is simply a substitute for f(x), we get from (3),

$$y'' = \frac{(y^2 - x)(y' - 2x) - (y - x^2)(2yy' - 1)}{(y^2 - x)^2}$$

$$= \frac{(2x^2y - y^2 - x)y' + (x^2 - 2xy^2 + y)}{(y^2 - x)^2}.$$

We can get rid of the y' by replacing it by its value from (3). Doing this, the result simplifies to

$$y'' = \frac{2xy\ (3xy - x^3 - y^3 - 1)}{(y^2 - x)^3}$$

Now remember that x and $y(=f(x))$ always satisfy the identity $x^3 + y^3 = 3xy$. Hence

$$y'' = \frac{-2xy}{(y^2 - x)^3} .$$

Test 3 is now trivial; since x, y, and $y^2 - x$, which equals $\sqrt[3]{16} - \sqrt[3]{2}$, are all positive, y'' is negative and we have a maximum.

Another application of the differentiation of implicit functions is to extremum problems of the type of Example 7 of Section 6-1. Here we have a quantity to be maximized that depends on two variables, r and h, namely $A = 2\pi rh$. We also have an equation $r^2 + h^2 = R^2$ relating these two variables. The technique was to solve the equation for one of the variables as a function of the other and substitute in the expression for A, thus expressing A as a function of one variable:

$$h(r) = \sqrt{R^2 - r^2} , \quad A(r) = 2\pi rh(r) = 2\pi r\sqrt{R^2 - h^2} .$$

Suppose now that we do not actually solve for the function $h(r)$ explicitly but regard it as defined implicitly by the equation $r^2 + h^2 = R^2$. Then, differentiating with respect to r, we get

(4) $$2r + 2hh' = 0.$$

Similarly, regarding the h in $A = 2\pi rh$ as this same function, we get

$$A'(r) = 2\pi h + 2\pi rh' .$$

For a critical value of r we must have $A(r) = 0$, or

$$(5) \qquad 2\pi h + 2\pi rh' = 0.$$

Between (4) and (5) we can eliminate h', getting $r^2 = h^2$, or, since neither r nor h can be negative, $r = h$. This result in itself is interesting, telling us that for the maximum area the altitude of the cylinder should equal its diameter. If we want the actual values of r and h in terms of R we have only to solve $r = h$ and $r^2 + h^2 = R^2$ simultaneously.

Note the advantages and disadvantages of this implicit method. It may be extremely difficult, or perhaps even impossible, to solve the side condition for one variable in terms of the other; in this case the implicit method is helpful. On the other hand, this method leads eventually to the solution of two simultaneous equations, and this can be a difficult job. In any problem one must use his judgement as to which method to follow.

PROBLEMS

1. Assuming that $y = f(x)$ and that $y' = f'(x)$ exists, find y' in each of the following.

   (a) $xy = 1$

   (b) $x^3 + 6xy + 5y^3 = 3$

   (c) $y = \dfrac{\sqrt{x} + 1}{\sqrt{x} - 1}$

   (d) $y^2 + 2x = y$

   (e) $x + \sqrt{xy} = 2y$

   (f) $2x^2 - y^2 = 1$

   (g) $x^3 + x^2y^2 = x + 2y - 1$

   (h) $x^{1/2} + y^{1/2} = 1$

   (i) $x \sin y = y \sin x$

2. Find y'' in parts (a), (d), (h), and (i) of Problem 1. Simplify your answers if possible.

3. (a) Find the local extrema of the curve

   $$y^2 + xy + x^2 = y - x.$$

   (b) Now think of this equation as defining x as one or more functions of y and find the local extrema of these functions. What do these mean geometrically?

(c) Use the results of (a) and (b) to graph the curve.

4. (a) Find the equation of the line tangent to the parabola $y^2 = ax$ at the point $P = (c, \sqrt{ac})$.

(b) Find the point Q where this line intersects the x-axis.

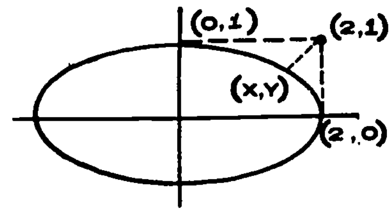
(c) If F is the point (a/4, 0), the focus of the parabola, show that PFQ is an isosceles triangle.

(d) Show that rays of light emitted by a source at F will be reflected by the parabola parallel to the x-axis, and, reversely, rays entering the parabola parallel to the axis are reflected to a focus at F. These properties are the reason for the use of parabolic mirrors for searchlights and for astronomical telescopes.

5. An artist has a 1000 lb. lump of clay. He wishes to make from it a cube and a sphere and paint their surfaces. In order to maximize the total surface to be painted (the cube will, of course, be supported on one corner) how should he divide his clay?

Ans. In the ratio 6 to π, or 656 and 344 lbs.

6. We wish to find the point (x, y) of the ellipse $x^2 + 4y^2 = 4$ nearest to the point $(2,1)$. Show that x must satisfy the equation $9x^4 - 48x^3 + 32x^2 + 192x - 256 = 0$. This equation has two real roots, about 1.637 and -1.982. Without using the picture determine which one gives the desired point.



7. Use Test 2 to show that $x = \sqrt[3]{2}$ is a local maximum point in Exercise 2. [Hint: Show that for a sufficiently small change in x from $x = \sqrt[3]{2}$ the change in x^2 is larger than the change in y. Use the result of the last part of Section 6-4.]

3. Inverse Functions.

The derivative of the cube-root function, $f(x) = \sqrt[3]{x}$, can be obtained, as was done in Chapter 5 for $\sqrt{x}$, by evaluating

$$\lim_{y \to x} \frac{\sqrt[3]{x} - \sqrt[3]{y}}{x - y}$$

by means of an algebraic trick. The same method can be used for any root of x, but a much easier way to proceed is to use implicit differentiation. If $y = \sqrt[3]{x}$ then $y^3 = x$, and

$$3y^2y' = 1,$$

$$y' = \frac{1}{3y^2}$$

$$= \frac{1}{3}y^{-2}$$

$$= \frac{1}{3}x^{-2/3}$$

More generally, let $y = x^{p/q}$, where p and q are positive integers. Then $y^q = x^p$, and

$$qy^{q-1}y' = px^{p-1},$$

$$y' = \frac{p}{q}x^{p-1}y^{-q+1}$$

$$= \frac{p}{q}x^{p-1}x^{(p/q)(-q+1)}$$

$$= \frac{p}{q}x^{p-1}x^{-p+(p/q)}$$

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

Thus the formula

(1) $$D_x(x^n) = nx^{n-1},$$

which we derived in Chapter 5 for the case where n is a positive integer, holds also if n is any positive rational number.

Finally, let n be a negative rational number, and let m=-n. Then

$$D_x(x^n) = D_x\left(\frac{1}{x^m}\right)$$

$$= \frac{x^m(D_x 1) - 1(D_x x^m)}{x^{2m}}$$

$$= \frac{0 - mx^{m-1}}{x^{2m}}$$

$$= -mx^{-m-1}$$

$$= nx^{n-1} .$$

Hence (1) holds for all rational values of n.

Returning now to our cube-root function, $f(x) = \sqrt[3]{x}$, we see that what we did was to consider another function $g(y) = y^3$. f and g are said to be inverses of each other since each "undoes" what the other does.

Let us visualize a function $x \to y$ as a machine which operates on an input x to give an output y. Two functions are then said to be inverses of each other if whenever we put a number in the input hopper of one machine and then feed the output into the second machine, the final result is just what we started with. This is illustrated below with the cubing function and the cube-rooting function.

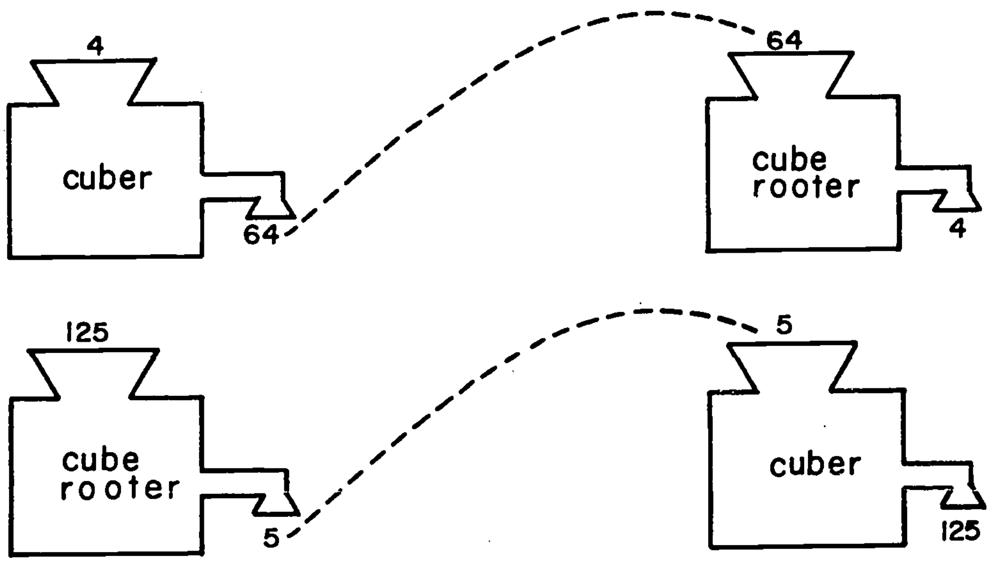


Figure 3-1

If we hooked these two machines together in either order to form a composite machine, then the net effect would be to have a machine for which the output is the same as the input - not a particularly useful piece of machinery.

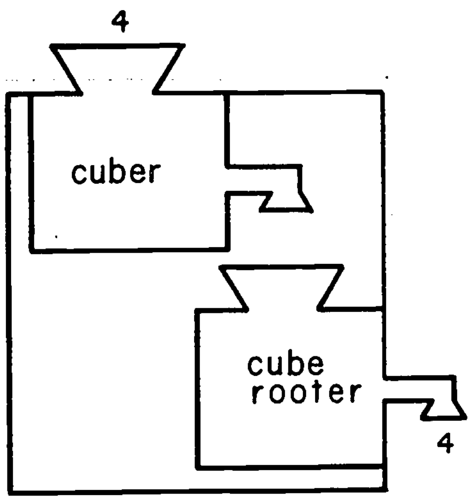


Figure 3-2

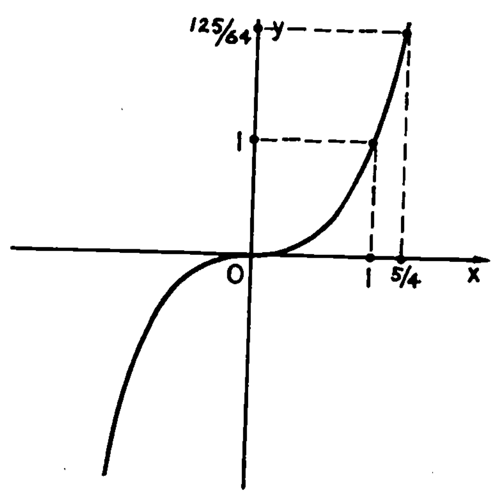


Figure 3-3

It is very helpful in studying inverse functions to look at their graphs - because the same graph will do for both functions. Thus in Figure 3-3 we see the graph of

$$y = x^3 \quad \text{or} \quad x = \sqrt[3]{y} .$$

This will serve as the graph of $f(x) = x^3$ and also as the graph of $g(y) = \sqrt[3]{y}$ provided that we lie on our ear and regard the x-axis as the output axis. Thus we see that

$$f(5/4) = (5/4)^3 = 125/64,$$

$$g(125/64) = \sqrt[3]{125/64} = 5/4;$$

so that

$$g(f(5/4)) = g(125/64) = 5/4,$$

and

$$f(g(125/64)) = f(5/4) = 125/64.$$

And in general we will have

$$g(f(x)) = x \text{ and } f(g(y)) = y.$$

These relations characterize inverse functions.

Suppose that we consider the inverse of the function

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

If g is to be the inverse function then, since $g(f(1)) = 1$ and $f(1) = 1$, we must apparently have $g(1) = 1$. But we also have $f(-1) = 1$, and since necessarily $g(f(-1)) = -1$, we get also $g(1) = -1$. This contradiction leads to the inescapable conclusion that this function f has no inverse.

You know the way out of this impasse, and have used it for years. We restrict the domain of f to non-negative values only, to give us the function in Figure 3-5. Here there is no trouble, for $g(1) = 1$ without question. $g(x)$ is denoted by $\sqrt{x}$, and the above process is usually expressed in high school by some less sophisticated statement like "An indicated square root is always assumed to have the plus sign."

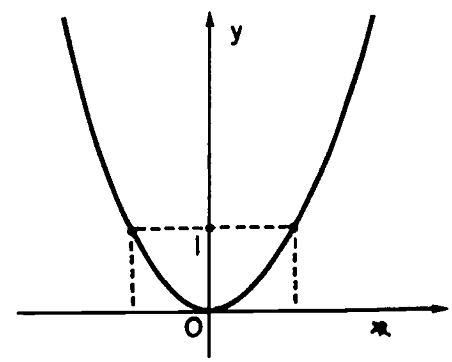


Figure 3-4

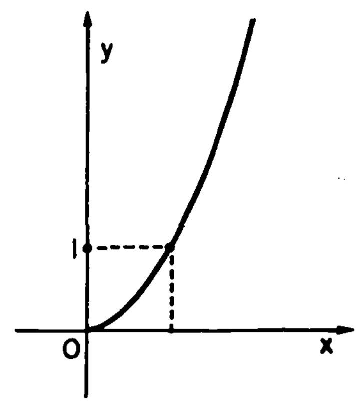


Figure 3-5

We can see from this example that in order for a function, f, to possess an inverse it is necessary that f be one-to-one. That is, not only does each input value determine a unique output, but also different inputs must yield different outputs. In the case that the function f is continuous (the only case we will be interested in) this means that f must be strictly monotone (i.e. either strictly increasing or strictly decreasing.) We leave it as an exercise for you to verify the truth of this assertion.

A more difficult case than the square-root function is the sine function. If g is to be the inverse function then what, for example, is g(1/2)? Is it $-7\frac{\pi}{6}$, $\frac{\pi}{6}$, $5\frac{\pi}{6}$, or what? Here again we must make a drastic reduction in the domain of the function before we can talk about an inverse. Now the values of sin x run from -1 to 1 inclusive; i.e. the range of f is [-1,1]. This, then, should be the domain of g, since in g(y) we should allow y to assume any value taken by f(x). In accordance with the previous paragraph we therefore want a monotone piece of the y = sin x curve running from y = -1 to y = 1. There are obviously plenty of these available; over the inrervals $[-\frac{3\pi}{2},-\frac{\pi}{2}]$, $[-\frac{\pi}{2},\frac{\pi}{2}]$, $[\frac{\pi}{2},\frac{3\pi}{2}]$, or in general $[(n+\frac{1}{2})\pi,(n+\frac{3}{2})\pi]$, where n is any integer. If we have to standardize by choosing one of these the most reasonable one seems to be $[-\frac{\pi}{2},\frac{\pi}{2}]$.

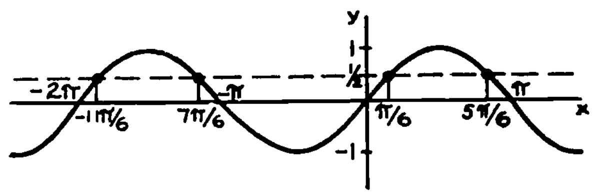


Figure 3-6

Making this choice, we define

$g(x) = \arcsin x, \quad -1 \leq x \leq 1$

as that inverse of the sine function that satisfies

$$-\frac{\pi}{2} \leq \arcsin x \leq \frac{\pi}{2}.$$

The graph of

$y = \arcsin x,$

is shown in Figure 3-7.

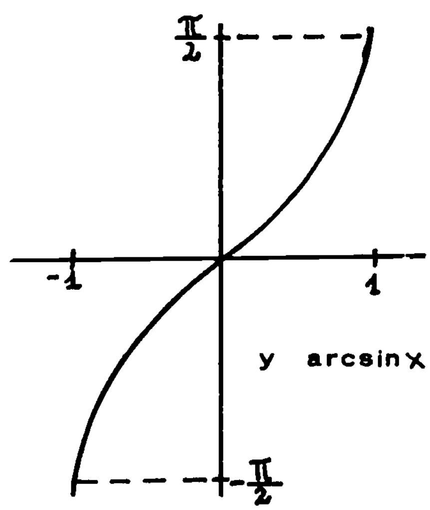


Figure 3-7

Another common notation for the function inverse to the sine function is $\sin^{-1}x$. The exponent -1 is used here to designate the "inverse of the function sin" in the same way as 2^{-1} designates the "inverse of the number 2." This is a reasonable notation, but unfortunately it can get confused with the common notation $\sin^2 x$, which means $(\sin x)^2$. One tends to think that $\sin^{-1}x$ means $(\sin x)^{-1}$ which is most definitely not the case. In this book we shall stick to the "arcsin" notation.

In a similar way we can define the inverses of the other trigonometric functions, but arccos and arctan are the only others that are ordinarily used. Arctan is easily handled and we leave it as an exercise to show by a graph that for

$$-\frac{\pi}{2} < \arctan x < \frac{\pi}{2}$$

and all values of x we have an inverse of the tangent function over the domain $(-\frac{\pi}{2}, \frac{\pi}{2})$.

The sine and cosine curves are very similar and hence the inverse functions are much alike. The choice of a monotone piece of the cosine curve is perhaps not quite as obvious as for the sine but still the one between 0 and π seems the logical choice. This gives the restriction

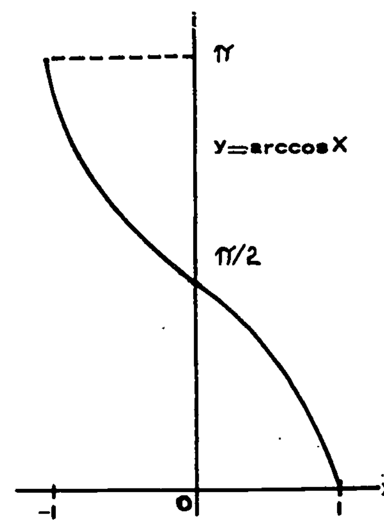


Figure 3-8

$$0 \leq \arccos x \leq \pi,$$

with $-1 \leq x \leq 1$ as before. The graph is shown in Figure 3-8.

Now we come to the differentiation. From

$$y = \arcsin x$$

we get

$$\sin y = x$$

and so

$$(\cos y)y' = 1$$

$$y' = \frac{1}{\cos y} .$$

Now since y satisfies $-\frac{\pi}{2} \leq y \leq \frac{\pi}{2}$ it is an angle in the first or fourth quadrants and so cos y is positive. Hence

$$\cos y = \sqrt{1 - \sin^2 y} = \sqrt{1 - x^2} ,$$

and we get finally

$$D_x \arcsin x = \frac{1}{\sqrt{1 - x^2}} .$$

In a similar manner we derive

$$D_x \arccos x = -\frac{1}{\sqrt{1 - x^2}}$$

$$D_x \arctan x = \frac{1}{1 + x^2} .$$

Consider, in general, two functions f and g that are inverses of each other. If f(a) = b we must have g(b) = a.

Or, in other words, if the curve y = f(x) goes through the point (a,b), then y = g(x) goes through (b,a). (Figure 3-9). These two points are symmetric with respect to the line y=x, and hence the two curves y = f(x) and y = g(x) are likewise symmetric since their points can be paired off in this symmetric fashion. Also, the tangent lines to the two curves at corresponding points, being the limits of secant lines through corresponding points, are symmetric. We leave to the reader the proof that two lines, symmetric to y=x, have reciprocal slopes unless one is vertical and the other horizontal. Thus $g'(b) = \frac{1}{f'(a)}$.

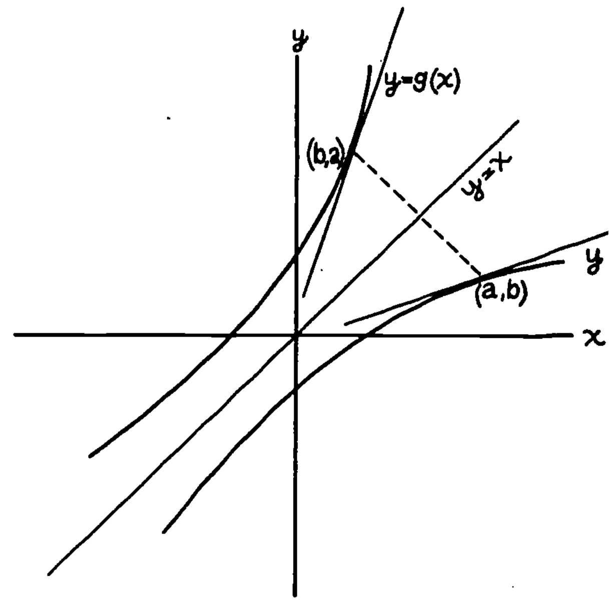


Figure 3-9

An analytic proof of this result follows at once from the chain rule. For we have

$$g(f(x)) = x$$

for all x in a suitable domain and so

$$g'(f(x))f'(x) = 1 \ .$$

For x=a, f(x) = b; hence

$$g'(b)f'(a) = 1$$

or

$$g'(b) = \frac{1}{f'(a)}$$

as before.

Bear in mind that the results of this section, since they are based on the differentiation of implicit functions, are subject to our general assumption that the desired functions exist and are continous and differentiable. For the special case of inverse functions it is not difficult to prove this, and a proof, together with a more careful general discussion of implicit functions, is given in Appendix A of this Chapter.

PROBLEMS

1. Derive the formulas for the derivative of arccos x and arctan x.

2. Prove that two lines symmetric with respect to the line $y = x$ have reciprocal slopes if neither of them is horizontal.

3. Differentiate the following and simplify the result when possible.

   (a) $\arcsin 2x$

   (b) $\arccos 5x$

   (c) $\arctan \frac{1}{x}$

   (d) $\arcsin \sqrt{x}$

   (e) $\arcsin \sqrt{1 - t^4}$

   (f) $(1 + \arcsin 3x)^2$

   (g) $x \arcsin x + \sqrt{1 - x^2}$

   (h) $\arctan \sqrt{x^2 - 1}$

   (i) $\arcsin x - x\sqrt{1 - x^2}$

   (j) $\sqrt{\arcsin 3x}$

   (k) $\arctan \frac{1 + 2x}{2 - x}$

   (l) $x \arccos x$

   (m) $\arccos \sqrt{y}$

   (n) $y^2 \arccos 2y$

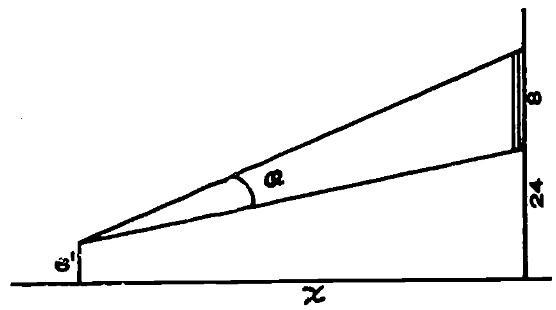
   (o) $\arctan (3 \tan x)$

   (p) $\arctan \left[\frac{3 \sin x}{4 + 5 \cos x}\right]$

4. Are sin(arcsin x) and arcsin(sin x) always equal to x?

5. Let $f(x)$ = arcsin x + arccos x and find $f'(x)$. What can you conclude about $f(x)$? Be as specific as possible.

6. A cathedral window 8 ft high has its bottom 24 ft from the floor. A tall tourist whose eyes are 6 ft from the floor wants to view the window so that it subtends the largest angle to his eye. How far from the wall should he stand?



7. (a) Find the first five derivatives of arctan x. (See Problem 2(b), Section 1).

(b) These derivatives suggest that the n-th derivative of arctan x is of the form

$$P_n(x)(x^2 + 1)^{-n}.$$

If this is true for the n-th derivative show that it will be true of the (n + 1)th derivative provided

$$P_{n+1}(x) = (x^2 + 1)P_n'(x) - 2nx\, P_n(x).$$

(c) Use this recursion formula to find two more derivatives of arctan x.

(d) Use your results to write out Taylor's Theorem for arctan x, with $a = 0$ and $n = 6$.

4. Related Rates.

In Chapter 5 we spoke of the velocity of a moving point as "rate of change of distance" and later on of acceleration as "rate of change of velocity." In general, when any quantity is a function of time we may refer to the derivative of this function as the "rate of change" of the quantity. Some writers even extend the use of the phrase to include any derivative, using such expressions as "the rate of change of the area of a circle with respect to its radius." We shall not do this.

If two related quantities are changing with time, we can often determine the rate of change of one if we know that of the other. The analytic machinery for doing this is somewhat like that used in the preceding section, but here we do not have to worry about the existence of our functions since this is presumably assured by the physical situation.

Example 1. A point is moving on the curve $x^2 + y^2 = 25$ in such a way that when it is at (3,4) its projection on the x-axis is moving towards the origin at the rate of 12 units/sec. How fast and in what direction is its projection on the y-axis moving? (Figure 4-1).

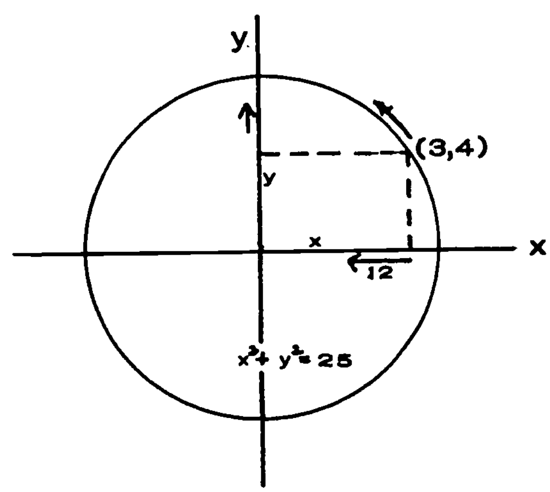


Figure 4-1

Here we have $x = f(t)$, $y = g(t)$, such that $f(t)^2 + g(t)^2 = 25$, and for some t_0 we have $f(t_0) = 3$, $g(t_0) = 4$, $f'(t_0) = -12$. We want to get $g'(t_0)$. Differentiating the identity in t gives

$$2f(t)f'(t) + 2g(t)g'(t) = 0,$$

and putting $t=t_0$ and substituting the known values gives $g'(t_0) = 9$. That is, the projection on the y-axis is moving upward with velocity 9 units/sec.

As in Section 2, we usually simplify the notation, writing merely x, y, x', y' instead of $f(t)$, $g(t)$, $f'(t)$, $g'(t)$.

Example 2. In Figure 4-2 BAC is a link belt of length 12 in. that will bend around a pulley at A but will not compress where it lies on a table along AC. A cord BCW, tied to the belt at B and sliding over the end at C, supports a weight W. At the moment when B is 4" above the table and is being pulled up at the rate of 2 in./sec, does W go up or down and how fast?

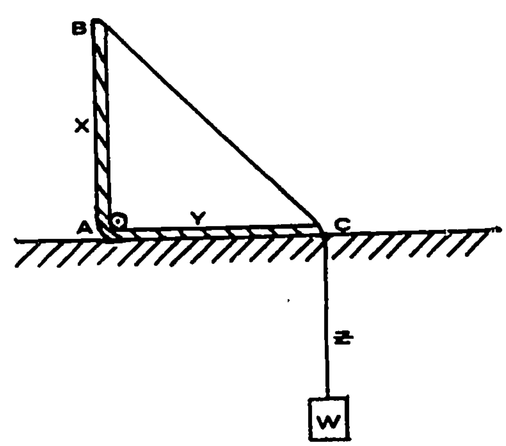


Figure 4-2

Letting x, y, z, stand for the values of AB, AC, CW, we have the equations

$$x + y = 12, \qquad \sqrt{x^2 + y^2} + z = L,$$

if L is the length of the string. Here we have three variables. We could eliminate y, in which we are not interested, but it is easier just to differentiate both equations and then eliminate y'. Thus:

$$x' + y' = 0, \quad \frac{1}{2}(x^2 + y^2)^{-1/2}(2xx' + 2yy') + z' = 0.$$

Now putting in the values x=4, x'=2 gives at once y=8, y'=-2, and then $z' = \frac{2}{\sqrt{5}} \approx .894$. Hence W is falling at about 0.9 in./sec.

Radicals are always annoying to differentiate, so we may prefer to get rid of the square root by transposing z in the second equation and squaring. This gives $x^2 + y^2 = (L-z)^2$, which differentiates to

$$2xx' + 2yy' = -2(L-z)z' = -2\sqrt{x^2 + y^2}\, z' ,$$

and the solution proceeds as before.

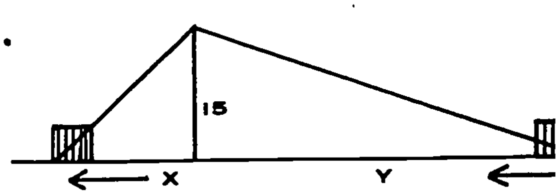
PROBLEMS

1. A balloon is being inflated at the rate of 8 in.^{3}/sec. Find the rate of change of the radius of the balloon when the radius of the balloon is $\frac{1}{2}$ ft.

2. Sand is being poured onto a conical pile at the rate of 9 ft^3/min. Due to the friction forces, it is known that the slope of the sides of the conical pile is always $\frac{2}{3}$. How fast is the altitude increasing when the radius of the base of the pile is 6 ft?

3. A ladder 25 ft. long is leaning against a wall, with the bottom of the ladder 7 ft. from the base of the wall. If the lower end is pulled away from the wall at the rate of 1 ft./sec. find the rate of descent of the upper end along the wall. Approximate this descent at the end of 8 sec.

4. A rope 35 ft. long runs over the top of a wall 12 ft high. Each end is attached to a heavy block which slides on the ground. One block is 16 ft. away from the foot of

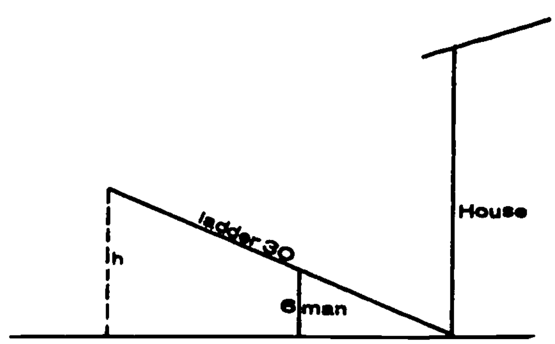


the wall and is being pulled farther away at the rate of 30 ft./min. How fast is the other block approaching the wall, assuming the rope is taut?

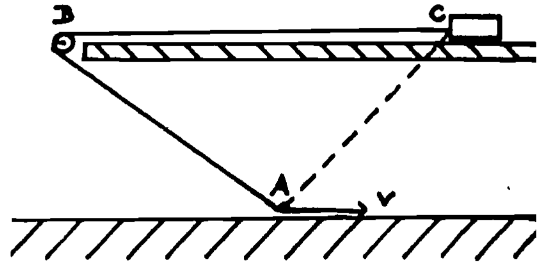
5. A man starts walking eastward at 5ft/sec from a point A. Ten minutes later a man starts walking west at the same rate from a point B, 3000 ft north of A. How fast are they separating 10 min. after the second man starts?

6. One way to lean a long light ladder against a house is to "walk-it-up". Propping the foot of the ladder against the house, you start at the far end, holding the rungs above your head, and walk towards the house, shifting your hands from rung to rung as you do so. When the ladder is against the eaves you steady it with one hand and with the other lift it slightly and move the other end a suitable distance away from the house.

A man raising a ladder 30 ft. long is holding the rungs 6 ft. above the ground and is walking towards the house at the rate of 2 ft/sec. At the moment that he is 8 ft. from the house, how fast is the high end of the ladder rising?



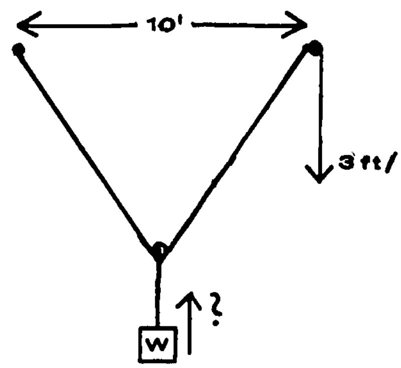
7. A rope ABC passes over a pulley at B, thereby enabling a man at end A to slide a weight at end C along a horizontal elevated platform. If A moves to the right with velocity v, express the speed of C and the rate of change of the distance AC in terms of v and the angles ABC and A



Ans. $(BC)' = -v \cos \angle ABC$

$(AC)' = -v \cos \angle ACB\ (1 + \cos \angle ABC)$.

8. (a) A weight is hung on a rope and pulleys as shown. If the free end of the rope is pulled down at 3 ft/sec, how fast is the weight rising when the lower pulley is 12 ft. below the level of the upper one?



(b) What answer do you get if you replace the 12 ft. in (a) by $\frac{1}{10}$ inch? Does this sound reasonable? What is wrong?

5. Some Notation.

Since elementary school you have been familiar with the use of letters to represent physical quantities. For example, if a body is falling we may use t to represent the time since it was dropped, s its height above ground, and v its downward velocity, all expressed in appropriate units. In a specific case these three quantities are not independent; in fact, any one of them determines the other two. There are therefore six functional relations of the type

(1) $$s = f(t), \quad v = g(s), \quad v = h(t), \quad \text{etc.},$$

these are illustrated in Figure 5-1. The functions are not independent either, but satisfy such relations as

$$h(t) = g(f(t)), \quad \text{etc.}$$

To avoid too much notation it is customary to use one symbol for each quantity, regardless of whether we are considering that quantity to be an independent variable, a function of some other quantity, or a particular value of one of these. Thus instead of giving names to the functions as in (1) and writing the chain rule as

$$h'(t) = g'(s)f'(t)$$

we can merely write directly

$$D_t v = D_s v \; D_t s$$

the subscripts telling us which function we are using in each differentiation.

If one wishes at some stage to emphasize that at that moment v is being considered as a function of s one can write v(s). Using this notation the chain rule could be written in still another form:

$$v'(t) = v'(s)\, s'(t).$$

Remember that notation is introduced for convenience, and it's up to you to pick the type that you regard as most convenient. Remember also that compact notation often conceals the true state of affairs; if you are in any doubt about its meaning, go back to the exact notation in terms of separately defined and designated functions.

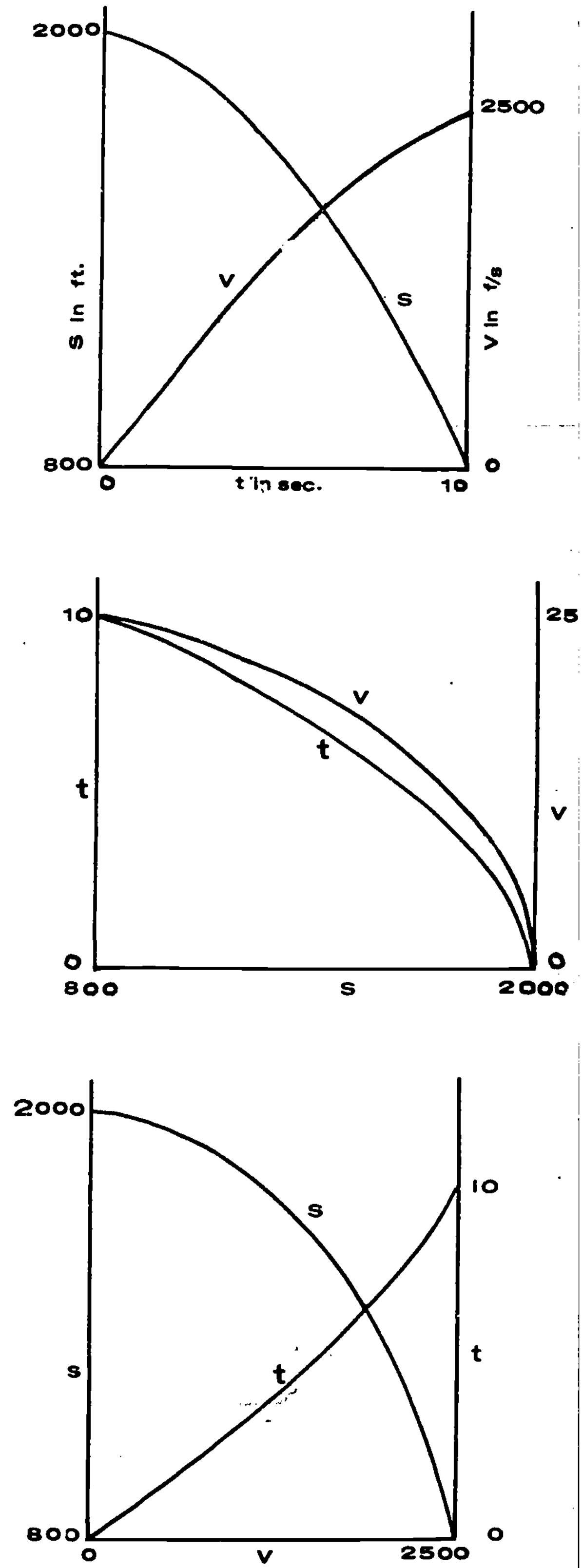


Figure 5-1

6. Differentials.

Much of the common notation in calculus is related to the notion of a "differential," and we start the study of this concept by introducing some of this notation.

If f is differentiable at x the Lemma of Section 1 can be written, in terms of "increment" notation, as

$$\Delta f(x) = f'(x)\Delta x + z(x + \Delta x)\Delta x,$$

where z(x) = 0. This is essentially the same form as Equation (1) at the end of Section 1. In Figure 6-1 (which pictures two cases, one for positive Δx and one for negative) $\Delta f(x)$ is the signed distance AQ. The signed distance AT is the linear approximation to $\Delta f(x)$ given by the first term on the right-hand side, $f'(x)\Delta x$. This approximation is linear in Δx and is, moreover, defined for all values of Δx, whereas the exact expression is defined only when $x + \Delta x$ lies within the range of definition of f(x). The concept of a differential is suggested by these properties of the linear approximation.

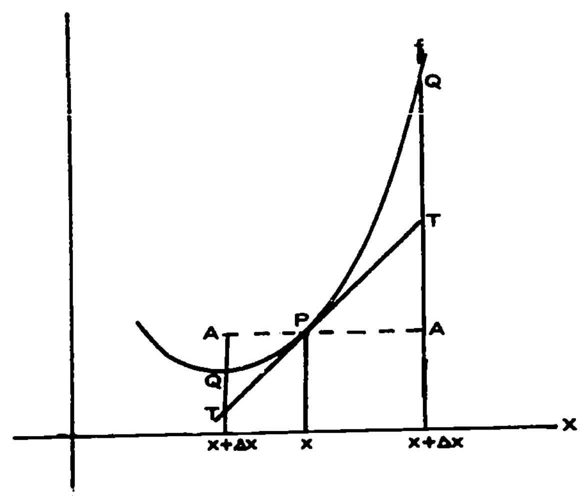


Figure 6-1

Definition. If f is differentiable at x, the differential of f at x is the function df defined by

$$h \to f'(x)h, \quad h \text{ in } (-\infty,\infty) .$$

If f and g are functions of the same variable x, and if $g'(x) \neq 0$, then the quotient of the two differentials, $\frac{df}{dg}$, is the mapping

$$h \to \frac{f'(x)h}{g'(x)h} = \frac{f'(x)}{g'(x)} .$$

That is, $\frac{df}{dg}$ is a constant function, its value not depending on h. We find it is convenient, and leads to no confusion, if we merely write

(1) $$\frac{df}{dg} = \frac{f'(x)}{g'(x)} .$$

Now consider the special case $g(x) = x$, that is, $g(x)$ is the identity function that maps $x \to x$. Then dg is also the identity function mapping $h \to h$. By rights we should have a special symbol, say I, for this function and write dI as its differential. However, it turns out to be much more useful to use dx as the differential of the identity function of the variable x. Then (1) becomes the important relation

(2) $$\frac{df}{dx} = f'(x) ,$$

which can also be written as

(3) $$df = f'(x)dx .$$

We are now ready to tackle the more complicated case of three related variables and to introduce the notation of Section 5. Consider the velocity v, which may be considered a function of either s or t. As long as functional notation is used, as in (2) or (3), we have no trouble, for if we write, say, $dv = v'(s)ds$ we know what function of v we are using and no confusion arises. But suppose the symbol dv appears in an equation with no indication of what v is a function of - what are we to assume? The interesting thing about differentials, the property that makes them so handy to manipulate, is that <u>it doesn't matter</u>. Whether v is considered a function of s or a function of t, or of any other related variable, dv is the same.

<u>Proof</u>. Let $v = v(s)$ and $s = s(t)$. Then v is a function of t, which we call $v_1(t)$, and $v_1(t) = v(s(t))$. Then

$dv_1 = v_1'(t)dt = v'(s)s'(t)dt = v'(s)ds = dv$

by the chain rule and repeated use of (3).

Equation (2) gives us another notation for the derivative: if $y = f(x)$ then

$$f'(x) = y' = D_x f(x) = D_x y = \frac{df}{dx} ,$$

to which we add

$$f'(x) = \frac{dy}{dx} = \frac{df(x)}{dx} = \frac{d}{dx}f(x) .$$

The last one is purely for convenience in indicating the derivative of a complicated function. The symbol $\frac{dy}{dx}$ is the one in most common use, since its relation to differentials gives it a flexibility that the others lack. As one example consider the expression for the chain rule: it is merely

$$\frac{dy}{dx} = \frac{dy}{dz}\frac{dz}{dx},$$

a triviality. (Of course this is not a <u>proof</u> of the chain rule, since the chain rule was used in proving the property of differentials that enables this equation to hold.) On the other hand, the $\frac{dy}{dx}$ notation is awkward if one wants to specify a particular value for x. Thus y'(2) must be written in some such form as $\left(\frac{dy}{dx}\right)_{x=2}$. One can, of course, switch from one notation to another, as convenient, but this is apt to confuse the reader and should be done only in moderation.

The basic formulas and the technique of differentiation can be reconsidered from the point of view of differentials. For example, Theorem 4 of Section 5-6,

$$D_x[f(x)g(x)] = f'(x)g(x) + f(x)g'(x),$$

can first be rewritten as

$$\frac{d(fg)}{dx} = \frac{df}{dx}g + f\frac{dg}{dx}$$

and then, on multiplying by dx,

$$d(fg) = gdf + fdg .$$

The other theorems in this group can be reformulated similarly

As an illustration of the application of these formulas consider Example 1 of Section 2. Instead of differentiating $x^3 + y^3 = 3xy$ with respect to x we can for the moment ignore the question of which is the independent variable and take differentials. This gives

$$3x^2dx + 3y^2dy = 3(xdy + ydx) .$$

Now dividing by dx and solving for $\frac{dy}{dx}$ gives the same equation,

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

as before.

It is natural now to ask "What about second derivatives?" Here, alas, the differential notation fails us. It is possible to define something called the "second differential" but it lacks the one property that makes the "first differential" so useful, namely its constancy under change of independent variable. Second and higher differentials were commonly used in mathematics about eighty years ago but today they are almost obsolete.

However, we do need some notation for higher derivatives to go along with $\frac{dy}{dx}$. Of course

$$y'' = \frac{dy'}{dx} = \frac{d\left(\frac{dy}{dx}\right)}{dx} = \frac{d}{dx}\left(\frac{dy}{dx}\right)$$

and

$$y''' = \frac{dy''}{dx} = \frac{d}{dx}\left(\frac{d}{dx}\left(\frac{dy}{dx}\right)\right) ,$$

but this compounding process will hardly do for, say, $y^{(7)}$ and not at all for $y^{(n)}$. The way out has been to invent a symbol that is compact but suggests its origin. For y''' we write $\frac{d^3y}{dx^3}$, suggesting the three d's upstairs and the three dx's downstairs in the expression for y'''. It is important to remember that $\frac{d^3y}{dx^3}$ is a single symbol - it is not a quotient of something called d^3y by the cube of the differential dx.

If you need to handle a lot of higher derivatives (as, for example, in Taylor's Theorem) it is probably best to abandon the d-notation and use primes, or perhaps D_x. The latter works quite well since we can, for example, abbreviate $y''' = D_x(D_x(D_xy))$ to D_x^3y and $y^{(n)}$ to D_x^ny .

PROBLEMS

1. In each of the following find dy in terms of dx and any of the variables that appear.

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

   (b) $y = (x^2 + 1)/(x + 1)$,

   (c) $x^2 + xy + y^2 = 1$,

   (d) $x \cos y = y \sin x$,

   (e) $z = \cos x, \quad y = \sin z$,

   (f) $x = t \sin t, \quad y = t \cos t$,

   (g) $x^2 + u^2 = 4, \quad y^3 + u^3 = 8$,

   (h) $z^2 + xz - 2x = 0, \quad x^3 + y^3 + z^3 = 24$.

2. Given that s is a function of t, and $v = \frac{ds}{dt}$, show that:

   (a) $$\frac{d^2s}{dt^2} = v\frac{dv}{ds},$$

   (b) $$\frac{d^3s}{dt^3} = v\left(\frac{dv}{ds}\right)^2 + v^2\frac{d^2v}{ds^2}.$$

3. u, v, and w are three functions of the same variable, having the property that their product is constant.

   (a) Find a relation between the differentials of u, v, and w. [Hint. First consider uvw as (uv)w.]

   (b) If u is taken as the independent variable what is $D_u v$ in terms of $D_u w$?

7. Parametric Equations.

Having found the tangent to $x^2 + y^2 = 25$ at the point (3,4) in three different ways we shall now do it in a fourth way. Associated with any point (x,y) on the circle is an angle θ formed by the x-axis and the radius to (x,y), and we have

(1) $$x = 5\cos\theta,$$
$$y = 5\sin\theta.$$

The extra variable θ that we have introduced is called a parameter, and equations like (1), expressing x and y as functions of the parameter, are called parametric equations.

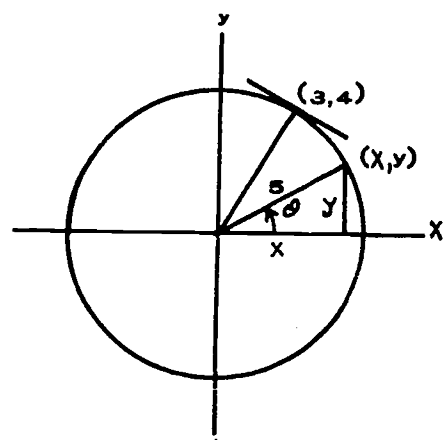


Figure 7-1

We now have a situation somewhat like the one discussed in Section 5. From (1) we get

$$dx = -5\sin\theta\, d\theta,$$
$$dy = 5\cos\theta\, d\theta\ ,$$

and so

$$D_x y = \frac{dy}{dx} = -\frac{5\cos\theta}{5\sin\theta} = -\frac{x}{y} = -\frac{3}{4}\ ,$$

as before. The division by dx is legitimate, since at (3,4) we have $dx = -4d\theta \neq 0$.

Many curves are more easily handled in parametric form than otherwise. Here is a classic case.

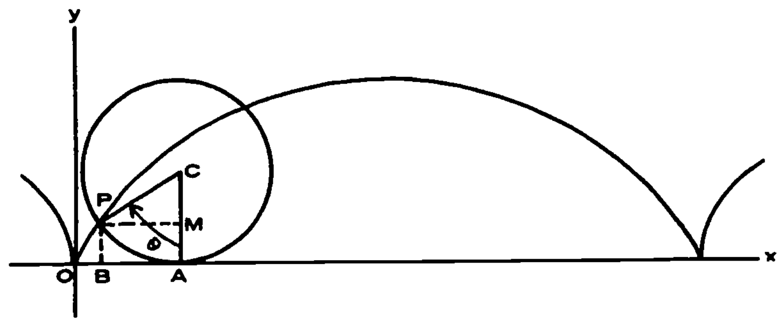


Figure 7-2

Example 1. A wheel of radius a is rolling without slipping along a straight road. The path of a point on the rim of the wheel is called a cycloid. To get equations for the cycloid we let the wheel roll along the x-axis, with the origin where the point P on the rim touched the axis, and take as parameter the angle through which the wheel has rolled from this point (Figure 7-2). Since there is no slipping,

$$OA = \text{arc } AP = a\theta ,$$

if a is the radius of the wheel. Then

$$x = OB = OA - PM = a\theta - a \sin \theta ,$$

$$y = BP = AC - CM = a - a \cos \theta .$$

The slope of the tangent line at P is

$$\frac{dy}{dx} = \frac{a \sin \theta \quad d\theta}{(a - a \cos \theta) d\theta}$$

$$= \frac{\sin \theta}{1 - \cos \theta} .$$

If we want $\frac{d^2y}{dx^2}$ we simply use

$$\frac{d^2y}{dx^2} = \frac{d(\frac{dy}{dx})}{dx}$$

$$= [\frac{(1 - \cos \theta)\cos \theta - \sin \theta \sin \theta}{(1 - \cos \theta)^2} d\theta] \div [a(1 - \cos \theta) d\theta]$$

$$= \frac{\cos \theta - 1}{a(1 - \cos \theta)^3}$$

$$= \frac{-1}{a(1 - \cos \theta)^2} = -\frac{a}{y^2} .$$

559

PROBLEMS

1. Prove: A line through a point of a cycloid, perpendicular to the tangent at that point, passes through the lowest point of the wheel.

2. Find $D_x y$ and $D_x^2 y$ in each of the following.

   (a) $x = t^2$, $\qquad y = t + 1$.

   (b) $x = \sec\theta$, $\qquad y = \tan\theta$.

   (c) $x = \sin^3 t$, $\qquad y = \cos^3 t$.

   (d) $x = \dfrac{t^2 - 1}{t^2 + 1}$ $\qquad y = \dfrac{2t}{t^2 + 1}$

   What curve is this?

   (e) $x = 2\sin t - \cos 2t$, $\qquad y = 2\sin t - \sin 2t$.

3. If $x = f(t)$, $y = g(t)$;

   (a) Show that $\dfrac{dy}{dx} = \dfrac{g'(t)}{f'(t)}$.

   (b) Show that $\dfrac{d^2y}{dx^2} = \dfrac{f'(t)g''(t) - g'(t)f''(t)}{[f'(t)]^3}$.

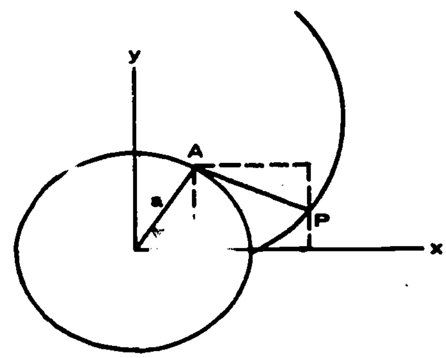
4. Using the result of 3(b) find $\dfrac{d^2y}{dx^2}$

(i) for the circle,

(ii) for the cycloid.

Compare the former with the result obtained from implicit differentiation.

5. (a) A string, wrapped around a cylinder of radius a, is held at a point P and unwound, being kept taut in the process. Find parametric equations for the path of P, find $D_x y$ and $D_x^2 y$, and determine any interesting geometric properties you can. This curve is called the <u>involute</u> of the circle. [Suggestion: Use the angle θ for the parameter.]



(b) Describe the involute of a square.

6. (a) The equation $x^3 + y^3 = 3xy$ of Example 1, Section 2, can be parametrized by setting $y = tx$. Show that this leads to

$$x = \frac{3t}{1 + t^3}, \quad y = \frac{3t^2}{1 + t^3}.$$

(b) Find $D_x y$ as a function of t.

(c) By setting this expression for $D_x y$ equal to zero and solving, we get $t = \sqrt[3]{2}$ and $t = 0$. The former agrees with the result of Example 2, Section 2, but the root $t = 0$ seems to give us something new. Explain why we get this solution from the parametric equations but not from the implicit function approach.

7. In the parametric equations of the cycloid, solve for θ in terms of y and substitute to get x as a function of y. Find $D_x y$ from this equation and show that it agrees with the result of Example 1. Which method do you prefer for finding $D_x y$?

APPENDIX A

Theory of Inverse Functions

Let f be a strictly monotone and continuous function defined over the interval [a,b] and g its inverse function. According to the intermediate value theorem for continuous functions f(x) assumes all values between f(a) and f(b), so that the domain of the inverse function g, is the interval between f(a) and f(b). Thus for any pair of inverse functions the domain of one is the range of the other.

It is also easily checked that g is strictly monotone in the same sense as f(i.e. increasing if f is increasing, decreasing if f is decreasing).

Furthermore we can show that g is also continuous. In order to show this we must be able, for every number c between f(a) and f(b) and every number $\varepsilon > 0$, to find a number δ so that for all y within a distance δ of c we have g(y) within ε of g(c).

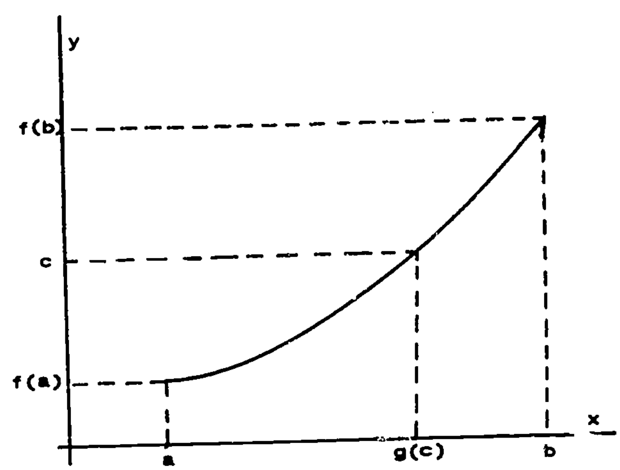


Figure A-1

The monotonicity of f and g makes it very simple to find what is actually the best possible value of δ. The method amounts to a four step algorithm.

I. Select a number c, between f(a) and f(b) and choose $\varepsilon > 0$. Locate g(c) on the x-axis.

II. Locate $g(c) - \varepsilon$ and $g(c) + \varepsilon$ on the x-axis. Call these numbers α and β for short.

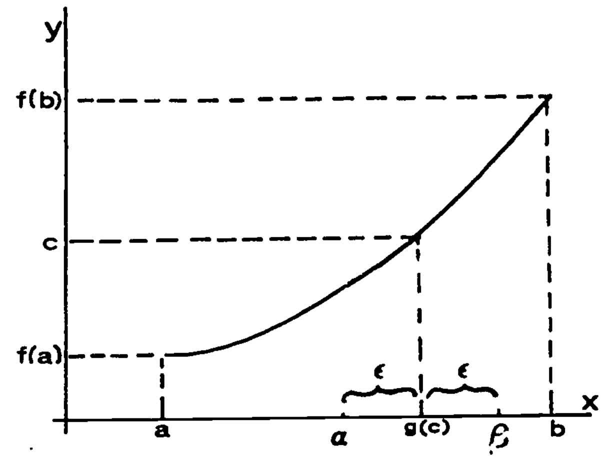


Figure A-2

III. Locate $f(\alpha)$ and $f(\beta)$ on the y-axis. Call them r and s for short. Note that, owing to the monotonicity of the function g, all numbers between r and s will be mapped by g into the

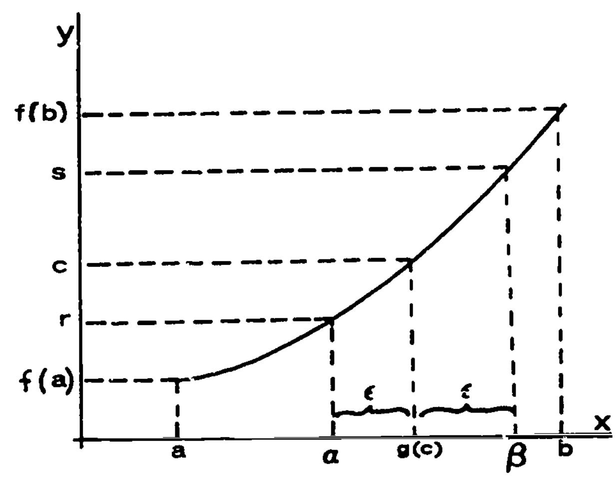


Figure A-3

interval $[\alpha,\beta]$ and hence lie within ε of $g(c)$.

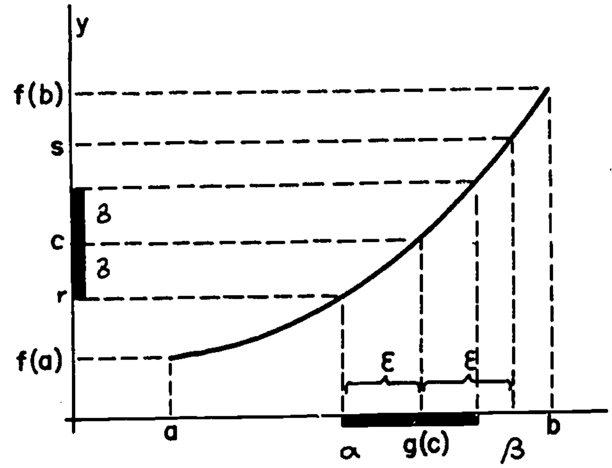


Figure A-4

iv. Letting δ be the smaller of the distances from c to r and from c to s we now see, in figure A-4 that if y is within a distance δ of c then $g(y)$ is within a distance ε of $g(c)$. This is what was to be proved.

(Darkened interval on x-axis is image under g of interval $[c-\delta, c+\delta]$.)

[More briefly but probably less clearly we could have said: "Let $\delta = \min\{|f(g(c)-\varepsilon) - c|, |f(g(c)+\varepsilon) - c|\}$"

Finally we can establish the required differentiation theorem.

Theorem. If f is strictly monotone and $f'(x_0)$ exists and is not zero, if g is the inverse of f, and if $y_0 = f(x_0)$, then $g'(y_0)$ exists and is equal to $\frac{1}{f'(x_0)}$.

Proof. Since f and g are inverse functions

$$\frac{g(y) - g(y_0)}{y - y_0} = \frac{g(y) - g(y_0)}{f(g(y)) - f(g(y_0))}$$

$$= \frac{g(y) - g(y_0)}{[g(y) - g(y_0)]\,[f'(g(y_0)) + z(g(y))]}$$

by the Lemma of Section 1. Now for $y \neq y_0$ we have $g(y) \neq g(y_0)$, since g is strictly monotonic. Hence as $y \to y_0$ the quantities $g(y) - g(y_0)$ are never zero and so can be removed as factors from the numerator and denominator. That is,

$$\lim_{y \to y_0} \frac{g(y) - g(y_0)}{y - y_0} = \lim_{y \to y_0} \frac{1}{f'(g(y_0)) + z(g(y))}$$

$$= \frac{1}{f'(g(y_0))}$$

$$= \frac{1}{f'(x_0)} ,$$

since, as we saw in an earlier proof,

$$\lim_{y \to y_0} z(g(y)) = z(\lim_{y \to y_0} g(y)) = z(g(y_0)) = 0.$$

Chapter 8

THE CONNECTION BETWEEN DIFFERENTIATION AND INTEGRATION

1. Anti-derivatives.

A rich source of mathematical problems from elementary physics is related to "falling body" problems. Consider an object or "body" dropped from a height and allowed to fall to the ground and let $v(t)$ represent its velocity at time t. As you are aware this velocity tends to increase during the time of fall, which is why it hurts more to jump off a ten story building than off a 5-foot [illegible].

The acceleration of a moving object is defined as the time rate of change of velocity, that is

$$a(t) = v'(t) .$$

In a very highly simplified model of Newton's theory of gravitation, the acceleration due to gravity is (for objects falling near the surface of the earth) the same for all objects and is independent of the time of fall; that is to say, it is constant. When distances are measured in feet and times in seconds the value of this constant is about 32; at least, this is the value we will use in our modeling of this situation.

We find then that $v'(t) = 32$. Now the question arises, "Can we, from this information, find the velocity function itself?"

This is just one instance of the more general problem: Given a formula for finding the values of $f'(x)$, say, like

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

can one find a formula for the values of $f(x)$?

At first the situation seems somewhat discouraging, for both the functions

$$g(x) = x^3 + 5x \quad \text{and} \quad h(x) = x^3 + 5x + [illegible]$$

have $3x^2 + 5$ as their derivatives. If many functions can have the same derivative then we cannot hope to determine a function completely from knowledge of its derivatives alone. The next two theorems, however, show that the indeterminacy is of a particularly simple kind.

<u>Theorem I.</u> If f is continuous in $[a,b]$ and $f'(x) = 0$ for all x in (a,b) then $f(x)$ is constant in $[a,b]$.

Proof. Let d be any point in (a,b) and let $c = f(d)$. For any x in [a,b], $x \neq d$, we can apply the MVT to either [d,x] or [x,d], depending on whether $x > d$ or $x < d$. In either case we get

$$f(x) - f(d) = (x - d)f'(\xi) ,$$

with ξ between x and d. Since ξ is in (a,b), $f'(\xi) = 0$, and so

$$f(x) = f(d) = c$$

for any x in [a,b].

Theorem 2. If f and g are continuous in [a,b] and $f'(x) = g'(x)$ for all x in (a,b) then f(x) and g(x) differ by a constant in [a,b].

Proof. Define a function h by

$$h(x) = f(x) - g(x) \quad \text{for all } x \text{ in } [a,b].$$

Then

$$h'(x) = f'(x) - g'(x) = 0 \quad \text{for all } x \text{ in } (a,b).$$

Thus according to Theorem 1 we see that $h(x) = c$ in [a,b]. Since $h(x) = f(x) - g(x)$ we have

$$f(x) - g(x) = c \quad \text{for all } x \text{ in } [a,b].$$

We see then that f' really does determine f except for an additive constant. Suppose we want to find the most general anti-derivative of, say,

$$3x^2 + 5.$$

(By an anti-derivative of $3x^2 + 5$ we mean a function whose derivative is $3x^2 + 5$). In order to solve this problem we first find a particular anti-derivative like

$$x^3 + 5x$$

and then tack on a constant

$$x^3 + 5x + c .$$

In particular, for the falling body problem where

$$v'(t) = 16$$

we see that $v(t)$ will be given by

$$v(t) = 16t + c .$$

The value of c in a particular case can be found by giving the velocity at a particular time, most often when $t = 0$. Thus

$$v(0) = 16 \cdot 0 + c = c .$$

If the body drops from rest then the initial velocity at $t = 0$ is 0, so that $v(0) = c = 0$. If the body is thrown up or down then c will have some other value.

Looking at the situation geometrically we see that the anti-derivatives of a given function consist of a family of curves obtained by graphing one anti-derivative and translating this curve up and down. This is illustrated for $f'(x) = x^2$ in Figure I-I, where $f(x)$ has the form $f(x) = \frac{x^3}{3} + c$.

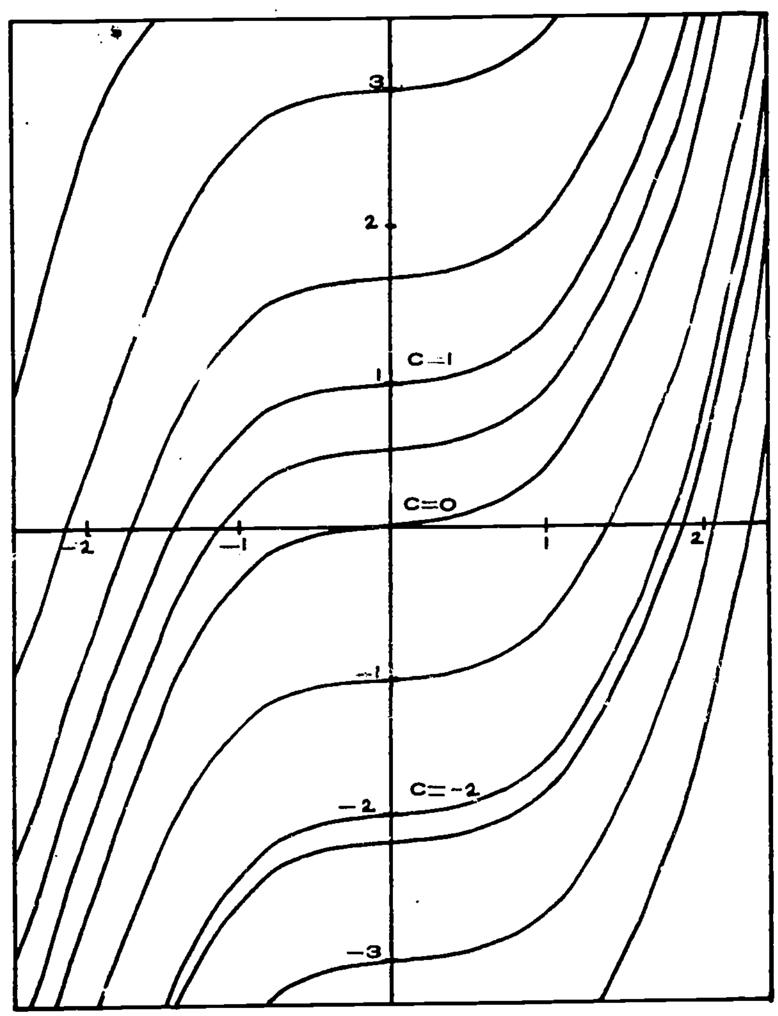


Figure I-I

As is easily checked by differentiation, an anti-derivative of

$$x^n \quad \text{is} \quad \frac{x^{n+1}}{n+1}$$

for any rational number n execpt n = -1. The most general anti-derivative of x^n is

$$\frac{x^{n+1}}{n+1} + c.$$

This observation enables us to find very easily the anti-derivatives of polynomials. For example the most general anti-derivative of

$$5x^3 + 4x^2 - 7x + 3$$

will be

$$\frac{5}{4}x^4 + \frac{4}{3}x^3 - \frac{7}{2}x^2 + 3x + c.$$

In fact, every differentiation formula can be reversed to give an anti-differentiation formula. For examnle, we have seen that the derivative of arcsin x is $1/\sqrt{1 - x^2}$; hence the anti-derivatives of $1/\sqrt{1 - x^2}$ are arcsin x + c.

The basic theorems on differentiation tell us that if functions are added, subtracted, or multiplied by constants their anti-derivatives behave the same way, but unfortunately there are no formulas for products, quotients, or composition of functions. Because of this we cannot in general find simple formulas for anti-derivatives, for instance, for $\sin(x^2)$, but in the next section we shall see that an anti-derivative can always be obtained as an integral.

We can, however, use the chain rule in a reverse way that vastly expands the set of functions we can handle. If

$$F(x) = G(u(x)),$$

where u is a function of x, we know from the chain rule that

$$F'(x) = G'(u(x))u'(x).$$

Now, if f is a function that can be written in the form

$$f(x) = g(u(x))u'(x),$$

for some g(x) and u(x), and if F and G are anti-derivatives of f and g, then we must have

$$F(x) = G(u(x)) + c.$$

Example 1. Although we cannot get a simple formula for an anti-derivative of $\sin(x^2)$ we can for $x \sin(x^2)$. For let $u(x) = x^2$. Then $u'(x) = 2x$, and

$$x \sin(x^2) = \frac{1}{2} \sin(u(x))u'(x).$$

Since $\frac{1}{2}\sin u$ has the anti-derivative $-\frac{1}{2}\cos u$, we get the anti-derivatives of $x \sin(x^2)$ to be $-\frac{1}{2}\cos(x^2) + c$.

This so-called "method of substitution" is much easier to handle in terms of differentials. Since $dF = F'(x)dx = f\,dx$, to find the anti-derivative of a function f amounts to the same thing as finding the "anti-differential" of dF. From this point of view the method of substitution takes the form

$$dF = f(x)dx = g(u)u'(x)dx = g(u)du;$$

i.e., dF, instead of being expressed in terms of x and dx, is expressed in terms of u and du. This accounts for the name of the method, i.e., u is substituted for x as the independent variable.

Example I now takes the shortened form;

$$dF = x \sin(x^2)dx = \frac{1}{2}\sin(x^2)d(x^2) = \frac{1}{2}\sin u\ du,$$

$$F = -\frac{1}{2}\cos u + c = -\frac{1}{2}\cos(x^2) + c.$$

In a simple case like this one usually omits all reference to u and jumps at once from the second expression for dF to the final answer.

PROBLEMS

1. In each of the following exercises, find an antiderivative of the given expression.

(a) $f(x) = 7 - 4x + 3x^2$

(b) $g(x) = 2x^3 - 3x^2$

(c) $f(x) = 3x^5 - 25x^3 + 60x$

(d) $g(t) = \frac{2}{3}(t + 2)^{3/2} - t$

(e) $f(x) = \frac{\sqrt{x}}{2} - \frac{2}{\sqrt{x}}$

(f) $f(x) = \sin x \cos x$

(g) $f(y) = \frac{\cos 2y}{\sqrt{\sin 2y}}$

(h) $f(\phi) = \sin 2\phi \cos 2\phi$

(i) $dF = (4x^3 - 5x + 7)\, dx$

(j) $dF = (x^3 + 2)^3\, dx$

(k) $dF = x^2(x^3 + 2)^3\, dx$

(l) $dF = \frac{x^3 + 2x^2 + 3}{x^2}\, dx$

(m) $dF = (y\sqrt{y} + \sqrt{y} - 5)\, dy$

(n) $dF = \sqrt[4]{x + 1}\ dx$

(o) $dF = \dfrac{\sqrt[3]{x^2} + \sqrt[4]{x}}{\sqrt{x}}\ dx$

(p) $dF = (\sqrt[3]{z} + 2z^{3/2})\ dz$

(q) $dF = 5(y - 6)^4\ dy$

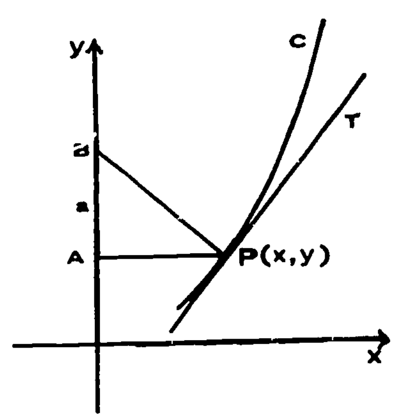
(r) $dF = \sin^7\theta \cos\theta\ d\theta$

2. An automobile tire rolls down an inclined plane 200 ft. long with an acceleration of 8 ft/sec^2. Find the position function of the tire if it is given no initial velocity. How long does it take the tire to reach the end of the plane?

3. A brick is thrown directly upward from the ground with initial velocity of 48 ft/sec. Assuming no air resistance, how high will the brick rise, and when will it return to the ground?

4. Starting from rest, with what constant acceleration must a car move to go 120 ft in 4 sec?

5. A curve C has the property illustrated in the figure. Here P is any point of the curve, PB is perpendicular to the tangent PT,



PA is parallel to the x-axis, and a is a constant distance between A and B.

(a) What are the possible equations of C?

(b) If C goes through the point (-2, -4) what are its possible equations?

2. The Fundamental Theorem of Calculus

Back in section 3-8 we obtained the formulas:

$$\int_0^x \cos t \, dt = \sin x \quad \text{and} \quad \int_0^x \sin t \, dt = 1 - \cos x.$$

These formulas were used in turn to derive inequalities such as

$$x - \frac{x^3}{3!} < \sin x < x - \frac{x^3}{3!} + \frac{x^5}{5!} \qquad \text{for } x > 0,$$

by means of which we were able to tabulate the sine and cosine functions.

This is just one of many uses of integrals with a variable upper limit of integration:

$$\int_a^x f(t)dt.$$

Here in this section as well as hereafter we will have much use of functions defined by such formulas as

$$F(x) = \int_0^x \sin t \, dt \qquad \text{or} \qquad F(x) = \int_1^x \frac{1}{t} dt,$$

or in general

$$F(x) = \int_a^x f(t)dt.$$

In the case that the integral, f(t), is positive and $x > a$ we have the area interpretation at our disposal. So

In Figure 2-1 we see that different values of x yield different areas, F(x).

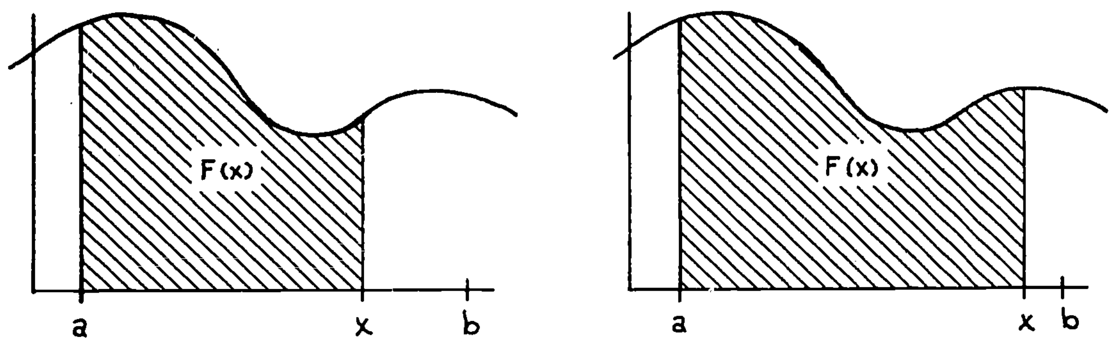


Figure 2-1

What we are most concerned with is the derivative of the function F. Accordingly let c be some number between a and b, and then from the definition of the derivative

$$(1) \qquad F'(c) = \lim_{x \to c} \frac{F(x) - F(c)}{x - c} .$$

Looking at the numerator of this difference quotient we see that

$$F(x) - F(c) = \int_a^x f(t)\,dt - \int_a^c f(t)\,dt$$

$$= \int_c^x f(t)\,dt$$

This is depicted in Figure 2-2.

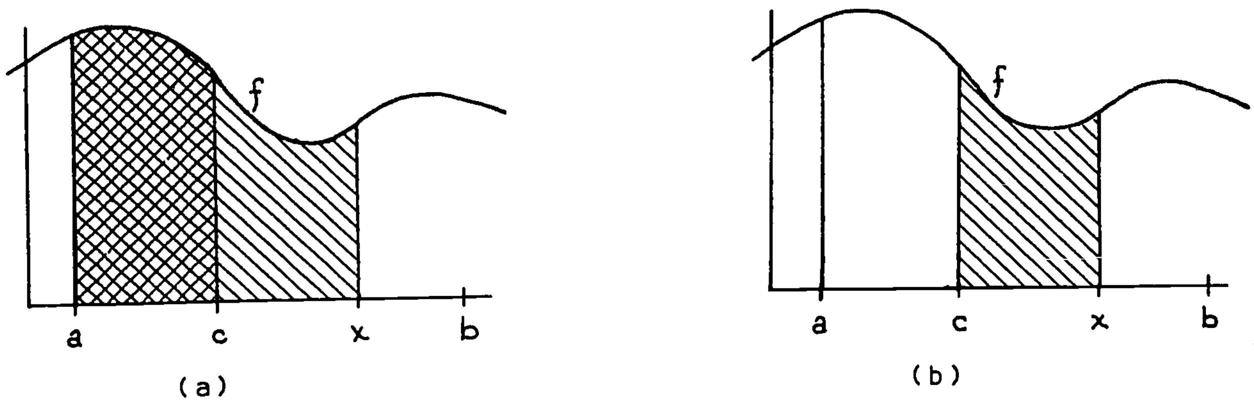


Figure 2-2

The entire shaded area in Figure 2-2(a) represents F(x) while the doubly cross-hatched area represents F(c). Thus the singly shaded area seen in this figure and again in Figure 2-2(b) represents

$$F(x) - F(c) = \int_c^x f(t)dt.$$

Let us now engage in some loose talk in order to get a feeling for what is going on here. Think of the area in Figure 2-2(b) as a sheet of "two dimensional ice" confined in a two dimensional container represented by the sides and bottom of this region.

Imagine that this ice is allowed to melt without changing its area to form a two dimensional liquid with the top surface horizontal.

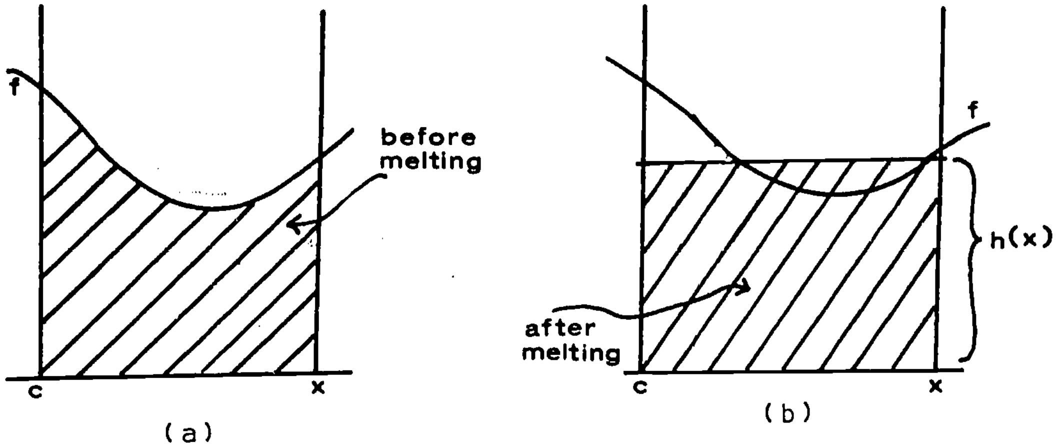


Figure 2-3

Thus the rectangle in Figure 2-3(b) has the same area as the region in Figure 2-3(a). Denoting the height of this rectangle by h(x) we see that

$$F(x) - F(c) = \int_c^x f(t)dt = (x - c)h(x).$$

Next we look at what happens to h(x) as x is taken closer and closer to c.

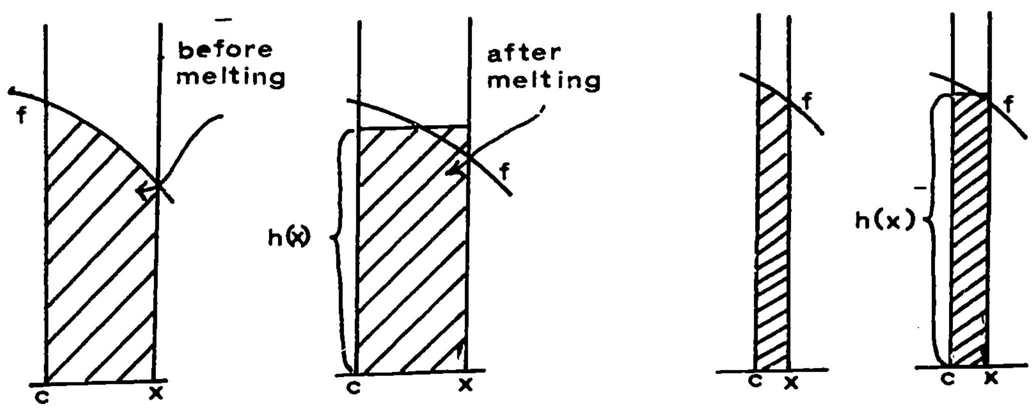


Figure 2-4

It seems evident that as x gets close to c, h(x) gets close to f(c) and in fact that $\lim_{x \to c} h(x) = f(c)$.

Armed with these intuitive observations we return to F'(c) as given in (1),

$$F'(c) = \lim_{x \to c} \frac{F(x) - F(c)}{x - c} = \lim_{x \to c} \frac{\int_c^x f(t)\,dt}{x - c}$$

$$= \lim_{x \to c} \frac{(x - c)h(x)}{x - c} = \lim_{x \to c} h(x) = f(c).$$

And so we come to the conclusion that the relation

$$F'(c) = f(c)$$

ought to hold true, and in general that

$$F'(x) = f(x).$$

And thus we find that starting with a function f, forming its integral,

$$\int_a^x f(t)\,dt,$$

and differentiating this integral

$$D_x \int_a^x f(t)\,dt$$

gives us back the function f again. Thus the operation of "integrating from a to x" is a sort of inverse of the operation of differentiation. Put slightly differently, the function

$$F(x) = \int_a^x f(t)dt$$

turns out to be an anti-derivative of f. This theorem is so important in Calculus that it is called "the Fundamental Theorem of Calculus." The proof of this theorem is quite short and follows fairly closely the lines of the above intuitive discussion, but it does use ε and δ. The idea is to show that for any $\varepsilon > 0$ there is a $\delta > 0$ so that

$$\text{if } 0 < |x-c| < \delta \quad \text{then} \quad \left|\frac{F(x) - F(c)}{x - c} - f(c)\right| < \varepsilon,$$

which is equivalent to

$$\lim_{x \to c} \frac{F(x) - F(c)}{x - c} = f(c).$$

Theorem 1. (The Fundamental Theorem of Calculus. First Form.) Suppose f is unicon in [a,b] and let $F(x) = \int_a^x f(t)dt$. Then, for each c in (a,b), $F'(c)$ exists and $F'(c) = f(c)$.

Proof: Let $\varepsilon > 0$. Since f is unicon in [a,b] there is a $\delta > 0$ so that

$$|f(s) - f(t)| < \varepsilon$$

whenever s and t are in [a,b] with $|s-t| < \delta$.

Now if $0 < |x - c| < \delta$, then for all t between c and x we have $|t - c| < \delta$ so that

$$|f(t) - f(c)| < \varepsilon$$

or

$$f(c) - \varepsilon < f(t) < f(c) + \varepsilon .$$

Hence $f(c) - \varepsilon$ and $f(c) + \varepsilon$ are respectively lower and upper bounds for the functional values over the interval $[c,x]$, assuming $x > c$. Hence

$$(x - c)(f(c) - \varepsilon)$$

$$\leq \int_c^x f(t)\,dt$$

$$\leq (x - c)(f(c) + \varepsilon),$$

and therefore

$$f(c)-\varepsilon \leq \frac{\int_c^x f(t)\,dt}{x - c} \leq f(c)+\varepsilon.$$

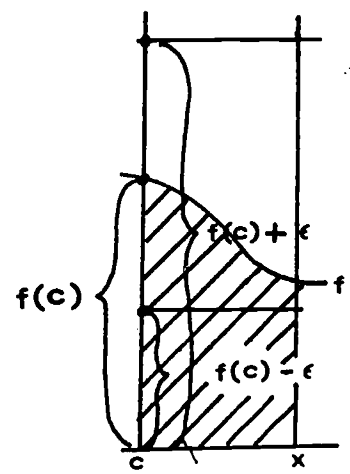


Figure 2-5

If $x < c$ a similar treatment of $\int_x^c f(t)\,dt$ produces the same result if this integral is replaced by $-\int_c^x f(t)\,dt$. Thus in either case, since

$$\int_c^x f(t)\,dt = F(x) - F(c),$$

we get

$$f(c) - \varepsilon \leq \frac{F(x) - F(c)}{x - c} \leq f(c) + \varepsilon,$$

which may be reexpressed in the forms

$$-\varepsilon < \frac{F(x) - F(c)}{x - c} - f(c) < \varepsilon$$

or

$$\left| \frac{F(x) - F(c)}{x - c} - f(c) \right| < \varepsilon .$$

This shows that

$$\lim_{x \to c} \frac{F(x) - F(c)}{x - c} = f(c),$$

which was to be proved.

A number of instances are at our disposal for verifying this theorem. Thus for example:

$$f(x) = \sin x,\ F(x) = \int_0^x \sin t\, dt = 1-\cos x,\ F'(x) = \sin x = f(x);$$

$$f(x) = x^2,\ F(x) = \int_0^x t^2 dt = \frac{x^3}{3},\ F'(x) = x^2 = f(x).$$

In the next section we will see another useful form of the fundamental theorem and some impressive applications.

3. Second Form of Fundamental Theorem

The Fundamental Theorem has told us that if f is unicon on [a,b] and if we define

$$(1) \qquad F(x) = \int_a^x f(t)\,dt \qquad \text{for all } x \text{ in } [a,b],$$

then

$$F'(x) = f(x) \qquad \text{for all } x \text{ in } (a,b).$$

As has already been observed, this means that the function F is an anti-derivative of f.

According to Theorem 2 of Section 1, any two anti-derivatives of f differ by a constant. Thus, if G is another anti-derivative of f we have

$$G(x) - F(x) = C \qquad \text{for all } x \text{ in } [a,b].$$

And the value of this constant is seen by substitution to have the form

$$C = G(a) - F(a).$$

The value of F(a) is seen from the definition in (1) to be

$$F(a) = \int_a^a f(t)\,dt = 0.$$

Hence we have

$$G(x) - F(x) = C = G(a) - F(a) = G(a),$$

whence

$$F(x) = G(x) - G(a),$$

586

so that

$$G(x) - G(a) = F(x) = \int_a^x f(t)dt.$$

Finally we may state the second form of the Fundamental Theorem of Calculus:

<u>Theorem 1.</u> Fundamental Theorem (Second Form). If f is unicon on [a,b] and F is an anti-derivative of f on this interval then

$$\int_a^x f(t)dt = F(x) - F(a) \qquad \text{for } x \text{ in } [a,b].$$

The content of this theorem can be made more meaningful by putting it in slightly different terms. The hypothesis that F is an anti-derivative of f can be expressed as

$$F'(x) = f(x) \qquad \text{for } x \text{ in } (a,b),$$

so that the conclusion of the theorem may be phrased

$$\int_a^x F'(t)dt = F(x) - F(a).$$

This means that differentiating the function F and then integrating this derivative gives us back the function F (minus a suitable constant). In other words (in a suitably general sense): Integration is the inverse operation of differentiation. The first form of the Fundamental Theorem told us that differentiation is the inverse operation of integration. Now we can see why we regard these two forms of the theorem as two faces of the same coin; the two forms taken together tell us that integration and differentiation are inverses of each other (again, in a suitably general sense.)

At the beginning of Section 3-4 it was promised that the seemingly useless dx in

$$\int_a^b f(x)\,dx$$

would eventually be explained. Now the reason for it is almost obvious. It is evident that when the notation was invented, integration was thought of as taking an anti-differential - that f(x)dx is the differential dF of the function F that we are seeking. Since, as we saw in Section 1, this point of view is convenient in the substitution method of finding anti-derivatives, the notation is still the one most commonly used, even though modern mathematics emphasizes functions and derivatives rather than differentials.

The first application of the Fundamental Theorem is a powerful method for evaluating integrals.

Example 1. Evaluate $\int_{\pi/4}^{\pi/3} \cos x \, dx$.

Noting that $D_x \sin x = \cos x$ we see by the Fundamental Theorem that

$$\int_{\pi/4}^{\pi/3} \cos x \, dx = \sin \frac{\pi}{3} - \sin \frac{\pi}{4}$$

$$= \frac{\sqrt{3}}{2} - \frac{\sqrt{2}}{2}$$

$$= \frac{\sqrt{3} - \sqrt{2}}{2} \approx .159.$$

To appreciate the power of this technique for evaluating integrals, recall the tedious process by which it was shown in Chapter 3 that

$$\int_0^x \cos t \, dt = \sin x.$$

On the other hand it was quite simple to see that

$$D_x \sin x = \cos x.$$

And from this easily derived derivative formula we see at once from the Fundamental Theorem that

$$\int_0^x \cos t \, dt = \sin x - \sin 0 = \sin x.$$

Further recall that in Chapter 2 we proved for n = 0,1,2 that

$$\int_0^x t^n dt = \frac{x^{n+1}}{n+1}$$

and then "borrowed" this formula for n = 3,4,5,... . This borrowed formula is now established at once by means of the Fundamental Theorem, since in a much more general form

$$D_x \frac{x^{n+1}}{n+1} = x^n$$

whence

$$\int_0^x t^n dt = \frac{x^{n+1}}{n+1} - \frac{0^{n+1}}{n+1} = \frac{x^{n+1}}{n+1}$$

for any rational value of n except n = -1. (This mysterious exception is the subject of the next chapter.)

This technique motivates the following definition of the "indefinite integral" which is really nothing more than a matter of nomenclature.

Definition. If F is an anti-derivative of f then we define the "indefinite integral" of f as

$$\int f(x)\,dx = F(x) + C.$$

This indefinite integral of f is nothing more than the most general form of the anti-derivative of f. The familiar definite integral

$$\int_a^b f(x)\,dx$$

can be evaluated by the process

1) evaluate the indefinite integral at b;
2) evaluate the indefinite integral at a;
3) subtract.

That is

$$\int_a^b f(x)\,dx = (F(b) + C) - (F(a) + C)$$

$$= F(b) - F(a).$$

We see that the "arbitrary constant" C drops out so there is no need to write it in this evaluation process.

<u>Example 2</u>. Evaluate $\int_2^5 x^2 dx$.

Since $$\int x^2 dx = \frac{x^3}{3} + c,$$

we have $$\int_2^5 x^2 dx = \frac{5^3}{3} - \frac{2^3}{3} = 39.$$

A further notational convention is the writing of

$$F(x)\Big|_a^b \text{ to mean } F(b) - F(a).$$

With this notation the solution in Example 2 can be developed in the convenient running form

$$\int_2^5 x^2 dx = \frac{x^3}{3}\Big|_2^5 = \frac{125}{3} - \frac{8}{3} = \frac{117}{3} = 39.$$

<u>Example 3</u>. Evaluate $\int_0^{\pi/2} \sin^2 x \cos x \, dx$.

Since none of our well-known derivative formulas leads to $\sin^2 x \cos x$, we look for some substitution that will simplify the function to be integrated (this is known as the <u>integrand</u>). Since $\cos x \, dx$ is the differential of $\sin x$, if we let $u = \sin x$ the integrand reduces to $u^2 du$. So

$$\int \sin^2 x \cos x \, dx = \int u^2 du = \frac{1}{3}u^3 = \frac{1}{3}\sin^3 x,$$

and

$$\int_0^{\pi/2} \sin^2 x \cos x \, dx = \frac{1}{3}\sin^3 x \Big|_0^{\pi/2} = \frac{1}{3} - 0 = \frac{1}{3}$$

Now there is a way of short-cutting this. Notice that we go from $\frac{1}{3}u^3$ to $\frac{1}{3}\sin^3 x$ to $\frac{1}{3}\sin^3 x \Big|_0^{\pi/2}$ to $\frac{1}{3}\sin^3 \frac{\pi}{2} - \frac{1}{3}\sin^3 0$. Why not simply say that when $x = 0$ and $\pi/2$, $u = 0$ and 1, correspondingly, and then go directly from $\frac{1}{3}u^3$ to $\frac{1}{3} \cdot 1^3 - \frac{1}{3} \cdot 0^3$? We can then write the whole process as follows

$$u = \sin x, \quad du = \cos x \, dx,$$

$$\int_0^{\pi/2} \sin^2 x \cos x \, dx = \int_0^1 u^2 du = \frac{u^3}{3}\Big|_0^1 = \frac{1}{3}.$$

The inverse trigonometric functions are of interest mainly because they arise naturally in the evaluation of some rather simple integrals. From the derivatives of arcsin and arctan we get the useful indefinite integrals.

$$\int \frac{1}{\sqrt{1 - x^2}} dx = \arcsin x + c, \quad -1 < x < 1,$$

$$\int \frac{1}{1 + x^2} dx = \arctan x + c, \quad -\infty < x < \infty.$$

59?

The substitution $u = ax$, $a > 0$, gives the somewhat more general forms

$$\int \frac{1}{\sqrt{a^2 - u^2}}\, du = \arcsin \frac{u}{a} + c, \quad -a < u < a,$$

$$\int \frac{1}{a^2 + u^2}\, du = \frac{1}{a} \arctan \frac{u}{a} + c, \quad -\infty < u < \infty.$$

In the first of each pair we can replace arcsin by -arccos, but this is rarely preferable.

Example 4. To evaluate $\int_{-2}^{2} \frac{dx}{\sqrt{25 - 4x^2}}$ we let $u = 2x$. This gives us

$$\int_{-4}^{4} \frac{\frac{1}{2}du}{\sqrt{25 - u^2}} = \frac{1}{2}\arcsin \frac{u}{5} \Big|_{-4}^{4}$$

$$= \frac{1}{2} \left(\arcsin \frac{4}{5} - \arcsin(-\frac{4}{5})\right).$$

Since sin is an odd function ($\sin(-x) = -\sin x$) it follows from the definition of $\arcsin x$ that arcsin is also an odd function; hence

$$\arcsin(-\frac{4}{5}) = -\arcsin \frac{4}{5} .$$

Thus our integral has the value

$$\frac{1}{2}\left[2 \arcsin \frac{4}{5}\right] = \arcsin 0.8 = 0.92727522,$$

from tables.

(Tan and arctan are also odd functions. On the other hand, although cos is an even function arccos is not. In fact, no even function can have an even inverse function. Why?)

Example 5. To evaluate

$$\int \frac{\sin \theta}{1 + \cos^2\theta} d\theta$$

we notice that $\sin \theta \, d\theta = -d(\cos \theta)$. Hence for $u = \cos \theta$ the integral becomes

$$\int \frac{-du}{1 + u^2} = -\arctan u + c$$

$$= -\arctan(\cos \theta) + c.$$

Example 6.

$$\int_0^2 \frac{dx}{x^2 - 2x + 4} = \int_0^2 \frac{dx}{(x - 1)^2 + 3}.$$

Letting $u = x - 1$ gives

59

$$\int_{-1}^{1} \frac{du}{u^2 + 3} = \frac{1}{\sqrt{3}} \arctan \frac{u}{\sqrt{3}}\Bigg|_{-1}^{1}$$

$$= \frac{1}{\sqrt{3}} \arctan \frac{1}{\sqrt{3}} - \arctan \frac{-1}{\sqrt{3}}$$

$$= \frac{2}{\sqrt{3}} \arctan \frac{1}{\sqrt{3}}$$

$$= \frac{2}{\sqrt{3}} \frac{\pi}{6} = \frac{\pi}{3\sqrt{3}} = .59453.$$

PROBLEMS

1. Evaluate each of the following integrals

(a) $F(x) = \int_0^x t^5 \, dt$

(b) $F(x) = \int_0^x (4t^3 - 2t^8 + t) \, dt$

(c) $F(x) = \int_a^x \frac{\sqrt{t}}{t} \, dt$

(d) $F(t) = \int_0^t \sqrt{x} \, dx$

(e) $F(y) = \int_{-10}^y x^3 \, dx$

(f) $G(u) = \int_{-1}^{u^2} (x + 1) \, dx$

(g) $H(z) = \int_{-1}^{1-z} (x^2 - 1)^2 \, dx$

2. Evaluate the given definite integrals

(a) $\int_{-1}^2 (x + 4) \, dx$

(b) $\int_1^4 (u^2 - 2u + 3) \, du$

(c) $\int_0^1 (\sqrt{x} + 1)^2 \, dx$

(d) $\int_3^6 (x - 2)^{1/2}\, dx$

(e) $\int_1^4 \frac{du}{\sqrt{u}}$

(f) $\int_1^3 \left(x - \frac{2}{x}\right)^2 dx$

(g) $\int_{-3}^{-1} (4x^3 - 3x^2 + 2)\, dx$

(h) $\int_0^{\pi/8} \sin 4x\, dx$

(i) $\int_0^{\pi/3} \sin^2\theta \cos\theta\, d\theta$

(j) $\int_4^{25} \frac{x^2 + 2}{x^2}\, dx$

(k) $\int_4^9 \left(\sqrt{x} - \frac{1}{\sqrt{x}}\right) dx$

(l) $\int_{-3}^{-1} \frac{z^2 - 1}{z^5}\, dz$

(m) $\int_1^5 \frac{1}{\sqrt{5t}}\, dt$

3. Evaluate the following definite integrals.

(a) $\int_{-2}^{1} \sqrt{2 - x}\, dx$

(b) $\int_0^1 y\sqrt{1 - y^2}\, dy$

(c) $\int_0^1 2x(x^2 + 2)^3 \, dx$

(d) $\int_{-3}^{-1} \frac{2 \, du}{(u - 1)^2}$

(e) $\int_1^2 (2x + 1)\sqrt{x^2 + x + 1} \, dx$

(f) $\int_0^1 \frac{(x^2 + 2x)\,dx}{\sqrt[3]{x^3 + 3x^2 + 1}}$

(g) $\int_1^4 \left(\frac{1}{\sqrt{x}} + \frac{1}{\sqrt{2}}\right) dx$

(h) $\int_{1/3}^{5/3} \sqrt{1 + (3\sqrt{x})^2} \, dx$

(i) $\int_1^2 \frac{1}{x^2}\sqrt{1 - \frac{1}{x}} \, dx$

(j) $\int_0^{\pi/4} \frac{\sin x}{\cos^3 x} \, dx$

4. Evaluate the following indefinite integrals.

(a) $\int 3(2x + 1)^5 \, dx$

(b) $\int \frac{u \, du}{\sqrt{u^2 - 1}}$

(c) $\int (x - 1)(2x + 1) \, dx$

(d) $\int \sqrt{\theta} \cos\theta^{3/2} \, d\theta$

(e) $\int \frac{u^2 + 2u - 1}{\sqrt{u}}\, du$

(f) $\int \frac{x^3 - x^2 + 2}{x^2}\, dx$

(g) $\int \frac{y^4}{\sqrt{y^5 + 2}}\, dy$

(h) $\int \frac{1}{x^3}\left(\frac{1}{x^2} - 1\right)^{10} dx$

(i) $\int \frac{1 - 5s + s^5}{s^4}\, ds$

(j) $\int x \sin(1 + x^2)\, dx$

(k) $\int \frac{1}{\sqrt{4 - t^2}}\, dt$

(l) $\int \frac{t}{\sqrt{1 - t^4}}\, dt$

(m) $\int \frac{1}{x^2 - 4x + 13}\, dx$

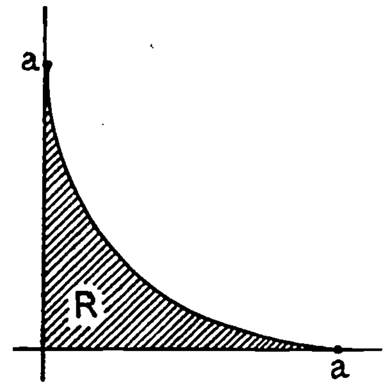
(n) $\int \frac{\sqrt{\arctan x}}{1 + x^2}\, dx$

(o) $\int \frac{\sin y\, dy}{\sqrt{10 - \cos^2 y}}$

(p) $\int \frac{du}{8u - u^2 - 25}$

(q) $\int \frac{r}{1 + 9r^4}\, dr$

5. Let R be the region bounded by the axes and the curve $\sqrt{x} + \sqrt{y} = \sqrt{a}$



(a) Find the area of R.

(b) Find the volume of the solid obtained by rotating R about the x-axis.

6. (a) Find the area bounded by the axes, the line $x = h$, where $h > 0$, and the curve $y = (1 + x^2)^{-1}$

(b) What happens to this area as h gets larger and larger without bound? Does this seem reasonable?

ANSWERS

Volume I

Chapter 0

4-3, page 24

a) $x < 1$, $x > -1$, b) $x < 5$, $x > -1$

5-1, page 30

a) $-5 \leq x \leq 1$, b) $x > 7$, $x < -3$, c) $-1 \leq x \leq 5$,

d) $x < 5$, $x > -1$, e) impossible, f) $x < \frac{5}{2}$,

g) $0 < x < 5$, $x \neq 2$, h) $x < 2$, $x \neq 0$

6-2, page 43

a) $2x$, b) $2 + x$, c) $|x| - 2x$, d) $x^3 + |x|$

6-3, page 43

a) x^2, b) $2x$, c) $-2x \cdot |x|$, d) $x^3 \cdot |x|$

6-4, page 43

Compare 2a and 3b.

7-6, page 55

Identity function

Chapter 0

<u>8-1, page 61</u>

Zeros of sine are $n\pi$

Zeros of cosine are $(2n+1)\frac{\pi}{2}$

<u>8-2, page 61</u>

Zeros are $\frac{1}{n\pi}$, n any integer $\neq 0$

<u>8-4, page 61</u>

a) $\sqrt{x}$, b) $\sqrt{x - 1}$, c) $-x$

<u>8-6, page 61</u>

Yes

<u>8-7, page 61</u>

f and its inverse are the reflection of each other on the line x = y.

Chapter 1

<u>3-1, page 88</u>

The effect is the adding of the latest SUM twice.

<u>3-2, page 88</u>

a) A = 13, B = 13; A = 7, B = 7, b) both, c) different

Chapter 2

2-2, page 141

a) $x \approx -1.117$, b) $x \approx -1.116$

c) roots are $\pm 1, 0, 0$

2-3, page 141

a) $x \approx 1.414$, b) $x \approx 5.$

c) $x \approx -.758$

3-1, page 145

a) $r^4 - 1$, b) $(r^4 - 1)/(r - 1)$, c) $(r^n - 1)/(r - 1)$

d) $1/(1 - r)$, e) no, you cannot.

3-2, page 145

a) $k(r^4 - 1)$, b) $k(r^4 - 1)/(r - 1)$, c) $k(r^n - 1)/(r - 1)$,

d) $k/1 - r$, e) no, you cannot.

3-3, page 146

a) 2, b) 1, c) 25/2, d) 21/5

3-5, page 147

1200 knashes

3-6, page 147

a) $\frac{1}{2} + \frac{1}{4} + \ldots \frac{1}{2^{40}}$ decays

$1 - (\frac{1}{2} + \frac{1}{4} + \ldots + \frac{1}{2^{40}})$ left

b) $3(\frac{1}{2} + \frac{1}{4} + \ldots + \frac{1}{2^{40}})$ decays

$3 - 3(\frac{1}{2} + \frac{1}{4} + \ldots + \frac{1}{2^{40}})$ lef

c) $3(\frac{1}{2} + \frac{1}{4} + \ldots + \frac{1}{2^{120}})$ decays

$3 - 3(\frac{1}{2} + \frac{1}{4} + \ldots + \frac{1}{2^{120}})$ left

d) $\frac{1}{2}$ decays, $\frac{1}{2}$ left

3-6, page 147 - con't.

e) $16(\frac{1}{2} + \frac{1}{4} + \ldots \frac{1}{2^{96}})$ decays

$16 - 16(\frac{1}{2} + \frac{1}{4} \ldots \frac{1}{2^{96}})$ left

f) $\frac{1}{2} + \frac{1}{4} + \ldots \frac{1}{2^{10^{10}}}$ decays

$1 - (\frac{1}{2} + \ldots + \frac{1}{2^{10^{10}}})$ left

g) $\frac{1}{2} + \frac{1}{4} \ldots \frac{1}{2^{10}}$ decays

$1 - (\frac{1}{2} + \frac{1}{4} \ldots \frac{1}{2^{10}})$ left

3-7, page 148

2 min.

4-1, page 160

a) 0, b) 0, c) 0, d) diverges,

e) 1, f) $\sqrt{3}$, g) 5

4-2, page 160

a) yes, b) no, c) no, d) yes, e) no

4-3, page 161

a) yes, b) yes, c) yes

4-7, page 163

Flow chart

6-1, page 184

a) 2, b) 8, c) 4, d) 1, e) 1/2, f) 3/4,

g) doesn't exist, h) 1

Chapter 2

6-2, page 184

$(-1)^n(-1)^n$

7-3, page 192

a) $4r^4 - r^3 - r^2 - r - 1$, b) $nr^n - r^{n-1} - \ldots - r - 1$,

e) $1/(r - 1)^2$

8-1, page 197

a) 1, b) 1/2, c) 0, d) 0, e) 0, f) k, g) 1/2

9-1, page 203

a) $\sqrt{5}$, b) $c_1 = 1$, $c_2 = 2.3333$, $c_3 = 2.2307$, $c_4 = 2.2363$,

c) $c_{n+2} = (9c_n + 20)/(4c_n + 9)$, d) $d_{n+1}/d_n = -1/(4c_n + 9)$,

f) $n > 7$

9-3, page 204

c) $1/2n$, d) 501, e) .693

Chapter 3

2-3, page 227

Correct analytical answers:

a) $\frac{2}{3}$

b) $\frac{2}{3}(2\sqrt{2} - 1)$

c) $\frac{2}{3}(2\sqrt{2} - 1)$

d) 1

e) 2

f) $\frac{1}{3}$

g) $\frac{\pi}{4}$

h) sin 1

i) 1

j) $\frac{1}{4}$

5-1, page 251

The maximum error:

a) .003

b) .002

c) .0014

d) .002

e) .005

f) .002

g) .002

h) .001

i) .002

j) .004

6-1, page 259

a) .05

b) .005

c) .0005

6-6, page 259

a) $\delta = \frac{\varepsilon}{6}$

b) $\delta = \frac{\varepsilon}{20}$

c) $\delta = \frac{\varepsilon}{2 \max(|a|, |b|)}$

Chapter 3

7-2, page 268

a) $43\frac{1}{3}$, b) $2\sqrt{3} - \frac{4}{3}\sqrt{2}$, c) $\frac{7}{2} + 2 \log 2$, d) $\frac{24}{5}$

8-1, page 280

a) $\frac{25}{2}$, b) 25, c) 40, d) 10, e) 55, f) $\frac{\sqrt{3}}{2}$, g) $\frac{\sqrt{3}}{2}$

8-5, page 282

a) 4, b) 0, c) 18, d) 64, e) 1, f) -1, g) $12\frac{2}{3}$,
h) 2, i) 129.

10-2, page 297

a) $\frac{.002}{\pi}$, b) $\frac{.001}{\pi}$, c) .0002, d) .0028, e) $\left(\frac{.001}{3}\right)^2$
f) $\frac{1}{4}(.00001)$, g) $\frac{1}{4}(.00001)$

Chapter 4

3-1, page 351

a) 45π, b) 6π, c) $\frac{2}{5}\pi$, d) $\frac{49}{8}\pi$, e) $\frac{1}{6}\pi$, f) $\frac{3}{10}\pi$,

g) 24π, h) 18π, i) π, j) $\frac{\pi^2}{2} - \pi$

3-2, page 351

No. Translate a square vertically.

3-3, page 351

Yes. Since $\int_a^b y^2(x)dx = \int_{a+h}^{b+h} y^2(x - h)dx$.

4-1, page 363

a) ≈ 792.5, b) 15,850, c) \$357

4-2, page 363

a) 2000, b) 117, c) 2, d) $\frac{1}{6}$

4-3, page 364

$\sqrt{10}$

4-4, page 364

b) All of them, c) yes

4-5, page 365-366

a) 14,000, b) 24.5×10^5, c) 10, d) 1.5×10^{-4}

e) $\frac{25}{4}$, f) 8×10^9

1-1, page 381

a) $\frac{1}{2}$, b) $\delta \le \sqrt{24\varepsilon}$, c) 4.2×10^{-6}, d) .002

1-3, page 38

160°/min.

1-4, page 382

a) 2280 ft., b) 2 ft/sec. (For faster convergence use $\frac{1}{n^4}$ in place of $\frac{1}{n}$.) c) $t_1 = 12$ sec., d) -382 ft/sec.

2-1, page 395

a) 1, b) $\frac{1}{2}$, c) -1, d) 1, e) undefined, f) 1, g) 4,

h) undefined, i) 0, j) $\frac{1}{2\sqrt{3}}$, k) $\frac{1}{2\sqrt{5}}$, l) $\frac{1}{6}$,

m) $\frac{1}{3}$, n) 4, o) $\lim_{x \to 2} \frac{x - tx - 2 + 2t}{x - 2} = 1 - t$

2-3, page 396

a) 1 for all values, b) 6, 34, $2t$, $2(\pi^2 + 1)$,

c) 48, 3, $3t^2$, d) +1, -1, $\begin{cases} +1, & t > 0 \\ \text{undefined}, & t = 0, \\ -1, & t < 0 \end{cases}$ undefined

e) $\frac{1}{2\sqrt{3}}$, $\frac{1}{2\sqrt{5}}$, f) $-\frac{\sqrt{2}}{2}$, $-\frac{\sqrt{3}}{2}$, $-\sin t$

2-6, page 397

a) $x = \pm 2$, b) $x = n\frac{\pi}{2}$, $n = \pm 1, \pm 2, \ldots$

c) $x = n\pi$, $n = 0, \pm 1, \pm 2, \ldots$

Chapter 5

3-1, page 405

a) 11, b) 8, c) 4, d) 3, e) 0, f) $\sqrt{7}$, g) $\frac{1}{4}$

3-2, page 405

a) $x = 0$, b) $x_1 = -1$, $x_2 = 3$, c) $x = 0$, d) $x = 1$

3-3, page 405-406

a) 0 for all n, b) not defined

4-1, page 414

a) 4, -2, 6, 2π, 10, b) $-\frac{1}{4}$, -1, $-\frac{1}{9}$, $-\frac{1}{(t+1)^2}$

4-2, page 414

$-\frac{1}{36}$

4-3, page 414

a) $y = 0$, b) No

4-4, page 414

0, $3c^2$

Chapter 5

4-5, page 414-415

a) $y = 24x - 48$, $y = 0$, b) $y = (14 + 8\sqrt{3})x - 111 - 64\sqrt{3}$

$y = (14 - 8\sqrt{3})x - 111 + 64\sqrt{3}$

c) There is no line which will pass through the point (a,b)

5-1, page 421

a) 1, b) $2x$, c) $3x^2$, d) $-\sin x$, e) $\frac{1}{2\sqrt{x}}$

5-2, page 421

$(\frac{3}{2}, 6)$

5-5, page 421

a) d, b) c, c) -32, d) $S(t) = -16t^2 + 20t + 200$

5-6, page 422

192 ft.

5-7, page 422

96 ft/sec.

5-8, page 422

a) $25 \sin\phi$ sec., b) $10000 \cos\phi \sin\phi$ ft, c) $\frac{\pi}{4}$

Chapter 5

6-1, page 434

a) $x = 0, \quad D_x x|x| = 0$

$x > 0, \quad D_x x|x| = 2x$

$x < 0, \quad D_x x|x| = -2x$

b) $x > 0, \quad D_x|x| = 1$

$x < 0, \quad D_x|x| = -1$

$x = 0, \quad D_x|x|$ is undefined

c) Yes, No

6-2, page 434

a) $D_x y = 5x^4 + 48x^3 - 9x^2 + 2$

b) $D_x y = 42x^5 + 12x^2$

c) $D_x y = 0$

d) $D_x y = \dfrac{2 - x}{2\sqrt{x}(x + 2)^2}$

e) $D_x y = -\csc^2 x$

f) $D_x y = 5x^4 + 18x^2 + 2x + 8$

g) $D_x y = x \sec^2 x + \tan x$

h) $D_x y = 3x^2 + 4x - 23$

i) $D_x y = \dfrac{6}{(x + 4)^2}$

j) $D_x y = \sin x (\sec^2 x + 1)$

k) $D_x y = -\cos x (\csc^2 x + 1)$

l) $D_x y = -\dfrac{1}{\sqrt{x}(x + 4) - 4x}$

m) $D_x y = -\dfrac{6x}{(x^2 + 1)^2}$

n) $D_x y = 18x^2$

o) $D_x y = -(\sin x + \cos x)$

p) $D_x y = \sec^2 x - 3x^2$

q) $D_x y = \dfrac{8x - 3x^2}{(x^3 - 4x^2 + 1)^2}$

r) $D_x y = \dfrac{128x^3}{(x^4 + 16)^2}$

s) $D_x y = -\dfrac{5}{x^2} - \dfrac{4}{x^3}$

t) $D_x y = 36x^3 + 12x$

u) $D_x y = 2x - \dfrac{2}{x^3}$

v) $D_x y = -10x^{-2} + 3x^{-4}$

w) $D_x y = \frac{x^2 + 10x - 18}{x^4}$ x) $D_x y = -\frac{1}{2x\sqrt{x}} + \frac{1}{18\sqrt{x}}$

6-3, page 435

$P_1(x,y) = (-3, -57)$

$P_2(x,y) = (-4, -56)$

6-4, page 435

$x = 0, 2\pi, 2k\pi$, where k = integers

6-5, page 436

b) No

6-6, page 436

b) $\delta = 3$

6-9, page 437

a) $f' = 1$, $f'' = f''' = f^{(4)} = 0$

b) $f' = 2x$, $f'' = 2$, $f''' = f^{(4)} = 0$

c) $f' = 3x^2$, $f'' = 6x$, $f''' = 6$, $f^{(4)} = 0$

d) $f' = 4x^3$, $f'' = 12x^2$, $f''' = 24x$, $f^{(4)} = 24$

e) $f' = 5x^4$, $f'' = 20x^3$, $f''' = 60x^2$, $f^{(4)} = 120x$

f) $f' = 20x^4$, $f'' = 80x^3$, $f''' = 240x^2$, $f^{(4)} = 480x$

g) $f' = \cos x$, $f'' = -\sin x$, $f''' = -\cos x$, $f^{(4)} = \sin x$

h) $f' = -\sin x$, $f'' = -\cos x$, $f''' = \sin x$, $f^{(4)} = \cos x$

i) $f' = 2 \cos x + 3 \sin x$, $f'' = -2 \sin x + 3 \cos x$,

$f''' = -2 \cos x - 3 \sin x$, $f^{(4)} = 2 \sin x - 3 \cos x$

<u>6-10, page 437</u>

a) 3 ft, b) 3 ft/sec, c) π sec, d) 3 ft/sec,

e) 0,

Chapter 6

2-1, page 449

a) Local: Max at $x = 0$.

b) Local: Min at $x = 1$, $x = -1$.

c) No extrema.

d) Local extrema: Max at $x = 5$,
Min at $x = 1$.

e) Local extrema: Max at $x = -1$,
Min at $x = 1$.

f) Local extrema: Max at $x = \sqrt{1/3}$,
Min at $x = -\sqrt{1/3}$.

g) Local: Min at $x = 3$.

h) Local: Max at $x = \frac{5\pi}{3}$

Chapter 6

i) Local extrema: Max at $x = .46$, $x = .46 - 2\pi$, $x = 2\pi$

Min at $x = .46 + \pi$, $x = .46 - \pi$, $x = -2\pi$.

j) Local extrema: Max at $x = 0$, $x = 2\pi$,

Min at $x = \frac{3}{2}\pi$, $x = -\frac{\pi}{2}$, $x = -2\pi$.

2-2, page 450

Length = 50 ft, Width = 25 ft

2-3, page 450

Length: 22.5 inch.

Width: 22.5 inch.

2-4, page 450

$x = \frac{H}{2}$, $y = \frac{A}{2}$, max. area of the rectangle: $\frac{A \cdot H}{4}$.

2-5, page 451

$r = \frac{2}{3}R\sqrt{2}$.

2-6, page 451

$x = \frac{a}{\sqrt{2}}$, $y = \frac{b}{\sqrt{2}}$

2-7, page 451

8, 8

2-9, page 451

$x = 6 - 2\sqrt{3}$

Chapter 6

<u>2-10, page 452</u>

$x = 13\frac{1}{3}$, $y = 26\frac{2}{3}$, $z = 5\frac{1}{3}$

<u>2-11, page 452</u>

$h = r\sqrt{2}$

<u>2-12, page 452</u>

$r = h$

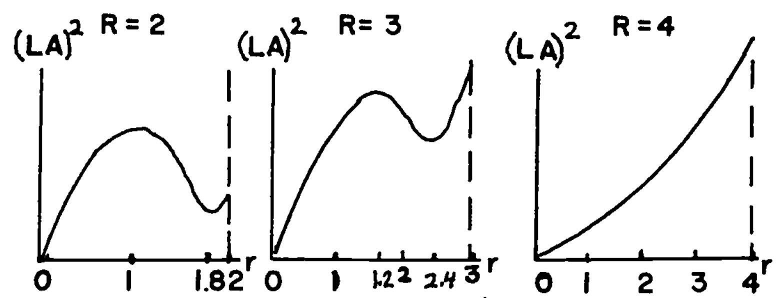
<u>2-13, page 453</u>

$$r = \frac{3H^2R - HR\sqrt{H^2 - 8R^2}}{4(R^2 + H^2)} \quad \text{provided that} \quad H^2 \geq 8R^2$$

a) $R = 2$

b) $R = 3$

c) $R = 4$



<u>2-14, page 453</u>

$\theta = 30^\circ$

<u>2-15, page 453</u>

Min at $x = 4.913$

Max at $x = 2.029$

<u>2-16, page 454</u>

b) x has to be in $[b-a, b+a]$.

<u>4-1, page 468</u>

a) $\xi = \frac{5}{2}$, b) $\xi = \frac{121}{4}$

4-1, page 468 (cont.)

c) $\xi = \sqrt{\frac{13}{3}}$, d) $\xi = 1$, e) $\xi = \frac{4}{3}$.

5-1, page 477

.81 cu. inch.

5-2, page 477

a) 2.03, $|E| \leq .00064$

b) 1.070, $|E| \leq .00244$

c) .5407, $|E| \leq .0011$

5-3, page 477

2.2361, $|E| \leq .00004$

6-1, page 488-489

a) Local extrema: Min at $x = -1$, Min at $x = 2$,
Max at $x = \frac{1}{2}$

Critical point: $x = \frac{1}{2}$

$f(x)$ is increasing on $[-1, \frac{1}{2}]$ and decreasing on $[\frac{1}{2}, 2]$.

b) Local extrema: Max at $x = 3$
Min at $x = -1$

Critical points: $x = 3$, $x = -1$

$f(x)$ is decreasing on $(-\infty, -1]$, increasing on $[-1, 3]$ and decreasing on $[3, \infty)$.

c) Local extrema: Min at $x = -3$, $x = 3$

Max at $x = 0$

Critical point: $x = 0$

f(x) has to be restricted to the interval [-3,3].

f(x) is increasing on [-3,0] and decreasing on [0,3].

d) Local extrema: Max at $x = 0$

Min at $x = 1$

Critical point: $x = 1$

f(x) has to be restricted to the interval [0,∞).

f(x) is decreasing on [0,1] and increasing on [1,∞).

e) Local extrema: Max at $x = -4$, $x = -1$

Min at $x = -2$

Critical point: $x = -2$

f(x) is decreasing on [-4,-2] and increasing on [-2,-1]

f) Critical point: Min at $x = 0$

f(x) is decreasing on (-∞,0] and increasing on [0,∞).

g) Critical points = end points.

Max at $x = -4$, Min at $x = 0$

f(x) is decreasing on [-4,0].

h) Local extrema = critical points.

Max at $x = 0, \pi, 2\pi$, Min at $x = \frac{2\pi}{3}, \frac{4\pi}{3}$.

$f(x)$ is decreasing on $[0, \frac{2\pi}{3}]$, increasing on $[\frac{2\pi}{3}, \pi]$, decreasing on $[\pi, \frac{4\pi}{3}]$ and increasing on $[\frac{4\pi}{3}, 2\pi]$.

6-2, page 489

Critical points $x = \pi + 2n\pi$, $n = 0,1,2,3,\ldots$.

no extrema.

6-3, page 489

$y = -\frac{1}{2}x^2 + x + \frac{3}{2}$

Chapter 7

1-1, page 504

a) $-10(1 - x)^9$, b) $-3x^{-4}(2 - x^{-3})^{-2}$, c) $15x^2(x^3 - 4)^4$

d) $\frac{5}{2}(x + 1)^{3/2}$, e) $-\frac{15}{2} x \left[(1 + x^2)^{3/2} \left(1 + [1 + x^2]^{5/2}\right)^{-5/2} \right]$

f) $2(3 - 2x)^4(3x^2 + 4)^2(-33x^2 + 27x - 20)$

g) $\frac{3(x^3 + 2x + 1)^2(x^4 + x^2 - 2x + 2)}{(x^2 + 1)^4}$, h) $2\frac{1}{x^{1/2}(1 - x)^{3/2}}$

i) $\frac{1}{(1 - x^2)^{3/2}}$, j) $\frac{3}{2}\sqrt{x}$, k) $\frac{1}{4\sqrt{x}(\sqrt{x} + 1)}$ l) $-\frac{\sqrt{\frac{1}{x} + 4}}{2x + 8x^2}$

m) $-3 \sin 3x - 2 \cos x$

page 505

n) $(2x^2 - 4x)\sin x^2 + 4x \sin x^4 - \cos x^2 + 8x^5 \cos x^4$,

o) $-\sin 2x \cos(\cos 2x)$.

p) $\cos[\sin(\sin x)][\cos(\sin x)](\cos x)$,

q) $(x + 3)(x^2 + 2x + 1)^2(x^2 + 4)^3[2(x^2 + 2x + 1)(x^2 + 4)$
$+ 3(2x + 2)(x + 3)(x^2 + 4) + 8x(x + 3)(x^2 + 2x + 1)]$,

r) $\frac{x^2 - 8x - 33}{(x - 4)^2}$, s) $\frac{\frac{1}{2}x^4 - 3x^2 + 8x}{(x^3 + 4)^{3/2}}$,

t) $\dfrac{(x^3 + 3x - 4)(-x^5 + 2x^4 + 5x^3 + x^2 + 16x - 11)}{2(x + 3)^3}$,

u) $3(x \;\; + 5x - 6x^{-1})\,(4x^3 + 5 + 6x^{-2})$, v) 0,

w) $\sin x + x \cos x$, x) $\dfrac{\cos x}{2\sqrt{\sin x}}$, y) $\dfrac{\;}{2\sqrt{x}} \cos\sqrt{x}$,

z) $-\dfrac{2 \cos x}{\sin^3 x}$

1-2, page 505

a) $f' = x(x^2 + 1)^{-1/2}$

$f'' = (x^2 + 1)^{-3/2}$

b) $f' = \dfrac{-(3x^2 - 1)}{(x^2 + 1)^3}$, $f'' = \dfrac{12x(x^2 - 1)}{(x^2 + 1)^4}$

c) $f' = 2 \cos 2x$

$f'' = -4 \sin 2x$

d) $f' = 3 \sec^2 3x$

$f'' = 18 \sec^2 3x \tan 3x$

e) $f' = 2x(\sin 2x + x \cos 2x)$, $f'' = 2(1 - 2x^2)\sin 2x + 8x \cos 2x$

f) $f' = -\dfrac{\sqrt{2} \cot 2x}{(\sin 2x)^{1/2}}$, $f'' = \dfrac{\sqrt{2 \sin 2x}\,(2 + \frac{1}{2} \sin^2 4x)}{(\sin 2x)^3}$

1-6, page 506

a) Local extrema: Max at $x = \sqrt{1/2}$, $x = -1$

Min at $x = -\sqrt{1/2}$, $x = 1$

$f(x)$ restricted on $[-1, 1]$

Critical points $x = \sqrt{1/2}$, $x = -\sqrt{1/2}$.

b) Critical point $x = 2$

1-9, page 508

$\theta = \arcsin\sqrt[3]{\frac{W}{L}}$

2-1, page 516

a) $y' = -\frac{y}{x}$, b) $y' = -\frac{(2y + x^2)}{2x + 5y^2}$, c) $y' = \frac{-1}{\sqrt{x}\,(\sqrt{x} - 1)^2}$,

d) $y' = \frac{2}{1 - 2y}$, e) $y' = \frac{-y - 2\sqrt{xy}}{x - 4\sqrt{xy}}$, f) $y' = \frac{2x}{y}$,

g) $y' = \frac{3x^2 + 2xy^2 - 1}{2 - 2x^2y}$, h) $y' = -\frac{\sqrt{y}}{\sqrt{x}}$, i) $y' = \frac{y}{x}\,\frac{\cos x - \sin y}{\cos y - \sin}$

2-2, page 516

a) $y'' = \frac{2y}{x^2}$, d) $y'' = \frac{-8}{(1 - 2y)^3}$, h) $y'' = \frac{1}{2}x^{-3/2}$

i) $[(x \cos y - \sin x)(-y \sin x + y' \cos x - y' \cos y)$

$-(y \cos x - \sin y)(-xy' \sin y + \cos y - \cos x)](x \cos y - \mathrm{si}$

2-3, page 516

a) $x = \frac{-3 \pm \sqrt{3}}{3}$, b) $y = \frac{3 \pm \sqrt{3}}{3}$

Chapter 7

2-4, page 517

a) $y = \frac{c}{\sqrt{ac}} x + k$, b) $Q(x,y) = (-c,0)$

3-3, page 530

a) $\frac{2}{\sqrt{1 - 4x^2}}$, b) $-\frac{5}{\sqrt{1 - 25x^2}}$, c) $-\frac{1}{x^2 + 1}$,

d) $\frac{1}{2\sqrt{x(1 - x)}}$, e) $-\frac{2t}{\sqrt{1 - t^4}}$, f) $\frac{6(1 + \arcsin 3x)}{\sqrt{1 - 9x^2}}$,

g) $\arcsin x$, h) $\frac{1}{x\sqrt{x^2 - 1}}$, i) $\frac{2x^2}{\sqrt{1 - x^2}}$,

j) $\frac{3}{2\sqrt{\arcsin 3x(1 - 9x^2)}}$, k) $\frac{1}{x^2 + 1}$, l) $\frac{\sqrt{1 - x^2} \arccos x - x}{\sqrt{1 - x^2}}$,

m) $-\frac{\sqrt{y(1 - y)}}{2y(1 - y)}$, n) $\frac{2y(\sqrt{1 - 4y^2} \arccos 2y - y)}{\sqrt{1 - 4y^2}}$,

o) $\frac{3 \sec^2 x}{1 + 9 \tan^2 x}$, p) $\frac{3}{5 + 4 \cos x}$.

3-4, page 531

yes

3-5, page 531

$f'(x) = 0$

3-6, page 531

$x = 6\sqrt{13}$ ft.

4-1, page 536

$\frac{1}{18\pi}$ in./sec.

4-2, page 536

$\frac{1}{4\pi}$

4-3, page 536

$-\frac{3}{4}$ ft/sec.

4-4, page 536

5 ft./min.

4-5, page 537

$3\sqrt{10}$ ft./sec.

4-6, page 537

$\frac{8}{\sqrt{209}}$ ft/sec.

4-8, page 538

a) $\frac{13}{8}$ ft/sec.

6-1, page 547

a) $dy = \frac{-x}{\sqrt{1 - x^2}}dx$, b) $dy = \frac{x^2 + 2x - 1}{(x + 1)^2}dx$,

c) $dy = \frac{-2x + y}{x + 2y}dx$, d) $dy = \frac{\cos y - y \cos x}{\sin x + x \sin y}dx$,

e) $dy = -\cos z \sin x \, dx$, f) $dy = \frac{\cos t - t \sin t}{\sin t + t \cos t}dx$,

g) $dy = \frac{ux}{y^2}dx$, h) $dy = -\frac{1}{y^2}(x^2 + \frac{2z^2 - z^3}{x + 2z})dx$

6-3, page 547

a) $dw = -\frac{w}{u}\,du - \frac{w}{v}\,dv$

$du = -\frac{u}{w}\,dw - \frac{u}{v}\,dv$

$dv = -\frac{v}{w}\,dw - \frac{v}{u}\,du$

b) $D_u v = -\frac{v}{w}D_u w - \frac{v}{u}$

7-2, page 551

a) $D_x y = \frac{1}{2t}$, $D_x^2 y = -\frac{1}{4t^3}$, b) $D_x y = \csc\theta$, $D_x^2 y = -\cot^3\theta$

c) $D_x y = -\cot t$, $D_x^2 y = \frac{1}{3}\csc^4 t \sec t$

d) $D_x y = -\frac{x}{y}$, $D_x^2 y = -\frac{1}{y^3}$,

e) $D_x y = \frac{\cos t - \cos 2t}{\sin t - \cos 2t}$,

$$D_x^2 y = \frac{2 + 2\cos t[\sin 2t - \cos 2t] - \sin t(\sin 2t + \cos 2t)}{2(\cos t + \sin 2t)^3}$$

7-5, page 552

a) $x = a\cos\theta + a\theta\sin\theta$

$y = a\sin\theta - a\theta\cos\theta$

Chapter 8

1-1, page 567

a) $x^3 - 2x^2 + 7x + c$, b) $\frac{1}{2}x^4 - x^3 + c$,

c) $\frac{1}{2}x^6 - \frac{25}{4}x^4 + 30x^2 + c$, d) $\frac{4}{5}(t + 2)^{5/2} - \frac{t^2}{2} + c$

e) $\frac{1}{3}x^{3/2} - 4x^{1/2} + c$, f) $\frac{1}{2}\sin^2 x + c$, g) $(\sin 2y)^{1/2} + c$,

h) $\frac{1}{4}\sin^2 2\phi + c$, i) $x^4 - \frac{5}{2}x^2 + 7x + c$,

j) $\frac{1}{10}x^{10} + \frac{6}{7}x^7 + 3x^4 + 8x + c$, k) $\frac{1}{12}(x^3 + 2)^4 + c$,

l) $\frac{x^2}{2} + 2x - \frac{3}{x} + c$, m) $\frac{2}{5}y^{5/2} + \frac{2}{3}y^{3/2} - 5y + c$,

n) $\frac{4}{5}(x + 1)^{5/4} + c$, o) $\frac{6}{7}x^{7/6} + \frac{4}{3}x^{3/4} + c$,

p) $\frac{3}{4}z^{4/3} + \frac{4}{5}z^{5/2} + c$, q) $(y - 6)^5 + c$, r) $\frac{\sin^8\theta}{8} + c$

1-2, page 568

$5\sqrt{2}$ sec

1-3, page 568

The brick rises 36 ft and returns after 3 sec to the ground.

1-4, page 568

$c = 15$

1-5, page 568

a) $y = \frac{x^2}{2a} + c$, b) $y = \frac{x^2}{2a} - \frac{2}{a} - 4 + c$

Chapter 8

3-1, page 588

a) $\frac{1}{6}x^6$, b) $x^4 - \frac{2}{9}x^9 + \frac{1}{2}x^2$, c) $2(\sqrt{x} - \sqrt{a})$, d) $\frac{2}{3}x^{3/2}$,

e) $\frac{1}{4}(y^4 - 10^4)$, f) $\frac{1}{2}(u^4 + 2u^3 + 1)$,

g) $\frac{1}{15}[3(1 - z)^5 - 10(1 - z)^3 + 15(1 - z) + 8]$

3-2, page 588-589

a) $13\frac{1}{2}$, b) 15, c) $\frac{17}{6}$, d) $4\frac{2}{3}$, e) 2, f) $3\frac{1}{3}$,

g) -102, h) $\frac{1}{4}$, i) $\frac{1}{8}\sqrt{3}$, j) function not continuous at zero

k) $10\frac{2}{3}$, l) $-\frac{49}{162}$, m) $2(1 - \frac{1}{\sqrt{5}}$

3-3, page 589-590

a) $\frac{14}{3}$, b) $\frac{1}{3}$, c) $\frac{1}{4}(3^4 - 2^4)$, d) $\frac{1}{2}$, e) $\frac{14}{3}\sqrt{7} - 2\sqrt{3}$

f) $\sqrt[3]{5} - 1$, g) $2 + \frac{3\sqrt{2}}{2}$, h) $\frac{112}{27}$, i) $-\frac{\sqrt{2}}{6}$, j) $\frac{3}{16}$.

3-4, page 590-591

a) $\frac{1}{4}(2x + 1)^6 + c$, b) $\sqrt{u^2 - 1} + c$, c) $\frac{2}{3}x^3 - \frac{1}{2}x^2 - x + c$,

d) $\frac{2}{3}\sin\theta^{3/2} + c$, e) $\frac{2}{5}u^{5/2} + \frac{4}{3}u^{3/2} - 2u^{1/2} + c$,

f) $\frac{1}{2}x^2 - x - \frac{2}{x} + c$, g) $\frac{2}{5}\sqrt{y^5 + 2} + c$, h) $-\frac{1}{22}(\frac{1}{x^2} - 1)^{11} + c$,

i) $-\frac{1}{3}s^{-3} + \frac{5}{2}s^{-2} + \frac{1}{2}s^{2} + c$,   j) same as d,   k) $\arcsin \frac{t}{2} + c$,

l) $\frac{1}{2}\arcsin t^{2} + c$,   m) $\frac{1}{3}\arctan \frac{x - 2}{3} + c$,

n) $\frac{2}{3}(\arctan x)^{3/2} + c$,   o) $-\arcsin\left(\frac{\cos y}{\sqrt{10}}\right) + c$,

p) $\frac{1}{3}\arctan \frac{u - 4}{3} + c$,   q) $\frac{1}{6}\arctan 3r^{2} + c$.

<u>3-5, page 592</u>

a) $\frac{1}{6}a^{2}$,   b) $\frac{\pi}{15}a^{3}$

<u>3-6, page 592</u>

a) arctan h

b) $\lim_{h\to\infty} \arctan h = \frac{\pi}{2}$.

INDEX VOLUME I

Absolute convergence, 157
Absolute value, 17,33
Acceleration, 419
Algorithm, 65
  for area, 219 ff.
  for Fibonacci sequence, 108,158
  numerical, 73
Anti-derivatives, 562 ff.
Approximation
  of extrema, 455 ff.
  of function value, 470 ff.
  of integrals, 211 ff.
  of limit of sequence, 125 ff.
  of root of equation, 136
Area, 207,325 ff.
Arithmetic units, 103
Assignment box, 79
Assignment statement, 79
Axiom(s)
  Archimedean, 7
  completeness, 202
  of real numbers, 6

Bounded sequence, 173

Chain rule, 499
Circle, unit, 50
  area, 125,211,347
Coefficient
  of polynomial, 45
Completeness axiom, 202
Composition theorem, 399
Computer, 65
  concepts, 65
  language of, 105 ff.
  memory of, 77,94 ff.
  model of, 76 ff.
  word, 95
Compiling, 105
Control unit, 103
Convergence
  of sequences, 132,149, 153,154

Continuity, 386,387
Continuous functions, 386 ff.
Coordinate(s), 37
  system, 36
Cores, 95
Curve
  slope of, 410
Cycloid, 549

Decision box, 68
Definite integrals, 582
Derivative(s), 408 ff.
  chain rule for, 499
  of composite functions, 495 ff.
  formulas, 418,426,428, 429,527
  of implicit functions, 509
Differential(s), 541 ff.
Differentiation, 407 ff.
  explicit, 509
  implicit, 509

Error, 471
  in approximation, 219
  bound, 473
  roundoff, 118,438
Extended mean value theorem, 464

Floating point, 116
Flow chart, 67 ff.
  for maxima, 458
  for root of an equation, 140
  for trapezoidal rule, 226
Formulas for volumes and areas, 449
Fraction, 3
Function(s), 31 ff.
  circular, 49
  composition of, 57

Function(s) - con't.
constant, 38,542
continuous, 386
decreasing, 53
derived, 416
domain of, 33
graph of, 37
greatest integer, 47
identity, 39
increasing, 53
inverse trigonometric, 525 ff.
linear, 46
monotone, 53,479 ff.
piecewise monotone, 229
polynomial, 44
range of, 34
rational, 46
roots of, 46
strictly decreasing, 53
strictly increasing, 53
unicon, 254
Fundamental theorem of calculus, 575,579

Greatest integer function, 47

Identities, trigonometric, 275
Inequalities, 12,13
triangle inequality, 20
Input, 89
Input box, 90
Integer, 3
Integral
approximation of, 221 ff.
definition of, 235,317
definite, 582
indefinite, 582
Integration, 234 ff.
by substitution, 566
Intermediate value theorem, 403
Interval, 25 ff.

Law of the Mean, see Mean Value Theorem
Limit(s)
of a function, 374 ff.
of a sequence, 156
Lipschitz(ian)
coefficient, 293
condition, 293
function, 293,431
Local maximum tests, 446, 485,486
Local maximum point, 443
approximation of, 455

Machine language, 105
Maximum theorem, 441
Mean
arithmetic, 28
geometric, 28
Mean value theorem, 462
applications, 470
Modulus, unicon, 290 ff.

Numbers, 3 ff.

Ordinates, 36

Parametric equations, 548
Piecewise monotone functions, 229
Polynomial, 44
Pythagorean theorem, 4

Rate of change, 533
Rational numbers, 3
Real numbers, 5
Related rates, 533
Rolle's theorem, 463
Round-off error, 118,438

Samos, 94,103 ff.
Sequence, 125 ff.,149
   bounded, 173
   convergence of, 132,153,
      154,156
   Fibonacci, 158
   limit of, 156
Slope of a curve, 410
Solids of revolution, 340
Speed, 367
Squeeze theorem, 189,394
Step function, 48

Tangent line, 407
   equation of, 410 ff.
   slope of, 410 ff.
Taylor's theorem, 465
Trapezoidal rule, 266
Trigonometric functions,
      51 ff.
Trigonometric identities,
      275

Variable, 36,76
Velocity, 368
Volume, 337
   formulas for, 449
   of solid of revolution,
      340

Weighted average, 187
Work, 360

CALCULUS

PART 2

by
WARREN STENBERG
University of Minnesota
ROBERT J. WALKER
Cornell University

with the assistance of

GEORGE H. ANDREWS
Oberlin College
LAURA COOPER
Florida State University
URSULA EBERHARD
Florida State University
RAY L. GRAHAM
Appalachian State University
PAUL OHME
Florida State University
EMILE ROTH
Florida State University
EDWARD F. WISHART
University of Nevada
ROBERT J. WISNER
New Mexico State University

A COMPUTER ORIENTED PRESENTATION

CRICISAM

An Experimental Textbook Produced by
THE CENTER FOR RESEARCH IN COLLEGE
INSTRUCTION OF SCIENCE AND MATHEMATICS

635

Financial support for this project has been provided by the National Science Foundation (Grant #GY-3696).

Center for Research in College Instruction
of Science and Mathematics
Room 212 Diffenbaugh
Florida State University
Tallahassee, Florida 32306

First Printing, 1968-1969 in five parts
Second Printing, May 1969 in five parts
Third, Revised Printing, June 1970

CRICISAM is an organization devoted to the improvement of undergraduate instruction in science and mathematics. Its governing board is:

The National Advisory Committee for this calculus project is:

Richard D. Anderson, Louisiana State University
William F. Atchison, University of Maryland
Murray S. Klamkin, Ford Motor Company
Henry O. Pollak, Bell Telephone Laboratories
Warren Stenberg, University of Minnesota
Robert J. Walker, Cornell University
Robert J. Wisner, New Mexico State University
Gail S. Young, Tulane University, Chairman

The Local Coordinator at the grant site and host institution for this project is:

E. P. Miles, Director, Computing Center, Florida State University

The Director of CRICISAM is:

Guenter Schwarz, Florida State University

An expression of thanks is extended to the following persons who assisted in the preparation of this book:

G. J. Burks
Susan Carey
Kati B. Carus
Cynthia Corbin
Dorothea Driggs
Robert Ducharme
Tim Garland
J. B. Johnson
Susan McClellan
Judith Monette
Elizabeth Moore
Ilba L. Rosete
Linda Thompson
Patricia Vidzes

Special credit is given to Linda D. Thompson for her dedication in typing the manuscript and in coordinating activities concerned with the distribution of the text.

TABLE OF CONTENTS

VOLUMES I - II

Chapter 0 PRELIMINARIES

1	Introduction	1
2	Numbers	3
3	Inequalities	12
4	Absolute Value	17
5	Intervals	25
6	Functions	31
7	More on Functions	44
8	Composition and Inverses	57

Chapter 1 BASIC COMPUTER CONCEPTS

1	Algorithms and Flow Charts	65
2	A Numerical Algorithm	73
3	A Model of a Computer	76
4	Input	89
5	Computer Memory	94
6	Arithmetic and Control Units of SAMOS	103
7	Machine Language	105
8	Odds and Ends	116
9	Iteration Boxes	120

Chapter 2 SEQUENCES

1	Sequences of Approximations	125
2	Approximating Solutions of Equations	136
3	Problem Solution Using Sequences	142
4	Definition of Convergence	149
5	The Simplest Limit Theorems	165
6	Product and Quotient Theorems	172
7	The Squeeze Theorem	187

8 A Geometric Limit 194
9 Completeness 198

Chapter 3 AREA AND INTEGRAL

1 Area 207
2 An Algorithm for Area 219
3 Non-Monotone Functions 229
4 Integrals 234
5 Theory of Integration 241
6 Unicon Functions 252
7 Unicon Functions and Integrals 261
8 Formulas for Integrals 269
9 Tabulating the Sine and Cosine Functions 283
*10 The Unicon Modulus 290
A Moduli for Combinations of Unicon Functions 298
B Dispensing with Area 317

Chapter 4 APPLICATIONS OF INTEGRALS

1 Calculating Areas 325
2 Calculating Volume 337
3 Solids of Revolution 340
4 Another Characterization of the Integral 352

Chapter 5 LOCAL PROPERTIES OF FUNCTIONS

1 Speed and Limits 367
2 Limit Theorems and Continuity 384
3 Composition and Intermediate Value Theorems 398
4 Derivatives 407
5 The Derived Function 416
6 Derivative Theorems 424

Chapter 6 MAXIMA. THE MEAN VALUE THEOREM

1	An Example	439
2	The Maximum Theorem	441
3	Numerical Methods	455
4	The Mean Value Theorem	462
5	Approximations	470
6	Monotone Functions	479
*7	Existance of a Maximum	490

Chapter 7 THE CHAIN RULE

1	Differentiation of Composite Functions	495
2	Implicit Functions	509
3	Inverse Functions	519
4	Related Rates	533
5	Some Notations	539
6	Differentials	541
7	Parametric Equations	548
A	Theory of Inverse Functions	554

Chapter 8 THE CONNECTION BETWEEN DIFFERENTIATION AND INTEGRATION

1	Anti-derivatives	559
2	The Fundamental Theorem of Calculus	570
3	Second Form of Fundamental Theorem	578

VOLUME II

Chapter 9 THE LOGARITHMIC AND EXPONENTIAL FUNCTION

1	Logarithms and Exponentials	593
2	Differential Equations	615

Chapter 10 TWO APPLICATIONS OF DERIVATIVES

1	Convex Set and Functions	631
2	Convex Functions and Integration	645
3	Newton's Method	656
4	Evaluation of Limits	666

Chapter 11 TECHNIQUE OF INTEGRATION

1	Introduction	685
2	Algebraic Manipulation	691
3	Substitution	699
4	Integration by Parts	717
5	Partial Fractions	729
6	Integral Tables and Guessing	740
7	Definite Integrals	746
8	Improper Integrals	764
9	Numerical Methods	776
A	Algorithmic Treatment of Partial Fractions	796

Chapter 12 APPLICATIONS OF INTEGRATION

1	Cylindrical Shells	809
2	Generalization of the Integral	823
3	Work	832
4	Length of a Curve	835
5	Mean Value	846
6	Centroids	856

Chapter 13 INFINITE SERIES

1	Taylor Series	871
2	Convergence of Series	888
3	Tests for Convergence	905

4	Infinite Series and Improper Integrals	919
5	Power Series	932

Chapter 14 DIFFERENTIAL EQUATIONS

1	Numerical Solution	957
2	Graphical Solution	962
3	The Fundamental Theorem	968
4	Euler's Numerical Methods	975
5	Applications	992
6	Systems of Equations	1016
7	Separation of Variables	1034
8	Linear Equations	1054

Chapter 9

THE LOGARITHMIC AND EXPONENTIAL FUNCTIONS

1. Logarithms and Exponentials

We have just seen that $\int_a^b t^n dt$ is easy to evaluate for all rational values of n except $n = -1$. Obviously the formula

$$\int_a^b t^n dt = \frac{t^{n+1}}{n+1}\Bigg|_a^b$$

cannot be used when $n = -1$. On the other hand, since t^{-1} is a continuous function if the interval $[a,b]$ does not contain zero, the first form of the Fundamental Theorem tells us that there is a continuous function

$$L(x) = \int_a^x t^{-1} dt$$

with $L(a) = 0$ and $L(b) = \int_a^b t^{-1} dt$. What is this function $L(x)$? Do we already know it or is it something entirely new? To answer these questions we shall start by developing properties of $L(x)$ directly from its definition. To be entirely unambiguous we define

$$L(x) = \int_1^x t^{-1} dt, \qquad x > 0.$$

The first obvious things to notice about this function are:

(i) $L(1) = 0$;

(ii) If $x > 1$, then $L(x) > 0$;

(iii) If $0 < x < 1$, then $L(x) < 0$;

(iv) If $x > 0$ then $L'(x) = \frac{1}{x}$.

Property (i) is the simple computation

$$L(1) = \int_1^1 \frac{1}{t}dt = 0.$$

For (ii), the integral

$$\int_1^x \frac{1}{t}dt$$

can be represented by the shaded area in Figure 1-1(a) which is of course positive. If $0 < x < 1$, then the shaded area in Figure 1-1(b) is represented by

$$\int_x^1 \frac{1}{t}dt,$$

so

$$\int_x^1 \frac{1}{t}dt > 0.$$

Therefore, in this case

Figure 1-1

$$\int_1^x \frac{1}{t}dt = -\int_x^1 \frac{1}{t}dt < 0,$$

making (iii) true. Finally, (iv) follows at once from the first form of the Fundamental Theorem.

With this information we can estimate the shape of the curve $y = L(x)$. It goes through (1,0), is strictly increasing since $L'(x) > 0$, but the rate of increase decreases since $L''(x) = -1/x^2 < 0$. This gives us a shape somewhat like that shown in Figure 1-2 but we need more information before we can settle such items as asymptotes.

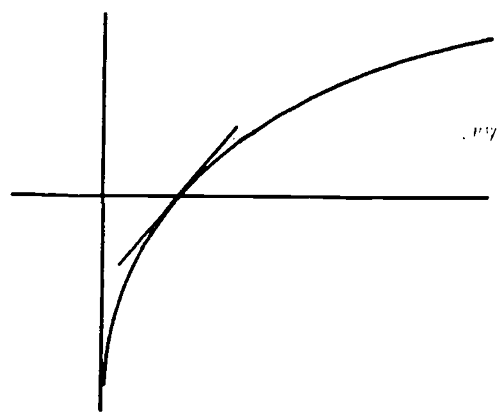


Figure 1-2

The next properties of $L(x)$ are not as obvious as the first four

(v) For any $a > 0$, $b > 0$,
$L(ab) = L(a) + L(b)$.

(vi) For any $a > 0$ and any rational number n,
$L(a^n) = nL(a)$.

Proof. (v) Consider the derivative of the function $L(ax)$. By the chain rule

$$D_x L(ax) = \frac{1}{ax} D_x(ax) = \frac{1}{ax} a = \frac{1}{x} = D_x L(x).$$

Since L(ax) and L(x) have the same derivative, by Theorem 2 of Section 8-1 they must differ only by a constant, i.e.

(1) $$L(ax) = L(x) + C.$$

The constant can be determined by putting $x = 1$. This gives, from (1),

$$L(a) = 0 + C, \quad \text{or} \quad C = L(a).$$

Putting this value for C in (1) and setting $x = b$ gives the desired result.

(vi) is proved similarly. We have

$$D_x L(x^n) = \frac{1}{x^n} nx^{n-1} = n\frac{1}{x} = D_x(nL(x)).$$

Hence

$$L(x^n) = nL(x) + C,$$

and setting $x = 1$ gives

$$0 = 0 + C,$$

so that the result follows on putting $x = a$.

Properties (v) and (vi) are the one we commonly associate with logarithms, and so the suspicion arises: Is L(x), after all, the same as log x, or perhaps closely associated to it? To show that the answer is "yes" we must investigate the inverse function of L(x), since log x is defined in terms of its inverse function: i.e. $y = \log_a x$ if $x = a^y$.

We have seen that L is strictly increasing on $(0,\infty)$ and so it has an inverse function which we call the <u>exponential function</u>, abbreviated by exp. That is, exp(L(x)) = x. Now the domain of exp is the range of L, and we should first of all determine this.

Let a be any number greater than 1. Then by (ii), $L(a) > 0$. Now $L(a^n) = nL(a)$, and since n can be arbitrarily large we see that L(x), for $x = a^n$, can be arbitrarily large. Similarly L(x) can be arbitrarily small. Thus the range of L, and the domain of exp, is $(-\infty, \infty)$. The range of exp, being the domain of L, is $(0,\infty)$.

Each of the properties (i) to (vi) of L can be translated into a property of exp, as follows:

(i') exp 0 = 1;

(ii') If $x > 0$ then $\exp x > 1$;

(iii') If $x < 0$ then $\exp x < 1$;

(iv') $D_x \exp x = \exp x$.

(v') $\exp(a + b) = (\exp a)(\exp b)$

(vi') $\exp(na) = (\exp a)^n$ for any rational n.

We shall prove (iv') and (v') as s mples, leaving the rest to the reader.

<u>Proof.</u> (iv') Since L and exp are inverse functions we have $L(\exp x) = x$. Differentiating gives

$$L'(\exp x)\, D_x(\exp x) = 1.$$

Since, by (iv),

$$L'(\exp x) = \frac{1}{\exp x}$$

this gives $D_x(\exp x) = \exp x$.

(v') Let $\exp a = A$, $\exp b = B$. Then $a = L(A)$, $b = L(B)$ and, by (v),

$$L(AB) = L(A) + L(B) = a + b$$

This, in turn is equivalent to

$$\exp(a + b) = AB = (\exp a)(\exp b).$$

Define

$$e = \exp 1.$$

Then if x is a rational number

(1) $\exp x = \exp(x \cdot 1) = (\exp 1)^x = e^x$.

Now if $x = \frac{p}{q}$ is rational, e^x has a definite meaning, namely $\sqrt[q]{e^p}$, (we can assume $q > 0$). For irrational values of x, e.g. $x = \sqrt{2}, \pi$, etc., e^x has not been defined. We now take equation (1) to be the definition in this case, so that for all values of x we have $\exp x = e^x$. Thus $\exp x$ and e^x are simply two ways of writing the same function. e^x is the more common, and makes it easier to apply properties (i') to (vi'). Exp x is common in computer work, and is convenient when x is a complicated expression, e.g.

$$\exp\left(\frac{1 - \exp\sqrt{1 - x^2}}{1 + \exp\sqrt{1 - x^2}} + 1\right).$$

Now since $y = L(x)$ is the same as $x = e^y$ we see that $L(x) = \log_e x$, the natural logarithm of x. Hence we will abandon the notation $L(x)$ in favor of $\log x$, the base e being understood. The important new property of log x is

(iv) $$D_x \log x = \frac{1}{x}.$$

Note also that from the definition of e,

$$\log e = 1.$$

How large is e? From Figure 1-3 we can see that $e < 4$. For the shaded area is larger than the sum of the areas of the rectangles,

$$\log 4 = \int_1^4 \frac{dx}{x}$$

$$> \frac{1}{2} + \frac{1}{3} + \frac{1}{4} > 1.$$

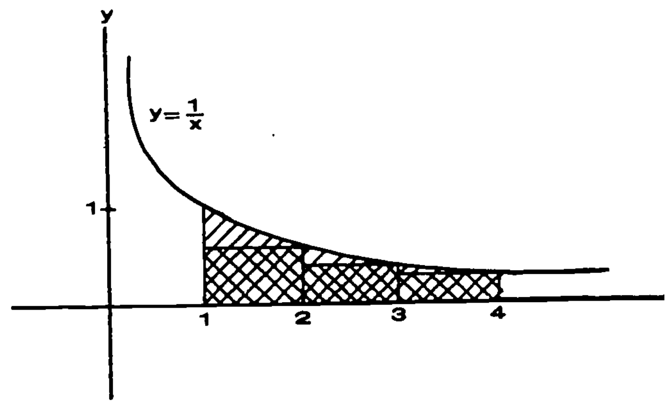


Figure 1-3

Hence, since exp is an increasing function, $\exp(\log 4) = 4 > \exp 1 = e$.

We now apply Taylor's Theorem (Section 6-3). If $f(x) = e^x$ then $f'(x) = e^x$, $f''(x) = e^x$, etc. Putting $a = 0$ and $b = 1$ in Taylor's Theorem gives us, since $e^0 = 1$,

$$e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \ldots + \frac{1}{n!} + R_n,$$

where

$$R_n = \frac{1}{(n+1)!} e^{\xi}, \quad 0 < \xi < 1.$$

Since e^x is increasing we therefore have, from the above bound on e,

$$0 < R_n < \frac{4}{(n+1)!}$$

Taking $n = 10$, for example, gives

$$e \approx 1 + 1 + \frac{1}{2!} + \dots + \frac{1}{10!} + \frac{2}{11!} \approx 2.71828185,$$

with an error of at most $\frac{2}{11!} < 5 \times 10^{-8}$.

We can now make a careful graph of $\log x$ and e^x, as in Figure 1-4. The two curves, being graphs of inverse functions, are symmetric with respect to the line $y = x$.

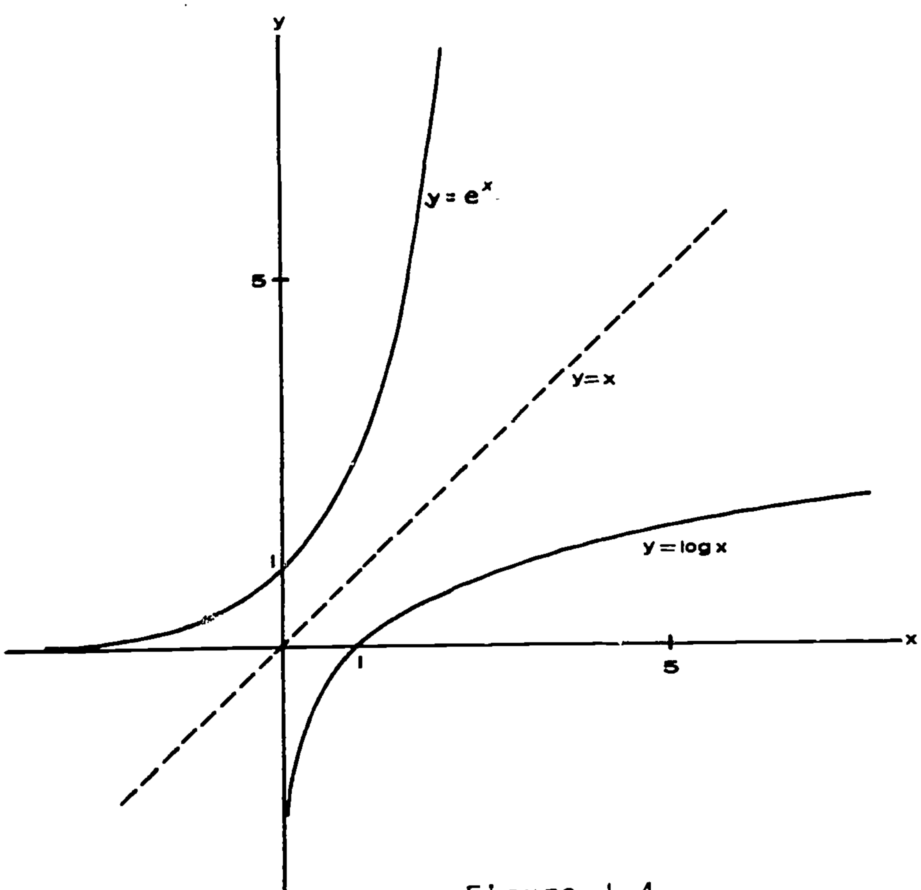


Figure 1-4

The negative x-axis is an asymptote to $y = e^x$, and the negative y-axis to $y = \log x$, but there are no other asymptotes; and, of course, no extrema since both functions are strictly increasing.

Example !. A function of the form

$$y = Ae^{-ht} \sin(\omega t + \theta),$$

where A, h, ω, and θ are constants, is said to define damped harmonic motion. A is the initial amplitude, h the damping factor, $\frac{\omega}{2\pi}$ the frequency, and θ the phase angle. The graph of such a function is shown in Figure 1-5. Evidently the graph crosses the x-axis when $\sin(\omega t + \theta) = 0$, and it is easy to show that it is tangent to one of the curves

$$y = Ae^{-ht} \quad \text{and} \quad y = -Ae^{-ht}$$

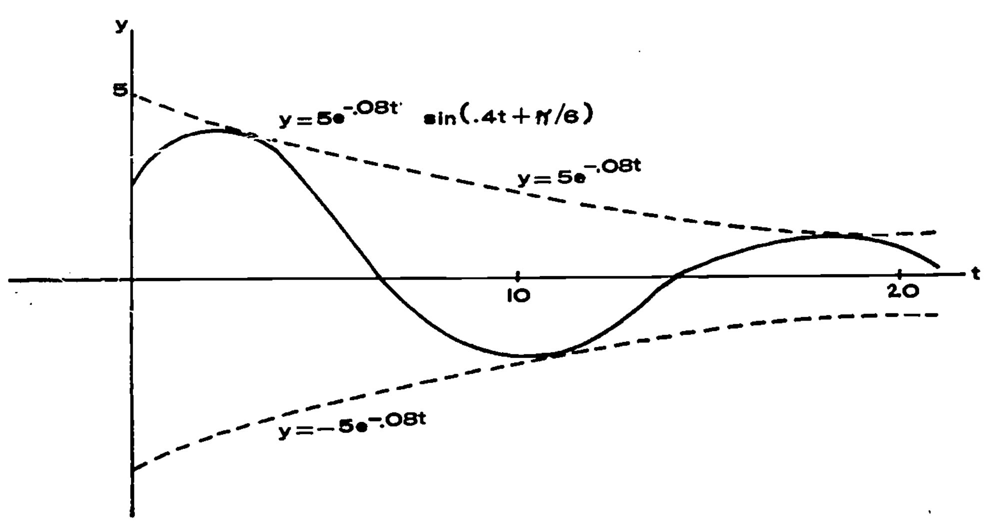


Figure 1-5

at each critical point of $y = \sin(\omega t + \theta)$, (Problem 17). This information makes it relatively easy to draw the curve.

The method of defining e^x for irrational values of x can be extended to a^x for any positive number a. If x is rational then

$$\log a^x = x \log a$$

and hence

$$a^x = e^{x \log a} .$$

We therefore take this equation as the definition of a^x for irrational values of x. It then follows (Problem 2) that

$$\log a^b = b \log a$$

and

$$(a^b)^c = a^{bc}$$

for any positive a and any values of b and c.

Now we finally complete our formula for the derivative of x^n.

<u>Theorem 1.</u> $D_x x^n = n x^{n-1}$ for any n and any positive x.

Proof. Let $y = x^n$. Then

$$\log y = n \log x,$$

and

$$\frac{1}{y} y' = n \frac{1}{x},$$

$$y' = n \frac{y}{x} = nx^{n-1}.$$

We have already proved that this formula also holds for all values of x if $n = \frac{p}{q}$ where p is any integer and q is an odd integer.

The corresponding anti-derivative, or indefinite integral formula is:

$$\int x^n dx = \begin{cases} \frac{1}{n+1} x^{n+1} + c, \text{ if } n \neq -1, \text{ for } \begin{cases} \text{all } x, \text{ if n} \\ \text{p and q int} \\ \text{q odd;} \\ x > 0 \text{ other} \end{cases} \\ \log x + c, \text{ if } n = -1, \text{ for } x > 0 \\ \log(-x) + c, \text{ if } n = -1, \text{ for } x < 0. \end{cases}$$

The last case is easily checked. Some writers combine the last two cases in the form

$$\log |x| + c, \quad \text{for all } x \neq 0.$$

This is convenient to use but is subject to the danger that one may forget the restriction $x \neq 0$ and jump from positive to negative values of x with disastrous results. It is safer, and often more illuminating, to consider what the sign of x is and use the appropriate formula for the integral. We shall see some examples of this in the next section.

We can add here another basic indefinite integral.

$$\int e^x dx = e^x + c$$

This follows from the fact that $D_x e^x = e^x$.

Our method of getting $D_x x^n$ is of enough general use to be given a name, "logarithmic differentiation." It is particularly useful for functions expressed as products, quotients, and powers.

Example 2. Differentiate $e^{2x}\sqrt{\dfrac{\sin 2x}{x \log x}}$.

Setting y equal to the given expression and taking logs gives

$$\log y = 2x + \tfrac{1}{2}[\log \sin 2x - \log x - \log \log x].$$

Hence

$$\frac{y'}{y} = 2 + \frac{1}{2}\left[\frac{1}{\sin 2x}(\cos 2x)2 - \frac{1}{x} - \frac{1}{\log x}\,\frac{1}{x}\right]$$

and so

$$y' = \left[2 + \cot 2x - \frac{1}{2x}\left(1 + \frac{1}{\log x}\right)\right]e^{2x}\sqrt{\frac{\sin 2x}{x \log x}} \,.$$

Example 3. Differentiate x^x. If $y = x^x$ then

$$\log y = x \log x$$

$$\frac{y'}{y} = \log x + x \cdot \frac{1}{x}$$

$$= \log x + 1$$

$$y' = x^x(\log x + 1).$$

PROBLEMS

1. Prove (i'), (ii'), (iii'), and (vi').

2. Prove:

   (a) $\log a^b = b \log a$,

   (b) $(a^b)^c = a^{bc}$.

3. Prove: $\log_a x = \dfrac{\log x}{\log a}$.

4. Derive:

   (a) $D_x\, a^x = a^x \log a$,

   (b) $D_x \log_a x = \dfrac{1}{x \log a}$.

5. Are $2 \log x$ and $\log x^2$ identical functions?

6. Show that if $f(x) = Ae^{kx}$ then $f'(x) = kf(x)$.

7. Prove that if $0 < a < b$ then

   $a^x < b^x$ if $x > 0$,

   $a^x > b^x$ if $x < 0$.

8. Differentiate each of the following and simplify if possible:

   (a) e^{ax}, a constant

(b) $(1 + e^x)^2$

(c) $x^2 e^{x^2}$

(d) $\log (x + 3)$

(e) $\frac{e^x - 1}{e^x + 1}$

(f) $\frac{\log t}{t}$

(g) $\exp \sqrt{z^2 + 1}$

(h) $\frac{e^{3x}}{\log (e^{3x} + 1)}$

(i) $e^{-2x} \sin 2x$

(j) $3^x \cdot x^3$

(k) $x(\log x - 1)$

(l) $\log \sin x$

(m) $\log (\sec x + \tan x)$

(n) $\log (x + \sqrt{x^2 + 1})$

(o) $\arctan (\frac{e^x - e^{-x}}{2})$

(p) $\log \log (1 + x^2)$

(q) $(x^2 - 2x + 2) e^x$

(r) $e^{ax} \sin bx$

(s) $e^{3x} (4 \sin 4x + 3 \cos 4x)$

9. Find $\frac{dy}{dx}$ in each of the following and simplify if possible:

(a) $y = x^2 \sqrt{1 + x^2}\, e^{-2x}$

(b) $y = (x + 1)(\quad)^2(x + 3)^3(x + 4)^4$

(c) $y = \dfrac{e^{-x} \sin \quad 2x}{x^2 \sin 3x \cos 4x}$

(d) $x^y = y^x$

(e) $\log \sqrt{x^2 + y^2} + \arctan \frac{y}{x} = 10.$

Ans. $\frac{y - x}{y + x}$.

10. Find the local maximum and minimum points of each of the following functions, and draw their graphs:

(a) $y = xe^x$

(b) $y = x^2e^{-x}$

(c) $y = 10\, e^{-x^2/2}$

(d) $y = \dfrac{e^x + e^{-x}}{e^x - e^{-x}}$

(e) $y = x \log x$

(f) $y = \dfrac{\log x}{x}$

11. Evaluate the following integrals:

(a) $\int e^{-5x}\, dx$

(b) $\int x^2 e^{x^3}\, dx$

(c) $\int \frac{dx}{x - 1}$

(d) $\int \frac{t^2 + 1}{t}\, dt$

(e) $\int \frac{t}{t^2 + 1}\, dt$

(f) $\int \frac{e^x}{e^x + 1}\, dx$

(g) $\int \frac{1}{1 + e^{-x}}\, dx$

(h) $\int \frac{dx}{2 + 3e^{4x}}$

(i) $\int \frac{e^z}{1 + e^{2z}}\, dz$

(j) $\int \frac{\cos x}{\sin x}\, dx$

(k) $\int \tan x\, dx$

(l) $\int_{-1}^{1} e^{-x}\, dx$

(m) $\int_0^2 \frac{du}{e^u}$

(n) $\int_0^5 xe^{-x^2}\, dx$

661

(o) $\int_0^2 \frac{dx}{x + 1}$.

(p) $\int_0^1 \left(\frac{1}{x^2 + 1} - \frac{x}{x^2 + 1}\right) dx$

(r) $\int_0^2 \frac{dx}{x - 1}$

(s) $\int_0^{\pi} \tan x \, dx$

12. Show from a graph that

$$\int_1^x \log t \, dt = \int_0^{\log x} (x - e^t) \, dt,$$

and evaluate the latter integral thereby getting the indefinite integral of log x.

13. It is convenient to introduce some combinations of the exponential function as special functions:

$\sinh x = \frac{e^x - e^{-x}}{2}$ (hyperbolic sine)

$\cosh x = \frac{e^x + e^{-x}}{2}$ (hyperbolic cosine)

$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ (hyperbolic tangent)

$\coth x = \frac{e^x + e^{-x}}{e^x - e^{-x}}$ (hyperbolic cotangent)

(a) Find the derivatives of these four hyperbolic functions.

(b) Find the indefinite integrals of these four functic [Hint. tanh x and coth x are of the form $f'(x)/f(x)$.]

(c) Prove the identities:

(i) $\cosh^2 x - \sinh^2 x = 1$

(ii) $\sinh (x + y) = \sinh x \cosh y + \cosh x \sinh y.$

(iii) $\cosh^2 x + \sinh^2 x = \cosh 2x.$

14. Use Taylor's Theorem to compute e^2 and e^{-2} to 3 decimal place accuracy (i.e., with error less than .0005). Is their product 1, as it should be?

15. (a) Show that

$$D_x^n (\log(1 + x)) = (n - 1)!(1 + x)^{-n}(-1)^{n-1}.$$

(b) Use Taylor's Theorem, with a = 0, to compute log 1.2 to four decimal places.

(c) Make a flow chart for a computer program to compute log (1 + x) with error at most ε for $0 \leq x \leq .5$.

(d) Write the program and make a table of log(x + 1) f x = 0(.01).5 to 5D, (i.e. from x = 0 to .5, at intervals of .01, with error at most 5×10^{-6}).

16. (a) $-\log 2 + 2\log 3 - \log 5 = \log \frac{9}{10} = \log(1 - .1)$,

$3\log 2 + \log 3 - 2\log 5 = \log \frac{24}{25} = \log(1 - .04)$,

$3\log 3 - 2\log 5 = \log \frac{27}{25} = \log(1 + .08)$.

Compute the logs on the right hand side to 4D by hand and solve the three simultaneous equations for log 2, log 3, and log 5.

(b) Use 49/50, 99/100, and 1001/1000 to get log 7, log 11, and log 13.

This method has been used to get the logs of prime numbers, and from them the logs of other integers, to very many (e.g., 50) decimal places.

17. Given the curve C:

$$y = Ae^{-ht}\sin(\omega t + \theta).$$

(a) Show that C is tangent to one of

$$y = Ae^{-ht}, \quad y = -Ae^{-ht},$$

at any critical point of

$$y = \sin(\omega t + \theta).$$

(b) Find the local maximum and minimum points of C. In particular, find them to two decimal places for the extrema in Figure 1-5.

Ans. $t = \omega^{-1}(-\theta + \arctan(\omega/h) + n\pi)$, n, an integer.

(c) If h is very small compared with ω, (this condition is often written $h << \omega$) a critical point t_c of C is close to a critical point t_s of $y = \sin(\omega t$ · Show that to a first approximation $t_c - t_s = -h/\omega^2$. [Hint. In the answer to (b) change arctan (ω/h) to $\pi/2$ - arctan (h/ω), (justify this) and get the linear approximation to t as a function of z, where $z = h/\omega$.]

(d) Show that y satisfies the relation

$y'' + 2hy' + (h^2 + \omega^2)y = 0.$

2. Differential Equations.

Example 1. Many savings accounts today pay 5%, compounded quarterly. This means that one's balance is increased by 1.25% every three months. At the end of a year the balance is $(1.0125)^4 = 1.05095$ times its original value; the equivalent, therefore, of 5.095% compounded once a year. A few banks try to impress the public by compounding monthly; this increases the equivalent yearly compounding rate only to 5.116%. Probably no bank has advertized continuous compounding, but we can easily see how it would work.

Suppose $B(t)$ is our balance at any time t, and that t and $t + \Delta t$ are two consecutive times of compounding. Then

$$B(t + \Delta t) - B(t) = B(t)(.05)\Delta t,$$

or

$$\frac{B(t + \Delta t) - B(t)}{\Delta t} = .05\, B(t).$$

Taking the limit as $\Delta t \to 0$ gives

(1) $$B'(t) = .05\, B(t).$$

From Problem 6 of Section 1 we see that $B(t) = Ae^{.05t}$ satisfies this equation for any constant A. But are there perhaps other solutions? To answer this we write (1) in the form

$$\frac{1}{B(t)} B'(t) = .05. \tag{2}$$

This, of course, cannot be done if $B(t) = 0$, so we must restrict the range of $B(t)$ either to $(0,\infty)$ or $(-\infty,0)$. For the present we use $0 < B(t) < \infty$. (This is the only one that makes sense in our interest problem.) Now the left-hand side of (2) is the derivative of $\log B(t)$, and since the right-hand side is the derivative of $.05t$, the basic property of anti-derivatives tells us that

$$\log B(t) = .05t + c.$$

This can be written in exponential form;

$$\begin{aligned} B(t) &= e^{.05t + c} \\ &= e^{c}e^{.05t} \\ &= Ae^{.05t}, \end{aligned}$$

where $A = e^{c}$ is a positive constant. Thus we get nothing new from this.

For the case $-\infty < B(t) < 0$ we use the fact that $D_x \log(-B) = \frac{B'}{B}$ to end up with the same result except that in this case $A = -e^{c}$ is negative. Thus $B(t) = Ae^{.05t}$ includes all solutions for positive B and for negative B; and also the solution $B(t) = 0$ for all t, obtained by putting $A = 0$.

It is now easy to finish our interest problem. For we have

$$B(0) = Ae^{0} = A$$

$$B(1) = Ae^{.05} \approx 1.0512927\ A,$$

corresponding to a yearly compounded interest rate of about $5\frac{1}{8}\%$.

<u>Example 2</u>. "The earth's population is increasing 2% per year." This says

$$\Delta P = P(t + \Delta t) - P(t) = .02P(t)$$

if $\Delta t = 1$ year. We cannot assume that if $\Delta t = 1$ month we have

$$\Delta P = \frac{.02}{12}P(t),$$

for, as we saw in Example 1, this would lead to an annual increase greater than 2%. We ask ourselves the question : "What rate of continuous increase will give an annual increase of 2%?

If the rate of continuous increase is k then, as in Example 1, we are led to the equation

$$P'(t) = kP(t),$$

with the general solution

(3) $$P(t) = Ae^{kt}.$$

Our opening quotation says that

$$P(1) = 1.02\ P(0),$$

and so

$$Ae^{k} = 1.02A$$

$$e^{k} = 1.02$$

$$k = \log 1.02 \approx .01980263 \approx 1.98\%$$

We must be cautious in applying formulas like (3). Suppose we ask for the present-day increase in population in one-tenth of a second. Since

$P(0) = A \approx 3.4 \times 10^{9}$, and $t \approx .1 \times (3.1 \times 10^{7})^{-1} \approx 3.1 \times$

years, we get using the values of A, k, t and P(0) given ab

$$P(t) - P(0) \approx 3.4 \times 10^{9} [\exp(.0198 \times 3.1 \times 10^{-9}) - 1]$$
$$\approx .21 \text{ people}$$

Of course this is nonsense. In fact, in the strict sense there if no such thing as "continuous increases" of a quant whose values are integers. Our model of population growth is therefore, like all models,,an approximation that is applicable in suitable circumstances; in this case, only when we are dealing with large numbers of people.

Equations like (1) are often more easily handled in terms of differentials. Using simply B for B(t), we can write (1) in the form

$$\frac{dB}{dt} = .05\ B,$$

or, using differentials,

$$dB = .05\ Bdt.$$

The solution then proceeds as follows:

(4) $$\frac{1}{B}dB = .05\ dt,$$

and, taking anti-differentials,

$$\log B = .05t + c, \quad \text{if } B > 0,$$

$$\log (-B) = .05t + c, \quad \text{if } B < 0,$$

as before.

The use of differentials is particularly convenient in some complicated cases since they enable us to treat the two sides of the equation as separate problems in anti-differentials, or in indefinite integration. For example,

(5) $$\frac{dy}{dx} = \frac{y^2 + 1}{xy}$$

becomes

(6) $$\frac{y}{y^2 + 1}dy = \frac{1}{x}dx.$$

We integrate the left-hand side with the substitution $u = y^2 + 1$, giving

$$\int \frac{y}{y^2 + 1}dy = \int \frac{1}{2}\,\frac{1}{u}du = \frac{1}{2}\log u = \frac{1}{2}\log(y^2 + 1).$$

Hence, assuming $x > 0$,

$$\frac{1}{2}\log(y^2 + 1) = \log x + c,$$

$$\log(y^2 + 1) = 2\log x + 2c$$

$$= \log x^2 + 2c$$

$$y^2 + 1 = x^2 e^{2c}$$

$$= Ax^2,$$

$$y = \pm\sqrt{Ax^2 - 1}.$$

Note that since $y^2 + 1 > 0$ for all y we need not worr about $\log(y^2 + 1)$. In the case $x < 0$, $\log x$ is replaced b $\log(-x)$ in the first two lines of the solution but the res is unchanged.

Because of this interpretation and technique involving differentials, equations like (1), (4), (5) and (6) were called <u>differential equations</u>. The name has now been extended to any equation involving a function and one or more of its derivatives, even though in many cases it would be extremely awkward, to say the least, to express it in terms of differentials.

The technique we have been using for these "first order" equations (i.e., involving only the independent variable, the function, and its first derivative) is called "separation of variables" for the obvious reason. Not all first order equations can be handled this way; for example, the simple equation $y' = x + y$. In a later chapter we shall make a more thorough study of first order equations.

You have probably noticed that each of our solutions involved an arbitrary constant, C or A. This will evidently be the case whenever we use the separation of variables technique, since it involves an anti-differentiation. In most applied problems this constant is determined by some "side condition", usually the requirement that the function have a specified value of the independent variable for a given value of the dependent variable; in symbols, we require that $y(x_0) = y_0$ where x_0 and y_0 are given numbers. Such a side condition

will usually settle any other indeterminacy that may arise in the solution of the differential equation. Thus the condition $y(2) = -3$ on Equation (5) gives the solution

$$y = -\sqrt{\frac{5}{2}x^2 - 1}\ .$$

Example 3. A body moving with velocity v_0 enters a resistin medium which decelerates the body at a rate proportional to v^α, where $v = v(t)$ is the velocity at time t and α is a positive constant. When and where does the body stop?

We are given

$$\frac{dv}{dt} = -kv^\alpha,$$

where k is a positive constant. (It is usually preferable to adjust signs, if possible, so that physical quantities are positive). Separating variables gives

$$v^{-\alpha}dv = -kdt$$

$$\frac{v^{1-\alpha}}{1-\alpha} = C - kt.$$

Let us get rid of the C at once by putting in the initial condition: $v = v_0$ when $t = 0$.

$$C = \frac{v_0^{1-\alpha}}{1 - \alpha},$$

$$v^{1-\alpha} = v_0^{1-\alpha} - (1 - \alpha)kt.$$

Evidently the behavior depends on the sign of $1 - \alpha$. If $\alpha < 1$ then

$$v = 0 \quad \text{when} \quad t = \frac{v_0^{1-\alpha}}{(1 - \alpha)k}.$$

But if $\alpha > 1$ we have

$$v^{\alpha-1} = \left[\frac{1}{v_0^{\alpha-1}} + (\alpha - 1)kt\right]^{-1}$$

and no value of t makes $v = 0$. We can only say that $v \to 0$ as $t \to \infty$.

So much for the time. To investigate the distance traversed we have two choices. Since $v = \frac{ds}{dt}$ we can integrate $v\,dt$ to get the change in s. Considering the complexity of v as a function of t this is not appealing. The other choice is to go back to the differential equation and replace $\frac{dv}{dt}$ by $v\frac{dv}{ds}$. (See Problem 2 in Section 7-6.) This gives us

$$v\frac{dv}{ds} = -kv^{\alpha},$$

$$v^{1-\alpha}dv = -k\,ds,$$

$$\frac{v^{2-\alpha}}{2 - \alpha} = C - ks.$$

Determining C from $v(0) = v_0$, as before, we get

$$v^{2-\alpha} = v_0^{2-\alpha} - (2 - \alpha)ks.$$

Here $\alpha = 2$ is the dividing line. If $\alpha < 2$ the velocity becomes zero if and only if

$$s = \frac{v_0^{2-\alpha}}{2 - \alpha}.$$

If $\alpha > 2$ the body moves arbitrarily far into the resisting medium.

The most interesting case is $1 < \alpha < 2$. The body never stops but it never gets as far as $s = v_0^{2-\alpha}/(2 - \alpha)$. It is left to the reader (Problem 13) to show that it gets arbitrarily close to this point. The two critical cases, $\alpha = 1$ and $\alpha = 2$, are also left to the reader.

Since $\frac{dy}{dx}$ is the slope of the curve $y = f(x)$ at the point (x,y), a property of a curve involving the tangent line can lead to a differential equation. Figure 2-1 shows some of the geometric quantities that might be involved

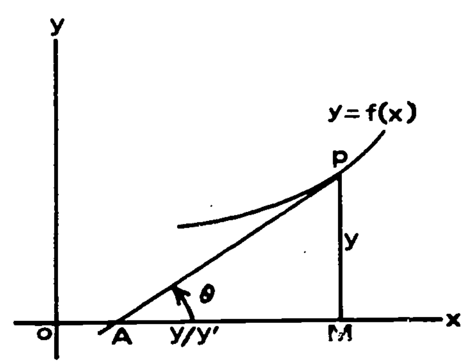


Figure 2-1

675

in such relations. Note that

$$OM = x, \quad MP = y, \quad \tan\theta = \frac{dy}{dx} = y', \quad AM = y/y'.$$

<u>Example 4</u>. A curve passing through the point (1,1) has the property that the tangent line at any point bisects the segment between the origin and the foot of the perpendicular to the x-axis from that point. What is its equation?

At first we ignore the point (1,1); this is the side condition that will eventually determine the constant. Referring to Figure 2-1, the condition given is that OA = AM, or $OM = 2AM$, or $x = 2y/y'$, or, finally, $y' = 2y/x$. We solve this in the successive steps:

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

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

(7) $\log y = 2 \log x + c = \log x^2 + c$

(8) $y = e^c x^2 = Ax^2.$

In (7) we use log y and log x as the integrals since both y and x are positive at our given point (1,1). Putting this point in (8) gives, finally, $A = 1$, and so $y = x^2$ is our desired curve.

PROBLEMS

1. Solve the given differential equation, subject to the given initial condition.

(a) $\frac{ds}{dt} = 32t + 5$ $\quad s = 100$ when $t = 0$

(b) $r^2 \frac{dr}{d\theta} = \sin\theta$, $\quad r = 1$ when $\theta = \frac{\pi}{4}$

(c) $yy' = x + 1$ $\quad y = 3$ when $x = 0$

(d) $\frac{dy}{dx} = \frac{x}{y}$ $\quad y = 5$ when $x = 2\sqrt{6}$

(e) $e^{y-x}\, y' + 1 = 0$ $\quad y = 2$ when $x = 2$

(f) $\frac{dz}{dt} = \sqrt{zt}$ $\quad z = 100$ when $t = 9$

2. Find the general solution of

$$\frac{dy}{dx} = ay + b,$$

where a and b are constants.

[Hint: There are two cases, depending on the value of a.]

3. (a) Using the data from Example 2, what was the population of the earth in 895 AD?

(b) What will it be in 3569 AD? If the average mass of a person is 120 pounds compare the total mass of people with that of the earth ($\approx 1.2 \times 10^{25}$ lbs).

4. The population of the United States was approximately 131,000,000 in 1940 and 179,000,000 in 1960. Assuming that the rate of increase is proportional to the population, in what year will the population be 250,000,000? When will it reach 300,000,000?

5. Radioactive elements decay at a rate proportional to the amount present. The rate of decay is usually measured by the "half-life", that is, the time required for an amount of the element to decay to half its present value.

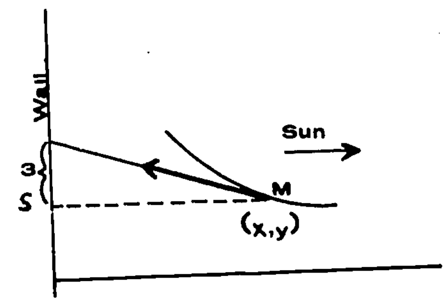
(a) Strontium 90, one of today's hazards, has a half-life of 5 years. Of an initial pound of ^{90}Sr how much will remain after t years?

(b) Of this initial pound how much will be left in 500 years? Be careful!

6. In Figure 2-1, if the triangle AMP has constant area 6, what is the equation of the curve if it goes through the point (2,2)? Sketch the curve.

7. In Figure 2-1, the line through P perpendicular to the tangent line is called the <u>normal</u> line. If the normal line meets the x-axis at N, the distance MN is called the <u>subnormal</u>. Find the general equation of a curve whose subnormal is constant.

8. The setting sun casts the little moron's shadow on a long wall. The moron wants to get ahead of his shadow, so he walks toward the wall, always heading for a point three feet ahead of his shadow. What is his path?



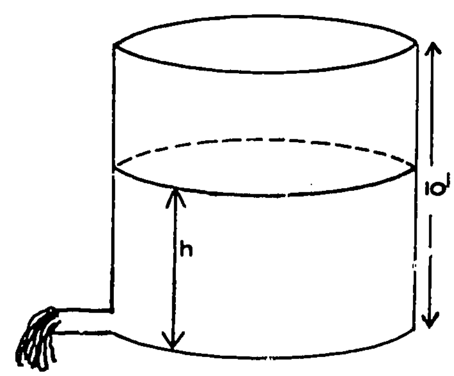
9. A spherical raindrop is losing water by evaporation. Assuming that the rate of los of water is proportional to the surface area find an expression for the radius as a function of time.

10. A 100-gallon tank is full of a solution of 2 pounds of salt per gallon of water. To flush the tank, fresh water is run in at 5 gal/min and solution is run out at the same rate. The fluid in the tank is stirred constantly, so that the salt concentration may be regarded as uniform throughout the tank at any moment. What is

the salt concentration at the end of one hour? How long would it take to get the concentration down to .01 lb/gal? [Hint. If s is the number of pounds of salt in the tank, what is $\frac{ds}{dt}$?]

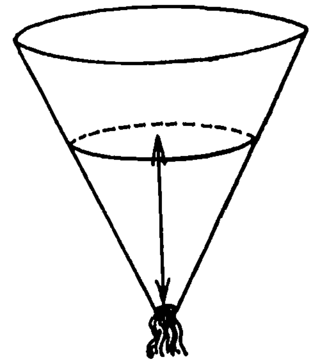
11. When the salt concentration in the tank in Problem 10 has been lowered to .01 lb/gal the process is reversed; brine containing 2 lb/gal is pumped in at 5 gal/min and the mixture runs out at the same rate. How long will it take for the concentration in the tank to reach 1.99 lb/gal?

12. (a) When water spurts from an orifice in an open tank its velocity is proportional to the square root of the distance from the orifice to the surface of the water. If it takes one hour for half the water to run out of an upright cylindrical tank how long will it take for .9 of the water to run out? For all of it to run out?



Ans. 2.335 hours; 3.414 hours.

(b) Answer the second question for an inverted conical tank.



Ans. 2.279 hours.

13. (a) In example 3, show that for $1 < \alpha < 2$,

$$s \to \frac{v_0^{2-\alpha}}{2-\alpha} \quad \text{as} \quad t \to \infty.$$

(b) Investigate the cases $\alpha = 1$ and $\alpha = 2$.

14. The differential equation

$$\frac{dy}{dx} = 3y^{2/3}$$

has <u>four</u> solutions on $(-\infty, \infty)$ that satisfy $f(0) = 0$. Can you find them?

Chapter 10

TWO APPLICATIONS OF DERIVATIVES

1. Convex Sets and Functions.

A set of points in the plane is said to be <u>convex</u> provided that you can travel from any point of the set to any other point of the set in a straight line without leaving the set. More precisely: If S is such a set of points that for any two points P,Q of S, the segment $\overline{PQ}$ is also contained in S, then we say that S is convex.

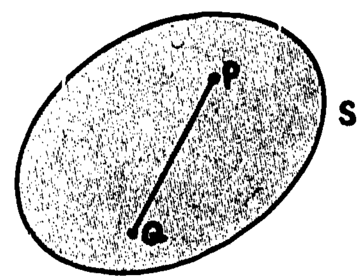


Figure 1-1

We show some examples of convex and non-convex sets in Figure 1-2.

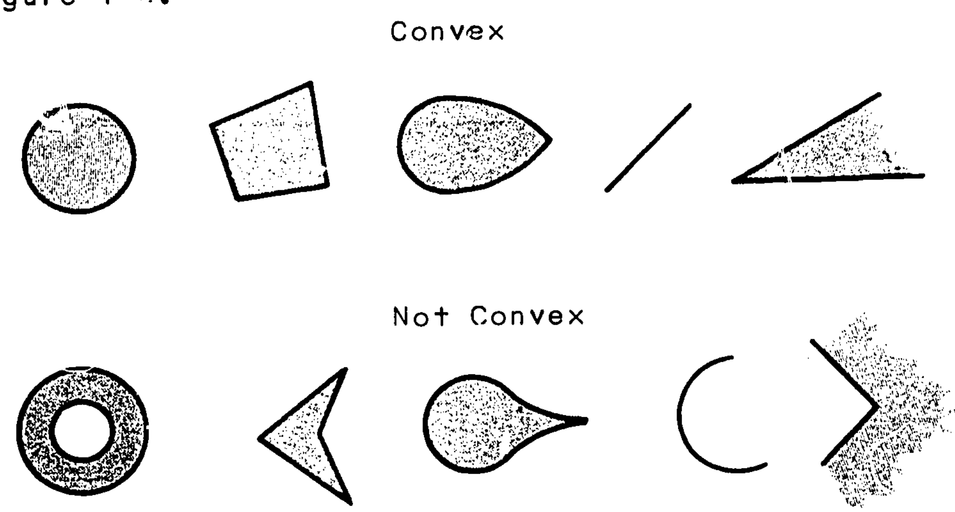


Figure 1-2

One of the basic properties of convex sets is that the intersection of two convex sets is in turn convex. To see this, suppose that P and Q are points in the intersection $S \cap T$ of the convex sets S and T, (Figure 1-3(a)). That P and Q are in $S \cap T$ means that these points are in S and also in T. Since they are both in S then the segment $\overline{PQ} \subset S$ and since they are both in T then the segment $\overline{PQ} \subset T$. $\overline{PQ}$, being contained in both S and T is contained in their intersection. [Note that the union of two convex sets is not necessarily convex as illustrated in the previous figure.].

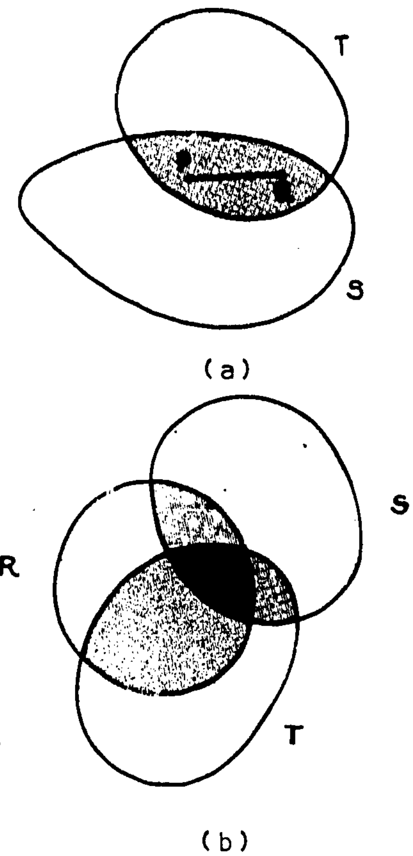


Figure 1-3

By the same reasoning we can see in more generality that the intersection of any number of convex sets is convex. This is illustrated in Figure 1-3(b), where the intersection $R \cap S \cap T$ is shown in black.

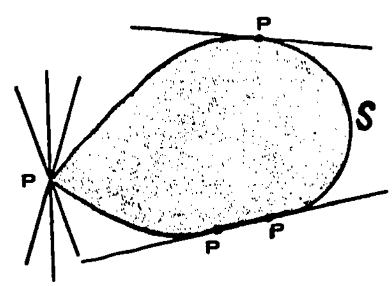
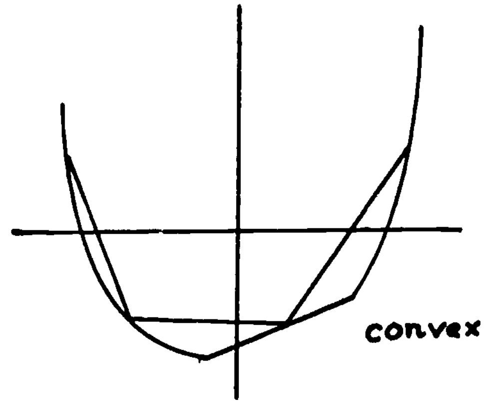


Figure 1-4

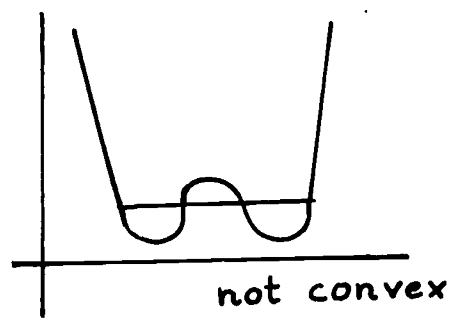
Of the many other properties of convex sets we shall mention (but not prove) only one. First a definition: If P is a point on the boundary of a convex set S, a line through P is called a <u>support</u> of S if S lies entirely on one side of the line. Figure I-4 shows some supports of S at various boundary points P.

<u>Theorem I.</u> A convex set has at least one support through every boundary point.

This seems obvious, but like many obvious-looking geometric theorems its proof is complicated. Since we do not use Theorem I in later proofs we shall not attempt to give a proof of it in this general form.



(a)



(b)

Figure I-5

A function is said to be convex provided that for any two points of its graph the arc of the graph joining these two points lies below the line segment (or chord) joining the points (Figure 1-5). This is equivalent to saying that the set of points lying above the graph is convex (Figure 1-6). [Here we use the words above and below in the sense of including the possibility of points on the curve or line. Hence by "the set of points above the graph of f" we mean the set of (x,y) for which $y \geq f(x)$. We will use the phrases "strictly above" or "strictly below" when we wish to exclude the possibility of a point being on the curve.]

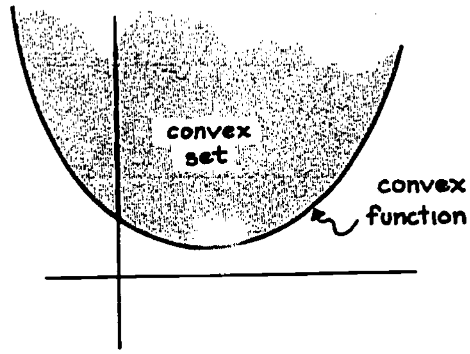


Figure 1-6

A function is said to be "concave" if the graph lies above the chord (i.e., the set of points below the graph is convex.) The

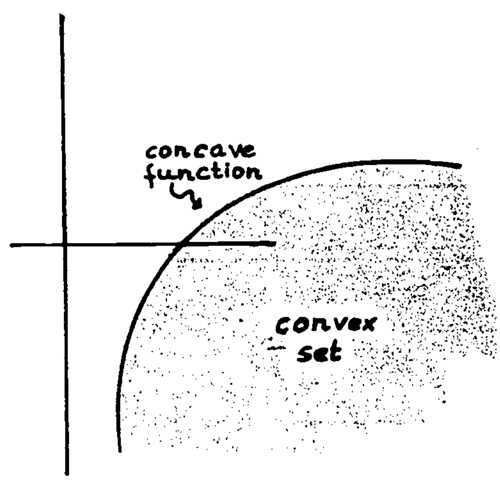


Figure 1-7

negative of a concave function is convex; since the properties of concave functions mirror those of convex functions, our study here will largely be confined to convex functions.

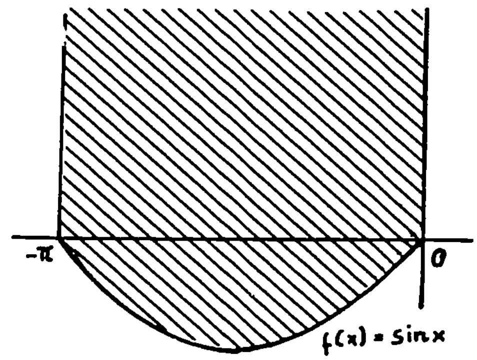


Figure 1-8

We speak of a function as convex over an interval if the function is convex when its domain is restricted to that interval. (Figure 1-8). Thus a function may be convex over some intervals and concave over other intervals. (Figure 1-9). When we say "the domain of convexity of a function", or some similar phrase, we mean an interval over which the function is convex. The function may be defined over a larger domain and may even be convex over some other intervals but this does not concern us.

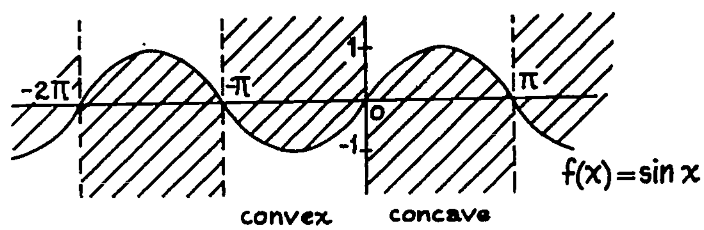


Figure 1-9

To pro · things about convex functions we shall use some properties of chords of their graphs. Let $a < x_1 < x_2$

be in the domain of the convex function f, and let A, P_1, P_2 be the corresponding points on the graph of f. Let m_1 and m_2 be the slopes of AP_1 and AP_2. Then $m_1 \leq m_2$.

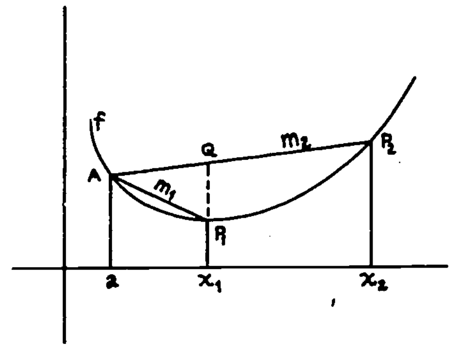


Figure 1-10

Proof. Let $Q:(x_1, y_1)$ be the point on the line AP_2 with abscissa x_1. Then, by the convexity of f, $f(x_1) \leq y_1$. Hence

$$m_1 = \frac{f(x_1) - f(a)}{x_1 - a} \leq \frac{y_1 - f(a)}{x_1 - a} = m_2.$$

This property can be expressed by saying that for fixed a and variable x the slope of the line joining (a,f(a)) and (x,f(x)) is an increasing function of x for $x > a$. It is not hard to generalize the property to the case $x \neq a$ but we shall not need this. One can also prove the converse; if f is a curve for which this property holds for every a in the domain then f is convex.

We are now in a position to prove a special case of Theorem 1.

Theorem 2. If f is convex and if $f'(a)$ exists then the tangent line at a is a support of f.

Proof. We wish to show that for any $x \neq a$, the point $(x, f(x))$ is above the tangent line. Since the equation of the tangent line is

$$y = f(a) + (x - a)f'(a)$$

this amounts to proving that

$$f(x) \geq f(a) + (x - a)f'(a).$$

Consider first the case $x > a$. Let $x = x_1, x_2, x_3, \ldots$ be a decreasing sequence with limit a. Then the slopes

$$m_n = \frac{f(x_n) - f(a)}{x_n - a}$$

also form a decreasing sequence, by the property we just proved. Since $\lim_{n\to\infty} m_n = f'(a)$, we therefore have $f'(a) \leq m_n$ for any given m_n; in particular for m_1. Since $x_1 = x$ this gives

$$f'(a) \leq \frac{f(x) - f(a)}{x - a},$$

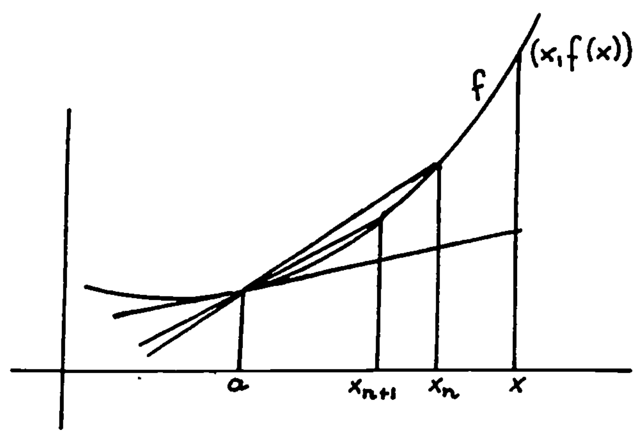


Figure 1-11

from which

$$f(x) \geq f(a) + (x - a)f'(a),$$

as desired.

For the case $x < a$ the easiest procedure is just to replace a by $-a$ and x by $-x$. The function g defined by $g(x) = f(-x)$ is convex, being just the reflection of f in the y-axis, and the above argument holds for $-x$ and the tangent at $-a$. Since vertical distances are unchanged in the reflection this gives us what we want.

We come finally to the more useful facts concerning convex functions.

Theorem 3. If f is convex and differentiable then f' is an increasing function.

Proof. Let $x_1 < x_2$. Since the tangent at x_1 is a support we have,

$$f(x_2) \geq f(x_1) + (x_2 - x_1)f'(x_1).$$

Since the tangent at x_2 is a support,

$$f(x_1) \geq f(x_2) + (x_1 - x_2)f'(x_2).$$

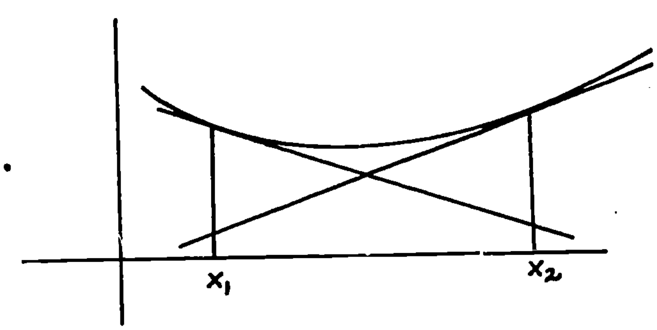


Figure 1-12

Since $x_2 - x_1 > 0$ the first inequality gives

$$f'(x_1) \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} ;$$

and the second,

$$f'(x_2) \geq \frac{f(x_2) - f(x_1)}{x_2 - x_1} .$$

Hence $f'(x_1) \leq f'(x_2)$, as was to be proved.

<u>Corollary 1</u>. If f is convex and if f"(a) exists then f"(a) $\geq$ 0.

<u>Proof</u>. For f"(a) to be defined f'(x) must exist in some neighborhood of a, and by Theorem 3, f' is increasing. Hence its derivative f"(a) must be $\geq$ 0.

One nice thing about these last two results is that their converses are true.

<u>Theorem 4</u>. If f' is increasing in the domain of f then f is convex.

<u>Proof</u>. Let P_1 and P_2 be any two points on the graph of f and Q any point on the segment P_1P_2. Let A be the point of the graph with the same abscissa, a, as Q. We have to show that Q is above A. By the Mean Value Theorem there

is a value ξ strictly between a and x_1, the abscissa of P_1, such that

$$f'(\xi) = \frac{f(a) - f(x_1)}{a - x_1} .$$

Since $\xi < a$ and f' is increasing we have

$$f'(\xi) \leq f'(a),$$

and so

$$f'(a) \geq \frac{f(a) - f(x_1)}{a - x_1} ,$$

or

$$f(x_1) \geq f(a) + (x_1 - a)f'(a).$$

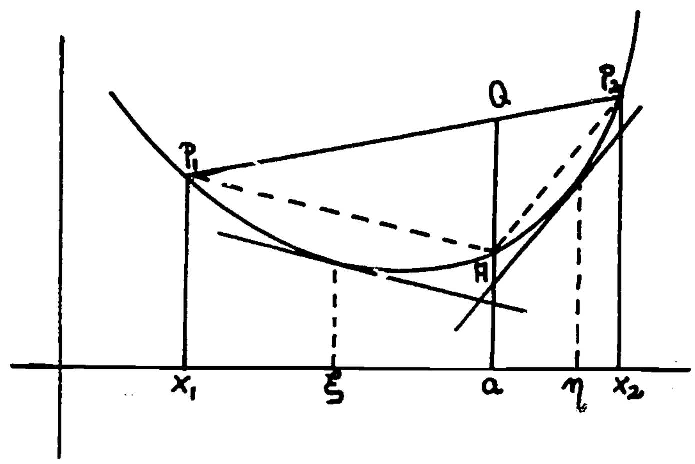


Figure 1-13

That is, P_1 is above the tangent line at A. In the same way we prove that P_2 is above the tangent line at A. Hence Q, being on the segment P_1P_2 must also be above this tangent line; which means, for Q, above A.

Corollary 2. If $f''(x) \geq 0$ in the domain of f then f is convex.

Proof. If $f''(x) \geq 0$ for all x then f' is an increasing function for all x.

It is convenient to combine Corollaries 1 and 2 into a single statement, which is important enough to be listed as a separate theorem.

Theorem 5. If f has a second derivative throughout its domain then f is convex if and only if this second derivative is always ≥ 0.

PROBLEMS

1. Which of these functions is convex or concave over the given interval? Justify your answer.

(a) $y = e^{-x}$, $(-\infty, \infty)$

(b) $y = \log x$, $(0, \infty)$

(c) $y = x e^{-x}$, $(-\infty, \infty)$

(d) $y = 2 - 3x$, $[-2, 2]$

(e) $y = |x - 1|$, $[0, 2]$

(f) $y = \sqrt{x}$, $[0, 10]$

(g) $y = |x - 1| + |x + 1|$, $[-2, 2]$

(h) $y = x^2 - 2$, $[-2, 2]$

(i) $y = |x^2 - 2|$, $[-2, 2]$

2. Suppose f and g are each convex on an interval [a,b]. For each of the following, either prove the statement or give an example contradicting it. [Assume f" and g" exist.]

(a) $h(x) = f(x) + g(x)$ is convex.

(b) $h(x) = f(x) - g(x)$ is convex.

(c) $h(x) = f(x)g(x)$ is convex.

3. Suppose f is convex on [a,b] and g is linear, i.e. $g(x) = cx + d$.

(a) Is $f(g(x))$ necessarily convex?

(b) Is $g(f(x))$ necessarily convex?

4. (a) Prove that $e^x \geq x + 1$ for any x. [Hint: Consider the tangent to $y = e^x$ at (0,1).]

(b) Prove that $\log x \leq x - 1$ for any $x > 0$.

5. If a is a point in the domain of f at which f'(a) exists and such that f is convex on one side of a and concave on the other, then the point (a,f(a)) of the graph of f is called a _point of inflection_, or a _flex_, of the graph.

(a) Prove that if (a,f(a)) is a flex and if f"(a) exists then f"(a) = 0. Is the converse true?

(b) Prove that the tangent line crosses the curve at a flex.

6. Find the points of inflection of the following curves and draw the curves.

(a) $y = x^3 + 3x^2$

(b) $y = x^4 - 3x^3 + 4$

(c) $y = \dfrac{1}{1 + x^2}$

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

(e) $y = x e^{-x}$

(f) $y = x^3\sqrt{1 - x^2}$

(g) $y = x \sin x$, $[0, \pi/2]$

(h) $y = x^2 - 2 + 2 \cos x$, $[-\pi/2, \pi/2]$.

7. Show that if a solution of the differential equation

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

has any points of inflection they must lie on the curve $x = y^2 + \dfrac{1}{2y}$. [Hint: When is $\dfrac{d^2y}{dx^2} = 0$?]

2. Convex Functions and Integration.

Back in Chapter 3, working with monotone functions, we were able to construct algorithms for approximating integrals by means of the trapezoid rule and to find an upper bound for the error in our approximation. The error estimate was based on the following principle: If A represents the true value of the integral, and L and U are upper and lower sums for a certain partition, then $L \leq A \leq U$. Moreover the trapezoidal approximation T for the same partition satisfies $T = (L + U)/2$, and so we see geometrically that since A lies between L and U,

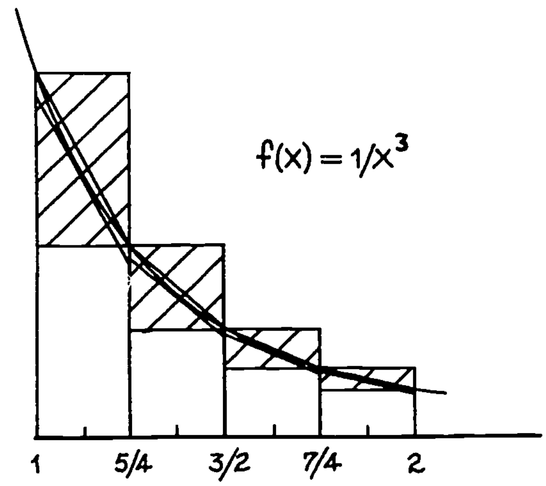


Figure 2-1

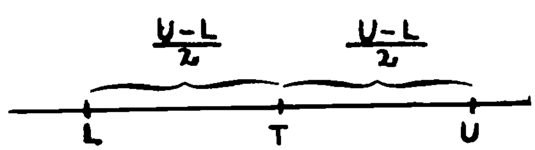


Figure 2-2

A must lie within a distance $\frac{U - L}{2}$ of T. Thus the error T-A satisfies

$$|T - A| < \frac{U - L}{2} .$$

In Figure 2-1 we have graphed $f(x) = x^{-3}$ over the interval [1,2] with 4 partitions. Here the shaded area represents U - L so that our bound for the error is half the shaded area. The actual error, represented by the sum of the four little areas between the chords and the curve, is barely visible on this graph.

It is clear that our error bounds were many times too large and this effect is even more pronounced when the number of partitions is large. In this section we will see for convex (or concave) functions how to approximate the integral with very much smaller bounds on the error.

It will be recalled that in Chapter 3 we discussed a second method of approximating areas in which we partitioned the interval, chose according to some rule a point ξ_k in the k^{th} subinterval for $k = 1, 2, \ldots, n$ and formed the sum

$$S = \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}).$$

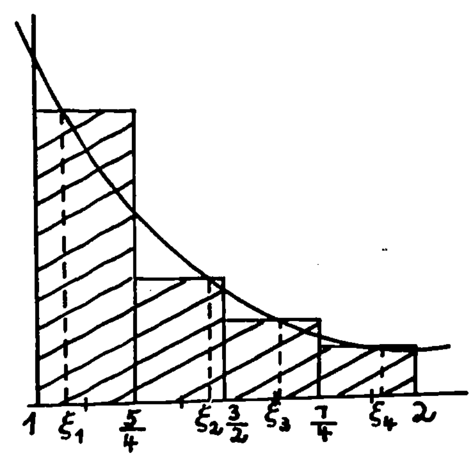


Figure 2-3

This is illustrated in Figure 2-3 where we have used the same function and the same partition as in Figure 2-1, the shaded area representing the value of the approximating sum S with $\xi_1, \xi_2, \xi_3, \xi_4$ chosen as indicated in the figure.

It would be reasonably natural to choose the points ξ_k at the midpoints of the subintervals as shown in Figure 2-4(a).

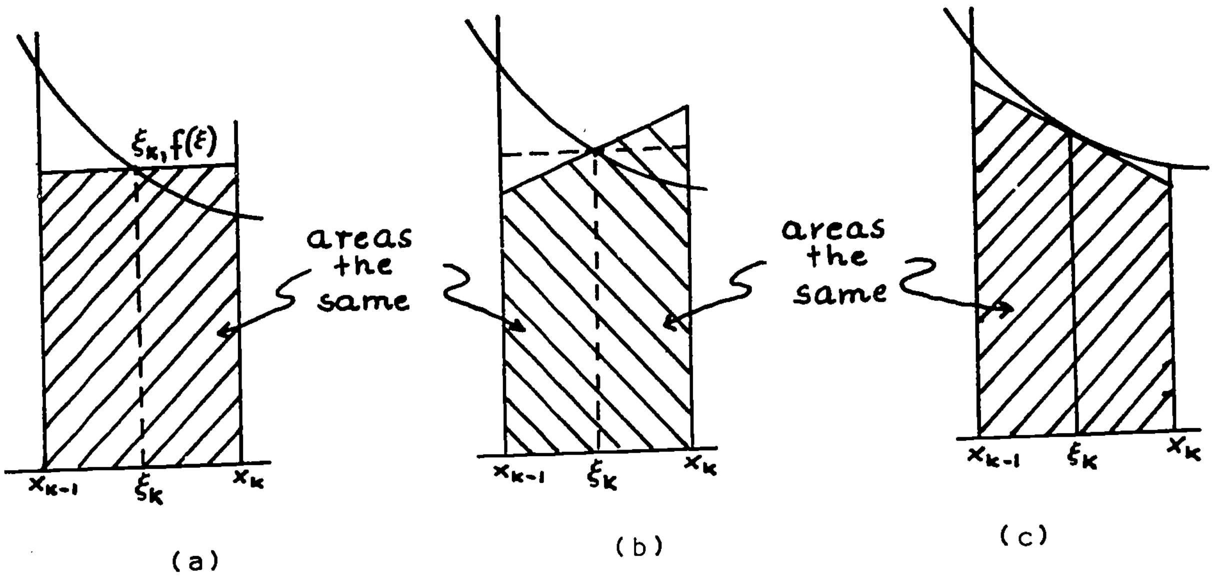


Figure 2-4

We find that this choice of the midpoint is quite special in that if we pivot the top edge of the rectangle in Figure 2-4(a) about its midpoint as in Figure 2-4(b) the area of the resulting trapezoid is the same as the area of the original rectangle. (The area "gained" on the right is equal to the area "lost" on the left.) In particular we can rotate this top edge so as to be tangent to the graph as seen in Figure 2-4(c). And now, if the function is convex, as depicted in Figure 2-4(c), the shaded area lies entirely below the curve.

And, on the other hand, the approximation given by the trapezoid rule is represented by an area which entirely contains the area under the curve as in Figure 2-5.

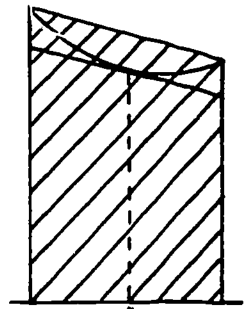


Figure 2-5

Since these observations apply for each subinterval in the partition, we find for convex functions the following inequalities between the integral, A, the midpoint approximating sum, M, and the trapezoidal approximating sum, T:

$$M = \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}) \leq A \leq \sum_{k=1}^{n} \frac{f(x_k) + f(x_{k-1})}{2}(x_k - x_{k-1}) = T.$$

Here $\xi_k = \dfrac{x_k + x_{k-1}}{2}$.

This is illustrated in Figure 2-6 where the areas represening M, A, and T are successively shaded. Note that nowhere in our argument was it required that f should be monotone.

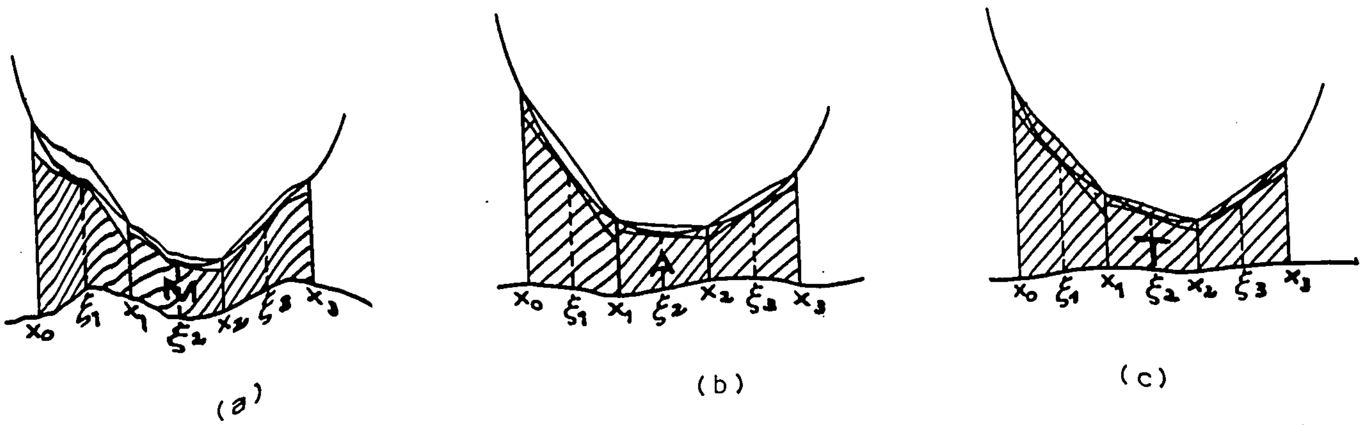


Figure 2-6

Since $M \le A \le T$ we see that if we use $\frac{T+M}{2}$ as an approximation for the integral, A, we obtain the error bound

$$\left|\frac{T+M}{2} - A\right| \le \frac{T-M}{2} .$$

Returning, for contrast, to the situation illustrated in Figure 2-1 we see in Figure 2-7 that the error estimate $\frac{T-M}{2}$ is represented by half the shaded area, which is hardly visible. We can see that a considerable improvement has been made on the error estimate illustrated geometrically in Figure 2-1.

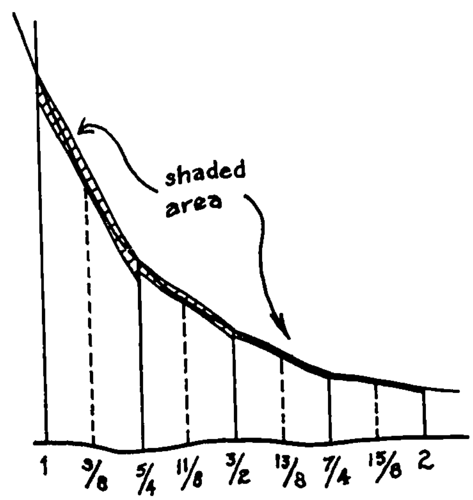


Figure 2-7

For concave functions the same analysis goes through with the inequalities reversed:

$$T \leq A \leq M.$$

The error estimate, written in the form,

$$\left|\frac{T + M}{2} - A\right| \leq \left|\frac{T - M}{2}\right|$$

will hold for concave as well as convex functions.

Now we wish to present an algorithm which computes integrals of convex and concave functions making use of these improved error estimates. We give the flow chart along with a partial explanation, leaving some of the details for you to work out in the following problem set.

Each time through the loop of Boxes 3 through 10 this algorithm successively halves the subinterval widths, w, and doubles the COUNT of the number of intervals (both in Box 8) and outputs (Box 7) the TRAPezoid and MIDpoint rule approximations of the integral. Boxes 3,4, and 5 sum up the functional values going into the midpoint rule calculation and the first line in Box 6 computes the midpoint approximation by multiplying this sum by the common width. The second assignment in Box 6 calculates the error estimate derived above.

701

Flow Chart

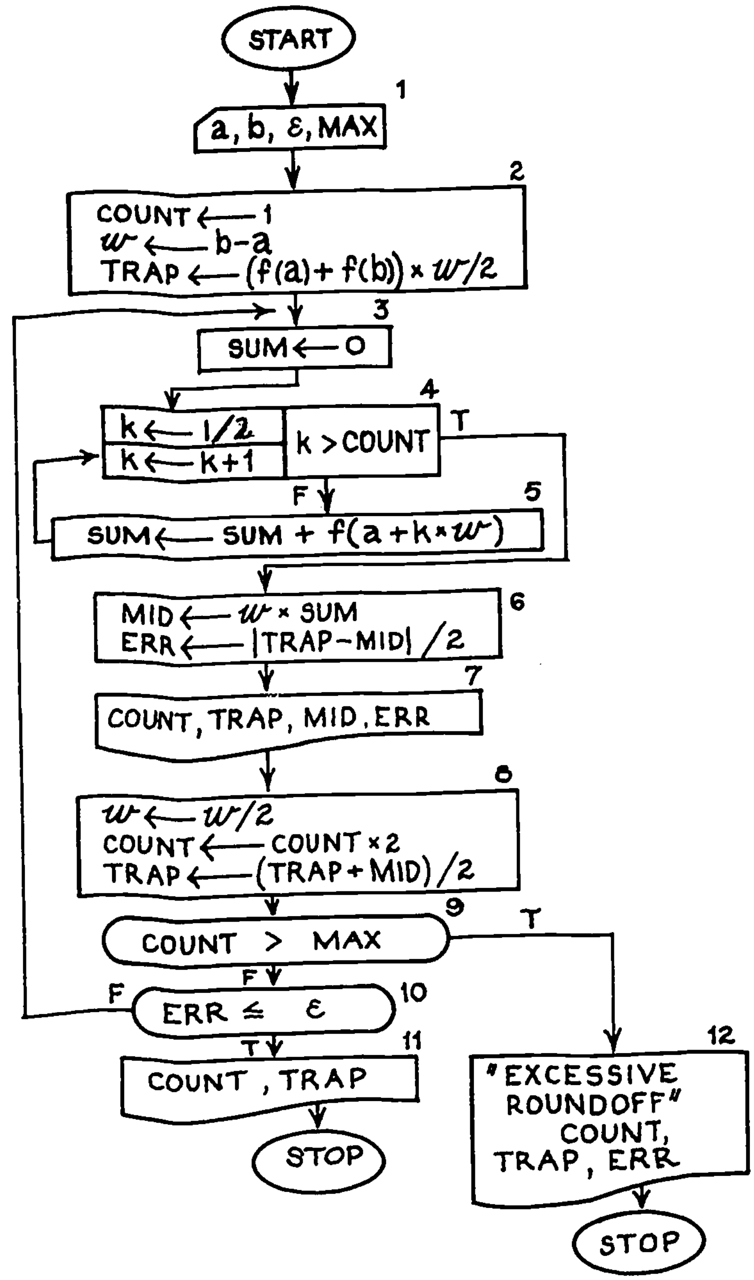


Figure 2-8

The value of TRAP is initialized in Box 2 with the appropriate value before the interval [a,b] has been subdivided at all. We leave it to you to work out how it is that the later values assigned to TRAP in Box 8 turn out to be correct.

The test in Box 9 is for roundoff effects. If COUNT is going to run into the high thousands or the millions one must worry about the accumulation of the errors that arise in the additions and the function evaluations in Box 5. The cumulative effect of these errors depends upon the kind of arithmetic used in the machine (fixed point or floating point - see Chapter 1), upon the behavior of the function f and the method of evaluating it, and possibly other factors. Knowing all these one can usually estimate an upper bound, MAX, to the number of terms that can be used in the summation without the total roundoff error exceeding ε. If ERR gets below ε before COUNT reaches MAX then the output is in error by at most 2ε. If COUNT exceeds MAX first then the roundoff error has possibly reached ε, and the possible error in the answer at this point is ERR + ε for the current value of ERR.

PROBLEMS

1. (a) For the function $f(x) = 1/x$ and the input

   $a = 1, \quad b = 2, \quad \varepsilon = 10^{-8}, \quad MAX = 4,$

   what should be the total output of the flow chart of Figure 2-8?

   (b) Carry through in detail all the steps of the flow chart for this case and see if you get what you should.

2. Write a program based on the flow chart of Figure 2-8. Check it with the case of Problem 1. Use it to approximate the following integrals.

   (a) $\int_0^1 \frac{2x}{1 + x^2}\, dx, \quad \varepsilon = 2 \times 10^{-6}$

   (b) $\int_0^5 \sqrt{25 - x^2}\, dx, \quad \varepsilon = 2 \times 10^{-4}$

3. In Figure 2-5 let $m = \xi_k$, $h = \frac{1}{2}(x_k - x_{k-1})$, so that $x_{k-1} = m - h$, $x_k = m + h$.

   (a) Use Taylor's Theorem to approximate $f(x)$ to terms of fourth degree, with $a = m$, $b = m + t$, to get

$$(1) \quad f(m + t) \approx f(m) + f'(m)t + \frac{1}{2}f''(m)t^2 + \frac{1}{6}f'''(m)t^3 + \frac{1}{24}f^{(4)}(m)t^4.$$

(b) The contribution of this one strip to the midpoint sum is $\Delta M = 2h\, f(m)$. Show that the contribution to the trapezoidal sum is

$$\Delta T = h[f(m + h) + f(m - h)]$$

$$\approx 2h\, f(m) + h^3 f''(m) + \frac{1}{12}h^5 f^{(4)}(m).$$

(c) Show that the approximation to the area of the strip obtained by integrating (1) is

$$\Delta A = \int_{-h}^{h} f(m + t)\,dt$$

$$\approx 2h\, f(m) + \frac{1}{3}h^3 f''(m) + \frac{1}{60}h^5 f^{(4)}(m).$$

(d) Show that

$$\Delta A \approx \frac{1}{3}\Delta T + \frac{2}{3}\Delta M$$

is good to terms of fourth degree in h, whereas

$$\Delta A \approx \frac{1}{2}(\Delta T + \Delta M)$$

is good only to terms of second degree.

4. Using the subdivision of Figure 2-7 for the function $f(x) = x^{-3}$, tabulated here, compute the following to four decimal places:

x	x^{-3}
1	1.00000
9/8	.70233
5/4	.51200
11/8	.38467
3/2	.29630
13/8	.23305
7/4	.18659
15/8	.15170
2	.12500

(a) The exact area under the curve, A;

(b) The trapezoidal approximation, T;

(c) The midpoint approximation, M;

(d) $A_1 = \frac{1}{2}(T + M)$;

(e) $A_2 = \frac{1}{3}T + \frac{2}{3}M$.

Discuss the advantages and disadvantages of A_1 and A_2 as an approximation to $\int_a^b f(x)dx$.

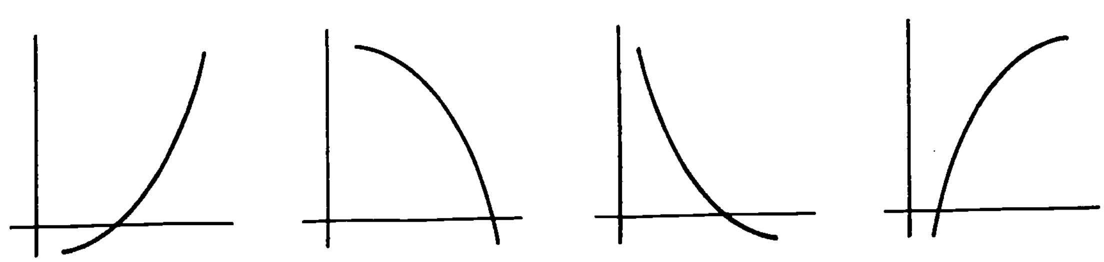
5. Can you think of any way of applying the results of this section to functions which are not convex or concave? For example, could they be applied to $\int_0^3 xe^{-x}dx$?

3. Newton's Method.

Back in Chapter 2 we encountered the "bisection algorithm" for the approximation of zeros of a function. We are going to develop in this section another algorithm, "Newton's method" for carrying out this task.

Newton's method is not so generally applicable as the bisection method. The bisection method required only that the function be continuous and assume opposite signs at the ends of our given interval. Newton's method requires both these conditions **and** as well that the function be strictly monotone and either convex or concave. Such conditions may be difficult to verify, even though they are generally satisfied in a sufficiently small neighborhood of the root. Weighed against these considerable disadvantages is the fantastically rapid convergence of Newton's Method when the favorable conditions obtain.

There are four cases to be considered, as shown in Figure 3-1.



(a) (b) (c) (d)

Figure 3-1

We will confine our attention to Case I (Figure I(a)) and observe that the other cases are handled similarly.

Suppose then that f is convex, increasing and differentiable in [a,b] with $f(a) < 0 < f(b)$. Note that these conditions imply that f will assume the value 0 at just one point of [a,b] which we call r.

It will be convenient to rename b as x_0. Draw the tangent to the graph of f at the point $(x_0, f(x_0))$ and note, since it is a line of support, that this tangent line will intersect the x-axis at a point, x_1, to the right of r as in Figure 3-2.

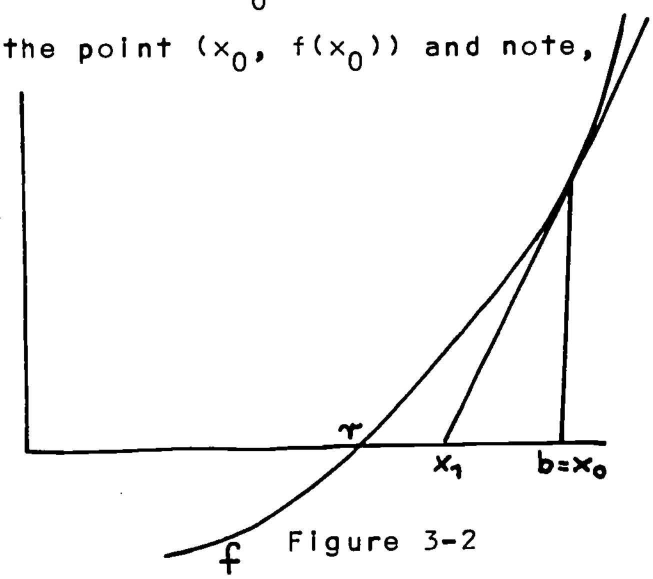


Figure 3-2

Since the slope of this tangent line is $f'(x_0)$ and can also be represented, as seen in Figure 3-2, as $\dfrac{f(x_0)}{x_0 - x_1}$, we can quickly solve for x_1:

$$\frac{f(x_0)}{x_0 - x_1} = f'(x_0)$$

$$x_0 - x_1 = \frac{f(x_0)}{f'(x_0)},$$

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Now we can iterate this process using x_1 in place of x_0.

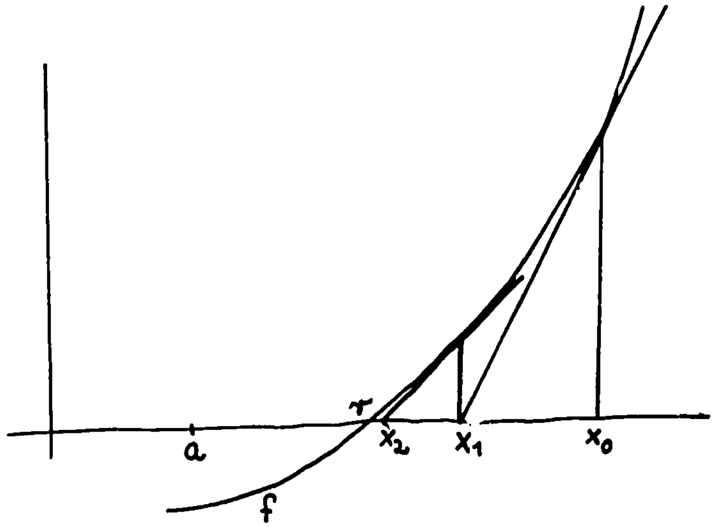


Figure 3-3

And so on.

As seen in Figure 3-3, these points seem to converge to r quite rapidly. Before considering <u>how</u> rapidly, let us see how to flow chart this process. We won't need successive names for the x_0, x_1, x_2, ... We will just create one variable XO which we allow to take on different values as seen in the flow chart fragment in Figure 3-4.

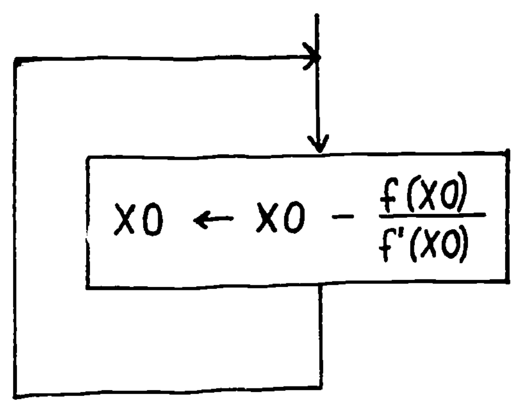


Figure 3-4

Of course we need to include a test for branching out of this loop to a stop. And that is where the error analysis comes in. Let us look first at the degree of reduction of the error after the first step. That is, let us compare

$$x_1 - r \quad \text{with} \quad x_0 - r.$$

Let us now assume that the second derivative of f is bounded by a number M on the interval [a,b] and also that $f'(a) > 0$. Consider now Figure 3-5 in which the tangent line at $(x_0, f(x_0))$ has been extended somewhat. The distance D is the deviation of the function f from the tangent line and so according to the Extended Mean Value Theorem,

(2) $D \leq \frac{M}{2}(r - x_0)^2.$

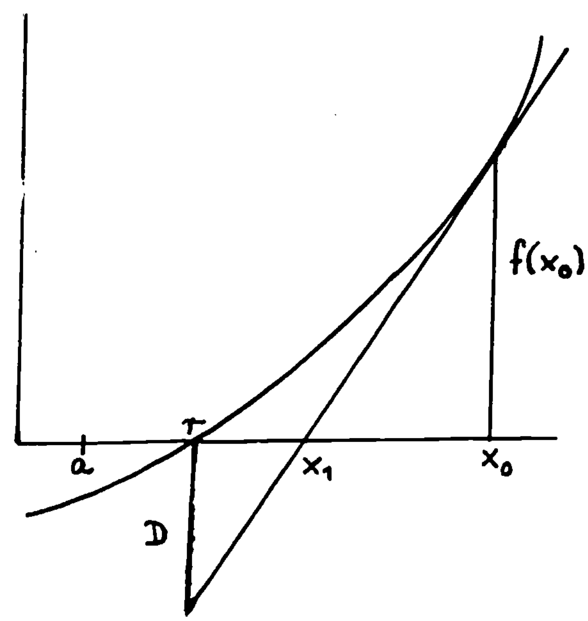


Figure 3-5

And since the slope of the tangent line is $f'(x_0)$ we see that

$$\frac{D}{x_1 - r} = f'(x_0).$$

Using this together with (2) yields

$$(x_1 - r) \cdot f'(x_0) \leq \frac{M}{2}(x_0 - r)^2,$$

or

$$x_1 - r \leq \frac{M}{2f'(x_0)}(x_0 - r)^2.$$

And since f' is an increasing function

$$f'(a) \leq f'(x_0),$$

so that

$$x_1 - r \leq \frac{M}{2f'(a)}(x_0 - r)^2.$$

Letting

$$Q = \frac{2f'(a)}{M},$$

this is most usefully written as

$$\frac{x_1 - r}{Q} \leq \left(\frac{x_0 - r}{Q}\right)^2. \tag{3}$$

Similarly, letting x_1 play the role of x_0,

$$\frac{x_2 - r}{Q} \leq \left(\frac{x_1 - r}{Q}\right)^2 \leq \left(\frac{x_0 - r}{Q}\right)^4,$$

and in general

$$\frac{x_n - r}{Q} \leq \left(\frac{x_0 - r}{Q}\right)^{2^n}. \tag{4}$$

To get an idea of what this means take a rather bad case, when f'(a) is small but f'(x) increases rapidly, so that M is large. Suppose that $Q \approx 1/10$. For formula (4) to insure convergence we would have to have $(x_0 - r)/Q < 1$, i.e., $(x_0 - r) < Q$. Suppose we take $x_0 - r = .05$. Then from (4) we get:

$$x_1 - r \leq (.5)^2/10 = 2.5 \times 10^{-2}$$

$$x_2 - r \leq (.5)^4/10 \approx 6.2 \times 10^{-3}$$

$$x_3 - r \leq (.5)^8/10 \approx 3.9 \times 10^{-4}$$

$$x_4 - r \leq (.5)^{16}/10 \approx 1.5 \times 10^{-6}$$

$$x_5 - r \leq (.5)^{32}/10 \approx 2.3 \times 10^{-11}$$

$$x_6 - r \leq (.5)^{64}/10 \approx 5.4 \times 10^{-21}$$

One thus gets some idea of the fantastic convergence. After only six iterations of the process we have an answer correct to 20 decimal places.

The above computation assumes exact arithmetic at each stage. On a computer accurate to, say, 15 places, the last step would certainly not hold, since the best we could ever hope for would be $x_6 - r \leq 5 \times 10^{-16}$. The earlier steps, however, would still be valid, the computation being too short for the accumulated roundoff errors to affect x_5 significantly.

Although the inequality (4) tells us the rate at which Newton's Method converges it does not tell us the actual accuracy at any step, since we don't know $x_0 - r$ to start with. However, we can very easily get a usable error bound from Figure 3-6. Here we have (again using convexity),

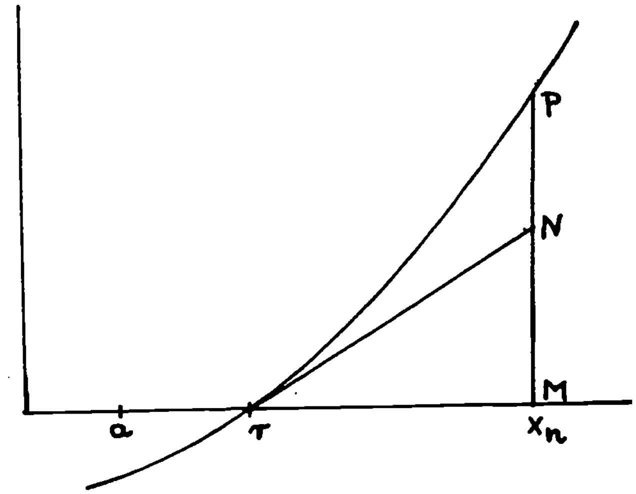


Figure 3-6

$$MN \leq MP \quad \text{or}$$

$$f'(r)(x_n - r) \leq f(x_n).$$

Hence

$$(5)\quad x_n - r \leq \frac{f(x_n)}{f'(r)} \leq \frac{f(x_n)}{f'(a)},$$

since, f' being increasing, $f'(a) \leq f'(r)$.

The flow chart is shown in in Figure 3-7. Its simplicity is striking. The absolute values are used in Boxes 2 and 4 to take care of the other three cases in Figure 3-1.

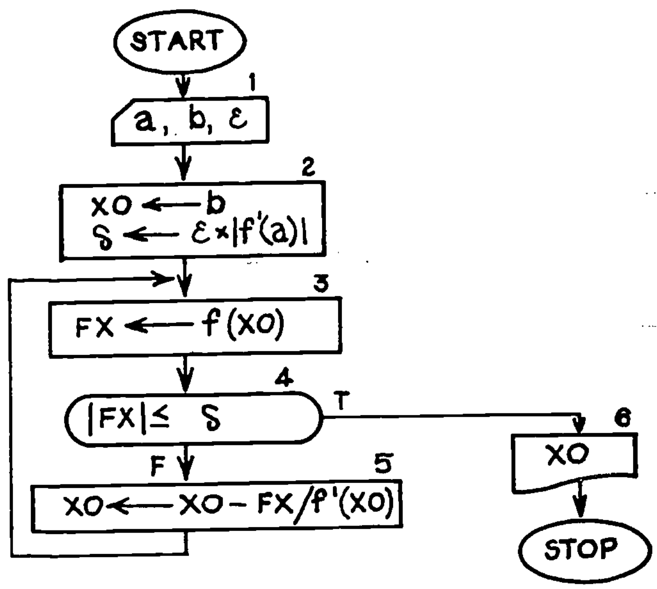


Figure 3-7

71

PROBLEMS

1. Carry out the following steps for Case IV, f concave and increasing, Figure 3-1(d).
   (a) Start with $x_0 = a$. Show that (1) holds unchanged.
   (b) Assuming $f'(b) < 0$, show that (3) and (4) hold if we replace each side of the inequality by its absolute value, and let $Q = 2f'(b)/M$.
   (c) Show that (5) holds if each side of the inequality is replaced by its absolute value and $f'(a)$ is replace by $f'(b)$.

2. In each of the following carry out two steps of Newton's Method and get a bound for the accuracy of your result.
   (a) $f(x) = x^3 - x - 1, \quad x_0 = 2.$
   (b) $f(x) = \cos x - x, \quad x_0 = 1.$

3. (a) Show that Newton's Method, applied to the function $f(x) = x^2 - c$, $c > 0$, reduces to the recursion formula

$$x_{n+1} = \frac{1}{2}(x_n + c/x_n).$$

   (b) Use this formula to find $\sqrt{8}$ accurate to 3 decimal places. Use an error bound to prove you have the desired accuracy.

7i

4. (a) In using the above recursion formula show that one can restrict the range of c to $.1 \leq c < 1$, as the square root of any other number can easily be computed from that of a number in this range.

(b) For a value of c in this range, starting with $x_0 = \frac{1}{2} + \frac{c}{2}$, how many iterations of the recursion formula will insure that the computed approximation to the square root will be accurate to 15 decimal places, neglecting roundoff? [Hint. Which value of c requires the most computation?]

(c) Is it safe to neglect roundoff in this case? Discuss the question.

5. How to divide by multiplication! Show that using Newton's Method on $f(x) = c - \frac{1}{x}$ gives

$$x_{n+1} = x_n(2 - cx_n).$$

This method has been used to do division on computers.

6. The flow chart in Figure 3-7 takes no account of roundoff errors. Where could such errors appear in a way to invalidate the output of the program? Modify the flow chart so as to correct this defect.

7. Write a program for the corrected flow chart, check it with the special cases of Problems 3 and 5, and use it to solve the following to as much accuracy as your computer permits.

(a) $x^3 = x + 1$

(b) $\cos x = x$

(c) $x^4 - 3x^3 + 7 = 0$ (2 roots)

(d) $2^x = x^2$ (3 roots)

4. Evaluation of Limits

In earlier chapters we have had to evaluate certain limits in order to derive integration and differentiation formulas. The two basic limits are

$$\lim_{n\to\infty} \frac{1}{n} = 0 \qquad \text{and} \qquad \lim_{x\to 0} \frac{\sin x}{x} = 1;$$

from these, others were obtained by using the properties of limits with respect to sums, products, etc.

We shall soon encounter limits arising from other processes, and also some extended versions of limits, for which these simple methods do not suffice to determine the value. In this section we shall see how some of the properties of the derivative, arising from the Mean Value Theorem, enable us to evaluate limits in a great many cases.

The most useful single technique is known as L'Hospital's (or L'Hôpital's) Rule.

Theorem 1. In the neighborhood of $x = a$ let f and g be continuous and differentiable, let $f(a) = g(a) = 0$, and let $g(x) \neq 0$ for $x \neq a$. If

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} = L$$

then

$$\lim_{x\to a} \frac{f(x)}{g(x)} = L.$$

The proof of this theorem is somewhat long and is deferred to the end of the section.

Example 1. $\lim_{x\to 1} \frac{\log x}{x^2-1}$.

We note that $f(x) = \log x$ and $g(x) = x^2 - 1$ satisfy the conditions of L'Hospital's Rule, so we consider

$$\lim_{x\to 1} \frac{f'(x)}{g'(x)} = \lim_{x\to 1} \frac{1/x}{2x} = \frac{1}{2}.$$

Hence the original limit is $\frac{1}{2}$.

Example 2. $\lim_{x\to 0} \frac{x - \sin x}{1 - \cos x}$.

The conditions are satisfied, so consider

$$\lim_{x\to 0} \frac{1 - \cos x}{\sin x}.$$

Here again both numerator and denominator are zero at $x = 0$, so we apply L'Hospital's Rule once more, to get

$$\lim_{x\to 0} \frac{\sin x}{\cos x}.$$

Here $\sin 0 = 0$ but $\cos 0 \neq 0$. We cannot apply L'Hospital's Rule but it is not needed. Since the limit of the denominator is not zero the limit of the quotient is the quotient of the limits, and our answer is zero.

Example 3. $\lim_{x\to 1} \frac{\sqrt{1 - x}}{\arccos x}$.

Since neither arccos x nor $\sqrt{1 - x}$ are defined for $x > 1$, Theorem 1 as stated does not apply here. However, the conditions are satisfied in a domain of the type: $0 < 1 - x < \delta$. A limit defined for this kind of domain is called a left-hand limit and designated by $\lim_{x\to 1-}$. Similarly we have right-hand limits $\lim_{x\to a+}$. The proof for L'Hospital's rule applies equally well to these one-sided limits.

Applying L'Hospital's rule to our problem leads us to

$$\lim_{x\to 1-} \frac{\frac{-1}{2\sqrt{1 - x}}}{\frac{-1}{\sqrt{1 - x^2}}} = \lim_{x\to 1-} \frac{\sqrt{1 - x^2}}{2\sqrt{1 - x}}$$

$$= \lim_{x\to 1-} \frac{1}{2}\sqrt{1 + x}$$

$$= \sqrt{2}/2 .$$

Another type of limit to which we can apply L'Hospital's Rule is $\lim_{x\to\infty}$. The definition of this type of functional limit is essentially the same as that of the sequential limit.

Definition 1. Let f(x) be defined in a domain of the type: $x \geq R$ for some R, (such a domain is called a neighborhood of infinity). Then $\lim_{x\to\infty} f(x) = L$ if for every $\varepsilon > 0$ there is an $N \geq R$ such that $|f(x) - L| < \varepsilon$ whenever $x > N$.

A definition of $\lim_{x\to-\infty} f(x)$ is left as an exercise. For limits as $x \to \pm\infty$ L'Hospital's Rule needs a restatement.

Theorem 2. In a neighborhood of infinity let f and g be continuous and differentiable, and $g(x) \neq 0$. Let $\lim_{x\to\infty} f(x) = \lim_{x\to\infty} g(x) = 0$. If

$$\lim_{x\to\infty} \frac{f'(x)}{g'(x)} = L$$

then

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = L.$$

Example 4. $\lim_{x\to\infty} x(\pi/2 - \arctan x)$. Here $x \to \infty$ and $(\frac{\pi}{2} - \arctan x) \to 0$, so we cannot say what the product will do. To use L'Hospital's Rule write it as

$$\lim_{x\to\infty} \frac{\pi/2 - \arctan x}{1/x};$$

then both numerator and denominator $\to 0$. We consider, then, the limit

$$\lim_{x\to\infty} \frac{-\frac{1}{1+x^2}}{-\frac{1}{x^2}} = \lim_{x\to\infty} \frac{x^2}{x^2+1}$$

$$= \lim_{x\to\infty} \frac{1}{1 + 1/x^2} = 1,$$

since $\lim_{x\to\infty} 1/x^2 = 0$.

Just as we have given a meaning to the replacement of a by ∞ in $\lim_{x \to a} f(x) = L$, so also we can give a meaning to the replacement of L by ∞.

Definition 2. If f(x) is defined in $0 < |x - a| < R$, for some $R > 0$, then $\lim_{x \to a} f(x) = \infty$ if for every M there is a $\delta > 0$ such that $f(x) > M$ whenever $0 < |x - a| < \delta$.

$\lim_{x \to a} f(x) = -\infty$ can be defined similarly, as also can $\lim_{x \to a+} f(x) = \infty$, $\lim_{x \to \infty} f(x) = \infty$, etc. L'Hospital's Rule holds for all of these if L is replaced by ∞ or by $-\infty$.

One must be careful with these infinite limits. Since ∞ and $-\infty$ are not numbers they cannot be combined with each other or with numbers by the rules of arithmetic. Even some very obvious-looking statements are not true; for instance it is not true that $\lim_{x \to 0} \frac{1}{x} = \infty$. What is true is that $\lim_{x \to 0+} \frac{1}{x} = \infty$, and $\lim_{x \to 0-} \frac{1}{x} = -\infty$, and $\lim_{x \to 0} \left|\frac{1}{x}\right| = \infty$. One useful property of infinite limits is the following substitute for Theorem 6 of Section 2-6.

Theorem 3. If $\lim f(x) = \infty$, and if g(x) has a positive lower bound then $\lim f(x)g(x) = \infty$.

The limit can be taken in any one of the types we have considered.

The proof is left as an exercise.

Example 5. (a) $\lim_{x\to\infty} x(2 - \cos x) = \infty$, since $2 - \cos x \geq 1$.

(b) $\lim_{x\to\infty} x(1 - \cos x)$ does not exist. $x(1 - \cos x)$ oscillates between zero values, at $x = 2\pi n$, and very large values, at $x = \pi(2n + 1)$, and so cannot have either ∞ or a number as a limit.

The final form of L'Hospital's Rule handles infinite limits of f and g. Like Theorem 3 it applies to all the different types of limits.

Theorem 4. In a suitable neighborhood let f and g be continuous and differentiable, and let $\lim |g(x)| = \infty$. If

$$\lim_{x\to\infty} \frac{f'(x)}{g'(x)} = L$$

then

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = L.$$

Here L may be a number, ∞, or $-\infty$.

Note that the theorem puts no direct requirement on f. Of course if f(x) is bounded and $\lim |g(x)| = \infty$ then

$$\lim \frac{f(x)}{g(x)} = \lim \frac{1}{g(x)} f(x) = 0.$$

without any further ado. But presumably f(x) could oscillate wildly, say like x(1 - cos x) in Example 5(b). The catch is that if f(x) behaved too badly so would f'(x), and $\lim f'(x)/g'(x)$ would not exist. It turns out that we are essentially restricted to the case when $\lim |f(x)| = \infty$ also.

Example 6. $\lim_{x\to 0+} x^a \log x,\ a > 0$. To bring this under Theorem 4 we write it as

$$\lim_{x\to 0+} \frac{\log x}{x^{-a}}.$$

To apply the theorem we take

$$\lim_{x\to 0+} \frac{1/x}{-ax^{-a-1}} = \lim_{x\to 0+} -\frac{1}{a}x^a = 0.$$

Example 7. $\lim_{x\to 0+} x^x$.

If $y = x^x$ then $\log y = x \log x$. By Example 6, $\lim_{x\to 0+} \log y = 0$. and so $\lim_{x\to 0+} y = e^0 = 1$.

Proof of L'Hospital's Rule.

To prove L'Hospital's Rule it is convenient first to prove a generalization of the Mean Value Theorem.

Generalized Mean Value Theorem. If f and g are functions continuous in $[a,b]$ and differentiable in (a,b) then there is a ξ in (a,b) such that

$$(f(b) - f(a))g'(\xi) = (g(b) - g(a))f'(\xi).$$

Notice that if $g(x) = x$ this equation reduces to the ordinary MVT. This is what the word "generalized" means; the original theorem is a special case of this one. The proof is a modification of the proof of the original MVT.

Proof. The function

$$h(x) = (f(b) - f(a))(g(x) - g(a)) - (g(b) - g(a))(f(x) - f(a))$$

has the following properties:

h is continuous in $[a,b]$,

h' exists in (a,b),

$h(a) = 0, \quad h(b) = 0.$

Hence by Rolle's Theorem, $h'(\xi) = 0$ for some ξ in (a,b). That is,

$$h'(\xi) = (f(b) - f(a))g'(\xi) - (g(b) - g(a))f'(\xi) = 0,$$

which proves the theorem.

Like the ordinary MVT, the GMVT has a geometric meaning. If

$$x = f(t), \quad y = g(t)$$

is a parametric representation of a curve from A: $(f(a), g(a))$ to B: $(f(b), g(b))$, as in Figure 4-1, then the GMVT says that at some point (there may be several) of the curve between A and B the tangent line is parallel to AB.

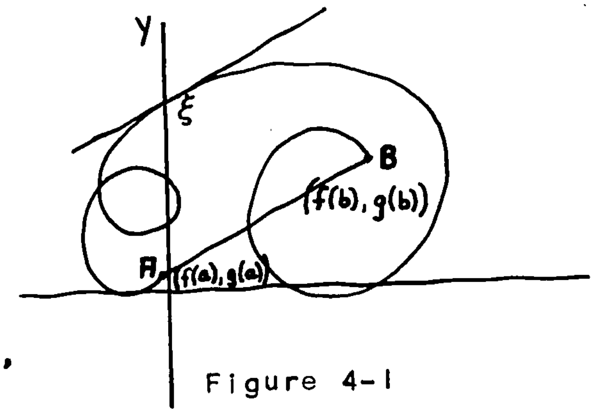


Figure 4-1

Proof of Theorem 1. Apply the GMVT with $b = x$. Since $f(a) = g(a) = 0$ we have

(1) $$f(x)g'(\xi) = g(x)f'(\xi),$$

with ξ between a and x. Since we are assuming that

$$\lim_{x \to a} \frac{f'(x)}{g'(x)}$$

exists there must be a $\delta_1 > 0$ such that for $0 < |x - a| \leq \delta_1$, we have $g'(x) \neq 0$. Since ξ lies between a and x we have $0 < |\xi - a| < |x - a| < \delta_1$, and so $g'(\xi) \neq 0$. We are given that for some $\delta_2 > 0$, $g(x) \neq 0$ for $0 < |x - a| < \delta_2$. Hence if we take $R = \min(\delta_1, \delta_2)$ we have that for $0 < |x - a| < R$ both $g(x) \neq 0$ and $g'(\xi) \neq 0$. Equation (1) can then be written

$$\frac{f(x)}{g(x)} = \frac{f'(\xi)}{g'(\xi)} .$$

[Comment. Division often takes a lot of justifying.]

Now, given $\varepsilon > 0$ there is a $\delta > 0$ such that if

(2) $$0 < |x - a| \leq \delta,$$

then

$$\left|\frac{f'(x)}{g'(x)} - L\right| \leq \varepsilon .$$

But, as we have seen, if x satisfies (2) so does ξ, and therefore

$$\left|\frac{f'(\xi)}{g'(\xi)} - L\right| \leq \varepsilon .$$

Since this is the same as

$$\left|\frac{f\ (x)}{g\ (x)} - L\right| \leq \varepsilon,$$

our theorem is proved.

For the case $L = \infty$ simply replace "given $\varepsilon > 0$" by "given $M > 0$" and replace all "$\leq \varepsilon$" by "$\geq M$". $L = -\infty$ is handled similarly.

This proof will apply also to the cases $x \to a+$ and $x \to a-$, since the essential point, that ξ satisfies the same inequality as x, is assured here also by the fact that ξ lies between a and x.

For the case $x \to \infty$ or $x \to -\infty$ we need the following fairly obvious fact.

Theorem 5. If $\lim_{x\to\infty} f(x) = L$, then $\lim_{y\to 0+} f(1/y) = L$.

Proof. Given $\lim_{x\to\infty} f(x) = L$ we want to prove that for any $\varepsilon > 0$ there is a $\delta > 0$ such that

$$|f(1/y) - L| < \varepsilon \quad \text{whenever} \quad 0 < y < \delta.$$

Now for the given ε, since $\lim_{x\to\infty} f(x) = L$, there is an N such that

$$|f(x) - L| < \varepsilon \quad \text{whenever} \quad x > N.$$

Take $\delta = 1/N$ if $N > 0$, otherwise $\delta = 1$. Then $0 < y < \delta$ implies that $1/y > N$, and so implies that $|f(1/y) - L| < \varepsilon$, as was to be proved. The converse is proved similarly.

Proof of Theorem 2. Letting $y = \frac{1}{x}$ we define

$$F(y) = \begin{cases} f(1/y) & \text{if } y \neq 0, \\ 0 & \text{if } y = 0, \end{cases} \qquad G(y) = \begin{cases} g(1/y) & \text{if } y \neq 0, \\ 0 & \text{if } y = 0. \end{cases}$$

Then

$$F'(y) = f'\left(\frac{1}{y}\right)\left(-\frac{1}{y^2}\right) = -x^2 f'(x), \quad G'(y) = -x^2 g'(x),$$

and by Theorem 5,

$$\lim_{y\to 0+} \frac{F'(y)}{G'(y)} = \lim_{x\to\infty} \frac{-x^2 f'(x)}{-x^2 g'(x)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)} = L.$$

Since F and G satisfy the conditions of Theorem 1 we then have

$$\lim_{y\to 0+} \frac{F(y)}{G(y)} = \lim_{y\to 0+} \frac{F'(y)}{G'(y)} = L.$$

Finally, applying Theorem 5 again,

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to 0+} \frac{F(y)}{G(y)} = L.$$

Proof of Theorem 4. We do the case for $\lim_{x\to a+}$. Given $\varepsilon > 0$ our task is to find a $\delta > 0$ such that

$$\left|\frac{f(x)}{g(x)} - L\right| < \varepsilon \quad \text{whenever} \quad 0 < x - a < \delta.$$

Since we are given that

$$\lim_{x\to a+} \frac{f'(x)}{g'(x)} = L,$$

it follows that there is a $\delta_1 > 0$ such that

$$(3) \qquad \left|\frac{f'(\xi)}{g'(\xi)} - L\right| < \frac{\varepsilon}{2} \quad \text{whenever} \quad 0 < \xi - a < \delta_1.$$

Now let x be any point in $0 < x - a < \delta_1$. By the Generalized Mean Value Theorem there is a point ξ, between x and $a + \delta_1$, such that

$$\frac{f(x) - f(a + \delta_1)}{g(x) - g(a + \delta_1)} = \frac{f'(\xi)}{g'(\xi)}.$$

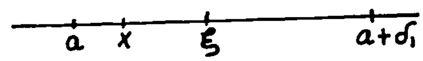


Figure 4-2

Since ξ also satisfies $0 < \xi - a < \delta_1$, we have, from (3),

$$(4) \qquad \left|\frac{f(x) - f(a + \delta_1)}{g(x) - g(a + \delta_1)} - L\right| < \frac{\varepsilon}{2}.$$

For convenience set $A = f(a + \delta_1)$, $B = g(a + \delta_1)$. Then (4) can be written in the form

$$\frac{f(x) - A}{g(x) - B} = L + \theta, \quad \text{where} \quad |\theta| < \frac{\varepsilon}{2}.$$

This gives

$$f(x) = A + g(x)(L + \theta) - B(L + \theta)$$

or

$$(5) \qquad \frac{f(x)}{g(x)} - L = \theta + \frac{A - B(L + \theta)}{g(x)} \leq \theta + \frac{|A| + |B|(|L| + \varepsilon/2)}{|g(x)|}$$

Now since $\lim_{x\to a+} |g(x)| = \infty$ we can find a δ_2 such that

$$|g(x)| > \frac{2}{\varepsilon}\Big(|A| + |B|(|L| + \varepsilon/2)\Big) \quad \text{whenever} \quad 0 < x - a < \delta_2.$$

If we take δ to be the smaller of δ_1 and δ_2 it then follows from (5) that

$$\left|\frac{f(x)}{g(x)} - L\right| < \varepsilon \quad \text{whenever} \quad 0 < x - a < \delta,$$

which is what we wanted to prove.

The proofs for $\lim_{x\to\infty}$ and the cases when L is ∞ or $-\infty$ are much the same as the one given for $\lim_{x\to a+}$. A similar direct proof for $\lim_{x\to a}$ does not work, for there is no way of insuring that $\xi \neq a$. However, it is easy to prove that if $\lim_{x\to a+} f(x) = \lim_{x\to a-} f(x) = L$ then $\lim_{x\to a} f(x)$ exists and also equals L, and from this the theorem can be proved.

PROBLEMS

1. Write complete definitions for the following:

   (a) $\lim_{x \to a-} f(x) = L.$

   (b) $\lim_{x \to -\infty} f(x) = L.$

   (c) $\lim_{x \to a} f(x) = -\infty.$

2. Prove Theorem 3.

3. Prove the following properties of infinite limits:

   (a) If $\lim f(x) = \infty$ then $\lim (-f(x)) = -\infty.$

   (b) If $\lim f(x) = A$ and $\lim g(x) = \pm\infty$ then $\lim (f(x) + g(x)) = \pm\infty.$

   (c) If $\lim f(x) = \infty$ and $\lim g(x) = \infty$ then $\lim (f(x) + g(x)) = \infty$ and $\lim f(x)g(x) = \infty.$

   (d) If $\lim |f(x)| = \infty$ then $\lim \frac{1}{f(x)} = 0.$

4. Evaluate each of the following:

   (a) $\lim_{x \to 0} \frac{\cos 2x - \cos x}{x^2}$

   (b) $\lim_{x \to 1+} \frac{\sqrt{x} - 1}{\sqrt{x - 1}}$

(c) $\lim_{x \to 1-} \frac{\log x}{\arccos x}$

(d) $\lim_{x \to 1} \left(\frac{1}{\log x} - \frac{1}{x - 1} \right)$

[Hint. Combine to make a single fraction.]

(e) $\lim_{x \to 1+} \left(\frac{1}{\sqrt{x - 1}} - \frac{1}{\sqrt{x} - 1} \right)$

(f) $\lim_{x \to \pi/2} (\sec x - \tan x)$

(g) $\lim_{x \to \infty} (\log(x - 1) - \log x)$

(h) $\lim_{x \to \infty} (\sqrt{x + 1} - \sqrt{x})$

[Hint. Use some algebra or else put $y = 1/x$.]

(i) $\lim_{x \to \infty} \sqrt{x}(\sqrt{x + 1} - \sqrt{x})$

(j) $\lim_{x \to 1-} \log x \log(1-x)$

5. Show that each of the following limits is zero:

(a) $\lim_{x \to \infty} \frac{x}{e^{ax}}$, $a > 0$

(b) $\lim_{x \to \infty} \frac{x^c}{e^{ax}}$, $a > 0$, $c > 0$.

[Hint. You can use the result of (a).]

(c) $\lim_{x \to \infty} \frac{\log x}{x^a}$, $a > 0$.

These three, along with Example 6

$$\lim_{x\to 0} x^{-a}\log x = 0, \qquad a > 0$$

occur frequently and are worth remembering.

6. Prove that if P and Q are polynomials, of degrees p and q respectively, then

$$\lim_{x\to\infty} \frac{P(x)}{Q(x)} = \begin{cases} 0 & \text{if } p < q, \\ \pm\infty & \text{if } p > q, \\ c \neq 0, & \text{if } p = q. \end{cases}$$

7. Prove: If g and h are differentiable in the neighborhood of 0 and if $g(0) = h(0) = 0$, then

$$\lim_{x\to 0} (1 + g(x))^{1/h(x)} = \exp\left(\lim_{x\to 0} \frac{g'(x)}{h'(x)}\right)$$

if the latter limit exists.

Evaluate the special cases:

(a) $\lim_{h\to 0} (1 + h)^{1/h}$

(b) $\lim_{n\to\infty} (1 + \frac{x}{n})^n$

(c) $\lim_{x\to 0} (1 + x^2)^{1/x}$

(d) $\lim_{x\to 0} (1 + x)^{1/x^2}$

73

8. (a) Given that f is continuous in the neighborhood of a and that $f'(a)$ exists, prove that

$$\lim_{h\to 0} \frac{f(a + h) - f(a - h)}{2h} = f'(a).$$

[Hint. Use the Lemma of Section 7-1.]

(b) Given that f and f' are continuous in the neighborhood of a and that $f''(a)$ exists, prove that

$$\lim_{h\to 0} \frac{f(a + h) - 2f(a) + f(a - h)}{h^2} = f''(a).$$

Chapter 11

TECHNIQUE OF INTEGRATION

1. Introduction.

We saw in Chapter 8 that an integral $\int_a^b f(x)dx$ can be evaluated as F(b) - F(a) provided we can find an indefinite integral, or anti-derivative, F, such that $F'(x) = f(x)$. This method of evaluation is often so much more valuable than a numerical approximation that one may go to great pains to effect it, if possible. This chapter is primarily concerned with the most useful devices for finding indefinite integrals.

What do we mean by "finding an indefinite integral"? The function $f(x) = \sqrt{1 + x^2}$ has the indefinite integral $F(x) = \int_0^x \sqrt{1 + t^2}dt$, but obviously this expression does not help us in evaluating $\int_0^1 \sqrt{1 + x^2}dx$; on the other hand,

$$F(x) = \frac{1}{2}[x\sqrt{x^2 + 1} + \log(x + \sqrt{x^2 + 1})]$$

does help, yielding

$$\int_0^1 \sqrt{1 + x^2}\, dx = F(1) - F(0) = \frac{1}{2}[\sqrt{2} + \log(1 + \sqrt{2})]$$

$$\approx 1.14779.$$

What we want, in fact, is a combination of the functions we have introduced - polynomial, trigonometric, logarithmic, and their inverses - by the arithmetical operations and by composition of functions. A function expressible, explicitly or implicitly, by such a combination is called an elementary function.

The differentiation formulas insure that the derivative of any elementary function is again an elementary function. This is not true of anti-differentiation. It can be proved, for instance, that the anti-derivative of e^x/x is not elementary. Hence our indefinite integration techniques will not always work. Even worse, there is no usable criterion for determining which functions have an elementary indefinite integral and which do not. We shall see that certain classes of functions can be integrated in elementary terms but beyond this it's simply a matter of trying until you decide it can't be done. There are ways of proving it can't be done in certain cases but these are topics for advanced analysis.

Two of the three basic processes for integration have already been introduced, namely algebraic simplification and substitution. Both of these aim to reduce the integrand to an expression, or a sum of expressions, whose integrals are known. It follows that the more integrals we know the easier is our task. On page 669 there is a very short table of integrals to serve as a basis for this chapter, but anyone expecting to make much use of integration should get one of the more extensive tables that are available. Most of the integrals in this short table have already been encountered (numbers 1,2,3,5,7,8,9,10,11). The others are added to give some completeness; for example, numbers 1 to 4 enable us to find

$$\int \frac{P(x)}{ax^2 + bx + c}\, dx,$$

where P(x) is any polynomial and a,b,c any constants.

The new integrals, numbers 4,6,12 and 13, can be checked by differentiating the answer to get the integrand. In fact, this is the ultimate test of any problem in indefinite integration. One perfectly good method of integration is to guess an answer and test it by differentiating. Usually the "guess" has some reasoning or past experience behind it. We shall discuss this possibility in Section 6.

In the short table of integrals we have used log $|u|$ for brevity in those formulas leading to logs. This may also be done in the problems. As suggested in Chapter 9 it is safer in applications to use either log u or log (-u) depending on whether $u > 0$ or $u < 0$. We have also omitted the arbitrary constant (this is common practice in integral tables) but this should be supplied in all problems.

A SHORT TABLE OF INTEGRALS

1. $\int x^n dx = \frac{1}{n+1} x^{n+1}$ if $n \neq -1$

2. $\int x^{-1} dx = \log |x|$

3. $\int \frac{1}{x^2 + a^2} dx = \frac{1}{a} \arctan \frac{x}{a}$

4. $\int \frac{1}{x^2 - a^2} dx = \frac{1}{2a} \log \left|\frac{x - a}{x + a}\right|$

5. $\int \frac{1}{\sqrt{a^2 - x^2}} dx = \arcsin \frac{x}{a}$, $a > 0$

6. $\int \frac{1}{\sqrt{x^2 \pm a^2}} dx = \log |x + \sqrt{x^2 \pm a^2}|$

7. $\int e^{ax} dx = \frac{1}{a} e^{ax}$

8. $\int \sin ax \, dx = -\frac{1}{a} \cos ax$

9. $\int \cos ax \, dx = \frac{1}{a} \sin ax$

10. $\int \tan ax \, dx = -\frac{1}{a} \log |\cos ax|$

11. $\int \cot ax \, dx = \frac{1}{a} \log |\sin ax|$

12. $\int \sec ax \, dx = \frac{1}{a} \log |\sec ax + \tan ax|$

13. $\int \csc ax \, dx = -\frac{1}{a} \log |\csc ax + \cot ax|$

You may find that different methods of doing an indefinite integral can lead to apparently quite different answers. By the basic theorem on antiderivatives any two answers, if correct, must differ by at most a constant, although it may take some algebraic ingenuity to show this. The same theorem can be used to simplify some answers; for instance, $\log \left|\frac{x + 3}{2}\right|$ can be replaced by $\log |x + 3|$ since these differ by $\log 2$, a constant.

In this chapter we shall frequently be referring to examples in other sections, so we temporarily adopt the notation that a reference to Example 3-4, for instance, is to Example 4 of Section 3.

Problems

1. Evaluate the following indefinite integrals.

   (a) $\int (x^3 - 3x^2 + 3x - 1)dx$

   (b) $\int (\sqrt{x} - \frac{1}{\sqrt{x}})dx$

   (c) $\int 2 \sin 3\theta \, d\theta$

   (d) $\int 4e^{-2x}dx$

   (e) $\int \frac{1}{x^2 + 5} dx$

   (f) $\int \frac{1}{x^2 - 5} dx$

   (g) $\int \frac{4}{\sqrt{x^2 + 3}} dx$

   (h) $\int \sec^2\theta \, d\theta$ [Hint. What function has $\sec^2\theta$ as its derivative?].

   (i) $\int \sec\theta \tan\theta \, d\theta$ [Same hint.]

2. Algebraic Manipulation.

The possibilities of writing the integrand in different forms are so varied, especially for trigonometric integrands, that only vague general rules can be given. One of these is that breaking the integrand into a sum of terms is often useful. Another one is the simplification of a quadratic polynomial by completing the square and making a substitution. On the whole, however, what one needs most is ingenuity. The following examples illustrate some standard tricks.

Example 1. $\int \frac{2x^2 + 3x - 7}{x + 2}\, dx.$

By long division

$$\frac{2x^2 + 3x - 7}{x + 2} = 2x - 1 - \frac{5}{x + 2},$$

and

$$\int \left(2x - 1 - \frac{5}{x + 2}\right) dx = x^2 - x - 5 \log |x + 2| + C.$$

Evidently any integrand of the form $P(x)/(ax + b)$, where P is a polynomial, can be integrated this way.

Example 2. $\int \frac{x^3}{2x^2 - 7x - 5}\, dx.$

As in Example 1 we first divide, to get

$$\frac{x^3}{2x^2 - 7x - 5} = \frac{1}{2}x + \frac{7}{4} + \frac{1}{8}\,\frac{59x + 35}{x^2 - \frac{7}{2}x - \frac{5}{2}} .$$

In the denominator of the fraction we complete the square, thus:

$$x^2 - \frac{7}{2}x - \frac{5}{2} = (x - \frac{7}{4})^2 - \frac{89}{16} .$$

In the fraction we now make the substitution $y = x - \frac{7}{4}$, $dy = dx$, to get

$$\int \frac{59x + 35}{x^2 - \frac{7}{2}x - \frac{5}{2}}\,dx = \frac{59}{2}\int \frac{2y\,dy}{y^2 - \frac{89}{16}} + \frac{553}{4}\int \frac{1}{y^2 - \frac{89}{16}}\,dy.$$

The first term integrates to

$$\frac{59}{2}\log\left|y^2 - \frac{89}{16}\right| , \qquad \text{formula 2,}$$

and the second one to

$$\frac{553}{2\sqrt{89}}\log\left|\frac{y - \sqrt{89}/4}{y + \sqrt{89}/4}\right| , \text{ formula 4.}$$

Hence, finally,

$$\int \frac{x^3}{2x^2 - 7x - 5}\,dx = \frac{1}{4}x^2 + \frac{7}{4}x + \frac{59}{16}\log\left|(x - \frac{7}{4})^2 - \frac{89}{16}\right| +$$

742

$$+ \frac{553}{16\sqrt{89}} \log \left| \frac{x - 7/4 - \sqrt{89}/4}{x - 7/4 + \sqrt{89}/4} \right| + C$$

$$= \frac{1}{16}\Big[4x^2 + 28x + 59 \log|2x^2 - 7x - 5|$$

$$+ \frac{553}{\sqrt{89}} \log \left| \frac{4x - 7 - \sqrt{89}}{4x - 7 + \sqrt{89}} \right| \Big] + C_1$$

where $C = C_1 + \frac{59}{16} \log 2$.

Example 3. $\int \sin ax \sin bx \, dx$.

We use the trigonometric identities,

$$\cos (A + B) = \cos A \cos B - \sin A \sin B,$$

$$\cos (A - B) = \cos A \cos B + \sin A \sin B.$$

Subtracting gives

$$\cos (A - B) - \cos (A + B) = 2 \sin A \sin B;$$

i.e., the product on the right hand side has been expressed as a sum. Our procedure is now clear:

$$\int \sin ax \sin bx \, dx = \int \tfrac{1}{2}[\cos(a - b)x - \cos(a + b)x]dx$$

$$= \frac{1}{2(a - b)} \sin (a - b)x$$

$$- \frac{1}{2(a + b)} \sin (a + b)x + C,$$

provided $a \neq \pm b$. If $a = b$ we get the important special case:

$$\int \sin^2 ax\, dx = \int \frac{1}{2}(1 - \cos 2ax)\,dx$$

$$= \frac{1}{2}x - \frac{1}{4a} \sin 2ax + C.$$

The cases $\int \cos ax \cos bx\, dx$ and $\int \sin ax \cos bx\, dx$ can be handled similarly, using the identities

$$2 \cos A \cos B = \cos (A - B) + \cos (A + B)$$

$$2 \sin A \cos B = \sin (A - B) + \sin (A + B).$$

The special case $\int \cos^2 ax\, dx = \frac{1}{2}x + \frac{1}{4a} \sin 2ax + C$ is also of special note.

Example 4. $\int \frac{1}{1 + \sin x}\, dx = \int \frac{1 - \sin x}{1 - \sin^2 x}\, dx$

$$= \int \frac{1 - \sin x}{\cos^2 x}\, dx$$

$$= \int (\sec^2 x - \sec x \tan x)\,dx$$

$$= \tan x - \sec x + C.$$

This is an example of a problem that just happens to succumb to the right trick. If the 1 in the denominator is replaced by 2 the trick doesn't work and the answer, obtained in Example 3-11, is much more compli ated.

For convenience we list here some of the lesser known trigonometric identities that are useful in integration problems.

$$\sec^2 x - \tan^2 x = 1$$

$$\csc^2 x - \cot^2 x = 1$$

or

$$(\csc x - \cot x)(\csc x + \cot x) = 1$$

$$\sin^2 \frac{x}{2} = \frac{1}{2}(1 - \cos x)$$

$$\cos^2 \frac{x}{2} = \frac{1}{2}(1 + \cos x)$$

$$\tan \frac{x}{2} = \frac{1 - \cos x}{\sin x} = \frac{\sin x}{1 + \cos x} = \sec x - \tan x$$

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

Problems

1. Evaluate the following indefinite integrals.

(a) $\int xe^{-2x^2}dx$

(b) $\int \frac{\cos 2x}{1 + \sin 2x}\,dx$

(c) $\int \frac{1}{t} \sin \log t \; dt$

(d) $\int \frac{x}{\sqrt{x^2 + 1}}\,dx$

(e) $\int \frac{x}{\sqrt{x^4 + 1}}\,dx$

(f) $\int \frac{e^{2x} - 1}{e^{2x} + 1}\,dx$ [There is a quick method.]

(g) $\int \frac{x^5}{(x^3 + 1)^2}\,dx$

(h) $\int \frac{\sin x}{1 + \cos^2 x}\,dx$

(i) $\int \sin 7x \sin 3x \; dx$

(j) $\int \cos 3x \cos 7x \; dx$

(k) $\int \sin 3x \cos 2x \; dx$

(l) $\int \frac{8x - 2}{4x^2 - 4x + 3}\,dx$

(m) $\int \frac{\sec^2 x}{1 + \tan x}\, dx$

(n) $\int \frac{1}{\sqrt{2x^2 + 4x + 3}}\, dx$

(o) $\int \frac{1}{\sqrt{5x - x^2}}\, dx$

(p) $\int \frac{(\log x)^3}{x}\, dx$

(q) $\int \frac{4x^3 + 5}{2x^2 + x}\, dx$

(r) $\int \frac{x^4 + 1}{(x + 1)^2}\, dx$

(s) $\int \frac{dx}{(x + 2)\sqrt{x^2 + 4x + 3}}$

(t) $\int \frac{1}{\cos x(2 \cos x + 3 \sin x)}\, dx$ [Compare with (m).]

(u) $\int \frac{\sqrt{x^2 + 2x}}{x + \text{[illegible]}}\, dx$

(v) $\int \sec^2 x \tan^2 x\, dx$

(w) $\int \sec^4 x\, dx$

(x) $\int \frac{1}{1 - \cos x}\, dx$

(y) $\int \frac{1 + \cos x}{1 - \cos x}\, dx$

(z) $\int \frac{1}{1 + 4 \sin^2 x}\, dx$

2. Evaluate

$$\int \sin x \sin 2x \sin 5x \, dx$$

3. Derive the formulas for cos A cos B and sin A cos B on page 675.

4. Let m and n be integers $\geq$ 0 and let

$$A_{mn} = \int_0^{2\pi} \sin mx \sin nx \, dx,$$

$$B_{mn} = \int_0^{2\pi} \cos mx \cos nx \, dx,$$

$$C_{mn} = \int_0^{2\pi} \sin mx \cos nx \, dx.$$

Show that A_{mn}, B_{mn}, C_{mn} are all zero except that

$$A_{nn} = B_{nn} = \begin{cases} \pi & \text{if } n \neq 0, \\ 2\pi & \text{if } n = 0. \end{cases}$$

5. (a) For any number n, prove that $\tan^n x = \tan^{n-2} x \sec^2 x - \tan^{n-2} x$.

(b) Establish the formula

$$\int \tan^n x \, dx = \frac{1}{n-1} \tan^{n-1} x - \int \tan^{n-2} x \, dx.$$

This is known as a "reduction formula".

(c) Use the reduction formula to evaluate

$$\int \tan^8 x \, dx \quad \text{and} \quad \int \tan^9 x \, dx.$$

3. Substitution.

In Chapter 8 we considered substitutions that reduced the integrand immediately to one of the standard forms. Usually this cannot be done; the substitution merely simplifies the integrand, making it ready for a further substitution or some other modification. There are several useful general types of substitutions.

1. If the integrand involves a function $g(x)$ to a negative or fractional power, or if $g(x)$ appears as the argument of a transcendental function, try the substitution $u = g(x)$ or $u^n = g(x)$ for a suitable n.

Example 1. $\int \frac{1}{e^x + 1}\,dx = \int (e^x + 1)^{-1}dx$.

Let $u = e^x + 1$, $du = e^x dx = (u - 1)dx$. Then

$$\int \frac{1}{e^x + 1}\,dx = \int \frac{1}{u}\,\frac{1}{u - 1}\,du = \int \frac{1}{u^2 - u}\,du.$$

This is now of the form discussed in Example 2-2.

is a quicker solution:

$$\begin{aligned}\int \frac{1}{e^x + 1}\,dx &= \int \frac{e^{-x}}{1 + e^{-x}}\,dx \\ &= \int \frac{-d(1 + e^{-x})}{1 + e^{-x}} \\ &= -\log(1 + e^{-x}) + C.\end{aligned}$$

749

<u>Example 2.</u> $\int \frac{x^5}{\sqrt{x^2+1}}\,dx.$

Let $u^2 = x^2 + 1$, $2udu = 2xdx$. Then

$$\int \frac{x^5}{\sqrt{x^2+1}}\,dx = \int \frac{(x^2)^2(xdx)}{\sqrt{x^2+1}}$$

$$= \int \frac{(u^2-1)^2 udu}{u}$$

$$= \int (u^4 - 2u^2 + 1)du$$

$$= \frac{1}{5}u^5 - \frac{2}{3}u^3 + u + C$$

$$= (\frac{1}{5}u^4 - \frac{2}{3}u^2 + 1)u + C$$

$$= \frac{1}{15}(3x^4 - 4x^2 + 8)\sqrt{x^2+1} + C.$$

<u>Example 3.</u> $\int \sin(\log x)dx.$

Let $u = \log x$, $x = e^u$, $dx = e^u du$. Then

$$\int \sin(\log x)dx = \int e^u \sin u\, du.$$

The new integral is a standard type that we shall examine later.

II. A trigonometric integrand that can be reduced to a function of sin ax and cos ax can often be simplified by using $u = \sin ax$ or $u = \cos ax$.

Example 4. $\int \frac{\sin 3\theta}{1 + \cos\theta}\, d\theta.$

First we have

$$\begin{aligned}\sin 3\theta &= \sin(2\theta + \theta)\\ &= \sin 2\theta \cos\theta + \cos 2\theta \sin\theta\\ &= 2 \sin\theta \cos^2\theta + (\cos^2\theta - \sin^2\theta)\sin\theta.\end{aligned}$$

Now let $u = \cos\theta$, $du = \sin\theta\, d\theta$. Then $\sin^2\theta = 1 - u^2$ and we get

$$\int \frac{\sin 3\theta}{1 + \cos\theta}\, d\theta = \int \frac{2u^2 + (u^2 - 1 + u^2)}{1 + u}\, du,$$

ready for the substitution $v = 1 + u$ or the method of Example 2-1. We could have saved one step by letting $u = 1 + \cos\theta$ to begin with but it is sometimes better not to do too much at once.

Example 5. $\int \sin^4 2x \cos^5 2x\, dx.$

Here the proper substitution is $u = \sin 2x$, $du = 2 \cos 2x\, dx$. Let us first see how it works, and then why. Changing the form of the integrand slightly, we have

$$\begin{aligned}&\int \sin^4 2x \cos^4 2x \cos 2x\, dx\\ &= \int u^4(1 - u^2)^2 \tfrac{1}{2} du,\end{aligned}$$

since $\cos^2 2x = 1 - \sin^2 2x = 1 - u^2$.

The reason this simplified so nicely is that after removing one factor of cos 2x to go with the dx to make du, we had an even power of cos 2x left, and so no square roots were introduced. The same phenomenon occurs in Example 4 with the roles of sin and cos interchanged.

It is easy to see that this process will work for

$$\int \sin^m ax \cos^n ax \, dx$$

if one or both of m and n are odd. The case when they are both even will be treated later.

Occasionally a substitution of another trigonometric function, $u = \tan ax$, $u = \sec ax$, etc. will simplify things.

III. An algebraic integrand involving the one irrationality $\sqrt{a^2 - x^2}$ can be reduced to a rational trigonometric integrand by the substitution $\theta = \arcsin x/a$ or $x = a \sin \theta$. (One can use cos instead of sin but there is rarely any reason for doing so). Then $\sqrt{a^2 - x^2} = a \cos \theta$.

Example 6. $\int \frac{2}{(x + 2)\sqrt{4 - x^2}} dx$.

$x = 2 \sin \theta$, $\sqrt{4 - x^2} = 2 \cos \theta$, $dx = 2 \cos\theta \, d\theta$.

$$\int \frac{2}{(x + 2)\sqrt{4 - x^2}} dx = \int \frac{4 \cos\theta \, d\theta}{(2 \sin \theta + 2) 2\cos \theta}$$

$$= \int \frac{d\theta}{\sin\theta + 1}$$

$$= (\tan\theta - \sec\theta) + C, \text{ by Example 2-4},$$

$$= \frac{2\sin\theta - 2}{2\cos\theta} + C$$

$$= \frac{x - 2}{\sqrt{4 - x^2}} + C.$$

Example 7. We can finally find the area of a circle by integration.

$$\int \sqrt{a^2 - x^2}\, dx = \int a \cos\theta \; a\cos\theta \; d\theta$$

$$= a^2 \int \cos^2\theta \; d\theta$$

$$= \frac{a^2}{2}\int (1 + \cos 2\theta)\, d\theta, \text{ see Example 2-3},$$

$$= \frac{a^2}{2}(\theta + \frac{1}{2}\sin 2\theta) + C$$

$$= \frac{a^2}{2}(\theta + \sin\theta\cos\theta) + C$$

$$= \frac{a^2}{2} \arcsin\frac{x}{a} + \frac{x}{2}\sqrt{a^2 - x^2} + C.$$

The area of the circle is

$$4\int_0^a \sqrt{a^2 - x^2}\, dx = 2a^2(\arcsin 1) = \pi a^2.$$

The irrationalities $\sqrt{x^2 - a^2}$ and $\sqrt{x^2 + a^2}$ can be handled similarly by the substitutions $x = a \sec \theta$ and $x = a \tan \theta$, with appropriate use of the identity $\sec^2\theta - \tan^2\theta = 1$.

<u>Example 8</u>. $\int \frac{1 + x}{x^3\sqrt{x^2 - 4}} dx$.

Here we use

$$x = 2 \sec \theta, \quad \sqrt{x^2 - 4} = 2 \tan \theta, \quad dx = 2 \sec \theta \tan \theta \, d\theta,$$

to get

$$\int \frac{2(1 + 2 \sec \theta)\sec \theta \tan \theta}{16 \sec^3\theta \tan \theta} d\theta = \frac{1}{8}\int(\cos^2\theta + 2 \cos \theta)d\theta.$$

Doing $\int \cos^2\theta \, d\theta$ as in Example 7, we get

$$\frac{1}{8}\left[\frac{1}{2}\theta + \frac{1}{2}\sin \theta \cos \theta + 2 \sin \theta\right] + C$$

$$= \frac{1}{16}\left[\arccos\frac{2}{x} + \frac{2\sqrt{x^2 - 4}}{x^2} + 4\frac{\sqrt{x^2 - 4}}{x}\right] + C$$

$$= \frac{1}{16}\left[\frac{4x + 2}{x^2}\sqrt{x^2 - 4} + \arccos\frac{2}{x}\right] + C.$$

The expressions for the functions of θ in terms of x are most easily seen from a diagram like Figure 3-1. In the final answer we used $\arccos\frac{2}{x}$ as being a more familiar function than $\text{arcsec}\frac{x}{2}$.

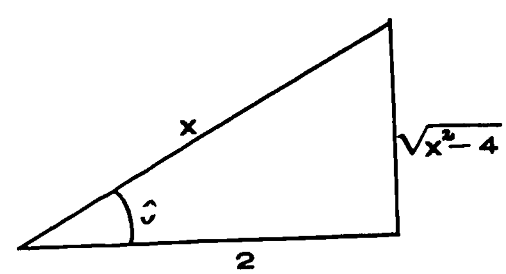


Figure 3-1

Example 9. $\int \frac{1}{x^2\sqrt{x^2 + a^2}}\, dx.$

Here we use

$$x = a \tan \theta, \quad \sqrt{x^2 + a^2} = a \sec \theta, \quad dx = a \sec^2\theta \, d\theta,$$

so

$$\int \frac{1}{x^2\sqrt{x^2 + a^2}}\, dx = \int \frac{a \sec^2\theta}{a^2 \tan^2\theta \; a \sec \theta}\, d\theta$$

$$= \frac{1}{a^2}\int \frac{\cos \theta}{\sin^2\theta}\, d\theta$$

$$= \frac{1}{a^2}\int \frac{d(\sin \theta)}{\sin^2\theta}$$

$$= -\frac{1}{a^2 \sin \theta}$$

$$= -\frac{\sqrt{x^2 + a^2}}{a^2 x},$$

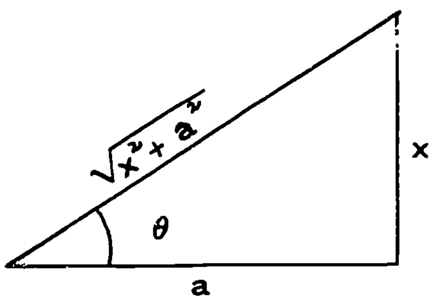


Figure 3-2

from Figure 3-2.

One can combine this type of substitution with the technique of completing the square, used in Example 2-2, to handle any irrationality of the form $\sqrt{ax^2 + bx + c}$.

<u>Example 10</u>. $\int \sqrt{\frac{x}{3x+4}}\,dx = \int \frac{x}{\sqrt{3x^2+4x}}\,dx$

$$= \frac{1}{\sqrt{3}} \int \frac{x}{\sqrt{x^2 + \frac{4}{3}x + \frac{4}{9} - \frac{4}{9}}}\,dx$$

$$= \frac{1}{\sqrt{3}} \int \frac{x}{\sqrt{(x + \frac{2}{3})^2 - \frac{4}{9}}}\,dx .$$

Put $y = x + \frac{2}{3}$, to get

$$\frac{1}{\sqrt{3}} \int \frac{y - 2/3}{\sqrt{y^2 - 4/9}}\,dy.$$

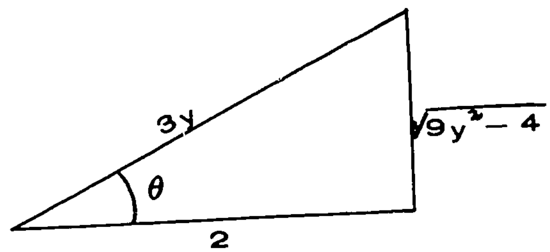


Figure 3-3

Now we use $y = \frac{2}{3}\sec\theta$.

$$\frac{1}{\sqrt{3}} \int \frac{\frac{2}{3}\sec\theta - \frac{2}{3}}{\frac{2}{3}\tan\theta}\,\frac{2}{3}\sec\theta\tan\theta\,d\theta$$

$$= \frac{2}{3\sqrt{3}} \int (\sec^2\theta - \sec\theta)\,d\theta$$

$$= \frac{2}{3\sqrt{3}} (\tan\theta - \log|\sec\theta + \tan\theta|) + C$$

$$= \frac{2}{3\sqrt{3}} \left[\frac{\sqrt{9y^2 - 4}}{2} - \log\left|\frac{3y + \sqrt{9y^2 - 4}}{2}\right|\right] + C$$

$$= \frac{1}{3\sqrt{3}} \left[\sqrt{9x^2 + 12x} - 2 \log \left| \frac{3x + 2 + \sqrt{9x^2 + 12x}}{2} \right| \right] + C$$

$$= \frac{1}{3} \left[\sqrt{3x^2 + 4x} - \frac{2}{\sqrt{3}} \log \left| 3x + 2 + \sqrt{3(3x^2 + 4x)} \right| \right] + C_1 .$$

IV. We saw in Examples 2-1 and 2-2 how any rational function $P(x)/Q(x)$ could be integrated if Q were of degree 1 or 2. In Section 5 we shall extend this result to a denominator of any degree. Because any rational function can be integrated there is interest in techniques that transform an integrand into a rational function, as was done in Examples 2,4, and 5. There is a standard method of doing this if our integrand is any rational function of $\sin \theta$ and $\cos \theta$. The method frequently gives the answer in a rather complicated form, and it should be used only when other methods fail.

The substitution arises from the parametrization of the unit circle (see Problem 2(d), Section 7-7),

$$x = \frac{1 - t^2}{1 + t^2}, \qquad y = \frac{2t}{1 + t^2} .$$

Since

$$x = \cos \theta, \qquad y = \sin \theta ,$$

is also a parametrization of the unit circle we might expect to get something interesting by putting

$$\cos\theta = \frac{1 - t^2}{1 + t^2} = -1 + \frac{2}{1 + t^2}, \quad \sin\theta = \frac{2t}{1 + t^2}.$$

We do! First of all, differentiating cos θ with respect to t gives

$$-\sin\theta \frac{d\theta}{dt} = \frac{-4t}{(1 + t^2)^2} = \frac{-2}{1 + t^2} \sin\theta,$$

so

$$d\theta = \frac{2}{1 + t^2} dt.$$

The relation $dt = \frac{d\theta}{1 + \cos\theta}$ is sometimes useful.

To solve for t in terms of θ notice that

$$\cos\theta + 1 = \frac{2}{1 + t^2} = \frac{\sin\theta}{t},$$

or

$$t = \frac{\sin\theta}{1 + \cos\theta} = \tan\frac{\theta}{2} = \frac{1 - \cos\theta}{\sin\theta}.$$

Armed with these formulas we can proceed to integrate.

<u>Example 11.</u> $\int \frac{1}{2 + \sin\theta} d\theta.$

This is the integral mentioned in Example 2-4 as not being integrable by a simple trick. If we use our substitutions for sin θ and dθ we get

This is of the type of Example 2-2. We have

$$\int \frac{1}{t^2 + t + 1}\,dt = \int \frac{1}{(t + \frac{1}{2})^2 + \frac{3}{4}}\,dt$$

$$= \frac{2}{\sqrt{3}} \arctan \frac{t + 1/2}{\sqrt{3}/2} + C$$

$$= \frac{2}{\sqrt{3}} \arctan \frac{2 \tan \frac{\theta}{2} + 1}{\sqrt{3}} + C.$$

If we prefer to have the answer in terms of $\sin \theta$ and $\cos \theta$ we can use

$$\tan \frac{\theta}{2} = \frac{1 - \cos \theta}{\sin \theta}$$

to get

$$\frac{2}{\sqrt{3}} \arctan \frac{2 - 2 \cos \theta + \sin \theta}{\sqrt{3} \sin \theta} + C.$$

Speaking of guessing: one would hardly guess that the derivative of this function is $\frac{1}{2 + \sin \theta}$.

Problems

Part I.

1. Evaluate each of the following.

(a) $\int \frac{dx}{x(1 + x^2)}$

(b) $\int \frac{dx}{x\sqrt{x^2 + 9}}$

(c) $\int \frac{dx}{e^{2x} - 1}$

(d) $\int \tan 2x \sec 2x \, dx$

(e) $\int \frac{x + 3}{\sqrt{2x - 1}} \, dx$

(f) $\int \sin (ax + b) dx$

(g) $\int \frac{x + 1}{\sqrt{x - 1}} \, dx$

(h) $\int \frac{dx}{x\sqrt{x^3 - 8}}$

(i) $\int x^3 (x^2 - 1)^{2/3} dx$

(j) $\int \frac{1}{x} e^{\log x} dx$

(k) $\int \frac{1}{1 + \sqrt{x}}\, dx$

(l) $\int x\, \sqrt[3]{x + 1}\, dx$

2. For what values of c is $\int \frac{x^c}{\sqrt{1 - x^2}}\, dx$ easily integrable?

Part II.

3. Evaluate each of the following.

(a) $\int \sin 2x \cos^4 2x \, dx$

(b) $\int \tan 2x \sec^2 2x \, dx$

(c) $\int \frac{\cos 2x}{\cos x} \, dx$

(d) $\int \sin^3 3x \cos^4 3x \, dx$

(e) $\int \tan^3 x \sec x \, dx$

(f) $\int \sec^2 x \tan^2 x \, dx$

(g) $\int \frac{\sin 2x}{a + b \cos x} \, dx$

(h) $\int \tan^2 x \tan 2x \, dx$

4. Finish Example 4.

5. Show by integrating that

$$\int \frac{\sin 2x + a \sin x}{\cos 2x + a \cos x} \, dx = -\frac{1}{2} \log \left| \cos 2x + a \cos x \right|$$

$$+ \frac{a}{2\sqrt{a^2 + 8}} \log \left| \frac{4 \cos x + a + \sqrt{a^2 + 8}}{4 \cos x + a - \sqrt{a^2 + 8}} \right| + C.$$

Part III.

6. Evaluate each of the following.

(a) $\int \frac{x^3}{\sqrt{4 - x^2}}\, dx$

(b) $\int \frac{x^2}{\sqrt{4 - x^2}}\, dx$

(c) $\int \frac{x^3}{\sqrt{4 + x^2}}\, dx$

(d) $\int \sqrt{\frac{x + 1}{x - 1}}\, dx$

(e) $\int \frac{x^2}{\sqrt{3 - 2x - x^2}}\, dx$

(f) $\int (x^2 - 4x + 8)^{-3/2}\, dx$

(g) $\int \frac{x}{(2x^2 + 2x - 1)^{3/2}}\, dx$

(h) $\int \frac{1}{x\sqrt{x^2 + 2ax}}\, dx$

7. Show that for any value of m and n,

$$\int x^m (a^2 - x^2)^{n/2}\, dx = a^{m+n+1} \int \sin^m \theta \cos^{n+1} \theta\, d\theta.$$

8. The hyperbolic functions sinh u and cosh u (Section 9-1, Problem 13) can be used instead of tan θ and sec θ for the integrands involving $\sqrt{x^2 + a^2}$ and $\sqrt{x^2 - a^2}$ respectively.

(a) Do Problem 6(c) by this method.

(b) Solve $v = \cosh u$ for u to get, for $u > 0$, $\text{arccosh } v = u = \log(v + \sqrt{v^2 - 1})$. [Hint. Reduce $v = \cosh u$ to a quadratic in e^u.]

(c) Do Problem 6(d) by this method.

Part IV. Many of these problems can be done by simpler methods than the substitution $t = \tan x/2$.

9. Evaluate each of the following.

(a) $\int \frac{\cos x}{1 + \cos x}\, dx$

(b) $\int \frac{1}{1 + \cos x + \sin x}\, dx$

(c) $\int \frac{1}{\cos x - \sin x}\, dx$

(d) $\int \frac{\cos x}{\sin x(1 + \cos x)}\, dx$

(e) $\int \frac{\cos x}{\sin x(1 + \sin x)}\, dx$

(f) $\int \frac{1}{\sqrt{(1 + \sin x)(1 + \cos x)}}\, dx$

(g) $\int \frac{1}{\sqrt{(1 + \sin x)(2 + \cos x)}}\, dx$

(h) $\int \frac{1}{a + \cos x}\, dx$. Consider all possible values of a.

10. Reduce each of the following to the integral of a rational function. Get as simple a result as you can.

(a) $\int \frac{\sin x}{1 + \sin x}\, dx$

(b) $\int \tan x \tan 2x\, dx$

(c) $\int \sec^3 x\, dx$

11. (a) Reduce $\int \frac{dx}{a + b \sin x + c \cos x}$ to

$$\int \frac{2dt}{(a - c)t^2 + 2bt + (a + c)}$$

(b) Evaluate $\int \frac{dx}{a + b \sin x + c \cos x}$

Consider the cases $a^2 \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} b^2 + c^2$ and the special case $a = c$.

4. Integration by Parts.

There remains one basic differentiation formula that we have not yet exploited in our integration, that is the product formula. In differential form this formula is

$$d(uv) = u\,dv + v\,du,$$

giving the corresponding formula for indefinite integrals,

$$(1) \qquad uv = \int u\,dv + \int v\,du.$$

Only in extremely rare cases can we expect to find functions u(x) and v(x) for which our given integral reduces to $\int u\,dv + \int v\,du$. Our use of (1) depends on putting it in the form

$$\int u\,dv = uv - \int v\,du,$$

and choosing u(x) and v(x) so that our given integral $J = \int f(x)dx$ has the form $J = \int u\,dv$ <u>and so that</u> $\int v\,du$ <u>is simpler than</u> J. Thus integration by parts, the name for this process, never gives us a final answer but only changes the form of the integral, hopefully so that one of the other methods can be applied.

The first thing one must do in integrating by parts is to express the given integrand in the form u dv. This is ordinarily done by writing the given integrand f(x)dx in the form u(x)w(x)dx and taking v as $\int w(x)\,dx$.

Example 1. $J = \int x^2 \log x \, dx.$

We write the integrand as $(\log x)(x^2 dx)$, taking

$$u = \log x, \qquad dv = x^2 dx,$$

from which

$$du = \frac{1}{x}dx, \qquad v = \frac{1}{3}x^3.$$

This

$$J = \frac{1}{3}x^3 \log x - \int \frac{1}{3}x^3 \frac{1}{x}\,dx$$

$$= \frac{1}{3}x^3 \log x - \frac{1}{3}\int x^2\,dx$$

$$= \frac{1}{3}x^3 \log x - \frac{1}{9}x^3 + C.$$

A given integrand can be expressed as u dv in many ways – in Example 1, for instance, as $(x \log x)(x\,dx)$ or $(x^2 \log x)(dx)$ or $(x^4)(\log x\,dx/x)$, etc. – and it is to our advantage to pick one that leads quickly and easily to the best form of $\int v\,du$ for further integration. The following observations may be helpful.

1. $dv = w\,dx$ should be easily integrable. The most common choices of w are x^n, sin ax, cos ax, e^{ax}. The possibility of more elaborate choices, such as xe^{x^2}, $\sin^n x \cos x$, etc., should not be overlooked.

2. The simplification in $\int v\,du$ usually arises from the differentiation of u, and one should be on the lookout for this. Polynomials, logs, and inverse trig functions are all simplified by differentiation.

Example 1 illustrates both these observations.

3. v is <u>any</u> function whose differential is dv. That is, the constant of integration obtained in passing from dv to v can be given any value we please. Usually we take it to be zero, as was done in Example 1. A different choice may affect the final constant, a matter of no concern, but it may also make the intermediate stages easier. Here is an example:

<u>Example 2</u>. $J = \int x^3 \arctan x\,dx$.

The main interest here is to get rid of the arctan x, so following observation 2 we take

$$u = \arctan x, \qquad dv = x^3 dx,$$

$$du = \frac{1}{x^2+1}\,dx, \qquad v = \frac{1}{4}x^4 + C_1.$$

The idea is to take C_1 so as to simplify v du. Since $x^4 - 1 = (x^2 - 1)(x^2 + 1)$, if we take $C_1 = -\frac{1}{4}$ we get

$$v\,du = \frac{1}{4}(x^2 - 1).$$

So

$$J = \frac{1}{4}(x^4 - 1) \arctan x - \frac{1}{4}\int (x^2 - 1)dx$$

$$= \frac{1}{4}(x^4 - 1) \arctan x - \frac{1}{12} x^3 + \frac{1}{4} x + C.$$

The choice $C_1 = 0$ would have left us with

$$\int \frac{x^4}{x^2 + 1}\,dx$$

to evaluate.

Example 3. $J = \int x^2 \sin 3x\,dx.$

We note that $\sin 3x$ is neither simplified nor complicated by either integration or differentiation. x^2 is simplified by differentiation, so we take

$$u = x^2, \qquad dv = \sin 3x\,dx,$$

giving

$$du = 2x\,dx, \qquad v = -\frac{1}{3}\cos 3x;$$

$$J = -\frac{1}{3} x^2 \cos 3x + \frac{2}{3}\int x \cos 3x\,dx.$$

We have achieved a simplification but must evidently repeat the process. This time we have

$$u = x, \qquad dv = \cos 3x\, dx,$$

$$du = dx, \qquad v = \frac{1}{3} \sin 3x;$$

$$J = -\frac{1}{3} x^2 \cos 3x + \frac{2}{3} \left[\frac{1}{3} x \sin 3x - \frac{1}{3} \int \sin 3x\, dx \right]$$

$$= -\frac{1}{3} x^2 \cos 3x + \frac{2}{9} x \sin 3x + \frac{2}{27} \cos 3x + C.$$

Some of the most useful applications of integration by parts leave us with an $\int v\, du$ that is as complicated as the original integral J, or, indeed, even identical with it. However, two occurrences of J in an equation is not fatal unless they cancel; otherwise we simply solve for J.

<u>Example 4</u>. $J = \int e^{-ht} \sin \omega t\, dt$.

The integrand is taken from Example 1 of Section 9-1. Here our two observations are no guide and we take, as a try,

$$u = e^{-ht}, \qquad dv = \sin \omega t\, dt,$$

$$du = -he^{-ht} dt, \qquad v = -\frac{1}{\omega} \cos \omega t,$$

$$J = -\frac{1}{\omega} e^{-ht} \cos \omega t - \frac{h}{\omega} \int e^{-ht} \cos \omega t\, dt.$$

The only significant change is the replacement of sin by cos in the integrand. Evidently if we apply the same scheme once more we shall get J back again.

$$u = e^{-ht}, \qquad dv = \cos \omega t \, dt,$$

$$du = -he^{-ht} dt, \qquad v = \frac{1}{\omega} \sin \omega t,$$

$$J = -\frac{1}{\omega} e^{-ht} \cos \omega t - \frac{h}{\omega} \left[\frac{1}{\omega} e^{-ht} \sin \omega t + \frac{h}{\omega} \int e^{-ht} \sin \omega t \, dt \right]$$

or

$$J = e^{-ht} \left[-\frac{1}{\omega} \cos \omega t - \frac{h}{\omega^2} \sin \omega t \right] - \frac{h^2}{\omega^2} J.$$

Solving for J gives

$$J = \frac{1}{\omega^2} e^{-ht} (-\omega \cos \omega t - h \sin \omega t) / (1 + \frac{h^2}{\omega^2}) + C$$

$$= \frac{1}{h^2 + \omega^2} e^{-ht} (-\omega \cos \omega t - h \sin \omega t) + C.$$

Procedures similar to this give us a very useful class of integration formulas known as "reduction formulas". These apply to integrands that depend on a constant, n, usually a positive integer. The reduction formula decreases the value of n, so that by successive applications we can get n sufficiently small, usually 0 or 1, that the integral can be evaluated. A few reduction formulas are obtained by algebraic manipulation (see Problem 5 in Section 2) but most come from integration by parts.

Example 5. Let $J_n = \int \sin^n ax\, dx$.

We take

$$u = \sin^{n-1} ax, \qquad dv = \sin ax\, dx,$$

$$du = (n - 1)a \sin^{n-2} ax \cos ax\, dx, \qquad v = -\frac{1}{a} \cos ax.$$

Then

$$J_n = -\frac{1}{a} \sin^{n-1} ax \cos ax + (n - 1) \int \sin^{n-2} ax \cos^2 ax\, dx$$

$$= -\frac{1}{a} \sin^{n-1} ax \cos ax + (n - 1) \int \sin^{n-2} ax (1 - \sin^2 ax) dx$$

$$= -\frac{1}{a} \sin^{n-1} ax \cos ax + (n - 1)(J_{n-2} - J_n).$$

Solving for J_n gives the reduction formula,

$$(2) \qquad \int \sin^n ax\, dx = -\frac{1}{an} \sin^{n-1} ax \cos ax + \frac{n-1}{n} \int \sin^{n-2} ax\, dx.$$

As an application:

$$\int \sin^6 2x\, dx = -\frac{1}{12} \sin^5 2x \cos 2x + \frac{5}{6} \int \sin^4 2x\, dx,$$

$$\int \sin^4 2x\, dx = -\frac{1}{8} \sin^3 2x \cos 2x + \frac{3}{4} \int \sin^2 2x\, dx,$$

$$\int \sin^2 2x\, dx = -\frac{1}{4} \sin 2x \cos 2x + \frac{1}{2} \int dx.$$

So

$$\sin^6 x\, dx = -\frac{1}{12}\sin^5 2x \cos 2x + \frac{5}{6}\left[-\frac{1}{8}\sin^3 2x \cos 2x + \frac{3}{4}\left(-\frac{1}{4}\sin 2x \cos 2x + \frac{1}{2}x\right)\right] + C$$

$$= \left(-\frac{1}{12}\sin^4 2x - \frac{5}{48}\sin^2 2x - \frac{15}{96}\right)\sin 2x \cos 2x + \frac{15}{48}x + C.$$

Such reduction formulas take care of the case left open in Example 3-5.

The reduction formula (2) is the type of recursion formula well adapted to flow charting and programming. At any stage of the process of finding any J_m we have an expression like

(3) $$J_m = c_m T_{m-1} + c_{m-2} T_{m-3} + \dots + c_{n+2} T_{n+1} + b_n J_n,$$

where the c's and b's are constants and $T_k = \sin^k ax \cos ax$. Applying (2) to J_n gives

$$J_m = c_m T_{m-1} + \dots + c_{n+2} T_{n+1} - \frac{1}{an} b_n T_{n-1} + \frac{n-1}{n} b_n J_{n-2}.$$

Since this is to give (3) carried one step further, namely,

$$J_m = c_m T_{m-1} + \dots + c_{n+2} T_{n+1} + c_n T_{n-1} + b_{n-2} J_{n-2},$$

we must have

$$c_n = -\frac{1}{an} b_n, \qquad b_{n-2} = \frac{n-1}{n} b_n.$$

These recursion formulas for the coefficients enable us to construct the flow diagram in Figure 4-1. We leave it as an exercise for the reader to put on the proper endings for the cases m odd and m even.

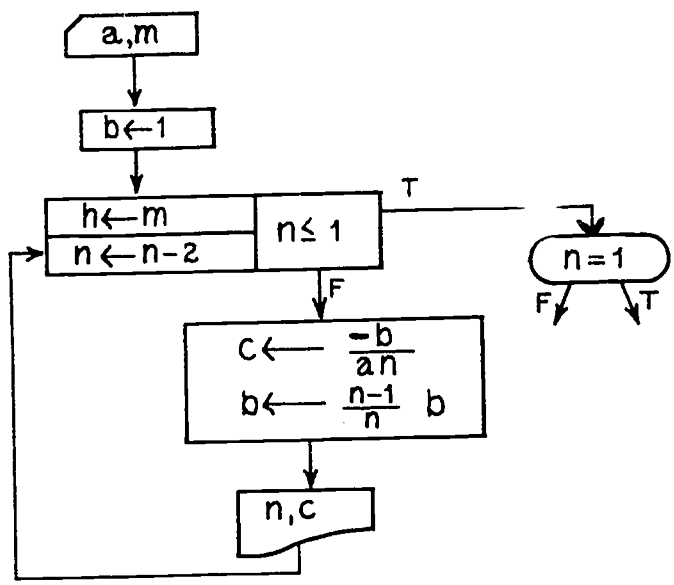


Figure 4-1

Problems

1. Evaluate each of the following.

(a) $\int x \sin x \, dx$

(b) $\int x e^{-x} \, dx$

(c) $\int \arcsin 2x \, dx$

(d) $\int x \sec^2 x \, dx$

(e) $\int 9x \tan^2 3x \, dx$

(f) $\int \arcsin \sqrt{x} \, dx$

(g) $\int x^2 e^x \, dx$

(h) $\int (\log x)^2 dx$

(i) $\int 3x^3 \sqrt{x^2 + 2} \, dx$

(j) $\int x \arctan x \, dx$

(k) $\int e^{ax} \cos bx \, dx$

(l) $\int x^2 \cos 2x \, dx$

(m) $\int \frac{x^3}{\sqrt{4 + x^2}} \, dx$

(n) $\int \frac{x \arctan x}{\sqrt{1 + x^2}} \, dx$

(o) $\int \log (x^2 + 1) dx$

(p) $\int \sin 2x \, e^{\sin x} dx$

(q) $\int \sin \sqrt{x} \, dx$

(r) $\int x^3 e^{x^2} dx$

(s) $\int \frac{\log x}{(x + 2)^2} \, dx$

2. Evaluate $\int \sec^3 x \, dx$.

3. Evaluate $\int xe^x \sin x$. [Hint. Take $dv = e^x \sin x \, dx$ and use the result of Example 4]

4. (a) Develop a reduction formula for

$$\int x^n \sqrt{x^2 + c} \, dx.$$

[Hint. Use $dv = x\sqrt{x^2 + c} \, dx$.]

(b) Evaluate $\int x^8 \sqrt{x^2 - 1}\, dx$.

5. (a) Develop a reduction formula for

$$\int x^n e^{x^2}\, dx$$

(b) Of the cases n = 10 and n = 11, one can be evaluated as an elementary function and not the other. Evaluate the one that you can.

5. Partial Fractions.

At the beginning of Section 2 it was stated that one method of integrating is to express a complicated quantity as a sum of simpler ones. The method of partial fractions is just this idea applied in a systematic way to rational functions.

You will recall that a rational function is one expressible in the form $f(x) = \frac{P(x)}{Q(x)}$, where P and Q are polynomials. P and Q can be any polynomials, with the one exception that Q(x) is not identically zero. However, we are not interested in the case Q(x) = constant, for then f is itself a polynomial and we know how to integrate it. So we assume that Q(x) involves x, and by dividing P(x) and Q(x) by the coefficient of the highest power of x in Q(x) we reduce Q(x) to the form

$$Q(x) = x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n, \qquad n \geq 1.$$

So far we have said nothing about the polynomial P. If the degree of P is less than n we leave P as it is. But if its degree is $\geq$ n we divide P(x) by Q(x) by long division, getting a quotient S(x) and a remainder $P_1(x)$, thus:

$$\frac{P(x)}{Q(x)} = S(x) + \frac{P_1(x)}{Q(x)},$$

where the degree of P_1 is < n. Since we can integrate the polynomial S without trouble we concentrate on $\frac{P_1(x)}{Q(x)}$. So from here on we shall discuss $\frac{P(x)}{Q(x)}$ under the assumption that the degree of P is < n. We can also assume that P(x) and Q(x) have no common factor, since such a one could be divided out.

Consider the possibility of writing our rational function in the form

$$\frac{P(x)}{Q(x)} = \frac{P_1(x)}{Q_1(x)} + \frac{P_2(x)}{Q_2(x)} + \dots ,$$

where each Q_i has lower degree than Q. For this to be possible each prime factor of Q(x) must appear in at least one $Q_i(x)$ to the same multiplicity that it appears in Q(x). This suggests that we take for the $Q_i(x)$ simply the prime factors of Q(x) to the powers to which they appear in Q(x). For example, we might expect to get

$$\frac{x^2 + 2}{(x - 1)(x + 2)^2(x^2 + 1)^3} = \frac{P_1(x)}{x - 1} + \frac{P_2(x)}{(x + 2)^2} + \frac{P_3(x)}{(x^2 + 1)^3}$$

The terms with multiple factors in the denominators can in turn be broken down still further; for example,

$$\frac{ax + b}{(x + 2)^2} = \frac{a(x + 2) + (b - 2a)}{(x + 2)^2} = \frac{a}{x + 2} + \frac{b - 2a}{(x + 2)^2}$$

There is more than one criterion for deciding when to stop but the most common case is stated in Theorem I, a proof of which is given in Appendix A.

<u>Theorem I.</u> Let Q(x) have the form

$$Q(x) = (x - a)^p (x - b)^q \ldots (x^2 + cx + d)^r (x^2 + ex + f)^s \ldots$$

where no two of $x - a$, $x - b$, ... , $x^2 + cx + d$, $x^2 + ex + f$, ... have a common factor. Then if degree of P < degree of Q,

$$\frac{P(x)}{Q(x)} = \frac{A_1}{x - a} + \frac{A_2}{(x - a)^2} + \ldots + \frac{A_p}{(x - a)^p}$$

$$+ \frac{B_1}{x - b} + \frac{B_2}{(x - b)^2} + \ldots + \frac{B_q}{(x - b)^q} + \ldots$$

$$+ \frac{C_1 x + D_1}{x^2 + cx + d} + \frac{C_2 x + D_2}{(x^2 + cx + d)^2} + \ldots + \frac{C_r x + D_r}{(x^2 + cx + d)^r}$$

$$+ \frac{E_1 x + F_1}{x^2 + ex + f} + \frac{E_2 x + F_2}{(x^2 + ex + f)^2} + \ldots + \frac{E_s x + F_s}{(x^2 + ex + f)^s}$$

$$+ \ldots \;.$$

The following examples illustrate the application of this theorem and two methods of determining the values of the constants A_1, B_1, ...

Example 1. We derive basic formula 4.

$$\frac{1}{x^2 - a^2} = \frac{A}{x - a} + \frac{B}{x + a}.$$

Clearing fractions gives

$$1 = A(x + a) + B(x - a).$$

Since this equation is an identity it will hold for any value of x.

For $x = a$: $1 = 2aA$ $\quad A = \frac{1}{2a}$.

For $x = -a$: $1 = -2aB$ $\quad B = -\frac{1}{2a}$.

Hence

$$\int \frac{1}{x^2 - a^2} dx = \int \frac{1/2a}{x - a} - \frac{1/2a}{x + a} dx$$

$$= \frac{1}{2a}\log |x - a| - \frac{1}{2a}\log |x + a| + C$$

$$= \frac{1}{2a}\log \left|\frac{x - a}{x + a}\right| + C.$$

Example 2. $\int \frac{x^4}{x^3 + 1} dx$.

First we must divide x^4 by $x^3 + 1$, to get $\frac{x^4}{x^3 + 1} = x - \frac{x}{x^3 + 1}$.

By Theorem 1,

$$\frac{x}{x^3 + 1} = \frac{x}{(x + 1)(x^2 - x + 1)} = \frac{A}{x + 1} + \frac{Bx + C}{x^2 - x + 1} .$$

Clearing fractions

$$x = A(x^2 - x + 1) + (Bx + C)(x + 1).$$

For $x = -1$: $\quad -1 = 3A, \quad A = -\frac{1}{3}$.

Since no real value of x will make $x^2 - x + 1$ equal zero (we prefer to avoid complex numbers at this point) we simply use any two values other than -1.

For $x = 0$: $\quad 0 = A + C, \quad C = -A = \frac{1}{3}$,

For $x = 1$: $\quad 1 = A + 2B + 2C, \quad B = \frac{1}{3}$.

Thus we have, finally,

$$\int \frac{x^4}{x^3 + 1}\, dx = \int \left[x - \frac{1/3}{x + 1} + \frac{1}{3} \frac{x + 1}{x^2 - x + 1} \right] dx.$$

The first two integrals are easy enough. For the last we complete the square in the denominator, $(x - 1/2)^2 + 3/4$, and then put $u = x - 1/2$, to get

$$\int \frac{u + 3/2}{u^2 + 3/4}\, du = \frac{1}{2} \log\,(u^2 + 3/4) + \frac{3}{2} \frac{2}{\sqrt{3}} \arctan \frac{2u}{\sqrt{3}} .$$

Hence

$$\int \frac{x^4}{x^3 + 1}\, dx = \frac{1}{2}x^2 - \frac{1}{3}\log|x + 1| + \frac{1}{6}\log(x^2 - x + 1)$$

$$+ \frac{1}{\sqrt{3}}\arctan\frac{2x - 1}{\sqrt{3}} + C.$$

Example 3. Let us do Example 2 by a different method. Starting with

$$x = A(x^2 - x + 1) + (Bx + C)(x + 1)$$

we multiply out the right hand side and collect terms, to get

$$x = (A + B)x^2 + (-A + B + C)x + (A + C).$$

Since this is an identity the coefficients of the various powers of x on the two sides must be equal; that is,

$$0 = A + B$$
$$1 = -A + B + C,$$
$$0 = A + C.$$

Solving these simultaneous equations gives $A = -1/3$, $B = 1/3$, $C = 1/3$, as before.

In some ways this is a more direct procedure than the substitution process of Example 2, but it leads to simultaneous equations of a more complicated sort. One can combine the two methods, first determining as many constants as possible by substitution and then expanding and equating coefficients.

Neither of these methods is suitable for very complicated problems. For these the algorithmic methods developed in Appendix A are to be preferred.

<u>Example 4.</u> $\int \frac{x^5 + 2}{(x^3 + 1)^2}\, dx.$

Here we have

$$\frac{x^5 + 2}{(x^3 + 1)^2} = \frac{A}{x + 1} + \frac{B}{(x + 1)^2} + \frac{Cx + D}{x^2 - x + 1} + \frac{Ex + F}{(x^2 - x + 1)^2},$$

or

$$x^5 + 2 = A(x + 1)(x^2 - x + 1)^2 + B(x^2 - x + 1)^2$$
$$+ (Cx + D)(x + 1)^2(x^2 - x + 1) + (Ex + F)(x + 1)^2.$$

By putting $x = -1$ we get $B = 1/9$. Further substitutions or equating of coefficients, (try each to see which you prefer) lead to five equations in the remaining five constants A, C, D, E, F. These can be solved by successive elimination to give

$$\frac{x^5 + 2}{(x^3 + 1)^2} = \frac{1}{9}\left[\frac{1}{(x + 1)^2} + \frac{7}{x + 1} + \frac{-9x + 6}{(x^2 - x + 1)^2} + \frac{2x + 4}{x^2 - x + 1}\right].$$

The first two of these fractions are easily integrated. For the last two we complete the square, $(x - 1/2)^2 + 3/4$, and substitute $u = x - 1/2$. The two fractions can then be further broken up, to give

$$(1) \qquad -\frac{9}{2}\int\frac{2u\,du}{(u^2+3/4)^2} + \frac{3}{2}\int\frac{du}{(u^2+3/4)^2} + \int\frac{2u\,du}{u^2+3/4} + 5\int\frac{du}{u^2+3/4}.$$

The first and third of these yield to the further substitution $v = u^2 + 3/4 = x^2 - x + 1$, giving

$$-\frac{9}{2}\int v^{-2}dv \quad \text{and} \quad \int v^{-1}dv.$$

The last is a standard form, so that leaves only the second. For this we have a recursion formula

$$\int\frac{dx}{(x^2+c)^n} = \frac{1}{2(n-1)c}\left[\frac{x}{(x^2+c)^{n-1}} + (2n-3)\int\frac{dx}{(x^2+c)^{n-1}}\right]$$

derived in Appendix A. This gives

$$\frac{3}{2}\int\frac{du}{(u^2+3/4)^2} = \frac{3}{2}\,\frac{2}{3}\left[\frac{u}{u^2+3/4} + \int\frac{du}{u^2+3/4}\right].$$

Combining these results and adding similar terms gives us finally

$$\int\frac{x^5+2}{(x^3+1)^2} = \frac{1}{9}\left[\frac{-1}{x+1} + 7\log|(x+1)| + \frac{x+4}{x^2-x+1} + \log|(x^2-x+1)| + \frac{12}{\sqrt{3}}\arctan\frac{2x-1}{\sqrt{3}}\right] + C.$$

Problems

1. Evaluate each of the following

(a) $\int \frac{1}{4x^2 + 12x + 9} dx$

(b) $\int \frac{6x^2 + 1}{2 - x - 6x^2} dx$

(c) $\int \frac{x^3}{x^2 - 2x - 3} dx$

(d) $\int \frac{dx}{(x - 1)^2(x - 2)}$

(e) $\int \frac{dx}{(x^2 - 4)(x^2 - 9)}$

(f) $\int \frac{x^2 + 2x + 3}{x^2 - 3x + 2} dx$

(g) $\int \frac{x^2 - 2}{(x + 1)(x - 1)^2} dx$

(h) $\int \frac{x^3 + 3}{x^2 + 4} dx$

(i) $\int \frac{2x^2 + 3x - 1}{(x + 3)(x + 2)(x - 1)} dx$

(j) $\int \frac{x^2 - 37}{x^2 - x - 12} dx$

(k) $\int \frac{x\,dx}{(x-1)(x-2)}$

(l) $\int \frac{x^2 - 2x + 3}{(x-1)^2(x^2+4)}\,dx$

(m) $\int \frac{2x^3 + x^2 + 5x + 4}{x^4 + 8x^2 + 12}\,dx$

(n) $\int \frac{x^2 - x + 1}{x^3 + 6x^2 + 11x + 6}\,dx$

(o) $\int \frac{x^3 - x^2 + 2x + 3}{(x^2 + 2x + 2)^2}\,dx$

(p) $\int \frac{3x^2 + 11x + 4}{x^3 + 4x^2 + x - 6}\,dx$

(q) $\int \frac{3x^3}{(x-1)^2(x^2+x+1)}\,dx$

(r) $\int \frac{6x^3}{(x^2+1)^2}\,dx$

(s) $\int \frac{dx}{x(x^2+1)}$

2. Finish Problem 10 in Section 3, namely

(a) $\int \frac{4t}{(t^2+1)(t+1)^2}\,dt,$ $\qquad t = \tan \frac{x}{2}.$

(b) $\int \frac{2u^2}{1-u^4}\,du,$ $\qquad u = \tan x,$

or

788

$$\int \frac{16t^2}{(t^2 + 1)(t^2 + 2t - 1)(t^2 - 2t - 1)} dt, \qquad t = \tan \frac{x}{2} .$$

(c) $\int \frac{1}{(1 - u^2)^2} du,$ $\qquad u = \sin x,$

or

$$\int \frac{2(1 + t^2)^2}{(1 - t^2)^3} dt, \qquad t = \tan \frac{x}{2}.$$

3. (a) $x^4 + 4$ can be factored in the form $(x^2 + ax + b)(x^2 + cx + d)$. Find a, b, c, d and evaluate $\int \frac{1}{x^4 + 4} dx$.

(b) Make a substitution to reduce

$$\int \frac{1}{x^4 + a^4} dx \quad \text{to} \quad K \int \frac{1}{u^4 + 4} du$$

and hence get a formula for this general integral.

6. Integral Tables and Guessing.

One needs to develop skill in handling complicated integrals but 95% of the integrals one meets can be found in a good set of tables. The word "found" is used loosely. No table will contain the integral

$$\int (x^6 - 2x^2 + 4)e^{-3x}dx,$$

but a typical table will have $\int e^x dx$, $\int x^2 e^x dx$, and a reduction formula for $\int x^n e^x dx$. Making the substitution $u = -3x$ and breaking the integral into three parts will enable one to use the table.

You should spend some time getting acquainted with your table, so that you know what to expect to find in it and where to look for a given type of integral. Otherwise you may spend more time hunting than you would in integrating the function by other methods. One table, for instance, does not give a reduction formula for

$$\int \frac{dx}{x^n \sqrt{x^2 + a^2}}$$

but does give one for

$$\int \frac{dx}{x^n (a + bx^2)^{1/2}} .$$

790

Some tables give an extensive set of integrals involving $\sqrt{ax^2 + bx + c}$ and others give none of these at all, expecting you to complete the square in all such cases.

At all times one should keep alert for short-cuts. In Example 5-4, for instance, we can write

$$\int \frac{x^5 + 2}{(x^3 + 1)^2}\,dx = \int \frac{x^3}{(x^3 + 1)^2}\,x^2\,dx + \int \frac{2}{(x^3 + 1)^2}\,dx$$

and the first integral yields at once to the substitution $u = x^3 + 1$. The second one must still be done by partial fractions, but is somewhat simpler than the original integral if the algorithmic method of Appendix A is used. Another example of this type is

$$\int \frac{x^{11}\,dx}{(x^2 + 1)^{7/2}} .$$

Even if one can find a reduction formula for it, it is easier to use

$$u^2 = x^2 + 1, \qquad 2u\,du = 2x\,dx,$$

to get

$$\int \frac{(u^2 - 1)^5 u\,du}{u^7} ,$$

which needs only a little algebra.

In Section 1 we spoke of "guessing" the result of an integration and then checking our guess by differentiating.

Usually our guess is based on considerable knowledge of the form of the answer, and the differentiating determines the details. An example will show how this works.

Example 1. $\int (x^3 - 2x^2 + 4x - 2)e^{-2x}dx$.

Integration by parts (see Problems 1b and g in Section 4) convince one that $\int x^n e^{ax}dx$ is of the form $P(x)e^{ax}$ where P is a polynomial of degree n. Adding several of these together, we must have (or we "guess" that)

$$\int (x^3 - 2x^2 + 4x - 2)\, e^{-2x}dx$$

$$= (ax^3 + bx^2 + cx + d)\, e^{-2x} + C.$$

Differentiating the "answer" gives

$$(3ax^2 + 2bx + c - 2ax^3 - 2bx^2 - 2cx - 2d)e^{-2x}.$$

Hence we want

$-2a = 1,$	$a = -\frac{1}{2},$
$-2b + 3a = -2,$	$b = \frac{1}{2}(3a + 2) = \frac{1}{4},$
$-2c + 2b = 4,$	$c = \frac{1}{2}(2b - 4) = -\frac{7}{4},$
$-2d + c = -2,$	$d = \frac{1}{2}(c + 2) = \frac{1}{8};$

giving

$$\int (x^3 - 2x^2 + 4x - 2)\, e^{-2x}dx = \frac{1}{8}(-4x^3 + 2x^2 - 14x + 1)e^{-2x}.$$

The easiest way to get

$$\int e^{ax}(A \sin bx + B \cos bx)dx$$

79

is not to integrate by parts as in Example 4-4, or to hunt for it in tables, but to assume that

$$\int e^{ax}(A \sin bx + B \cos bx)dx = e^{ax}(M \sin bx + N \cos bx) + C,$$

differentiate, equate coefficients of sin bx and cos bx, and solve for M and N.

This technique if particularly good for combinations of polynomials, e^{ax}, sin bx, and cos bx.

Problems

1. Use the tables to evaluate the following

(a) $\int \frac{x^2}{(2x^2 - 1)^{3/2}}\, dx$

(b) $\int x^2 \sin^2 3x\, dx$

(c) $\int \frac{x^2}{(2x - 1)^{3/2}}\, dx$

(d) $\int \sin^4 2x \cos^2 2x\, dx$

(e) $\int \frac{x^2}{(2x^2 - x - 1)^{3/2}}\, dx$

(f) $\int \frac{x^2}{(2x - 1)^3}\, dx$

(g) $\int e^{2x} \sin^4 3x\, dx$

(h) $\int \frac{x^2}{(2x^2 - 1)^3}\, dx$

(i) $\int x \sec^4 3x\, dx$

(j) $\int \frac{x^2}{(x^2 + 4)^2}\, dx$

2. Solve the following problems by using a table of integrals.

Section 2, Problem 1b, d, h, i, l, n, o, q, r, s, u, v, y, z.

Section 3, Problem 1a, b, e, g, i

Problem 3a, c, d

Problem 6, all parts

Problem 9, a, b, c, d, e, h

Problem 10, a, c

Problem 11

Section 4, Problem 1a, b, c, g, h, i, j, k, l, m.

3. Evaluate $\int (x^4 - x)e^{-x}\, dx$ by assuming a form for the answer and differentiating.

4. Get a formula for

$$\int e^{ax}(c \sin bx + d \cos bx)dx$$

5. Assuming that

$$\int xe^{ax}(c \sin x + d \cos x)dx$$

$$= e^{ax}[(px + q)\sin x + (rx + s) \cos x]$$

find p, q, r, s in terms of a, c, d.

6. What do you estimate to be the form of

$$\int (ax^2 + bx + c)e^{hx} \sin kx\, dx?$$

7. Definite Integrals.

At this point we abandon the convention of using "integral" to mean "indefinite integral". From now on "integral" will have its original meaning, although we may still refer to a "definite integral" for emphasis. An indefinite integral will always be referred to as such.

So far our main interest in indefinite integrals has been as an aid to the evaluation of definite integrals, although we also used them in solving differential equations in Chapter 9. The second form of the Fundamental Theorem (Section 8-3) tells us that if F is an indefinite integral of f over the interval [a,b] then

$$\int_a^b f(x)dx = F(b) - F(a).$$

Example 1. $\int_0^{\pi/2} \frac{1}{2 + \sin\theta} d\theta.$

From Example 3-11 we have

$$\int_0^{\pi/2} \frac{1}{2 + \sin\theta} d\theta = \frac{2}{\sqrt{3}} \arctan\left(\frac{1 + 2\tan\theta/2}{\sqrt{3}}\right)\Bigg|_0^{\pi/2}$$

Since $\tan(\pi/4) = 1$ and $\tan 0 = 0$ we get

$$\frac{2}{\sqrt{3}}(\arctan \sqrt{3} - \arctan \frac{1}{\sqrt{3}})$$

$$= \frac{2}{\sqrt{3}}(\frac{\pi}{3} - \frac{\pi}{6})$$

$$= \frac{\pi}{3\sqrt{3}} .$$

If we use the form

$$\frac{2}{\sqrt{3}} \arctan \frac{2 - 2 \cos \theta + \sin \theta}{\sqrt{3} \sin \theta}$$

for the indefinite integral we run into trouble at $\theta = 0$. This is not serious, for we can evaluate

$$\lim_{\theta \to 0+} \frac{2 - 2 \cos \theta + \sin \theta}{\sqrt{3} \sin \theta} = \lim_{\theta \to 0+} \frac{2 \sin \theta + \cos \theta}{\sqrt{3} \cos \theta} = \frac{1}{\sqrt{3}}$$

by L'Hospital's Rule. A similar situation occurs in the following example.

<u>Example 2.</u> $\displaystyle\int_0^\pi \frac{1}{2 + \sin \theta} d\theta = \frac{2}{\sqrt{3}} \arctan \frac{1 + 2 \tan \theta/2}{\sqrt{3}} \Bigg|_0^\pi$

Now $\tan \pi/2$ is not defined, but

$$\lim_{\theta \to \pi-} \tan \frac{\theta}{2} = \infty,$$

and

$$\lim_{\theta \to \pi-} \frac{1 + 2 \tan \theta/2}{\sqrt{3}} = \infty,$$

and

$$\lim_{v\to\infty} \arctan v = \pi/2.$$

Hence we conclude that

$$\lim_{\phi\to\pi-} \int_0^\phi \frac{1}{2+\sin\theta}\, d\theta = \frac{2}{\sqrt{3}} \left[\frac{\pi}{2} - \frac{\pi}{6}\right] = \frac{2\pi}{3\sqrt{3}} .$$

Finally,

$$F(\phi) = \int_0^\phi \frac{1}{2+\sin\theta}\, d\theta$$

is continuous at $\phi = \pi$ (this follows from the argument in Section 8-2), and so

$$F(\pi) = \lim_{\phi\to\pi-} F(\phi) = \frac{2\pi}{3\sqrt{3}} .$$

Notice that the function $F(\phi)$ defined above is defined for all values of ϕ, and is > 0 for $\phi \geq 0$ since $\frac{1}{2+\sin\theta} > 0$ for all θ.

Example 3. Proceeding as before, we get

$$\int_0^{2\pi} \frac{1}{2+\sin\theta}\, d\theta = \frac{2}{\sqrt{3}} \arctan \frac{1+2\tan\theta/2}{\sqrt{3}} \Bigg|_0^{2\pi}$$

$$\frac{2}{\sqrt{3}} \left[\arctan\frac{1}{\sqrt{3}} - \arctan\frac{1}{\sqrt{3}}\right] = 0.$$

By the comment above this is certainly wrong. We might suspect that the trouble occurred when θ went past π, since this point caused the complications in Example 2.

This is correct. When θ varies from $\pi - \varepsilon$ to $\pi + \varepsilon$ the quantity $\arctan \dfrac{1 + 2 \tan \theta/2}{\sqrt{3}}$ varies continuously from just below $\pi/2$ to just above $\pi/2$, and in doing so jumps from one branch of the arctan curve to another. (Figure 7-1). The value of $\arctan 1/\sqrt{3}$ arising from $\theta = 2\pi$ must therefore be taken not as $\pi/6$ but rather as $\pi/6 + \pi$. This gives us the sensible answer $2\pi/\sqrt{3}$.

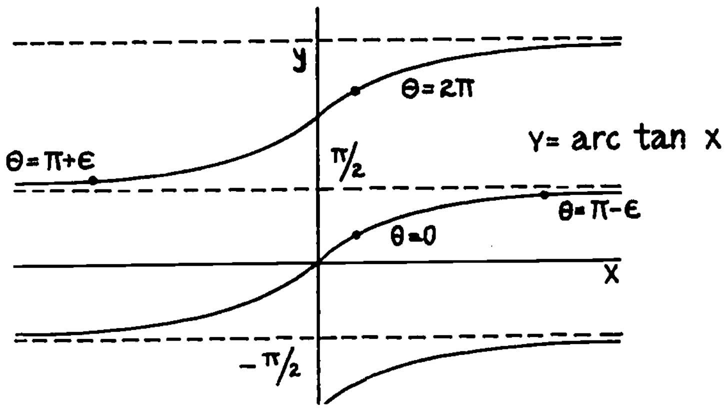


Figure 7-1

Another way of regarding this case is that the function

$$\frac{2}{\sqrt{3}} \arctan \frac{1 + 2 \tan \theta/2}{\sqrt{3}} ,$$

where arctan is restricted to the range $(-\pi/2, \pi/2)$ as agreed upon in Section 7-3, is discontinuous at $\theta = \pi$ and hence is an indefinite integral only over an interval that does not contain π (or 3π, or $-\pi$, or etc.). We could then proceed to evaluate $\int_0^{2\pi}$ as $\int_0^{\pi} + \int_{\pi}^{2\pi}$, using the technique of Example 2 on the limits π.

Considerations like this are distinctly annoying, especially when using integral tables. All answers involving inverse trig functions, particularly arctan, should be regarded with suspicion. For example

$$\int \frac{dx}{a^4 + x^4} = \frac{1}{4a^3\sqrt{2}} \log \frac{x^2 + ax\sqrt{2} + a^2}{x^2 - ax\sqrt{2} + a^2} + \frac{1}{2a^3\sqrt{2}} \arctan \frac{ax\sqrt{2}}{a^2 - x^2}$$

can be expected to cause trouble if the interval of integration contains either a or $-a$, values of x at which the argument of the arctan term becomes infinite.

In our examples so far we have obtained the indefinite integral and substituted the limits of the integral. If any substitutions are made in the evaluation of the integral, however, we generally change the limits to agree with the new variable and never go back to the original variable. This idea was introduced in Section 8-3.

Example 4. $J = \int_0^1 \frac{1}{(x + 2)\sqrt{1 - x^2}}\, dx.$

This calls for the substitution of Section 3-III,

$$x = \sin\theta, \quad \sqrt{1 - x^2} = \cos\theta, \quad dx = \cos\theta\, d\theta.$$

For $x = 0$ and $x = 1$ we have $\theta = 0$ and $\theta = \pi/2$. So we get our old friend,

$$J = \int_0^{\pi/2} \frac{1}{2 + \sin\theta}\, d\theta.$$

The first step in integrating this in Example 3-11 was the substitution $t = \tan\theta/2$. This gives the limits $t = 0$ and $t = \tan\pi/4 = 1$. So

$$J = \int_0^1 \frac{1}{t^2 + t + 1}\, dt.$$

The next substitution was $u = t + 1/2$, giving

$$J = \int_{1/2}^{3/2} \frac{1}{u^2 + 3/4}\, du$$

$$= \frac{2}{\sqrt{3}} \arctan \frac{2u}{\sqrt{3}} \Bigg|_{1/2}^{3/2}$$

$$= \frac{2}{\sqrt{3}} \left(\arctan \sqrt{3} - \arctan \frac{1}{\sqrt{3}} \right)$$

$$= \frac{\pi}{3\sqrt{3}} .$$

What happens if we try this process on Example 3? If we proceed mechanically we get $t = \tan \theta/2$, $t = 0$ when $\theta = 0$, $t = 0$ when $\theta = 2\pi$, and so

$$\int_0^{2\pi} \frac{1}{2 + \sin \theta}\, d\theta = \int_0^{0} \frac{1}{t^2 + t + 1}\, dt = 0,$$

the same wrong answer. Here it is obvious that the function $t = \tan \theta/2$ is discontinuous at $\theta = \pi$, and so this substitution cannot be used over any interval containing π. From this point of view the case of Example 3 will be examined in the next section.

In general, in evaluating definite integrals one must be suspicious of the numerical substitutions, making sure that the various functions used in the evaluation behave properly over the required intervals.

The substitutions of limits as one proceeds with the evaluation will generally simplify the writing of an integration by parts or a reduction formula, and sometimes even the formula itself.

Example 5. $\int_0^1 x^7 e^x \, dx.$

We get a reduction formula for

$$\int_0^1 x^n e^x \, dx.$$

Using

$$u = x^n, \qquad dv = e^x dx,$$

$$du = nx^{n-1} dx, \qquad v = e^x,$$

gives

$$\int_0^1 x^n e^x dx = x^n e^x \Big|_0^1 - n \int_0^1 x^{n-1} e^x dx$$

$$= e - n \int_0^1 x^{n-1} e^x dx.$$

Write this as a recursion formula:

$$J_n = e - nJ_{n-1}, \qquad J_0 = \int_0^1 e^x dx = e - 1.$$

Then

$$J_7 = e - 7J_6$$

$$= e - 7e + 7\cdot 6J_5$$

$$= e - 7e + 7\cdot 6e - 7\cdot 6\cdot 5J_4$$

$$= \ldots$$

$$= e(1 - 7 + 7\cdot 6 - 7\cdot 6\cdot 5 + \ldots + 7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2) - 7!(e - 1)$$

$$= 5040 - 1854e \approx .31.$$

Example 6. Definite integrals of the form

$$\int_0^{\pi/2} \sin^m x \cos^n x \, dx$$

occur frequently. Consider

$$\int_0^{\pi/2} \sin^n x \, dx.$$

In Example 4-5 we got the recursion formula

$$\int \sin^n x \, dx = -\frac{1}{n} \sin^{n-1} x \cos x + \frac{n-1}{n} \int \sin^{n-2} x \, dx.$$

Now for $n > 1$,

$$\sin^{n-1} x \cos x \Big|_0^{\pi/2} = 0,$$

since sin 0 = 0 and cos $\pi/2 = 0$. So we have

$$\int_0^{\pi/2} \sin^n x \, dx = \frac{n-1}{n} \int_0^{\pi/2} \sin^{n-2} x \, dx$$

as the recursion formula for the definite integral. This is simple enough that we can write a formula for the integral. Since

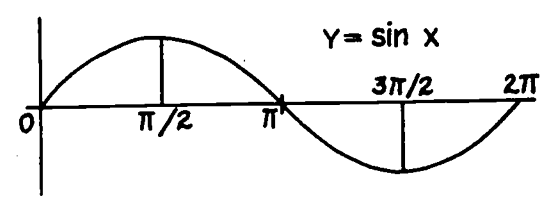
$$\int_0^{\pi/2} \sin x \, dx = 1 \text{ and } \int_0^{\pi/2} 1 \, dx = \pi/2,$$

we have

$$(1) \qquad \int_0^{\pi/2} \sin^n x \, dx = \begin{cases} \dfrac{(n-1)(n-3)\ldots 1}{n(n-2)\ldots 2}\dfrac{\pi}{2} & \text{if } n \text{ is even} \\[2ex] \dfrac{(n-1)(n-3)\ldots 2}{n(n-2)\ldots 3} & \text{if } n \text{ is odd.} \end{cases}$$

By interpreting definite integrals as areas we can apply (1) to other useful integrals. Since sin x assumes the same values in $[\pi/2,\pi]$ as in $[0,\pi/2]$ so does $\sin^n x$ for any n (Figure 7-2), and so

$$\int_{\pi/2}^{\pi} \sin^n x \, dx = \int_0^{\pi/2} \sin^n x \, dx,$$



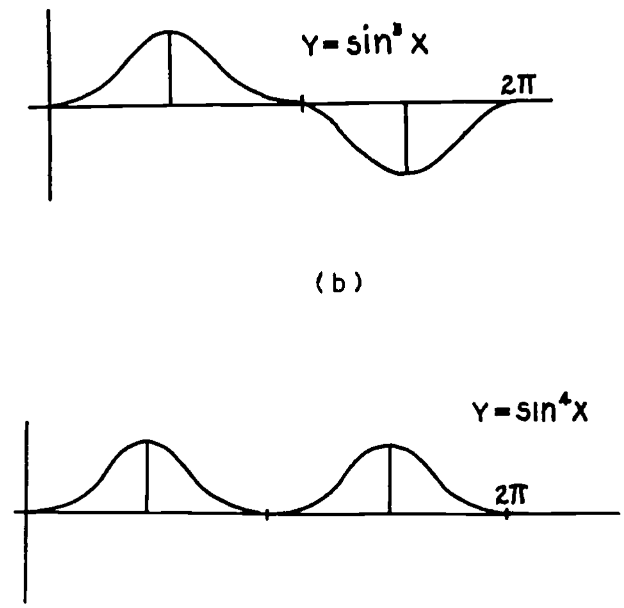
(a)

Figure 7-2

from which we get

$$\int_0^{\pi} \sin^n x \, dx = 2 \int_0^{\pi/2} \sin^n x \, dx.$$

On the interval $[\pi, 2\pi]$, the values of $\sin x$ are the negatives of those on $[0, \pi]$. The same is true for $\sin^n x$ if n is odd (Figure 7-2(b)) but for n even the signs are the same (Figure 7-2(c)). Hence for n even



(b)

(c)

Figure 7-2

$$\int_0^{2\pi} \sin^n x \, dx = 2 \int_0^{\pi} \sin^n x \, dx$$

$$= 4 \int_0^{\pi/2} \sin^n x \, dx,$$

but for n odd

$$\int_0^{2\pi} \sin^n x \, dx = 0.$$

By comparing the graph of $y = \cos x$ (Figure 7-3) with that of $y = \sin x$ we conclude similarly that

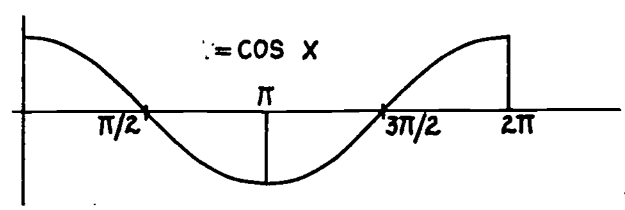


Figure 7-3

$$\int_0^{\pi/2} \cos^n x \, dx = \int_0^{\pi/2} \sin^n x \, dx;$$

$$\int_0^{\pi} \cos^n x \, dx = 2 \int_0^{\pi/2} \cos^n x \, dx \text{ if } n \text{ is even, otherwise } 0;$$

$$\int_0^{2\pi} \cos^n x \, dx = 2 \int_0^{\pi} \cos^n x \, dx.$$

These methods are useful in various places. One application is sufficiently general to justify formal statement. We recall that a function f defined on an interval $[-a,a]$ is <u>even</u> if $f(-x) = f(x)$ and <u>odd</u> if $f(-x) = -f(x)$ for every x in the interval.

<u>Theorem I.</u> If f is $\begin{Bmatrix} \text{even} \\ \text{odd} \end{Bmatrix}$ on $[-a,a]$ then

$$\int_{-a}^{0} f(x)\,dx = \begin{Bmatrix} + \\ - \end{Bmatrix} \int_0^a f(x)\,dx$$

and

$$\int_{-a}^{a} f(x)\,dx = \begin{Bmatrix} 2\int_0^a f(x)\,dx \\ 0 \end{Bmatrix}$$

Here either all the top terms in the three pairs of brackets are to be taken or all the bottom terms.

Proof. The statements are obvious from a graph but we shall give an analytic proof for variety. Putting $u = -x$ we get

$$\int_{-a}^{0} f(x)dx = \int_{a}^{0} f(-u)(-du) = \int_{0}^{a} f(-u)du = \begin{Bmatrix} + \\ - \end{Bmatrix} \int_{0}^{a} f(u)du.$$

This is the first half of the theorm. The second half follows from

$$\int_{-a}^{a} f(x)dx = \int_{-a}^{0} f(x)dx + \int_{0}^{a} f(x)\ dx.$$

Example 7. If we should wish to find $\int_{-\pi/4}^{\pi/4} x^2 \sin 3x\ dx$ the results of Example 4-3 are unnecessary. Since $x^2 \sin 3x$ is an odd function we know that the value of the integral is zero.

Example 8. An integral like

$$\int_{-1}^{1} \frac{x^3 + x + 1}{\sqrt{x^2 + 4}} dx$$

can be separated into odd and even parts;

$$\int_{-1}^{1} \frac{x^3 + x}{\sqrt{x^2 + 4}}\, dx + \int_{-1}^{1} \frac{1}{\sqrt{x^2 + 4}}\, dx.$$

The first is zero and the second is

$$2 \int_{0}^{1} \frac{1}{\sqrt{x^2 + 4}}\, dx = 2 \log (x + \sqrt{x^2 + 4}) \Big|_{0}^{1}$$

$$= 2 [\log (1 + \sqrt{5}) - \log 2]$$

$$= 2 \log \frac{1 + \sqrt{5}}{2} \approx .9624.$$

Many problems leading to definite integrals can conveniently be done in terms of parametric equations.

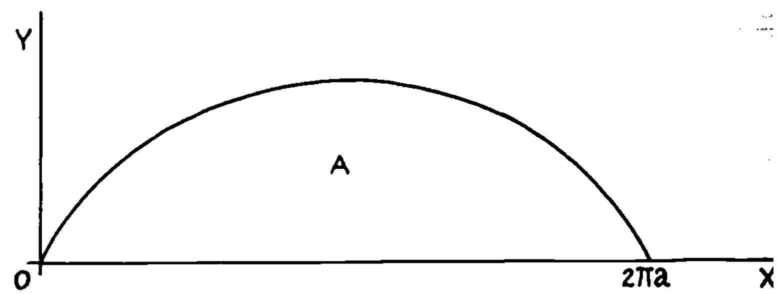


Figure 7-4

Example 9. To find the area of one arch of the cycloid

$$x = a(\theta - \sin \theta),$$

$$y = a(1 - \cos \theta)$$

we have

$$A = \int_0^{2\pi a} y\ dx$$

$$= \int_0^{2\pi} a(1 - \cos\theta)\ a(1 - \cos\theta)\ d\theta$$

$$= a^2 \int_0^{2\pi} (1 - 2\cos\theta + \cos^2\theta)\ d\theta$$

$$= a^2(2\pi + 4 \cdot \frac{1}{2}\frac{\pi}{2}), \text{ using Example 6,}$$

$$= 3\pi a^2.$$

This technique requires care, however. Given simply a parametrized curve,

(2) $x = f(t),\ y = g(t),$

$a \le t \le b,$

it may not be true that

$$\int y\ dx = \int_a^b g(t)f'(t)dt$$

represents an area. For instance, (2) may give us a curve like the one in Figure 7-5. What we have in mind when we write $A = \int y\ dx$

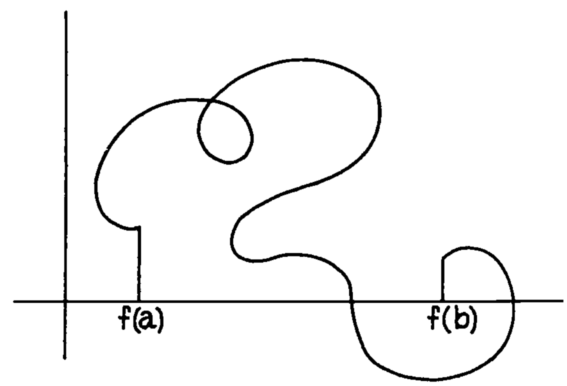


Figure 7-5

is someting like Figure 7-6; that is, (2) defines y as a function of x over the interval $f(a) \leq x \leq f(b)$. This will be so <u>if</u> $f(t)$ <u>is strictly monotone in</u> $[a,b]$. Then the function f has an inverse function h, so that $t = h(x)$ on $[f(a), f(b)]$, or on $[f(b), f(a)]$ if f is decreasing, and

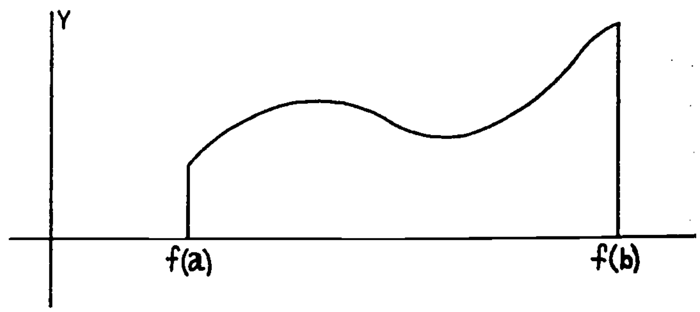


Figure 7-6

$$\int_a^b g(t)f'(t)dt = \int_{f(a)}^{f(b)} g(h(x))dx$$

$$= \int_{f(a)}^{f(b)} y(x)dx.$$

Problems

1. Evaluate the following definite integrals. Do not use tables of definite integrals.

(a) $\int_{-1}^{0} \frac{x^2 - 1}{x + 2}\, dx$

(b) $\int_{0}^{1} \arctan x\, dx$

(c) $\int_{1}^{2} \sqrt{x} \log x\, dx$

(d) $\int_{0}^{\pi/2} \frac{\cos \theta}{1 + \sin \theta}\, d\theta$

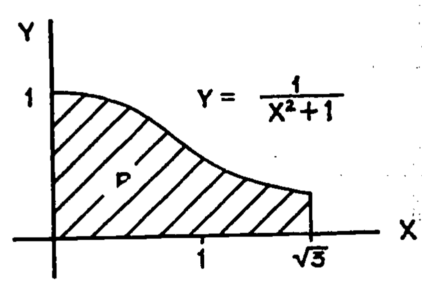
(e) $\int_{0}^{\pi} \frac{\cos \theta}{1 + \sin \theta}\, d\theta$

(f) $\int_{0}^{1} x^3(1 + x^2)^{3/2} dx$

(g) $\int_{0}^{\pi} e^{-x} \cos x\, dx$

(h) $\int_{0}^{1} \frac{1}{(x^2 - 4)\sqrt{x^2 + 1}}\, dx$ Ans. $-\frac{1}{4\sqrt{5}} \log \frac{13 + 4\sqrt{10}}{3} \approx -.24$

2. (a) The region in the first quadrant bounded by the axes, the line $x = \sqrt{3}$, and the curve $y = \frac{1}{x^2 + 1}$ is rotated about the x-axis. What is the volume of the resulting solid?



(b) What is the volume if the region is rotated about the y-axis?

Answers. $\pi(\sqrt{3}/8 + \pi/6)$, $\pi \log 4$.

3. (a) Find the area of an ellipse, using the parametric equations

$$x = a \cos \theta, \qquad y = b \sin \theta.$$

(b) Find the volume of the ellipsoid obtained by rotating the ellipse about the x-axis.

4. Find the volume obtained by rotating the area in Example 9 about the x-axis.

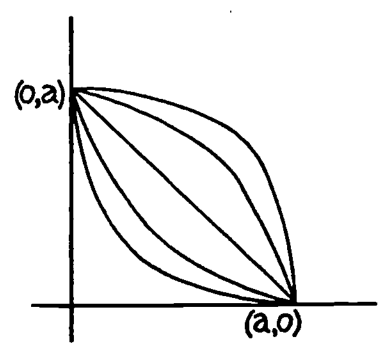
Ans. $5\pi^2 a^3$.

5. The curves

$$x^{2/p} + y^{2/p} = a^{2/p}, \quad p > 0,$$

have the general shapes shown in the figure. They can be parametrized as $x = a \cos^p\theta$, $y = a \sin^p\theta$.

We are interested in the area A_p bounded by such a curve and the axes.



(a) $A_1 = \frac{1}{4}\pi a^2$. Why?

(b) $A_2 = \frac{1}{2} a^2$. Why?

(c) Compute A_3 and A_4.

(d) Derive the reduction formula

$$A_p = \frac{1}{4} \frac{p}{p - 1} A_{p-2},$$

for $p > 2$. Check it with your answers to (c).

(e) What can you say about the shape of the curves for $p = 100$ and $p = .01$?

8. Improper Integrals.

The integral

$$\int_0^M e^{-x}dx$$

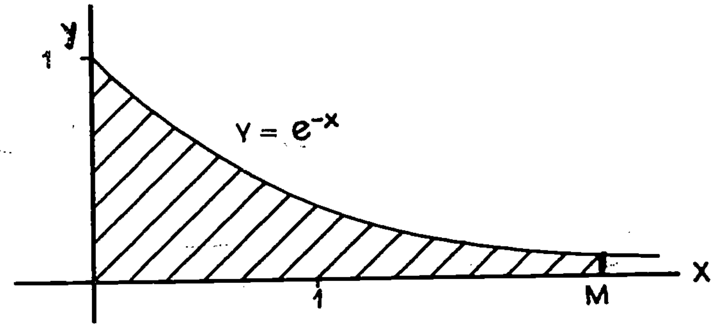


Figure 8-1

has for its value the area of the cross-hatched region in Figure 8-1. This value, as we saw in Chapter 8, is a function of M, F(M), and indeed we see that

$$F(M) = 1 - e^{-M}.$$

Now $\lim_{M \to \infty} F(M)$ exists, and is obviously equal to 1; that is,

$$\lim_{M \to \infty} \int_0^M e^{-x}dx = 1.$$

We summarize this situation by writing

$$\int_0^\infty e^{-x}dx = 1,$$

and saying that the area of the region bounded by the x-axis, the y-axis, and the curve $y = e^{-x}$ is 1. Notice that this is an extension of the concepts of "area" and "integral". To identify an integral of this new type we call it an <u>improper integral</u>.

In evaluating an improper integral we must go back to its definition, as a limit of a proper integral.

Example 1. $\int_0^\infty \frac{1}{1 + x^2}\, dx.$

We start with

$$\int_0^M \frac{1}{1 + x^2}\, dx = \arctan M.$$

Then

$$\int_0^\infty \frac{1}{1 + x^2}\, dx = \lim_{M\to\infty} \int_0^M \frac{1}{1 + x^2}\, dx$$

$$= \lim_{M\to\infty} \arctan M = \frac{\pi}{2}.$$

Example 2. $\int_0^\infty \frac{x}{1 + x^2}\, dx$

$$\int_0^M \frac{x}{1 + x^2}\, dx = \frac{1}{2} \log (1 + x^2) \Big|_0^M$$

$$= \frac{1}{2} \log (1 + M^2) - 0.$$

Hence

$$\int_0^\infty \frac{x}{1 + x^2}\, dx = \lim_{M\to\infty} \frac{1}{2} \log (1 + M^2) = \infty.$$

If the limit exists and is finite, as in Example 1, we say the integral converges, or exists. In a case like Example 2 we say the integral diverges or does not exist. We also apply the same terms to an integral like

$$\int_0^\infty \cos x \, dx = \lim_{M\to\infty} \sin M$$

which does not exist in any sense. If we wish to distinguish between the two types of divergence we can use the phrase "divergent to infinity" for the first kind.

We leave to the reader the definition of $\int_{-\infty}^a f(x)dx$. The doubly infinite integral $\int_{-\infty}^\infty f(x)dx$ is best treated by breaking into two single improper integrals,

$$\int_{-\infty}^a f(x)dx + \int_a^\infty f(x)dx,$$

for some convenient value of a.

A second type of improper integral arises from a discontinuity of the integrand at one end of the interval. Here again we define the improper integral as the limit of a proper integral.

Example 3. $\int_0^1 \frac{1}{\sqrt{1 - x^2}}\, dx.$

The integrand is defined only on [0,1) and not on [0,1] as we required for our theory in Chapter 3. So we take an interval [0,h], where h < 1, and let h → 1; i.e. take

$$\lim_{h \to 1-} \int_0^h \frac{1}{\sqrt{1 - x^2}}\, dx = \lim_{h \to 1-} \arcsin h = \frac{\pi}{2} .$$

This is our definition of the improper integral $\int_0^1 \frac{1}{\sqrt{1 - x^2}}\, dx$. This type of improper integral is not as easy to recognize as the first type, and one must always be on the watch for discontinuities of the integrand in the closed interval of integration.

Example 4. $\int_{-1}^2 x^{-3} dx.$

One is apt to write carelessly,

$$\int_{-1}^2 x^{-3} dx = -\frac{1}{2} x^{-2} \Big|_{-1}^{2} = -\frac{1}{8} + \frac{1}{2} = \frac{3}{8} ,$$

which is false. This is an improper integral because of the discontinuity at x = 0. Its value, if any, is

$$\lim_{h \to 0-} \int_{-1}^h x^{-3} dx + \lim_{k \to 0+} \int_k^2 x^{-3} dx$$

$$= \lim_{h\to 0-}\left(-\frac{1}{2h^2}+\frac{1}{2}\right)+\lim_{k\to 0+}\left(-\frac{1}{8}+\frac{1}{2k^2}\right).$$

The integral converges only if both limits exist. In this case neither limit exists, so the given improper integral does not exist.

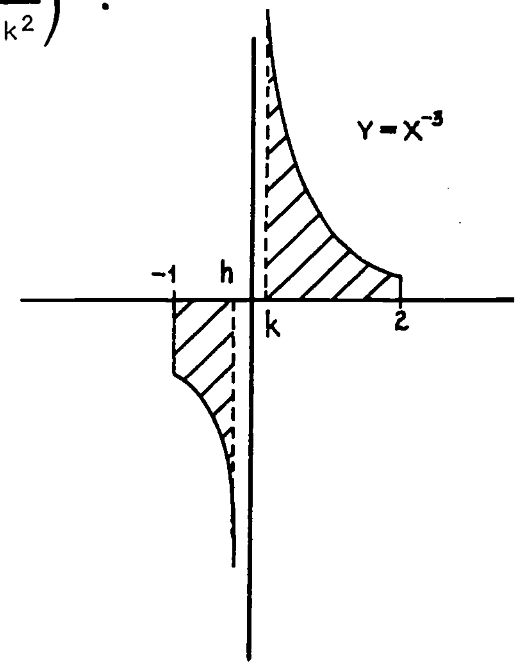


Figure 8-2

Example 5. $\int_2^\infty \frac{1}{x^2 - x}\,dx.$

Using partial fractions,

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

so

$$\int_2^\infty \frac{1}{x^2 - x}\,dx = \lim_{M\to\infty}\int_2^M \left(\frac{1}{x - 1} - \frac{1}{x}\right)dx$$

$$= \lim_{M\to\infty}\left[\log(x - 1) - \log x\right]_2^M$$

$$= \lim_{M\to\infty}\left[\log(M - 1) - \log M + \log 2\right].$$

Now neither $\log(M - 1)$ nor $\log M$ converges as $M \to \infty$, so we are at first inclined to compare this to Example 3 and say that the integral does not exist. Here, however, we are not interested in either of these two functions as such but in the function $\log(M - 1) - \log M$. As in Problem 4(g) of

818

Section 10-4 we find

$$\lim_{M\to\infty} (\log(M - 1) - \log M)$$

$$= \lim_{M\to\infty} \log \frac{M - 1}{M}$$

$$= \lim_{M\to\infty} \log (1 - \frac{1}{M}) = \log 1 = 0,$$

so our integral does converge to the value log 2.

In evaluating an improper integral by integration by parts we can carry out the appropriate limiting processes to the integrated part as soon as we obtain it.

Example 6. The gamma function, $\Gamma(x)$, is defined by

$$\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}dt$$

for all values of a for which the integral converges. (Also for some other values of a, by more advanced methods.) We have $\Gamma(1) = \int_0^\infty e^{-t}dt = 1$. If a is an integer greater than 1 we apply integration by parts:

$$u = t^{a-1}, \qquad dv = e^{-t}dt,$$

$$du = (a - 1)t^{a-2}dt, \qquad v = -e^{-t},$$

$$\Gamma(a) = -t^{a-1}e^{-t}\Big|_0^\infty + (a - 1) \int_0^\infty t^{a-2}e^{-t}dt$$

$$= -\lim_{M\to\infty} M^{a-1}e^{-M} + 0 + (a - 1)\Gamma(a - 1).$$

By Section 10-4, $\lim_{M\to\infty} M^{a-1}e^{-M} = 0$, so we get the recursion formula

$$\Gamma(a) = (a - 1)\Gamma(a - 1)$$

if a is an integer greater than 1. We shall see in a later chapter that the same formula holds for all a > 1.

Successive application of the recursion formula gives

$$\Gamma(a) = (a - 1)(a - 2)\Gamma(a - 2)$$

$$= \ldots$$

$$= (a - 1)(a - 2) \ldots 2\cdot 1$$

$$= (a - 1)!$$

Thus the gamma function is an extension to real values of x of the factorial function defined only for integer values.

The effect of a substitution on a definite integral may be to change it from proper to improper or vice versa. A simple example is

$$\int_0^1 \frac{x}{\sqrt{1 - x^2}}\, dx = \int_0^{\pi/2} \sin\theta \, d\theta,$$

where $x = \sin\theta$. The first integral is improper and the second is not. What we are actually doing here is the following

$$\lim_{h \to 1-} \int_0^h \frac{x}{\sqrt{1 - x^2}}\, dx = \lim_{\phi \to \pi/2-} \int_0^{\phi} \sin\theta \, d\theta$$

$$= \int_0^{\pi/2} \sin\theta \, d\theta.$$

The last step is valid because $\int_0^{\phi} \sin\theta \, d\theta$, as a function of ϕ, is continuous on the closed interval $[0,\pi/2]$ by virtue of the argument of Section 8-2, since $\sin\theta$ is unicon on $[0,\pi/2]$.

The substitution in the other direction, $\theta = \arcsin x$, is a little more subtle, since, starting with a proper integral, we may not notice that it has become improper. The critical factor is the differential $d\theta = dx/\sqrt{1 - x^2} = dx/\cos\theta$, which is not defined at $x = 1$ or $\theta = \pi/2$. Whenever this occurs the possibility of an improper integral must be considered.

For another example of what can happen consider

$$\int_0^\infty xe^{-x}dx = \int_0^1 -\log y \, dy$$

under the substitution $y = e^{-x}$, $x = -\log y$. Here an improper integral of one type is changed into one of the other type. We leave to the reader a careful analysis of just what is involved in this change.

It should now be clear that the troubles we had in Examples 7-2,3,4 arose from converting the proper integral

$$\int \frac{1}{2 + \sin\theta} d\theta$$

into an improper integral by the substitution $t = \tan\frac{\theta}{2}$. This is one of the drawbacks of this substitution.

A good set of tables contains many definite integrals, proper and improper, that cannot be evaluated by the methods of elementary calculus. Even for some common complicated cases that can be so evaluated a listing of the integral can save a lot of work.

Problems

1. Evaluate the given improper integral in case it converges.

(a) $\int_0^3 \frac{x\,dx}{\sqrt{9 - x^2}}$

(b) $\int_0^1 \frac{dx}{(1 - x)^2}$

(c) $\int_{-\infty}^{\infty} \frac{dx}{1 + x^2}$

(d) $\int_0^2 \frac{\log x}{x} dx$

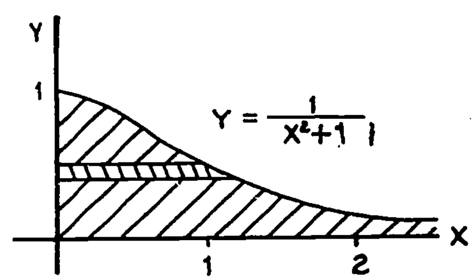
(e) $\int_0^{\infty} e^{-hx} \sin \omega x \, dx$

(f) $\int_{-1}^{1} \frac{dx}{x}$

(g) $\int_{-2}^{0} \frac{dx}{\sqrt[3]{x + 1}}$

(h) $\int_2^{\infty} \frac{dx}{x \log x}$

2. (a) If one attempts to find the adjacent cross-hatched area by horizontal strips he is led to the integral



$$\int_0^1 \sqrt{\frac{1 - y}{y}}\, dy = \int_0^1 \frac{1 - y}{\sqrt{y - y^2}}\, dy.$$

Evaluate this integral without using tables, indicating all the limiting processes involved.

(b) As a check on your answer find the area by using vertical strips.

3. (Refer to Problem 2 in Section 7)

(a) Find the volume of the solid obtained by rotating the region in Problem 2 above about the x-axis.

(b) About the y-axis.

4. Do the same as Problem 3 using the curve $y = e^{-ax}$ in place of $y = \frac{1}{x^2 + 1}$.

5. Is $\int_0^1 x \log x \, dx$ an improper integral? What is its value?

6. As stated in Problem 4 of Section 6-5 the weight of a body at height h is

$$w(h) = w(0) \frac{R^2}{(R + h)^2} ,$$

where w(0) is its weight at the surface, and R is the radius of the earth. How much work is done in lifting a one pound weight

(a) One mile,

(b) A thousand miles,

(c) To the moon,

(d) To infinity.

9. Numerical Methods.

The methods of the previous sections, extensive and useful as they are, do not always work, as was pointed out in Section I. Even when they work they are not always preferable.

Example I. To find the value of

$$J = \int_0^1 \frac{1}{x^3 + 5}\,dx$$

we can look up the indefinite integral in tables and substitute limits to get

$$(1) \quad J = \frac{1}{\sqrt[3]{25}}\left[\frac{1}{6}\log\frac{(\sqrt[3]{5}+1)^2}{\sqrt[3]{25}-\sqrt[3]{5}+1} + \frac{1}{\sqrt{3}}\arctan\frac{2-\sqrt[3]{5}}{\sqrt{3}\,\sqrt[3]{5}} - \frac{\pi}{6\sqrt{3}}\right].$$

This is the exact value of J but it isn't of much use as it stands. To use it to estimate J to 2D (2 decimal place) accuracy would be an annoying job, requiring the use of tables of cube roots, logs, and arctans, or some form of automatic computer. If we had started with the integral

$$K = \int_0^1 \frac{1}{\sqrt{x^3 + 5}}\,dx$$

none of our indefinite integral methods would have worked at all, for the indefinite integral of $(x^3 + 5)^{-1/2}$ is not an elementary function.

On the other hand, $f(x) = (x^3 + 5)^{-1}$ is easily shown to be a concave function ($f''(x) \leq 0$ on $[0,1]$) and so the method of Section 10-2 can be applied. Using $n = 2$ and the values of $f(x)$ from Table I (obtained with a slide rule), gives

x	f(x)
0	.2000
1/4	.1992
1/2	.1953
3/4	.1844
1	.1667

TABLE I.

$T = .1893$, $M = .1918$,

$M - T = .0025$.

So the approximation

$J \approx .1905 \approx .19$

is in error by less than .002. This is a much easier job of computation than the evaluation of (1), to say nothing of the derivation of (1) - by partial fractions - if our tables aren't handy.

Furthermore, we can evaluate K by simply taking square roots of the values of f(x) in Table I and repeating the simple computations of T and M.

Problems 5 and 6 illustrate other troubles that can arise in getting a numerical approximation to a definite integral through the use of an indefinite integral.

For direct numerical evaluation of a definite integral, "numerical quadrature" as it is commonly called, we cannot hope always to have a convex or concave integrand, or even to be able to divide the integral into pieces in each of which the function is convex or concave. What is needed is an error bound for the trapezoid rule or the midpoint rule that applies to all unicon functions, or at least to a very large class. The following theorem gives such a bound for the trapezoid rule. Its proof is given at the end of this section.

<u>Theorem 1.</u> Let $x_0 = a$, $x_n = b$, $x_i - x_{i-1} = h$, $i = 1, \ldots, n$. If f, f', and f'' are unicon on $[a,b]$ then

$$\int_a^b f(x)\,dx = h\left[\frac{1}{2}f(x_0) + f(x_1) + f(x_2) + \ldots + f(x_{n-1}) + \frac{1}{2}f(x_n)\right] + R_n,$$

where

$$R_n = -\frac{b-a}{12}h^2 f''(\xi), \qquad a \le \xi \le b.$$

<u>Example 2.</u> We wish to compute $\int_1^2 x^{-1}e^x\,dx$ accurate to 2D. We find

$$R_n = -\frac{1}{12}h^2 f''(\xi),$$

where

$$f''(\xi) = (\xi^2 - 2\xi + 2)\xi^{-3}e^{\xi}.$$

828

To find an upper bound on $|f''(\xi)|$ for $1 \leq \xi \leq 2$ a crude method is to take the maximum of each factor separately. This gives

$$|f''(\xi)| \leq 2 \times 1^{-3}e^{2} < 15.$$

For a little more finesse lump the rapidly changing factors, ξ^{-3} and e^{ξ}, together and find the maximum of $g(\xi) = \xi^{-3}e^{\xi}$. Since

$$g'(\xi) = (\xi - 3)\xi^{-4}e^{\xi}$$

is negative in [1,2] the maximum occurs at $\xi = 1$. This gives the better estimate

$$|f''(\xi)| \leq 2e < 5.4.$$

It is not worth the work involved to get a still lower bound by differentiating the whole function $f''(\xi)$.

Using this bound for $|f''(\xi)|$ we see that

$$|R_n| < .5\ h^2.$$

Hence h = .1, or n = 10, would give an error of < .005, insuring 2D accuracy. Computation of the trapezoid rule for this case gives 3.06066. Since the correct value of J to 8 places is 3.05911654, the actual error is .00154, showing that the computed bound is not a gross overestimate.

Even if the integrand is convex this method has the advantage of giving a bound for the value of n before any computation is done. This is important when high accuracy is desired, for high accuracy generally requires a large n, with a corresponding increase in roundoff error. We may find that n must be so large that the roundoff error would be greater than the required accuracy, in which case some more accurate method than the trapezoid rule would have to be used.

One such method is <u>Simpson's Rule</u>. Just as in deriving the trapezoid rule we approximated a portion of the graph by a line, i.e. a linear function $L(x) = a + bx$, so here we approximate a portion of the curve by a quadratic function $Q(x) = a + bx + cx^2$. Since Q involves three constants we can make it go through three points, P_0, P_1, P_2, on the graph, as in Figure 9-1. Let these points be $P_i\ (x_i, y_i)$, $i = 0,1,2$, where $y_i = f(x_i)$ and $x_1 - x_0 = x_2 - x_1 = h$.

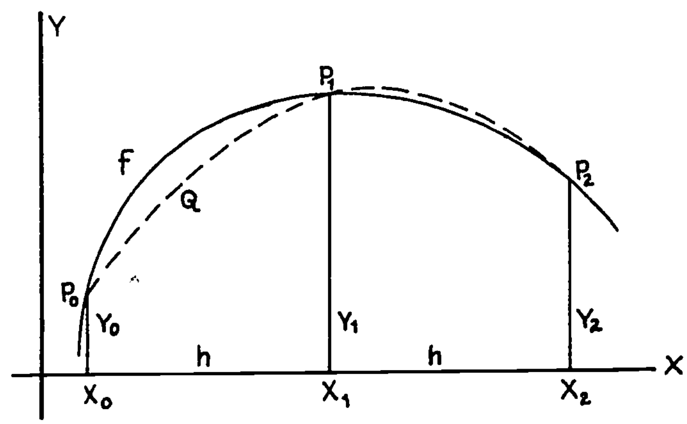


Figure 9-1

To get the area under the parabola it is helpful to change axes by the substitution $u = x - x_1$, to the case shown in Figure 9-2. In the new coordinates the equation of the parabola is still a quadratic, $y = \alpha u^2 + \beta u + \gamma$, so the area under the parabola is

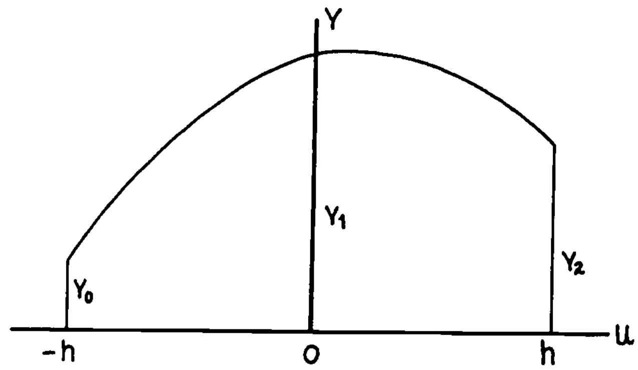


Figure 9-2

$$A = \int_{-h}^{h} (\alpha u^2 + \beta u + \gamma)du = \frac{2}{3}\alpha h^3 + 2\gamma h.$$

Now the parabola goes through the three points $(-h, y_0)$, $(0, y_1)$, (h, y_2), so we must have

$$\alpha h^2 - \beta h + \gamma = y_0,$$

$$\gamma = y_1,$$

$$\alpha h^2 + \beta h + \gamma = y_2.$$

Adding the first and third of these gives

$$2\alpha h^2 + 2\gamma = y_0 + y_2,$$

and it is now easy to see that

$$A = \frac{h}{3}(y_0 + y_2) + \frac{4h}{3} y_1 = \frac{h}{3}(y_0 + 4y_1 + y_2).$$

Referring back to our original curve this gives us the approximation,

$$\int_{x_0}^{x_2} f(x)\,dx \approx \frac{h}{3}\left[f(x_0) + 4\, f(x_1) + f(x_2)\right] .$$

To get <u>Simpson's Rule</u> we successively increase the subscripts by 2 and add the results, getting

$$(2) \qquad \int_a^b f(x)\,dx \approx \frac{h}{3}\left[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \ldots + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)\right] .$$

Of course n must be even to enable us to pair off the subintervals.

One can derive a bound for the error in Simpson's Rule, analogous to the one for the trapezoid rule given in Theorem 1. The derivation is considerably more complicated than the proof of Theorem 1 given in this section. We therefore merely state that if R_n is the difference of the left and right sides of (2) then

$$R_n = -\frac{b-a}{180} h^4 f^{(4)}(\xi), \qquad a \leq \xi \leq b,$$

provided $f, f', \ldots, f^{(4)}$ are all unicon on $[a,b]$.

To compare the accuracy of Simpson's Rule and the trapezoid rule apply this formula to Example 2. We find that

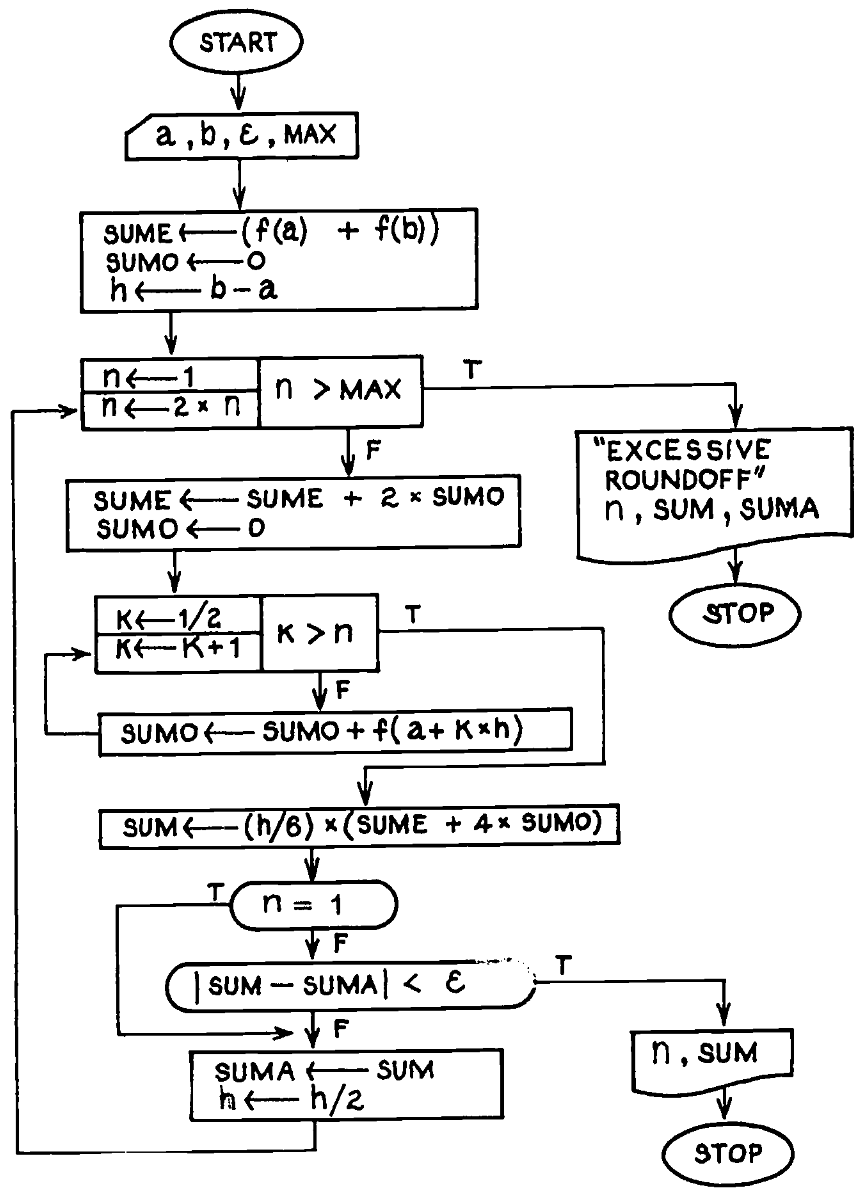
$$f^{(4)}(\xi) = (\xi^4 - 4\xi^3 + 12\xi^2 - 24\xi + 24)\xi^{-5}e^{\xi} .$$

As before, $\xi^{-5}e^{\xi}$ has maximum value e at $\xi = 1$, and the polynomial has maximum value 9, also at $\xi = 1$. (This is a little problem in finding extrema). So, for $h = .1$,

$$|R_n| \leq \frac{1}{180}(.1)^4 9e < .000014,$$

a gain in accuracy by a factor of 350. The Simpson Rule approximation to J computes to 3.0591200. The actual error is thus .0000035, again comparable to the computed bound.

It must be admitted, however, that in many cases f(x) is so complicated that it is well nigh impossible for one to compute its fourth derivative and get a bound for the absolute value. (See Example 3 in Section 12-4 for instance). In such cases one is usually satisfied to get several Simpson Rule approximations, successively doubling the number of subdivision, until two are obtained that differ by less than the permissible error. A flow chart for a program of this kind is shown in Figure 9-3.



START
a, b, ε, MAX
SUME ⟵ (f(a) + f(b))
SUMO ⟵ 0
h ⟵ b − a
n ⟵ 1
n ⟵ 2 × n
n > MAX
T
F
"EXCESSIVE ROUNDOFF"
n, SUM, SUMA
STOP
SUME ⟵ SUME + 2 × SUMO
SUMO ⟵ 0
K ⟵ 1/2
K ⟵ K+1
K > n
T
F
SUMO ⟵ SUMO + f(a+ K×h)
SUM ⟵ (h/6) × (SUME + 4 × SUMO)
n = 1
T
F
|SUM − SUMA| < ε
T
F
n, SUM
STOP
SUMA ⟵ SUM
h ⟵ h/2

Figure 9-3

By summing separately on the even and the odd points we can perform an efficient process of doubling the number n of intervals without unnecessary computation. MAX is a limit on the number of subintervals, imposed by round-off accumulation, as in Figure 2-8 in Chapter 10.

The numerical evaluation of improper integrals can sometimes be done fairly directly, but for efficient computation one usually looks for ingenious ways of simplifying the work. Some of the possible attacks are illustrated in the following example.

<u>Example 3</u>. $J = \int_1^\infty x^{-1}e^{-x}\,dx.$

(a) Easiest method. The function defined by $E_1(t) = \int_t^\infty x^{-1}e^{-x}\,dx$ is known as an "exponential integral" and is tabulated. $J = E_1(1) \approx .219383934.$

(b) Direct method.

$$\int_1^\infty x^{-1}e^{-x}dx = \int_1^M x^{-1}e^{-x}dx + \int_M^\infty x^{-1}e^{-x}dx,$$

and

$$\int_M^\infty x^{-1}e^{-x}dx < \int_M^\infty M^{-1}e^{-x}dx = M^{-1}(-e^{-x})\Big|_M^\infty = M^{-1}e^{-M}.$$

If we want to approximate J with error $< \varepsilon$ we can choose M so that $M^{-1}e^{-M} < \varepsilon/2$ and then take a numerical approximation of $\int_1^M x^{-1}e^{-x}dx$ of error $< \varepsilon/2$. For $\varepsilon = 5 \times 10^{-3}$, $M = 5$ will do; for $\varepsilon = 5 \times 10^{-5}$ we need $M = 9$.

(c) Transform into a proper integral. The substitution $y = e^{-x}$, $x = -\log y$ changes J into

$$- \int_0^{e^{-1}} \frac{1}{\log y}\, dy.$$

If we define $1/\log y$ to be zero when $y = 0$ then the function is unicon in the interval and this is a proper integral. There will be trouble in integrating it by either the trapezoid rule or Simpson's Rule, however, since even the first derivative becomes infinite as $x \to 0+$. This does not mean that these rules will not work but only that they will converge slowly as n increases. Taking this into account, method (b) is probably preferable.

(d) Integration by parts. In addition to its use to find indefinite integrals, integration by parts is a fine tool for changing integrals into hopefully more convenient forms. Apply it to $\int_M^\infty x^{-1}e^{-x}dx$, with

$$u = x^{-1}, \qquad dv = e^{-x}dx,$$
$$du = -x^{-2}dx, \qquad v = -e^{-x}.$$

We get

$$\int_M^\infty x^{-1}e^{-x}dx = -x^{-1}e^{-x}\Big|_M^\infty - \int_M^\infty x^{-2}e^{-x}dx$$

$$= M^{-1}e^{-M} - \int_M^\infty x^{-2}e^{-x}dx.$$

Applying the same upper bound argument as in (b) gives

$$\int_M^\infty x^{-2}e^{-x}dx < \int_M^\infty M^{-2}e^{-x}dx = M^{-2}e^{-M},$$

a smaller error than before. We could now, for instance, get the error less than 2.5×10^{-3} with $M = 4$ instead of $M = 5$.

The study of improper integrals is full of various tricks of this kind. An expert at integration has a large bag of them and a good intuition as to which ones to try on any given problem.

Proof of Theorem 1.

The following proof illustrates another interesting application of integration by parts.

Integrate $\int_c^d f(x)dx$ by parts, using

$$u = f(x), \qquad dv = dx,$$

$$du = f'(x)dx, \qquad v = x + p,$$

the constant p to be determined later. This gives

$$\int_c^d f(x)dx = (x+p)f(x)\Big|_c^d - \int_c^d (x+p)f'(x)dx.$$

Now repeat the process, with

$$u = f'(x), \qquad dv = x + p,$$

$$du = f''(x)dx, \qquad v = \frac{1}{2}x^2 + px + q,$$

giving

$$(3) \qquad \int_c^d f(x)dx = (x+p)f(x)\Big|_c^d - (\frac{1}{2}x^2 + px + q)f'(x)\Big|_c^d + \int_c^d (\frac{1}{2}x^2 + px + q)f''(x)$$

We wish to choose p and q so that the terms involving f' drop out. This will be so if $\frac{1}{2}x^2 + px + q$ is zero for $x = c$ and $x = d$; that is, if

$$\frac{1}{2}x^2 + px + q = \frac{1}{2}(x - c)(x - d).$$

This gives $p = -\frac{1}{2}(c + d)$, $q = \frac{1}{2}cd$, and (3) reduces to

$$(4) \qquad \int_c^d f(x)\,dx = (d - c)\left[\frac{1}{2}f(c) + \frac{1}{2}f(d)\right] - \frac{1}{2}\int_c^d (x - c)(d - x)f''(x)\,dx.$$

(The factors in the last integral are written this way to make them both positive). The form of the function $(x - c)(d - x)$ is shown in Figure 9-4. We find that

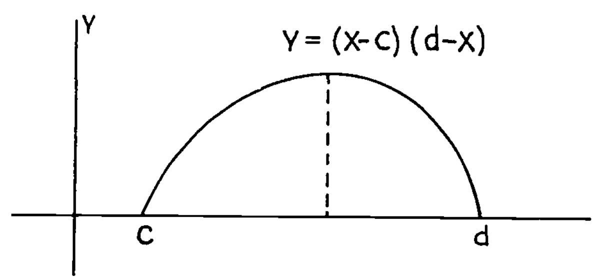


Figure 9-4

$$(5) \qquad \int_c^d (x - c)(d - x)\,dx = \frac{(d - c)^3}{6},$$

Now apply (4) successively to the cases $c = x_0$, $d = x_1$; $c = x_1$, $d = x_2$; ...; $c = x_{n-1}$, $d = x_n$; and add the results.

We get

$$\int_a^b f(x)\,dx = h\left[\frac{1}{2}f(x_0) + f(x_1) + \ldots + f(x_{n-1}) + \frac{1}{2}f(x_n)\right] + R_n,$$

where

$$R_n = -\frac{1}{2}\sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} (x - x_{i-1})(x_i - x)f''(x)\,dx.$$

To write R_n as one integral define a function Q on [a,b] by

$$Q_n(x) = (x - x_{i-1})(x_i - x)$$

if $x_{i-1} \leq x \leq x_i$.

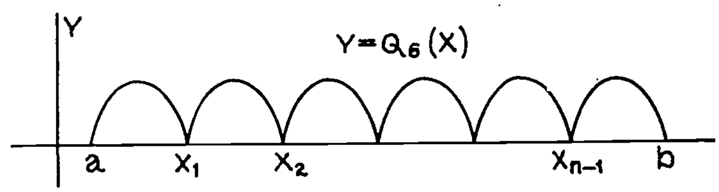


Figure 9-5

The graph of Q_6 is shown in Figure 9-5. Then we can write

$$R_n = -\frac{1}{2}\int_a^b Q_n(x)f''(x)\,dx.$$

For the next step we need the Mean Value Theorem for Integrals, which will be proved in the next chapter. It tells us that since $Q_n(x)$ is always ≥ 0, and f" is continuous,

$$\int_a^b Q_n(x)f''(x)\,dx = f''(\xi)\int_a^b Q_n(x)\,dx$$

for some ξ in [a,b].

Now

$$\int_a^b Q_n(x)dx = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} (x - x_{i-1})(x_i - x)dx$$

$$= n \frac{h^3}{6},$$

from (5), and $nh = b - a$. This gives us the required result,

$$R_n = -\frac{b - a}{12} h^2 f''(\xi), \qquad a \leq \xi \leq b.$$

Problems

1. (a) Write a flow chart to compute the trapezoid rule approximation to a given integral when the number n of subdivisions is specified.

   (b) Write a program from your flow chart.

2. For the function $f(x) = \dfrac{1}{1 + x^2}$:

   (a) Show that $|f''(x)|$ has its maximum value in $[0,1]$ at $x = 0$.

   (b) Find n so that for $h = 1/n$, $\frac{1}{12}h^2 \max |f''(\xi)| < 10^{-5}$ for ξ in $[0,1]$.

   (c) Using this value of n and your program from Problem 1, compute π to 4D.

3. (a) Write a flow chart for the Simpson Rule approximation when the number of subdivisions is specified.

   (b) Write a program from the flow chart.

   (c) Use your program to find

$$\log 10 = \int_1^{10} \frac{1}{x}\,dx$$

correct to 5D.

4. (a) Make a program from the flow chart in Figure 9-3, or some modification of it.

(b) Use your program to compute $\int_1^9 x^{-1}e^{-x}dx$ with error $< 3 \times 10^{-6}$. From this and the results of Example 3(b) compute $E_1(1)$ and compare with the value given in (a).

5. (a) As in Example 7-5, show that

$$\int_0^1 x^{10}e^{-x}\,dx = 3628800 - 9864101e^{-1}$$

What is the value of the integral correct to 2D?

(b) How would you find $\int_0^{.8} x^{10}e^{-x}dx$ correct to 2D using only a 5D table of e^x?

6. Consider the integral

$$J = \int \frac{1}{x^2 - 2.4682x + 1.5230}\,dx$$

where the coefficient of x is known to be exact but the constant term has been rounded off from 8D.

Show that:

(a) If 1.5230 is also exact then

$$J = \frac{1000}{2\sqrt{2.81}} \log \left| \frac{x - 1.2341 - \sqrt{2.81} \times 10^{-3}}{x - 1.2341 + \sqrt{2.81} \times 10^{-3}} \right|$$

(b) If 1.5230 was rounded off from 1.52300281, then

$$J = - \frac{1}{x - 1.2341}$$

(c) If 1.5230 was rounded off from 1.52300450, then

$$J = \frac{10000}{13} \arctan \frac{x - 1.2341}{.0013}$$

(d) How would you approximate $J \Big|_0^1$ to 3D?

7. Problem: To compute $\Gamma(\frac{3}{2}) = \int_0^\infty t^{1/2} e^{-t}\, dt$ to 5D accuracy. Because of the square root, leading to derivatives that become infinite as $t \to 0+$, direct application of the trapezoid or Simpson Rule is inadvisable.

(a) By a simple substitution reduce the integral to

$$\int_0^\infty 2u^2 e^{-u^2} du.$$

(b) By integrating by parts, show that

$$\int_0^\infty 2u^2 e^{-u^2} du = \int_0^\infty e^{-u^2} du.$$

(c) Using

$$\int_M^\infty e^{-u^2} du < \int_M^\infty \frac{u}{M} e^{-u^2} du,$$

show, by evaluating the latter integral, that

$$\int_{3.5}^\infty e^{-u^2} du < 10^{-6}.$$

(d) Evaluate $\int_0^{3.5} e^{-u^2} du$ numerically with error $< 4 \times 10^{-6}$ and so determine $\Gamma(\frac{3}{2})$ with error $< 5 \times 10^{-6}$.

8. Pro[illegible] 3 and 4 of Section 10-2 suggest that a good approximation to an integral might be $\frac{1}{3}(T + 2M)$ where T and M are the trapezoid and midpoint approximations. Show that this is just Simpson's Rule.

9. Prove that Simpson's Rule is exact if f is any polynomial of degree at most 3.

Appendix A

ALGORITHMIC TREATMENT OF PARTIAL FRACTIONS

The application of Theorem I of Section 5 to a given rational function $P(x)/Q(x)$ requires that $Q(x)$ be factored into powers of linear and quadratic factors. Can this always be done, and, if so, how? The answer to the first question is "yes", by virtue of the Fundamental Theorem of Algebra which says that every polynomial equation has a root (possibly complex). The proof of this theorem is a topic for more advanced mathematics. Assuming its truth, one can show that a factorization of the form in Theorem I is always possible, the linear factors arising from the real roots of $Q(x) = 0$ and the quadratic factors from the complex roots.

The second question, "How do we carry out the factoring?", has several answers. The most direct method is to find all the roots of $Q(x) = 0$, with their appropriate multiplicities. The real roots can be determined by Newton's Method, or a modification of it to take care of the multiplicity of the root if any exponent $p, q, \ldots$ is > 1. There are other similar methods for finding the complex roots, and hence the quadratic factors, but even a simple case like $Q(x) = x^6 - 3x^5 + 20$

which has two linear and two quadratic factors, is a major job without an automatic computer. Fortunately the factorization appears automatically in many problems: for instance the method of Section 3-IV applied to

$$\int \frac{\cos 2x}{2 + \sin x + \cos x}\, dx$$

gives directly

$$\int \frac{2t^4 - 12t^2 + 2}{(t^2 + 2t + 3)(t^2 + 1)^2}\, dt.$$

Let us assume, then, that this numerical work has already been carried out and $Q(x)$ is factored. Notice that the conditions of Theorem I do not require that the quadratic factors have complex roots but only that any pair of indicated factors are relatively prime. For example, $(x^2 - 1)^3$ could be taken as one of the parts of the denominator provided neither $x - 1$ nor $x + 1$ was a factor of any other part. Or one could carry $x^2 + 3x - 5$ along as a single factor instead of breaking it into $x + 3/2 + \sqrt{29}/2$ and $x + 3/2 - \sqrt{29}/2$. This is often convenient.

We give the details of an algorithm that simultaneously proves Theorem I and computes the constants A,B,C, etc. The algorithm applies to a P of any degree, but the computations are simpler if the degree is first lowered to less than that of Q. There are two kinds of steps in the algorithm, corresponding to the linear and the quadratic factors.

<u>Step 1</u>. If $x + a$ is not a factor of $Q_0(x)$ then for a given $P(x)$ there is a unique constant A and a unique polynomial $R(x)$ such that

$$(1) \qquad \frac{P(x)}{(x + a)^p Q_0(x)} = \frac{A}{(x + a)^p} + \frac{R(x)}{(x + a)^{p-1} Q_0(x)} .$$

<u>Proof</u>. Clearing fractions, (1) becomes

$$(2) \qquad P(x) = AQ_0(x) + (x + a)R(x).$$

Divide $Q_0(x)$ and $P(x)$ by $x + a$, to get

$$(3) \qquad Q_0(x) = (x + a)S(x) + m,$$

$$(4) \qquad P(x) = (x + a)T(x) + n.$$

Substituting in (2) and collecting terms having $x + a$ as a factor gives

$$(5) \qquad (x + a)[T(x) - AS(x) - R(x)] = Am - n.$$

Now the right-hand side is a constant and the left-hand side is divisible by $x + a$. Equality is possible if and only if both sides are zero. Hence (2), and therefore (1), holds if and only if,

(6) $A = n/m, \quad R(x) = T(x) - AS(x).$

Division by m is possible because $x + a$ is not a factor of $Q_0(x)$; this implies that $m \neq 0$.

Step I involves three major computational steps, the divisions by $x + a$ in (3) and (4) and the forming of $R(x)$ in (5). The divisions can be done by the following simple algorithm (essentially the same as "synthetic division"). If

$$H(x) = b_0x^n + b_1x^{n-1} + \dots + b_n,$$

and

$$K(x) = c_0x^{n-1} + c_1x^{n-2} + \dots + c_{n-1}.$$

Then

$$H(x) = (x + a)K(x) + r$$

becomes

$$c_0 = b_0,$$

$$c_i + ac_{i-1} = b_i,$$

or

$$c_i = b_i - ac_{i-1}, \qquad i = 1, \dots, n-1,$$

$$r = b_n - ac_{n-1}.$$

This is the simple loop shown in Figure A-1.

We leave to the reader the construction of a flow chart for the computation of $R(x) = T(x) - AS(x)$.

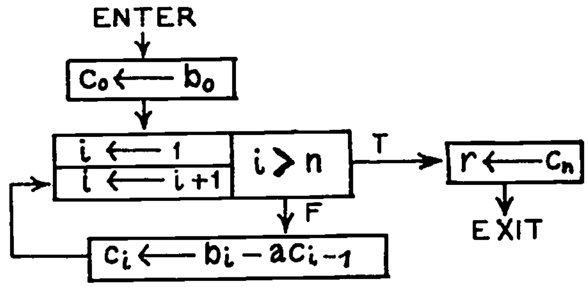


Figure A-1

Having carried out Step 1, if $p-1 > 0$ we proceed to treat $\frac{R(x)}{(x+a)^{p-1}Q_0(x)}$ in the same fashion. That is, we let $P(x) \leftarrow R(x)$, $p \leftarrow p-1$, and apply Step 1 again. Note, however, that we do not have to recompute $S(x)$ and m. This gives the flow chart in Figure A-2.

Finally, consider the true situation, where the denominator has the form

$$(x + a_1)^{p_1}(x + a_2)^{p_2} \ldots (x + a_m)^{p_m}Q^*(x),$$

where $Q^*(x)$ is the product of the quadratic factors in Theorem 1.

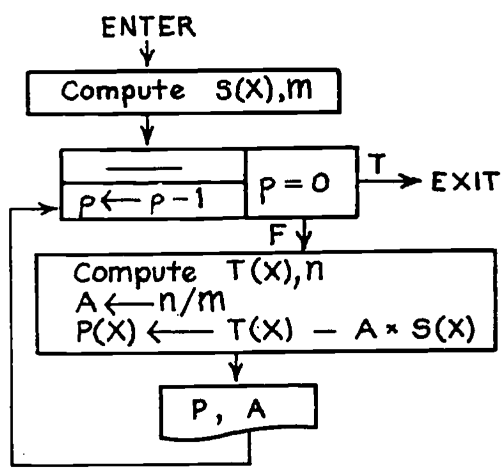


Figure A-2

To handle this we start with

$$a = a_1, \quad p = p_1, \quad Q_0(x) = (x + a_2)^{p_2} \ldots (x + a_m)^{p_m} Q^*(x)$$

and apply Figure A-2. When all the factors $x + a_l$ are removed we let

$$a = a_2, \quad p = p_2, \quad Q_0(x) = (x + a_3)^{p_3} \ldots (x + a_m)^{p_m} Q^*(x),$$

and proceed. This gives us Figure A-3. Of course all the polynomial algebra in these flow charts must in turn be broken down into algorithms on the coefficients, as we did for division by x + a in Figure A-1.

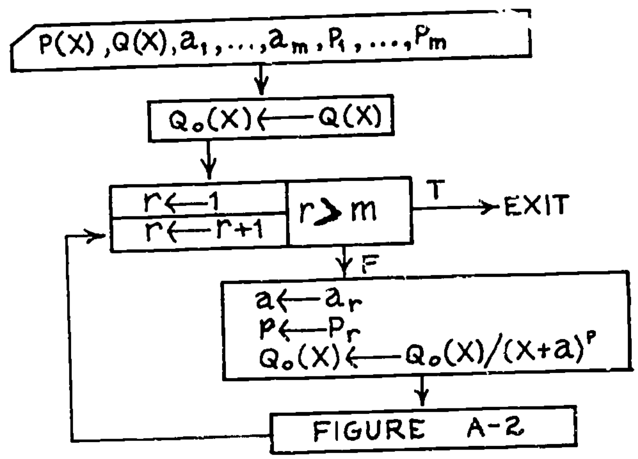


Figure A-3

<u>Example 1.</u>

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

It is better to work on the most complicated factor first, so as to make $Q_0(x)$ as simple as possible. So we take

1. $a = 1, \quad p = 3.$

$$Q_0(x) = x^3 - x^2 + x - 1 = (x + 1)(x^2 - 2x + 3) - 4;$$

$$S(x) = x^2 - 2x + 3, \qquad m = -4.$$

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

$$T(x) = x^4 - x^3 + x^2 - x + 1, \qquad n = -1.$$

$$A = (-1)(-4) = \frac{1}{4}, \quad R(x) = x^4 - x^3 + x^2 - x + 1 - \frac{1}{4}(x^2 - 2x + 3)$$

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

2. $$P(x) = x^4 - x^3 + \frac{3}{4}x^2 - \frac{1}{2}x + \frac{1}{4} = (x + 1)(x^3 - 2x^2 + \frac{11}{4}x - \frac{13}{4}) + \frac{7}{2}$$

$$A = (7/2)/(-4) = -\frac{7}{8}, \quad R(x) = x^3 - 2x^2 + \frac{11}{4}x - \frac{13}{4}$$

$$+ \frac{7}{8}(x^2 - 2x + 3)$$

$$= x^3 - \frac{9}{8}x^2 + x - \frac{5}{8}.$$

3. $$P(x) = x^3 - \frac{9}{8}x^2 + x - \frac{5}{8} = (x + 1)(x^2 - \frac{17}{8}x + \frac{25}{8}) - \frac{15}{4}.$$

$$A = (-15/4)/(-4) = \frac{15}{16}, \quad R(x) = x^2 - \frac{17}{8}x + \frac{25}{8} - \frac{15}{16}(x^2 - 2x + 3)$$

$$= \frac{1}{16}(x^2 - 4x + 5).$$

4. $a = -1$, $p = 1$.

$$Q_0(x) = x^2 + 1 = (x - 1)(x + 1) + 2;$$

$$S(x) = x + 1, \qquad m = 2.$$

$$P(x) = \frac{1}{16}(x^2 - 4x + 5) = \frac{1}{16}(x - 1)(x - 3) + \frac{2}{16}$$

$$A = (2/16)/2 = \frac{1}{16}, \quad R(x) = \frac{1}{16}(x - 3) - \frac{1}{16}(x + 1) = -\frac{1}{4}.$$

So

$$\frac{x^5}{(x-1)(x+1)^3(x^2+1)} = \frac{1}{4}\frac{1}{(x+1)^3} - \frac{7}{8}\frac{1}{(x+1)^2} + \frac{15}{16}\frac{1}{x+1}$$

$$+ \frac{1}{16}\frac{1}{x-1} - \frac{1}{4}\frac{1}{x^2+1} .$$

Step 2. If $x^2 + ax + b$ has no factor in common with $Q_0(x)$ then for a given $P(x)$ there are unique constants A and B and a unique polynomial $R(x)$ such that

$$\frac{P(x)}{(x^2+ax+b)^p Q_0(x)} = \frac{Ax+B}{(x^2+ax+b)^p} + \frac{R(x)}{(x^2+ax+b)^{p-1}Q_0(x)} .$$

Proof. The method is essentially the same as in Step 1, the only complicating factors arising from our now having two constants, A and B, to find. As before, we start with

$$P(x) = (Ax + B)Q_0(x) + (x^2 + ax + b)R(x),$$

(7) $Q_0(x) = (x^2 + ax + b)S(x) + ex + f,$

(8) $P(x) = (x^2 + ax + b)T(x) + gx + h.$

From these, corresponding to (5), we get

$$(x^2 + ax + b)(T(x) - (Ax + B)S(x) - R(x))$$

$$= eAx^2 + (fA + eB - g)x + fB - h$$

$$= eA(x^2 + ax + b) + (fA + eB - g - aeA)x$$

$$+ fB - h - beA,$$

or

$$(x^2 + ax + b)(T(x) - (Ax + B)S(x) - eA - R(x))$$

$$= (fA - aeA + eB - g)x - beA + fB - h.$$

Now the same argument can be used as in Step 1. The left-hand side is divisible by the quadratic $x^2 + ax + b$ but the right-hand side is only linear. Hence both must be zero, giving

$$(f - ae)A + eB = g,$$

$$-beA + fB = h,$$

(9) $R(x) = T(x) - (Ax + B)S(x) - eA.$

The first two equations have the unique solution

(10) $A = (fg - eh)/\Delta, \quad B = (fh - aeh + beg)/\Delta,$

$$\Delta = f^2 - aef + be^2,$$

provided $\Delta \neq 0$. We have therefore only to prove that $\Delta \neq 0$ to finish our proof of Step 2.

Now, from (7), e and f cannot both be zero or $x^2 + ax + b$ would be a factor of Q_0. If $e = 0$ then $\Delta = f^2 \neq 0$. So suppose $e \neq 0$. Then Δ can be written as

$$\Delta = e^2 \left((-\frac{f}{e})^2 + a(-\frac{f}{e}) + b \right) .$$

If $\Delta = 0$ then $-f/e$ is a root of $x^2 + ax + b = 0$, and $x + f/e$ is a factor of $x^2 + ax + b$. But, again from (7), $x + f/e$ would then be a factor of $Q_0(x)$, since $ex + f = e(x + f/e)$, contrary to the assumption that $x^2 + ax + b$ and $Q_0(x)$ have no common factors. Hence in all cases $\Delta \neq 0$ and Step 2 is valid.

The major computing steps in Step 2 are similar to those in Step 1, namely the two divisions by $x^2 + ax + b$ in (7) and (8) and the multiplication by $Ax + B$ in (9). Algorithms for these are similar to the ones in Step 1 but involve two multiplications per step instead of one. Hence Step 2 takes about twice the arithmetical work of Step 1, plus a little more for the computation of A and B in (10). Since it takes care of a quadratic factor it is therefore about equivalent to Step 1 in the amount accomplished per given amount of computation.

Details of these algorithms and the corresponding flow charts for Step 2 are left to the reader.

Example 2. Taking the function from Example 5-4 we do the quadratic part first. Usually one reserves Step 2 until the function has been simplified as much as possible by Step 1.

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

1. $a = -1, \quad b = 1, \quad p = 2.$

   $Q_1(x) = x^2 + 2x + 1 = (x^2 - x + 1) \times 1 + 3x;$

   $e = 3, \quad f = 0, \quad S(x) = 1; \quad \Delta = 9.$

   $P(x) = x^5 + 2 = (x^2 - x + 1)(x^3 + x^2 - 1) - x + 3;$

   $g = -1, \quad h = 3, \quad T(x) = x^3 + x^2 - 1.$

   $A = (-9)/9 = -1, \quad B = 6/9 = 2/3.$

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

2. $P(x) = x^3 + x^2 + x + 4/3 = (x^2 - x + 1)(x + 2) + 2x - 2/3;$

   $g = 2, \quad h = -2/3, \quad T(x) = x + 2.$

   $A = 2/9, \quad B = 4/9.$

   $R(x) = x + 2 - \frac{2}{9}x - \frac{4}{9} - \frac{6}{9} = \frac{7}{9}x + \frac{8}{9}.$

3. We have now completed all Steps 2 and are left with

$$\frac{1}{9}\,\frac{7x + 8}{(x + 1)^2}.$$

It is easier to abandon the algorithm now and do this as

$$\frac{7(x + 1) + 1}{(x + 1)^2} = \frac{7}{x + 1} + \frac{1}{(x + 1)^2}.$$

We have, therefore,

$$\frac{x^5 + 2}{(x^3 + 1)^2} = \frac{1}{9}\left[\frac{-9x + 6}{(x^2 - x + 1)^2} + \frac{2x + 4}{x^2 - x + 1} + \frac{1}{(x + 1)^2} + \frac{7}{x + 1}\right].$$

We need just one thing to complete the method of partial fractions and that is a reduction formula for the integration of the terms with the multiple quadratic denominators. As is shown in Example 5-4 these can be reduced to the one critical case

$$J_n = \int \frac{1}{(x^2 + c)^n}\, dx, \qquad c \neq 0.$$

As usual, we use integration by parts, taking

$$u = (x^2 + c)^{-n}, \qquad dv = dx,$$

$$du = -2nx(x^2 + c)^{-n-1}, \qquad v = x.$$

This gives

$$J_n = x(x^2 + c)^{-n} + 2n\int x^2(x^2 + c)^{-n-1}\, dx$$

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

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

Hence

$$J_{n+1} = \frac{1}{2nc}\left[\frac{x}{(x^2 + c)^n} + (2n - 1)J_n\right] .$$

Replacing n by n-1 gives us what we want:

$$\int \frac{1}{(x^2 + c)^n}\, dx = \frac{1}{2(n - 1)c}\left[\frac{x}{(x^2 + c)^{n-1}}\right.$$

$$\left. + (2n - 3)\int \frac{1}{(x^2 + c)^{n-1}}\, dx\right] .$$

Chapter 12

APPLICATIONS OF INTEGRATION

1. Cylindrical Shells.

In earlier chapters definite integrals have been used to express various geometrical or physical quantities. Now that our technique in evaluating the integrals has vastly improved it is time to consider other and more complicated applications.

We start by using a new technique on an old problem, namely the volumes of revolution. In Chapter 4 these were computed by the "parallel slice" method and we now want to use the "cylindrical shell" method. The two methods will be compared in a simple case.

Example 1. Find the volume of the cone generated by rotating a right triangular region of base B and altitude H around its altitude as an axis. (Figure 1-1).

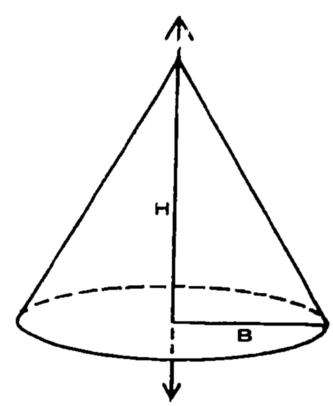


Figure 1-1

In the parallel slice method the volume was cut into pieces by planes perpendicular to the axis and upper and lower bounds were obtained for the volume of each slice. Considering one slice, (Figures 1-2 to 1-4) we have first, from Figure 1-3,

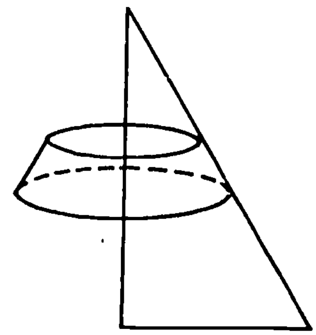


Figure 1-2

$$\frac{r_1}{B} = \frac{H - x}{H},$$

$$r_1 = B(1 - \frac{x}{H}).$$

Similarly,

$$r_2 = B(1 - \frac{x + \Delta x}{H}).$$

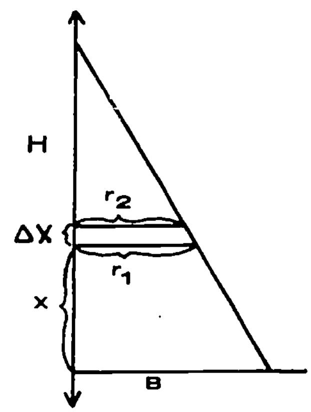


Figure 1-3

Evidently a cylinder of radius r_1 will completely contain the slice, and a cylinder of radius r_2 will be completely contained in the slice (Figure 1-4). The volume ΔV of the slice thus satisfies the bounds.

$$\pi B^2(1 - \frac{x + \Delta x}{H})^2 \Delta x \leq \Delta V \leq \pi B^2(1 - \frac{x}{H})^2 \Delta x.$$

More generally, if m and M are lower and upper bounds of the function

$$f(x) = \pi B^2(1 - \frac{x}{H})^2,$$

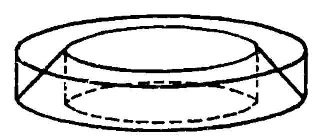


Figure 1-4

on the interval $[x, x + \Delta x]$ then

$$m\Delta x \leq \Delta V \leq M\Delta x.$$

Next we subdivide the interval $[0, H]$ into n parts with points $x_0 = 0, x_1, x_2, \ldots, x_n = H$ and consider cuts at the corresponding points along the altitude of the cone. If ΔV_i is the volume of the slice between x_{i-1} and x_i then

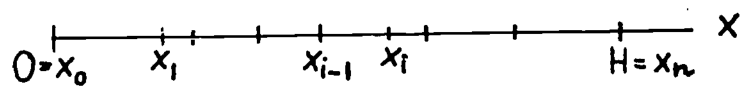


Figure 1-5

(1) $$m_i\Delta x_i \leq \Delta V_i \leq M_i\Delta x_i,$$

where $\Delta x_i = x_i - x_{i-1}$ and m_i and M_i are lower and upper bounds of the function $f(x) = \pi B^2(1 - \frac{x}{H})^2$ on the interval $[x_{i-1}, x_i]$.

From (1) we get

(2) $$\sum_{i=1}^{n} m_i\Delta x_i \leq V \leq \sum_{i=1}^{n} M_i\Delta x_i.$$

The theory in Chapter 3 then tells us that

$$V = \int_0^H f(x)dx$$

if f(x) is unicon on [0, H], and tells us also that f(x) <u>is</u> unicon. So, finally,

$$V = \int_0^H \pi B^2 (1 - \frac{x}{H})^2 dx$$

$$= -\frac{\pi B^2 H}{3}(1 - \frac{x}{H})^3 \Big|_0^H = \frac{1}{3}\pi B^2 H.$$

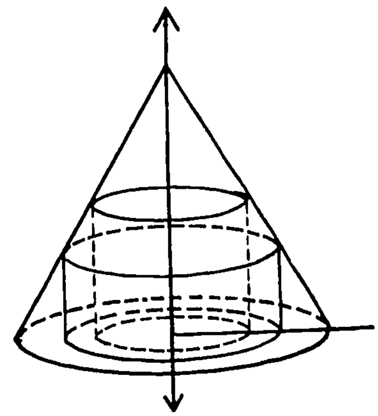


Figure 1-6

So much for review. For the cylindrical shell method we dissect the volume into pieces formed by rotating about the axis strips <u>parallel</u> to the axis. This gives us a solid like a piece of pipe with the top edge beveled outwards (Figure 1-6). To get upper and lower bounds for the volume ΔV of this shell we take cylindrical shells with flat tops containing it and contained in it (Figure 1-8). The smaller of

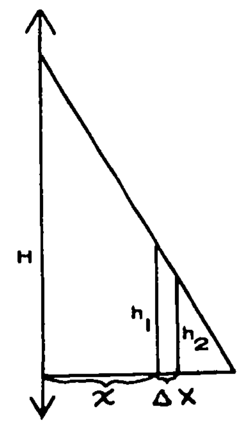


Figure 1-7

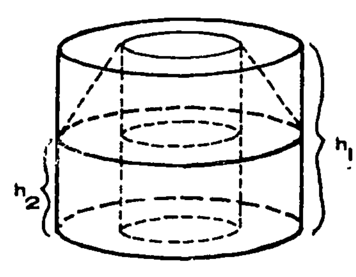


Figure 1-8

these is a cylinder of radius $x + \Delta x$ and height h_2, having a hole of radius x and height h_2. Its volume is therefore

$$\pi(x + \Delta x)^2 h_2 - \pi x^2 h_2 = 2\pi x h_2 \Delta x + \pi h_2 \Delta x^2$$

$$\geq 2\pi x h_2 \Delta x.$$

We take the lower bound in this form to avoid the term in Δx^2. Similarly, we find an upper bound

$$\pi(x + \Delta x)^2 h_1 - \pi x^2 h_1 = 2\pi x h_1 \Delta x + \pi h_1 \Delta x^2$$

$$= 2\pi(x + \Delta x) h_1 \Delta x - \pi h_1 \Delta x^2$$

$$\leq 2\pi(x + \Delta x) h_1 \Delta x.$$

From Figure 1-7 we get

$$\frac{h_1}{H} = \frac{B - x}{B},$$

$$h_1 = H(1 - \frac{x}{B}),$$

and similarly,

$$h_2 = H(1 - \frac{x + \Delta x}{B}).$$

So our bounds are

$$2\pi H x(1 - \frac{x + \Delta x}{B})\Delta x \leq \Delta V \leq 2\pi H(x + \Delta x)(1 - \frac{x}{B})\Delta x.$$

Here, however, we are stopped, for we cannot find a function f that behaves the way $\pi B^2(1 - \frac{x}{H})^2$ did in the previous case. That is, there is no function f such that if m and M are bounds for f(x) on [x, x + Δx] then necessarily $m\Delta x \le \Delta V \le M\Delta x$. The trouble is that the two factors x and $(1 - \frac{x}{B})$ reach their maxima and minima for different values of x. We must therefore consider not the bounds of

$$2\pi H x(1 - \frac{x}{B})$$

but those of

$$2\pi H x(1 - \frac{y}{B}),$$

where x and y vary independently in [x, x + Δx]. If m and M are such bounds then we have

$$m\Delta x \le \Delta V \le M\Delta x$$

As before, and again we can write the inequalities (2). Now, however, the theory in Chapter 3 no longer applies, since the m_i and M_i are bounds for a function of two variables. In the next section we shall show how the earlier theory can be modified to give us the expected result, that

$$V = \int_0^B 2\pi H x(1 - \frac{x}{B})\,dx$$

$$= 2\pi H \int_0^B (x - \frac{x^2}{B})\,dx$$

$$= 2\pi H(\frac{1}{2}x^2 - \frac{x^3}{3B})\Big|_0^B$$

$$= 2\pi H(\frac{1}{2}B^2 - \frac{1}{3}B^2) = \frac{1}{3}\pi B^2 H.$$

In any solid-of-revolution problem there are always these two methods that can be used. Consider the general set-up of Figure 1-9, where we wish to rotate the oval-shaped region about the axis AA. Strips perpendicular to the axis will generate the washer-shaped pieces of Figure 1-10, whose volumes are bounded by expressions of the form

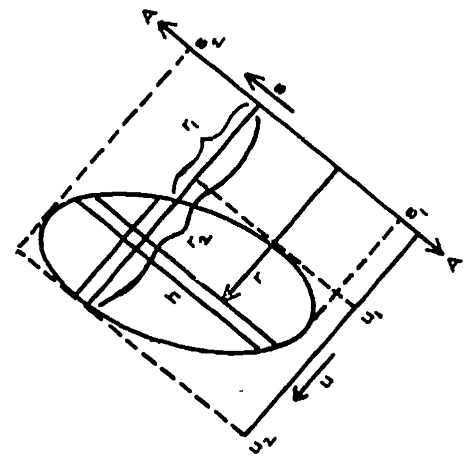


Figure 1-9

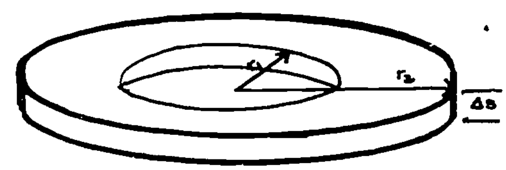


Figure 1-10

(3) $\pi(r_2^2 - r_1^2)\Delta s$,

leading to the integral

$$\int_{s_1}^{s_2} \pi[(r_2(s))^2 - r_1(s))^2]ds.$$

It is convenient to call expression (3) the "element of volume" and, as we just did, to by-pass the careful argument in Example 1 since it is evident that it is perfectly general.

Strips parallel to the axis of rotation generate pipe-shaped pieces (Figure 1-11) for which the element of volume is

$2\pi rh\Delta u.$

(This is most easily remembered by imagining the element cut along the section S and rolled out into a rectangular plate of dimensions $2\pi r \times h \times \Delta u$.) This gives the integral

$$\int_{u_1}^{u_2} 2\pi r(u)h(u)du$$

for the volume.

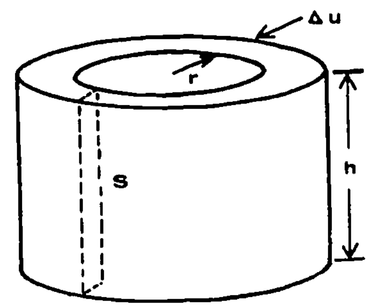


Figure 1-11

Note that in each case the independent variable, s or u, must be measured perpendicular to the strip, so that the width of the strip is an increment of that variable. Of course once the integral is set up we can make any substitutions we wish to evaluate it.

Example 2. The region bounded by the curve $y = x^2 - 1$ and the line $y = 2x + 2$ is rotated about the line $x = 4$. To get the volume of the resulting solid we obviously need to know the points of intersection of the two graphs, which are found, by solving the equations simultaneously, to be (-1, 0) and (3, 8).

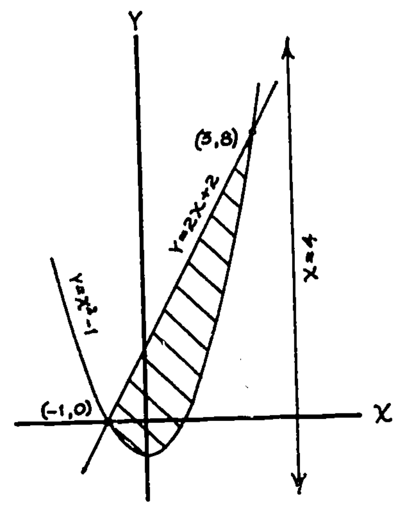


Figure 1-12

If we use the shell method our element of volume is

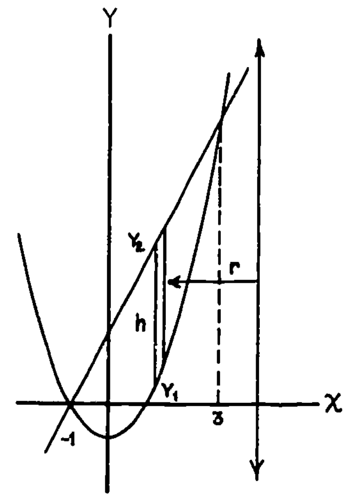


Figure 1-13

$$2\pi rh\Delta x = 2\pi(4 - x)(y_2 - y_1)\Delta x$$

$$= 2\pi(4 - x)[(2x + 2) - (x^2 - 1)]\Delta x.$$

Hence the volume is

$$\int_{-1}^{3} 2\pi(4 - x)(-x^2 + 2x + 3)dx$$

$$= 2\pi \int_{-1}^{3} (x^3 - 6x^2 + 5x + 12)dx$$

$$= 2\pi(\frac{1}{4}x^4 - 2x^3 + \frac{5}{2}x^2 + 12x)\Big|_{-1}^{3}$$

$$= 2\pi(\frac{81}{4} - 54 + \frac{45}{2} + 32 - \frac{1}{4} - 2 - \frac{5}{2} + 12)$$

$$= 64\pi.$$

The disc method is much clumsier for this problem. The element of volume is

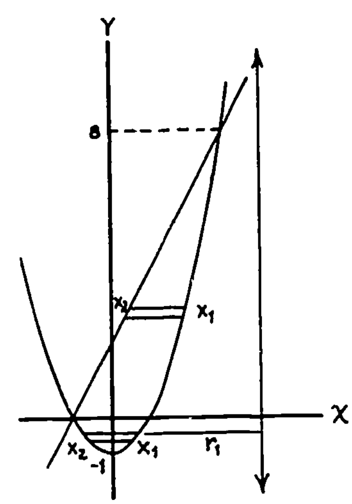


Figure 1-14

$$\pi(r_2{}^2 - r_1{}^2)\Delta y$$

$$= \pi[(4 - x_2)^2 - (4 - x_1)^2]\Delta y.$$

First of all, our independent variable is y, so we have to solve the bounding equations in terms of y. Secondly, we must do the problem in two parts, for the expression of x_2 as a function of y changes at y = 0. For $-1 \le y \le 0$ we have

$$x_1 = \sqrt{y+1}, \quad x_2 = -\sqrt{y+1},$$

and for $0 \le y \le 8$,

$$x_1 = \sqrt{y+1}, \quad x_2 = \frac{1}{2}y - 1.$$

Hence

$$V = \int_{-1}^{0} \pi[(4 + \sqrt{y+1})^2 - (4 - \sqrt{y+1})^2]\,dy$$

$$+ \int_{0}^{8} \pi[(4 - \frac{1}{2}y + 1)^2 - (4 - \sqrt{y+1})^2]\,dy.$$

The first integral reduces to

$$\pi \int_{-1}^{0} 16\sqrt{y+1}\,dy = \frac{32\pi}{3}(y+1)^{3/2}\Big|_{-1}^{0} = \frac{32}{3}\pi.$$

The second becomes

$$\pi \int_{0}^{8} (25 - 5y + \frac{1}{4}y^2 - 16 + 8\sqrt{y+1} - y - 1)\,dy$$

$$= \pi \int_{0}^{8} (8 - 6y + \frac{1}{4}y^2 + 8(y+1)^{1/2})\,dy$$

$$= \pi(8y - 3y^2 + \frac{1}{12}y^3 + \frac{16}{3}(y+1)^{3/2}\Big|_{0}^{8}$$

$$= \pi(64 - 192 + \frac{128}{3} + 144 - \frac{16}{3})$$

$$= \frac{160}{3}\pi.$$

So the volume is $\frac{32}{3}\pi + \frac{160}{3}\pi = 64\pi$, as before.

PROBLEMS

1. The region bounded by $y = x^3$, $y = 0$, and $x = 1$ is rotated about the y-axis. Find the volume of the solid generated.

2. The region the same as in Problem 1, but rotated about the x-axis.

3. The region bounded by $y = 0$ and the arch of $y = \sin x$ between 0 and π is rotated about the y-axis. Find the volume of the solid generated.

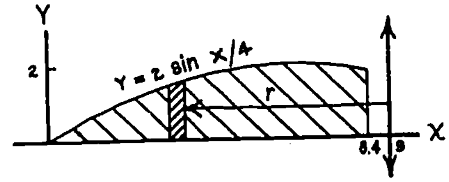
4. Find the volume of the solid generated when the region in Problem 3 is rotated about the line $x = -1$.

5. In the following problems the region bounded by the given pair of curves is rotated about the x-axis. Find the volume of the solid (a) by the shell method, and (b) by the disk method and show that the results are the same.

   (a) $y = 4x - x^2$ and $y = \frac{1}{2}x$

   (b) $y = \sqrt{x}$ and $y = x^3$

6. The base of a floor lamp is a piece of aluminum in the shape of the solid of revolution obtained by rotating about the line $x = 9$ the region bounded by the x-axis, the line $x = 8.4$, and the portion of the curve $y = 2 \sin x/4$ between 0 and 8.4. All measurements are in inches. If aluminum weighs .0975 lb/cu in. what is the weight of the lamp base?



Ans. 27.8 pounds.

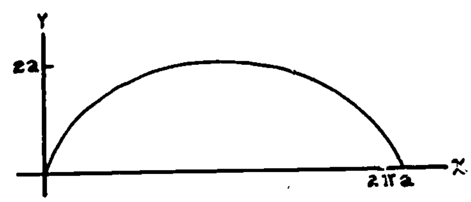
7. The region bounded by the x-axis and the first arch of the cycloid

$$x = a(\theta - \sin \theta),\ y = a(1 - \cos \theta)$$

is rotated

(a) about the x-axis,

(b) about the y-axis.



Find the volumes of the resulting solids.

Ans. $5\pi^2 a^3$, $6\pi^3 a^3$.

8. Find the volumes of the solids obtained by rotating the infinite region in the first quadrant bounded by the x-axis and the curve $y = xe^{-x}$

(a) about the x-axis,

(b) about the y-axis.

9. Consider the above problem for the region bounded by the x-axis, the line $x = 1$, and the curve $y = x^{-c}$. What conditions on c are needed in order that problems (a) and (b) have finite answers?

2. Generalization of the Integral.

To present the theory needed for the problems of Section I we must first introduce the notion of a function of two variables. Following the discussion in Chapter 0 we can define such a function F as a correspondence that assigns to each of certain pairs (x,y) of real numbers a unique real number z, and we write $z = F(x,y)$. Here and later, by the word "pair" we mean what is sometimes called an "ordered pair", i.e. the pair (1,2) is regarded as different from the pair (2,1).

Since a pair of real numbers represents a point in a plane, we can think of F as mapping a domain in the xy-plane into the real numbers, or the points of a line. The domain of F is then some set of points in the plane. Such a set of points can be extraordinarily complicated, and the samples given in Figure 2-1 are all very simple types. Fortunately such types will more than suffice for our present needs.

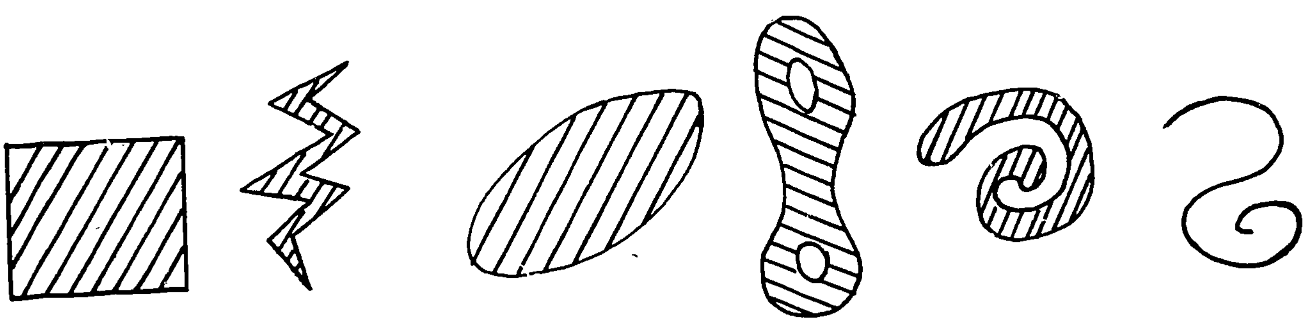


Figure 2-1

In talking about integrals it is not surprising that we want to consider unicon functions. The definition in Section 3-6 is easily extended, as follows;

A function of two variables defined in a domain D is unicon over D if for every $\varepsilon > 0$ there is a $\delta > 0$ such that

(1) $$|F(x_1,y_1) - F(x_2,y_2)| < \varepsilon$$

whenever (x_1,y_1) and (x_2,y_2) are in D and $|x_1 - x_2| < \delta$ and $|y_1 - y_2| < \delta$.

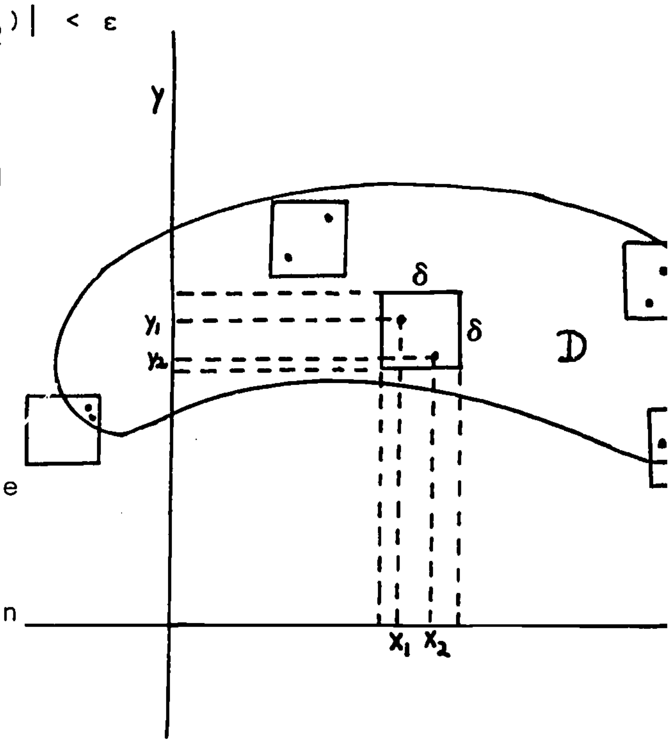


Figure 2-2

The picture in the xy-plane is illustrated by Figure 2-2. No matter where the $\delta \times \delta$ square is placed, two points in it and also in D must satisfy condition (1).

With this definition the whole theory of unicon functions, as given in Section 10 and Appendix A of Chapter 3, can be carried out for functions of two variables. (The case of composite functions is slightly more complicated but we shall have no use for this.)

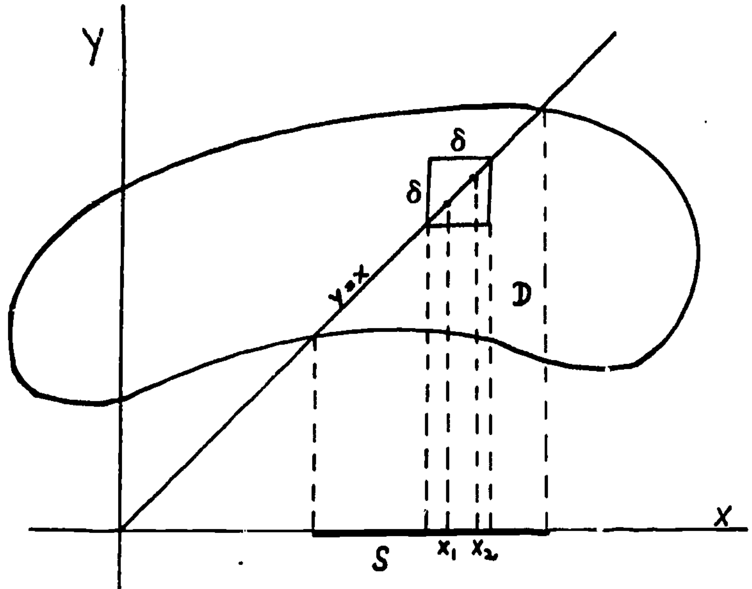


Figure 2-3

Theorem 1. If $F(x,y)$ is unicon over a domain D then $f(x) = F(x,x)$ is unicon over the domain which is the projection into the x-axis of the intersection of D and the line $y = x$. (Figure 2-3).

Proof. If $|x_1 - x_2| < \delta$ and we set $y_1 = x_1$, $y_2 = x_2$, then obivously $|y_1 - y_2| < \delta$, and so (1) is satisfied. That is

$$|f(x_1) - f(x_2)| =$$

$$|F(x_1,y_1) - F(x_2,y_2)| < \varepsilon,$$

which proves that f is unicon.

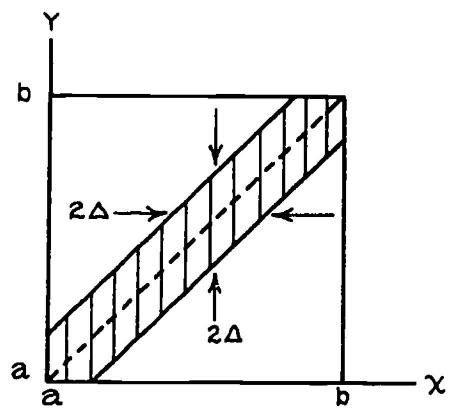


FIGURE 2-4

Figure 2-4

Now we are in a position to handle our integrals. Consider the domain D shown in Figure 2-4, defined by

$a \leq x \leq b,$

$a \leq y \leq b,$

$|x - y| \leq \Delta,$

for some positive number Δ. Let $F(x,y)$ be unicon in D. We subdivide the interval $[a,b]$ in the usual way with points

$a = x_0 < x_1 < x_2 < \dots < x_n = b,$

but such that $x_i - x_{i-1} = \Delta x_i < \Delta, \quad i = 1,2,\dots,n.$ Using these points we erect a covering of the diagonal segment from (a,a) to (b,b) by squares, as shown in Figure 2-5. On the square lying above the interval $[x_{i-1},x_i]$, $F(x,y)$, being unicon, is bounded below and above by some numbers m_i and M_i.

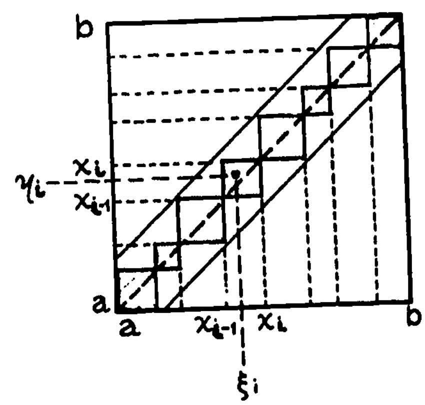


Figure 2-5

We can then construct lower and upper sums

$$L = \sum_{i=1}^{n} m_i \Delta x_i \quad \text{and} \quad U = \sum_{i=1}^{n} M_i \Delta x_i .$$

Corresponding to the definition and following discussion on page 235 we can prove a similar result.

Theorem 2. If $L_1, L_2, \ldots$ and $U_1, U_2, \ldots$ are sequences of lower and upper sums, as defined above, and if $\lim_{k\to\infty} |U_k - L_k| = 0$, then the two sequences have the common limit

$$\int_a^b F(x,x)\,dx.$$

Proof. All we have to show is that L_k and U_k are lower and upper sums of the function of one variable $F(x,x)$ and the result will follow from Appendix B of Chapter 3. But this is evident: since m_i, for instance, is less than $F(\xi_i, \eta_i)$ for all choices of ξ_i and η_i in $[x_{i-1}, x_i]$ it is certainly less than $F(\xi_i, \eta_i)$ for those particular choices for which $\eta_i = \xi_i$; that is, it is less than $F(\xi_i, \xi_i)$ for all choices of ξ_i in $[x_{i-1}, x_i]$.

We now need only one more result.

Theorem 3. If $F(x,y)$ is unicon then there exist sequences $L_1, L_2, \ldots, U_1, U_2, \ldots$ of upper and lower sums such that $\lim_{k\to\infty} |U_k - L_k| = 0$.

The proof follows, almost line for line, the similar proof for a function of one variable, on pages 261 to 265.

Returning for a moment to the cylindrical shells, we have the case where $F(x,y) = f(x)g(y)$, f and g being unicon functions. Now f(x) may be regarded as a function of x and y whose value varies only with x, and it is quite evident that if f is unicon when considered as a function of x only it is also unicon when considered as a function of x and y; similarly for g. F is then the product of two unicon functions and is therefore unicon. This fact and Theorems 3 and 2 above are all that is necessary to justify the expression of the volume as an integral.

As a by-product of the proof of Theorem 3, just as in the one-variable case in Chapter 3, we get the result that if $\delta(\varepsilon)$ is the unicon modulus of $F(x,y)$ and if our subdivision satisfies $\Delta_i < \delta(\varepsilon)$, $i = 1,\ldots,n$, then for any choice of ξ_i and η_i in $[x_{i-1},x_i]$ the sum

$$\sum_{i=1}^{n} F(\xi_i,\eta_i)\Delta x_i$$

approximates the integral $\int_a^b F(x,x)dx$ with error at most $(b - a)\varepsilon$. A corollary to this is the following, sometimes known as Duhamel's Theorem.

Theorem 4. Let $F(x,y)$ be unicon in a diagonal strip as in Figure 2-4, and let $\delta_1,\delta_2,\ldots$ be a sequence having limit 0. For each k, let

$$S_k = \sum_{i=1}^{n} F(\xi_i,\eta_i)\Delta x_i$$

in our usual notation, with $\Delta x_i < \delta_k$. n, ξ_i, η_i, and the x_i

will necessarily all vary with k. Then

$$\lim_{k\to\infty} S_k = \int_a^b F(x,x)dx.$$

For simplicity all this theory has been presented for the case of a function of two variables. Only trivial changes need be made, however, to generalize it to functions of 3,4, or N variables.

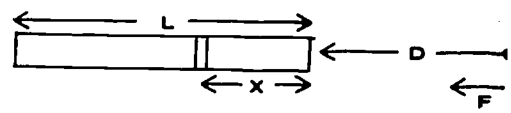
PROBLEMS

1. According to Newton's theory of gravitation two particles of masses m_1 and m_2 and a distance r apart attract each other with a force

$$F = \frac{Gm_1m_2}{r^2},$$

where G is a constant. To get the attraction of bodies that are not particles we must use integration.

(a) Consider a thin cylindrical rod of length L and of linear density (i.e. pounds per foot) $\rho(x)$ at a point x ft from one end. Let there be a particle of mass m on the axis of the rod at a distance D from this end. Show that the attractive force of the rod on the particle is

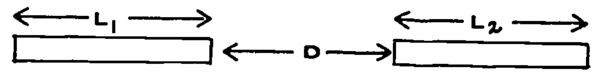


$$F = Gm \int_0^L \frac{\rho(x)}{(x + D)^2}\,dx.$$

(b) Find F for the case of constant density ρ. Is it the same as if the mass of the rod were concentrated at its center? If not, which is greater?

(c) If the rod in (b) were infinitely long would the force be infinite?

2. (a) If L_1 and L_2 are two rods of constant density ρ, arranged as in the figure, what is their mutual attractive force? [Hint. Use the result of Problem 1(b).]



(b) What happens if one of the rods becomes infinitely long? If both do?

3. Work.

At the end of Section 4-3 we met the concept of work done by a force acting in the direction of a displacement. There the only problems that were discussed involved objects that could be considered to be concentrated at a point, whose displacement could be specified by a single variable x. Now we wish to consider "bodies" that can be broken up, so that their different parts have different displacements.

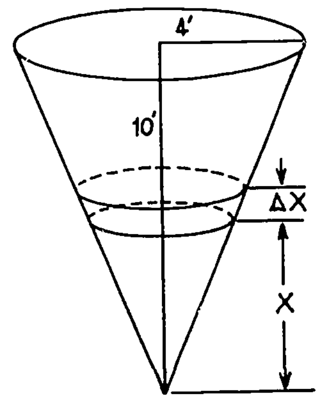


Figure 3-1

Example 1. An inverted cone of radius 4 ft and altitude 10 ft is full of water. How much work must be done to raise all the water to the level of the top of the cone? We imagine the water in the cone to be divided into a large number of thin horizontal layers and consider the work ΔW required to lift such a layer to the top. Let ΔV be the volume of the layer and ρ the density of the water in lbs/cu ft. To get

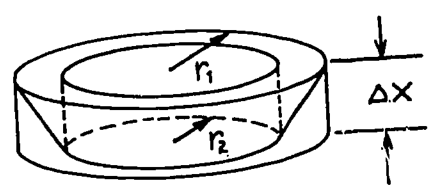


Figure 3-2

lower and upper bounds for ΔW we take cylinders contained in and containing the slice (Figure 3-2). Their volumes are respectively $\pi r_1{}^2 \Delta x$ and $\pi r_2{}^2 \Delta x$. From Figure 3-1 we have $x/10 = r/4$, so

$$r_1 = .4x, \qquad r_2 = .4(x + \Delta x),$$

and so

$$\pi(.4x)^2 \Delta x \leq \Delta V \leq \pi(.4(x + \Delta x))^2.$$

Individual molecules of the slice are lifted through distances varying from $10-(x + \Delta x)$ to $10-x$. If we assume them all taken from the top we get too little work done, a lower bound; whereas if we assume them all taken from the bottom we get an upper bound. Using this, and multiplying by ρ to change the volumes to weights, i.e. forces, we get $\pi(.4x)^2(10-(x + \Delta x))\rho\Delta x \leq \Delta W \leq \pi(.4(x + \Delta x))^2(10 - x)\rho\Delta x$. Thus we have an element of work of the type

$$\Delta W = \pi(.4x)^2(10 - x)\rho\Delta x,$$

and

$$W = \int_0^{10} \pi(.4x)^2(10 - x)\rho\, dx$$

$$= \frac{400}{3}\pi\ \rho \approx 26{,}000 \text{ ft-lbs},$$

since ρ is about 62.4 lbs/cu ft.

PROBLEMS

1. A vertical cylindrical tank 6 ft in diameter and 10 ft high is half full of water. Find the amount of work done in pumping the water out at the top of the tank.

2. A tank in the form of a cone has a base with a radius of 10 ft and is 20 ft high. If the axis is vertical and the vertex down, and if the tank is full of water, find the work required to pump the water to a point 15 ft above the top of the tank.

3. A ship is anchored, with the anchor 100 ft directly below the ship. The anchor weighs 3000 lbs and the anchor chain weighs 5 lbs/ft. How much work is done in bringing the anchor up?

4. (a) A cylindrical tank car, of length L and radius R, is full of oil of density ρ. How much work is required to empty it through a hole in the top? Express your answer in terms of the total weight of the oil.

   (b) Do the same problem for the case in which the car is only half full.

5. A water tank for a small community is a sphere 50 ft in diameter with its center 400 feet above the lake from which it is filled. How much work must be done to fill it?

4. Length of a Curve.

Having seen how to calculate areas and volumes of two and three dimensional figures we now take a step in the opposite direction and consider the length of a curve. First of all, what is a curve? We have used the term in an intuitive way in connection with graphs but we must now become more precise. Let us state a nice compact definition and then see what it means.

Definition. A plane curve is the range of a continuous mapping of a closed interval into the plane.

That is: If f and g are continuous functions on $[a,b]$ then the set of points (x,y) in the xy-plane defined by $x = f(t)$, $y = g(t)$ for all values of t in $[a,b]$ is said to constitute a plane curve. Since this is the only kind of curve we shall consider we shall omit the modifier "plane." (There are also space curves, and others.)

We are, in other words, defining a curve in terms of a parameter. Can this always be done for the things we intuitively think of as curves? Well, the graph of any continuous function f over a closed interval $[a,b]$ is a curve, with the parametrization

$$x = t, \quad y = f(t), \quad t \text{ in } [a,b].$$

Also it was suggested (but not proved) in the discussion of implicit functions that the graph of an equation of the form $F(x,y) = 0$ can be broken into pieces each of which is the graph of a continuous (implicit) function of x. Thus our definition of a curve covers all the cases we are likely to encounter.

In elementary geometry the length (circumference) of a circle is defined in terms of inscribed and circumscribed regular polygons. If p_n is the perimeter of an inscribed regular polygon of n sides, and P_n of the corresponding circumscribed polygon (Figure 4-1), then by arguments not unlike those we have used the ancient Greeks proved that the two sequences $p_3, p_4, \ldots$ and $P_3, P_4, \ldots$ have a common limit which is the circumference of the circle.

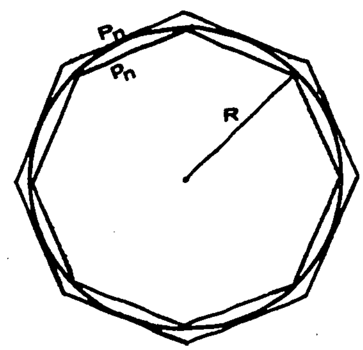


Figure 4-1

For our general curves the notion of inscribed and circumscribed regular polygons is of course inapplicable. We use the most easily applicable portion of the old process, the inscribed polygon.

Let

$t_0 = a,\ t_1, t_2, \ldots, t_n = b$

be the usual subdivision of the interval $[a,b]$ with $0 < t_i - t_{i-1} = \Delta t_i$. Corresponding to each value of i, from 0 to n, there is a point $P_i = (x(t_i),\ y(t_i))$

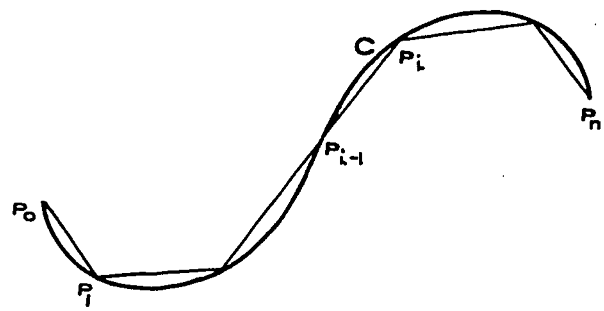


Figure 4-2

on the curve C. (For convenience we use $x(t)$, $y(t)$ as the notation for the parametrizing functions.) The set of line segments P_0P_1, P_1P_2, ..., $P_{n-1}P_n$ is called a <u>polygonal line</u> inscribed in C. The length of the polygonal line is just the sum of the lengths of the segments. Fairly obviously, if the Δt_i are very small the polygonal line will match the curve closely, and its length should be a good approximation to the length of the curve. This is the basis for the definition of length.

<u>Definition</u>. A number L is the <u>length</u> of C if for every $\varepsilon > 0$ there is a $\delta > 0$ such that every polygonal line inscribed in C for which all $\Delta t_i < \delta$ has a length which is an ε-approximation to L.

This is translated into symbols as follows; The length of the segment $P_{i-1}P_i$ is

$$\sqrt{(x(t_i) - x(t_{i-1}))^2 + (y(t_i) - y(t_{i-1}))^2}\ ,$$

and so the length of the inscribed polygonal line is the sum of these from $i = 1$ to n. The definition says that this sum should be close to L if the Δt_i are very small.

This begins to look like an integral in some respects. To actually reduce it to an integral we must make the assumption that the functions $x(t)$ and $y(t)$ are differentiable. Then the Mean Value Theorem tells us that

$$x(t_i) - x(t_{i-1}) = x'(\xi_i)\Delta t_i, \quad t_{i-1} < \xi_i < t_i,$$

$$y(t_i) - y(t_{i-1}) = y'(\eta_i)\Delta t_i, \quad t_{i-1} < \eta_i < t_i.$$

The length of the polygonal line can then be written as

$$\sum_{i=1}^{n} \sqrt{x'(\xi_i)^2 + y'(\eta_i)^2}\ \Delta t_i.$$

This is now in the form to apply Theorem 4 of Section 2. Assuming all the functions involved are unicon we get, finally,

$$(1) \qquad L = \int_a^b \sqrt{x'(t)^2 + y'(t)^2}\ dt.$$

Example 1. The length of the semicircle defined by

$$x = t, \quad y = \sqrt{r^2 - t^2}, \quad -r \le t \le r,$$

is

$$\int_{-r}^{r} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt$$

$$= \int_{-r}^{r} \sqrt{1 + \left(\frac{-t}{\sqrt{r^2 - t^2}}\right)^2}\, dt$$

$$= \int_{-r}^{r} \frac{r}{\sqrt{r^2 - t^2}}\, dt$$

$$= r \lim_{h \to -r+} \arcsin \frac{t}{r}\Big|_{h}^{0} + r \lim_{k \to r-} \arcsin \frac{t}{r}\Big|_{0}^{k}$$

$$= r(0 - (-\frac{\pi}{2})) + r(\frac{\pi}{2} - 0) = \pi r.$$

This is not a proof that the circumference of a circle is $2\pi r$ but merely a check on the consistency of our theory. We had already used the relation between the radius and the circumference in deriving properties of the trigonometric functions and of their inverses. However, we could at this point start afresh, by defining π to be $\int_{-1}^{1}(1 - x^2)^{-1/2}dx$, and defining sin x to be the inverse function of $\int_{0}^{x}(1 - t^2)^{-1/2}dt$, just as we defined e^x to be the inverse function of $\int_{1}^{x}t^{-1}dt$. From these definitions all the familiar properties of the trigonometric functions and their relation to the geometry of circles and triangles could be proved. This is sometimes done as an exercise in advanced calculus courses.

The use of the differential notation in this example may have suggested to you a simplification of the basic formula (1). Since

$$\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} = \frac{\sqrt{dx^2 + dy^2}}{dt}$$

we can write (1) in the very brief form

$$L = \int \sqrt{dx^2 + dy^2} \, .$$

This form is somewhat incomplete since we cannot say what the limits are until we have selected a parameter.

Example 2. Find the length of the arc of $y^2 = x^3$ between (0,0) and (4,8).

Method 1. $y = x^{3/2}$

$$\int \sqrt{dx^2 + dy^2} = \int \sqrt{dx^2 + \left(\frac{3}{2}x^{1/2}\right)^2 dx^2}$$

$$= \int_0^4 \sqrt{1 + \frac{9}{4}x}\, dx$$

$$= \frac{4}{9} \cdot \frac{2}{3}\left(1 + \frac{9}{4}x\right)^{3/2}\Bigg|_0^4 = \frac{8}{27}\left(10^{3/2} - 1\right).$$

Method 2. $2y\,dy = 3x^2\,dx$.

$$\int \sqrt{dx^2 + dy^2} = \int \sqrt{dx^2 + \frac{9x^4}{4y^2}\,dx^2}$$

$$= \int_0^4 \sqrt{1 + \frac{9x^4}{4x^3}}\,dx = \text{as before.}$$

Method 3. $x = t^2$, $y = t^3$, t in $[0,2]$.

$$\int \sqrt{dx^2 + dy^2} = \int \sqrt{(2t)^2 dt^2 + (3t^2)^2 dt^2}$$

$$= \int_0^2 \sqrt{4t^2 + 9t^4}\,dt$$

$$= \int_0^2 t\sqrt{4 + 9t^2}\,dt$$

$$= \frac{1}{2} \cdot \frac{1}{9} \cdot \frac{2}{3}(4 + 9t^2)^{3/2}\Big|_0^2 = \frac{8}{27}(10^{3/2} - 1).$$

In some parametrizations the parameter may go to infinity while the point of the curve remains bounded. This leads to improper integrals but may otherwise cause no trouble.

<u>Example 3</u>. The curve $x^3 + y^3 = 3xy$ has the parametrization (see Problem 6 in Section 7-7)

$$x = \frac{3t}{1 + t^3}, \quad y = \frac{3t^2}{1 + t^3}.$$

x and y are both positive if and only if t is positive. Hence the loop in the first quadrant is given by $0 \leq t < \infty$. For the piece of the loop defined by $0 \leq t \leq h$ we find

$$L(h) = \int_0^h \frac{3\sqrt{t^8 + 4t^6 - 4t^5 - 4t^3 + 4t^2 + 1}}{(1 + t^3)^3} dt.$$

As $t \to \infty$ the numerator behaves like t^4 and the denominator like t^6; hence the infinite integral converges like $\int t^{-2} dt$ and $L = \lim_{h\to\infty} L(h)$.

But this would be a ridiculous way to find the length of the loop. The improper integral converges quite slowly and we would have to take a very large value for h to get any kind of good approximation to the true value. Because of the symmetry of the curve, however, it is evident that we need only find the length of the lower half of the loop and multiply by two. The point of the loop on the line $x = y$ is evidently given by $t = 1$, so we need only evaluate the integral from 0 to 1. Using Simpson's Rule, the approximations with 8 and 16 subdivisions agree to 6D and so our value of 4.917488 for the length of the loop is almost surely correct to 5D. (In such a short computation the roundoff errors, which occur at the 12th significant figure, will not cause trouble).

PROBLEMS

1. What are the properties of unicon functions that assure us that if $x'(t)$ and $y'(t)$ are unicon so is $F(t,s) = \sqrt{x'(t)^2 + y'(s)^2}$?

2. Find the lengths of the following curves, either precisely (in terms of fractions, radicals, π, known functions etc.) or correct to 2 decimal places.

   (a) $y = x^2$ from $(0,0)$ to $(4,16)$.

   (b) $y = \sin x$ from $(0,0)$ to $(\pi,0)$.

   (c) The involute of a circle,

   $x = a(\cos\theta + \theta\sin\theta)$, $y = a(\sin\theta - \theta\cos\theta)$,

   (see Problem 5 in Section 7-7)

   for one revolution of the point A.

   (d) $y = \log x$ from $(1,0)$ to $(a, \log a)$

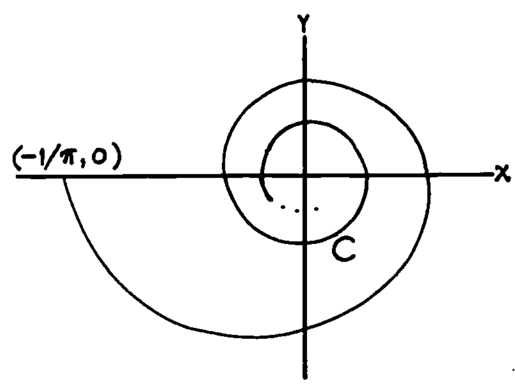
   (e) $y = e^x$ from $(0,1)$ to (c, e^c).

3. Show that the length of one arch of a cycloid is 8a.

4. The two functions

$$x(t) = \begin{cases} t\cos\frac{1}{t}, & 0 < t \le \frac{1}{\pi}, \\ 0 & t = 0; \end{cases}$$

$$y(t) = \begin{cases} t\sin\frac{1}{t}, & 0 < t \le \frac{1}{\pi}, \\ 0 & t = 0; \end{cases}$$



are easily seen to be continuous at $t = 0$, and so they define a curve C with endpoints at $(-1/\pi, 0)$ and $(0,0)$.

(a) Show that the length of the portion of C defined by t in $[h, 1/\pi]$, $h > 0$, is

$$\int_h^{1/\pi} \frac{\sqrt{1 + t^2}}{t}\, dt.$$

(b) What happens as $h \to 0+$? Does C have a length?

5. (a) Using the parametrization

$$x = b \cos \theta, \quad y = a \sin \theta,$$

for an ellipse, show that the length of one quadrant of the ellipse is

$$a \int_0^{\pi/2} \sqrt{1 - k^2 \sin^2 \theta}\, d\theta,$$

where $k^2 = (a^2 - b^2)/a^2$.

(b) Because of its appearance in this problem the integral above (omitting the factor a) is known as an <u>elliptic integral</u>. More precisely, this is the complete elliptic integral of the second kind, designated by $E(k, \pi/2)$. Write and run a program to make a table of $E(\sin \alpha, \pi/2)$ for $\alpha = 0(5^\circ)90^\circ$, accurate to 5D.

(c) Show that the answer to Problem 2(b) is $2\sqrt{2}\, E(\sin 45^\circ, \pi/2)$, and hence get it to 5D.

6. (a) For $n > 0$ the curve $x^n + y^n = 1$ has an arc going from $(0,1)$ to $(1,0)$. Show that the length of this arc is

$$L_n = \int_0^1 \sqrt{1 + \left(\frac{x^n}{1 - x^n}\right)^{2 - 2/n}}\, dx.$$

(b) The cases $n = 1$ and $n = 2$ are familiar. Show that $L_{2/3} = 3/2$.

(c) Find at least one other case that can be integrated exactly, and do so.

5. Mean Value.

To get the average grade on an examination we add together the grades of the individual students and divide by the number of students. This is the most common meaning of the word "average". The word is often used in other ways, however, and to avoid ambiguity this type of average is called in mathematics the arithmetic mean of the grades, or simply the mean or the mean value. It is customary to denote the mean by putting a bar over the appropriate symbol; thus

$$\bar{g} = \frac{1}{n} \sum_{i=1}^{n} g_i$$

is the mean of the grades $g_1, g_2, \ldots, g_n$.

Figure 5-1, taken from Chapter 4, gives the electric power consumption of a town for a 24-hour period. What should we mean by the "average consumption throughout this period"? As an approximation we might take the mean of 24 values at the end of each hour,

$$\bar{c} = \frac{1}{24} \sum_{i=1}^{24} c(i),$$

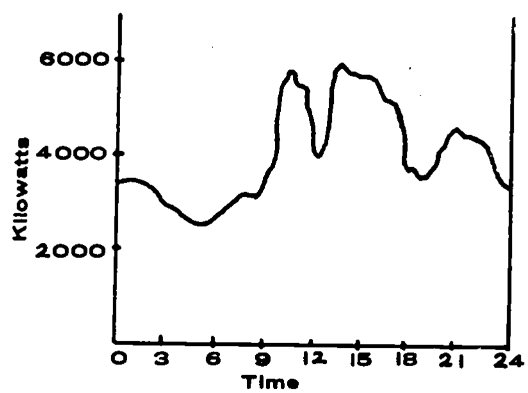


Figure 5-1

but this does not reflect any sudden surges that might have occured within any hour. A better approximation would be the mean of values taken every minute;

(1) $$\bar{c} \approx \frac{1}{1440} \sum_{i=1}^{1440} c(i/60)$$

Better yet would be values taken every second, and so on. The physical process breaks down at about this point but the mathematical model can continue merrily onwards.

Let us rewrite (1), slightly modified, as

(2) $$\bar{c} \approx \frac{1}{24} \sum_{i=1}^{1440} c(\xi_i)\Delta t_i ,$$

where $\Delta t_i = 1/60$ and ξ_i is taken somewhere within the i-th minute. This is obviously as good as (1) in giving us what we want. But now the "and so on" process is clear. For any $\delta > 0$ choose an $n > 1/\delta$ and let $t_i = i/n$, $i = 0,1,\ldots,24n$. Then $\Delta t_i = t_i - t_{i-1} = 1/n < \delta$. Let ξ_i be in $[t_{i-1}, t_i]$, and define

$$\bar{c}_\delta = \frac{1}{24} \sum_{i=1}^{24n} c(\xi_i)\Delta t_i .$$

Then by Section 3-7, page 265, we have

$$\lim_{\delta \to 0} \bar{c}_\delta = \frac{1}{24} \int_0^{24} c(t)\,dt.$$

This is then a reasonable number to take as $\bar{c}$, the mean consumption.

This procedure, which can be applied to many similar cases, leads to the following general definition.

__Definition.__ If a function f is defined over a closed interval [a,b] the __mean value of__ f __on__ [a,b] is

$$\frac{1}{b-a}\int_a^b f(x)dx \text{ ,}$$

provided this integral exists.

__Example 1.__ The mean value of sin x over the interval [0,π] is

$$\frac{1}{\pi}\int_0^\pi \sin x\, dx = \frac{1}{\pi}(-\cos x)\Big|_0^\pi = \frac{2}{\pi} \approx .636.$$

__Example 2.__ A rope 20 ft. long runs over a pulley C at the top of a pole CD 10 ft. high. One end is fastened to a heavy block B, and the other end A is moved horizontally away from the pole, thus raising B to the top of the pole. What is the mean length of AC in this process?

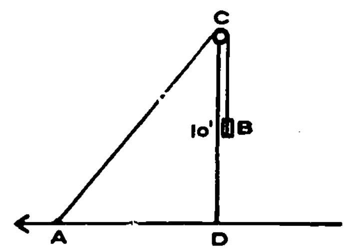


Figure 5-2

Method 1. Let AC = y, AD = x. Then $y = \sqrt{x^2 + 100}$ and x varies from 0 to $10\sqrt{3}$. Hence

$$\bar{y} = \frac{1}{10\sqrt{3}}\int_0^{10\sqrt{3}} \sqrt{x^2 + 100}\, dx$$

$$= \frac{1}{10\sqrt{3}}\left[\frac{1}{2}x\sqrt{x^2 + 100} + 50 \log(x + \sqrt{x^2 + 100})\right]_0^{10\sqrt{3}}$$

$$= 10 + \frac{5}{\sqrt{3}} \log(\sqrt{3} + 2) \approx 13.80 \text{ ft.}$$

Method 2. Let AC = y, BD = z. Then y = 20 - z, z varies over [0,10], and so

$$\bar{y} = \frac{1}{10} \int_0^{10} (20 - z)dz = 15 \text{ ft.}$$

Obviously something is wrong. The trouble lies in the formulation of the question, which asks for the mean of a variable, not of a function. In the two methods we have expressed the variable y as two different functions of different independent variables x and z, and have obtained two different results. In short, the mean value is not independent of the function expressing the variable to be averaged. To cover cases like this one uses the phrase "the mean value of one variable with respect to another." Thus the mean value of y with respect to x is 13.80 ft. and with respect to z is 15 ft.

There is a generalization of the notion of mean value that is often useful; in fact, we shall use it in the next section. Returning to our problem of averaging grades, suppose that an examination was given in 20 sections and that the mean grades in these sections are $g_1, g_2, \ldots, g_{20}$. To get the overall mean we do not merely take the mean of $g_1, \ldots, g_{20}$ but we "weight" these means with the number of students in the section; so that

$$\bar{g} = \frac{1}{N} \sum_{i=1}^{20} n_i g_i, \qquad \text{with} \qquad N = \sum_{i=1}^{20} n_i,$$

if there are n_i students in the i-th section. A moment's thought will convince you that this is the same $\bar{g}$ that we would get if we applied the original method to the whole set of N grades.

The passage from the finite to the continuous case can be made just as before. We are given two functions f and w on an interval [a,b], with the restriction that w(x) is continuous, is not identically zero, and is either always ≥ 0 or always ≤ 0 on [a,b]. Such a w is called a <u>weight function</u>. Then the <u>mean value of</u> f <u>on</u> [a,b] <u>with respect to the weight</u> w is

$$\frac{\int_a^b w(x)f(x)\,dx}{\int_a^b w(x)\,dx}$$

The conditions on w insure that the denominator is not zero (see Problem 1).

One property that any kind of average should have is to lie between the extremes of the averaged quantities. Let us prove that the weighted mean does so. Let m and M be lower and upper bounds of f on [a,b]. Then for every x in [a,b]

$$m \leq f(x) \leq M.$$

Consider the case $w(x) \geq 0$ for all x. We get

$$w(x)m \leq w(x)f(x) \leq w(x)M.$$

(For $w(x) \leq 0$ the inequalities are reversed.) From one of the basic properties of integrals it follows that

(3) $$\int_a^b w(x)m\,dx \leq \int_a^b w(x)f(x)dx \leq \int_a^b w(x)M\,dx.$$

Since m and M are constants and $\int_a^b w(x)dx > 0$ this gives us

(4) $$m \leq \frac{\int_a^b w(x)f(x)dx}{\int_a^b w(x)dx} \leq M.$$

(If $w(x) \leq 0$ the inequalities are again reversed at this point, and we get the same result.)

This result is true even if f is not continuous, as long as it is bounded and $\int_a^b w(x)f(x)dx$ exists. If f is continuous it has a minimum and a maximum on [a,b] and these can be taken to be m and M, thus giving the desired result.

A useful theorem comes at once from these inequalities.

Mean Value Theorem for Integrals. If f is continuous on [a,b] and if W is either always ≥ 0 or always ≤ 0 on [a,b] then there is a ξ in [a,b] such that

$$\int_a^b w(x)f(x)dx = f(\xi)\int_a^b w(x)dx$$

provided these integrals exist.

Proof. (For the case $W(x) \geq 0$). We start with the inequalities (3). If $\int_a^b w(x)dx = 0$ these become

$$0 \leq \int_a^b w(x)f(x)dx \leq 0.$$

Hence

$$\int_a^b w(x)f(x)dx = 0 = f(\xi) \int_a^b w(x)dx$$

for <u>any</u> ξ in [a,b].

Having settled this trivial case we can assume that $\int_a^b W(x)dx$ is not zero, and hence is positive since $W(x) \geq 0$. We can then pass to the inequalities (4). Let m and M be the minimum and maximum of f on [a,b] and let f(c) = m, f(d) = M. Then (4) says that the mean value of f lies between f(c) and f(d), and by the Intermediate Value Theorem there is a ξ in [c,d], and hence in [a,b], such that

$$\frac{\int_a^b w(x)f(x)dx}{\int_a^b w(x)dx} = f(\xi),$$

which gives the desired result.

Corollary. If f is unicon on [a,b] then there is a ξ in [a,b] such that

$$\int_a^b f(x)dx = (b - a)f(\xi).$$

PROBLEMS

1. Theorem. If w is continuous on [a,b], is not identically zero, and is never negative, then

$$\int_a^b w(x)dx > 0.$$

Prove the following:

   1. There is a point c in [a,b] such that $w(c) > 0$.
   2. There is a $\delta > 0$ such that $w(x) > w(c)/2$ for $|x - c| \leq \delta$.
   3. $$\int_{c-\delta}^{c+\delta} w(x)dx > \delta w(c)$$
   4. $$\int_a^b w(x)dx > 0.$$

2. Find the mean value of each given function over the given interval. Draw the graph of $y = f(x)$ and the line y = mean value, and see if the result looks reasonable.

   (a) $f(x) = \sin x$, $[0,\pi/2]$.

   (b) $f(x) = \sin x$, $[0,2\pi]$.

   (c) $f(x) = \sin^2 x$, $[0,2\pi]$.

   (d) $f(x) = \dfrac{1}{1 + x^2}$, $[-1,1]$.

   (e) $f(x) = x \log x$, $[0,2]$.

3. On the diameter of a semicircle 10,000 equally spaced points are taken and used in pairs as two corners of an inscribed rectangle.

(a) What is the mean area (approximately) of these 5000 rectangles?

(b) What is their mean altitude?

(c) What is their mean base?

4. As in 3, but this time the points are equally spaced along the arc of the semicircle.

5. Another useful kind of average is the root mean square, defined to be the square root of the mean of the square of the function, i.e.

$$RMS(f) = \sqrt{\frac{1}{b - a} \int_a^b f(x)^2 dx}.$$

(a) - (e) Find the RMS for each of the functions of Problem 2.

(f) In considering power consumption in alternating currents it is convenient to use the RMS of the voltage rather than the amplitude. (The mean voltage is of course zero). Show that the RMS of the voltage $V = V_0 \sin \omega t$ over one cycle is $V_0/\sqrt{2}$.

(g) A more useful quantity than the RMS of the function f itself is often the RMS of $f(x) - \bar{f}$, where $\bar{f}$ is the mean value of the function. Show that over any interval [a,b],

$$\text{RMS}(f(x) - \bar{f}) = \sqrt{(\text{RMS } f(x))^2 - \bar{f}^2}$$

6. Centroids.

We know from childhood the Law of the Lever: A weight of w_1 pounds at a distance s_1 from the fulcrum will balance a weight of w_2 pounds at a distance s_2 if $w_1 s_1 = w_2 s_2$.

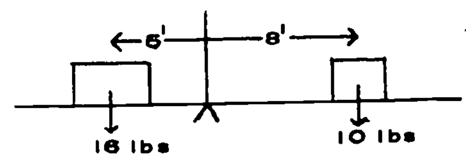


Figure 6-1

A more sophisticated way of stating the same law is the following: Let x_1 be the directed distance of w_1 from any point A of the lever and x_2 the distance of w_2 from A. The weights will balance if the fulcrum is at the weighted mean,

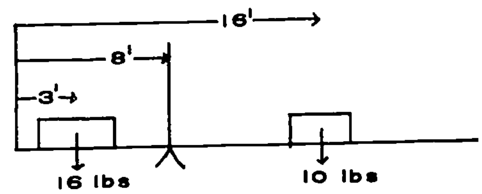


Figure 6-2

$$\bar{x} = \frac{w_1 x_1 + w_2 x_2}{w_1 + w_2},$$

from A.

The advantage of this formutation is that it applies as well to n weights as to two. The point at distance $\bar{x}$ from A is called the <u>center of gravity</u> of the weights, or, if one wishes to consider only the masses of the bodies and leave out any considerations of forces, the <u>center of mass</u>.

Given a distribution of n masses m_i on a line, to find their center of mass we have to introduce a coordinate system, that is a point A from which to measure distance and a unit of distance to do the measuring. Does the center of mass depend upon our choice of coordinate system? Presumably not, or we wouldn't have used the word "center," but this is something we should prove.

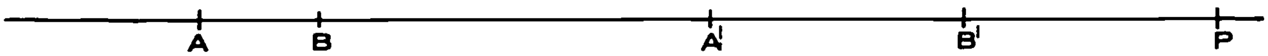


Figure 6-3

Let A be the origin of one coordinate system with unit distance AB, and A' the origin of another with unit distance A'B'. The coordinates of a point P in the two systems are

$$x = \frac{AP}{AB}, \qquad x' = \frac{A'P}{A'B'}.$$

(Of course these are all directed distances). Then

$$(1) \qquad x' = \frac{AP - AA'}{A'B'} = \frac{AB}{A'B'}x - \frac{AA'}{A'B'} = \alpha x + \beta,$$

where $\alpha = \frac{AB}{A'B'}$ and $\beta = -\frac{AA'}{A'B'}$ are constants independent of P. Now if masses $m_1, m_2, \ldots, m_n$ are at points $P_1, P_2, \ldots, P_n$ their center of mass is determined by

$$\bar{x} = \frac{1}{M} \sum_{i=1}^{n} m_i x_i, \qquad M = \sum_{i=1}^{n} m_i,$$

in the AB system and by

$$\bar{x}' = \frac{1}{M} \sum_{i=1}^{n} m_i x_i'$$

in the A'B' system. From (1) we get

$$\overline{x}' = \frac{1}{M} \sum_{i=1}^{n} m_i (\alpha x_i + \beta)$$

$$= \frac{1}{M}\left[\alpha \sum_{i=1}^{n} m_i x_i + \beta \sum_{i=1}^{n} m_i\right]$$

$$= \alpha \overline{x} + \beta.$$

Thus the point $\overline{P}$ with coordinate $\overline{x}$ in the AB system has coordinate $\overline{x}'$ in the A'B' system. In other words, the center of mass $\overline{P}$ does not depend on the coordinate system used to define it.

The case of continuous mass distribution is handled just as in the preceding section.

<u>Example 1.</u> The density of the material in a cylindrical rod varies exponentially, at one end being double its value at the other end. At what point does the rod balance?

Take the length of the rod as the unit of length, and let the densities at the two ends be ρ_0 and $2\rho_0$. Measuring from the lighter end the density ρ at point x is $\rho = Ae^{kx}$, and we have

$$\rho_0 = Ae^0 \qquad 2\rho_0 = Ae^k.$$

Hence $A = \rho_0$, $e^k = 2$, $k = \log 2$. An element of mass is of the form $\rho c \Delta x$, where c is the cross-sectional area of the rod. Hence

$$\overline{x} = \frac{\int_0^1 x\rho_0 e^{kx} c \, dx}{\int_0^1 \rho_0 e^{kx} c \, dx}.$$

ρ_0 and c drop out, and evaluating the integrals gives

$$\bar{x} = \frac{(x/k - 1/k^2)e^{kx}\Big|_0^1}{(1/k)e^{kx}\Big|_0^1}$$

$$= \frac{(1/k - 1/k^2)e^k + 1/k}{(1/k)(e^k - 1)}$$

$$= 2 - \frac{1}{k} = 2 - \frac{1}{\log 2} \approx .5588.$$

The balance point is thus about .56 of the way from the lighter end.

<u>Example 2</u>. The cross-hatched region in Figure 6-4 represents a flat uniform plate, the upper edge being an exponential curve. We wish to balance it in a vertical position at a point on the lower edge. What point?

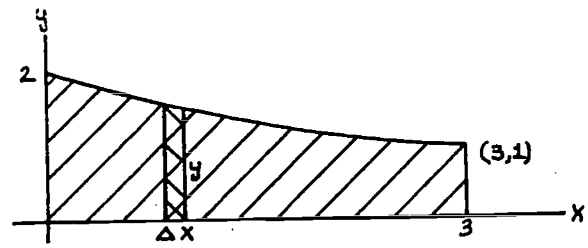


Figure 6-4

It is fairly evident that this is the same problem as in Example 1. The varying density ρ is replaced by the varying height y, and the element of mass is $\rho y \Delta x$, where ρ is the constant areal density of the plate, in pounds per square unit. The answer is then $\bar{x} \approx 1.32$, that is, .56 of the length from the small end.

<u>Example 3</u>. For a plane region, as in Figure 6-4, we can also compute $\bar{y}$, the mean of y weighted with respect to horizontal strips.

This would be the balancing point if the region were turned so that the y-axis is horizontal.

In computing $\bar{y}$ we must do the integration in two parts, since the expression for the length of the strip needs two different formulas. For $0 \le y \le 1$ the length is simply 3. For the equation of the curve we have, as in Example 1,

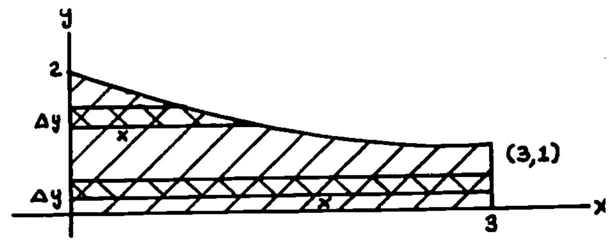


Figure 6-5

$$y = Ae^{kx}, \quad 2 = Ae^{0}, \quad 1 = Ae^{3k},$$

giving

$$A = 2, \quad k = \frac{1}{3}\log\frac{1}{2} = -\frac{1}{3}\log 2.$$

Hence for $1 \le y \le 2$,

$$x = \frac{1}{k}\log\frac{y}{A} = -\frac{3}{\log 2}\log\frac{y}{2}.$$

So

$$\bar{y} = \frac{\int_0^1 3y\,dy + \int_1^2 \left(-\frac{3}{\log 2}\right) y \log\frac{y}{2}\,dy}{\int_0^1 3\,dy + \int_1^2 \left(-\frac{3}{\log 2}\right)\log\frac{y}{2}\,dy} = \frac{\frac{9}{4\log 2}}{\frac{3}{\log 2}} = \frac{3}{4}.$$

(The constant density factor ρ cancels out so we did not bother to put it in).

With any plane region we can associate, as in Examples 2 and 3, a point $(\overline{x},\overline{y})$ with respect to a pair of axes. We have seen that this point is unchanged by change of scale or by translation of the axes. Is it also unchanged by rotation of axes? If so, we are justified in regarding it as a property of the region itself and not of its analytic representation. The answer is "yes", but we are not yet in a position to prove this. For the present we assume this property without proof.

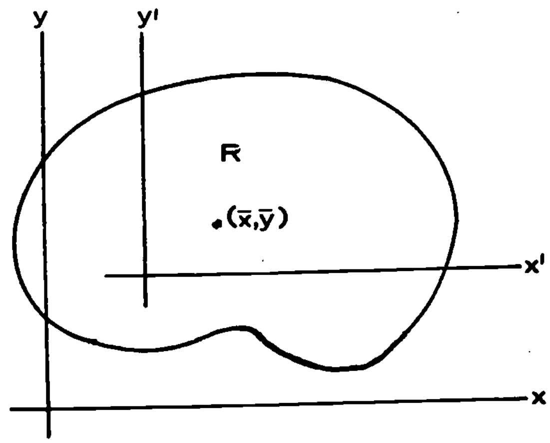


Figure 6-6

This point is sometimes called the <u>center of area</u> of the region but more often the <u>centroid</u>. It is found, in general, by the method of Examples 2 and 3, with one simplification. In the basic expression for both $\overline{x}$ and $\overline{y}$ the denominator is just the area of the region; hence there is no need to compute it twice, and if it is a region whose area is known we need not compute it by integration at all.

<u>Example 4</u>. To find the centroid of a quandrant of a circle (Figure 6-7) we have

$$\overline{x} = \frac{\int_0^a x\sqrt{a^2 - x^2}\,dx}{\frac{1}{4}\pi a^2}$$

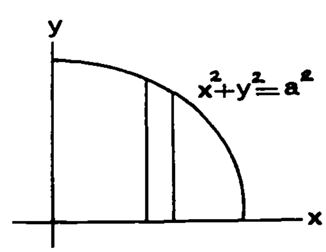


Figure 6-7

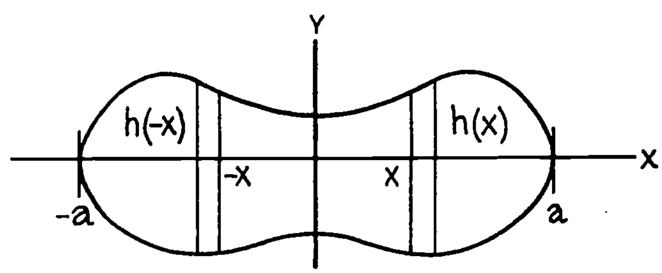
$$= \frac{\frac{1}{3}a^3}{\frac{1}{4}\pi a^2} = \frac{4a}{3\pi}.$$

Because of the symmetry of the figure it is evident that $\bar{y}$ has the same value.

Symmetry is a big help in finding centroids, by virtue of the following property.

Theorem 1. If a region is symmetric with respect to an axis its centroid lies on that axis.

Proof. Take the axis of symmetry as the y-axis. If $h(x)$ is the length of a vertical strip then because of the symmetry $h(-x) = h(x)$ and the limits of integration are $-a$ to a. Then $xh(x)$ is an odd function, and



$$\bar{x} = \frac{1}{\text{Area}}\int_{-a}^{a} xh(x)\,dx = 0.$$

Figure 6-8

(By Theorem 1, Section 11-7).

Corollary. If a region has two axes of symmetry their intersection is the centroid.

This corollary takes care of a lot of familiar figures: rectangle, circle, any regular polygon, etc. Note that the two axes need not be perpendicular; for instance, the corollary applies to an equilateral triangle.

Another general theorem that is often useful in locating centroids is the following:

<u>Theorem 2.</u> If a region is composed of two or more subregions its centroid is the weighted mean of the centroids of the subregions, the weights being the areas.

<u>Suggestion of proof.</u> Let R be the union of R_1, R_2, and R_3, as in Figure 6-9. If A, A_1, A_2, A_3 are the areas of the respective regions and $\bar{x}, \bar{x}_1, \bar{x}_2, \bar{x}_3$ the abscissas of their centroids then

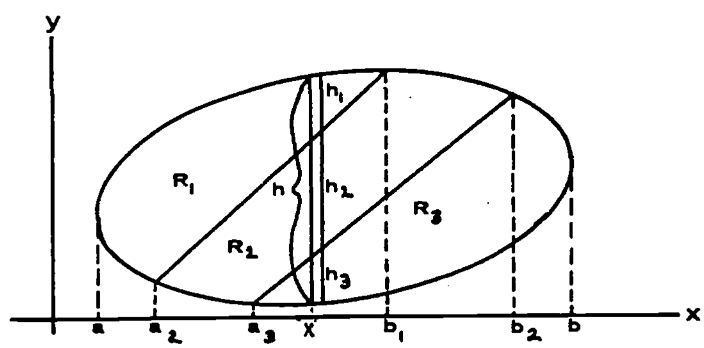


Figure 6-9

$$Ax = \int_a^b xh\,dx = \int_a^b x(h_1 + h_2 + h_3)dx$$

$$= \int_a^b xh_1 dx + \int_a^b xh_2 dx + \int_a^b xh_3 dx,$$

where we define $h_i = 0$ if the vertical line through x does not intersect R_i. The last expression is then the same as

$$\int_{a}^{b_1} xh_1 dx + \int_{a_2}^{b_2} xh_2 dx + \int_{a_3}^{b} xh_3 dx$$

$$= A_1\bar{x}_1 + A_2\bar{x}_2 + A_3\bar{x}_3 .$$

Hence

$$\bar{x} = \frac{A_1\bar{x}_1 + A_2\bar{x}_2 + A_3\bar{x}_3}{A_1 + A_2 + A_3} ,$$

and $\bar{y}$ can be treated similarly.

It is evident that this argument can be extended to the general situation.

Example 5. Figure 6-10 is a semicircle of radius 3 in. on top of a 12 × 3 in. rectangle. We introduce coordinates as shown. For the rectangle,

$$A = 36, \quad \bar{x} = 6.$$

and for the semicircle,

$$A = 9\pi/2, \quad \bar{x} = 9.$$

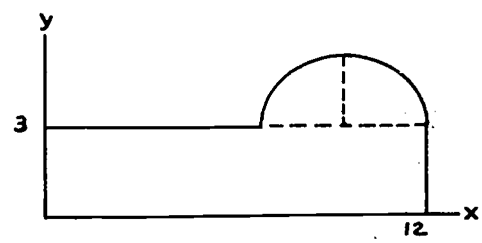


Figure 6-10

Hence for the whole figure,

$$\bar{x} = \frac{6(36) + 9(9\pi/2)}{36 + 9\pi/2} = \frac{48 + 9\pi}{8 + \pi} = 6.846.$$

To get y we must know $\bar{y}$ of the semicircle. We have found that for a quadrant of a circle it is $\frac{4a}{3\pi}$. Since it is the same for both halves of the semicircle (Figure 6-11 it will be the same for the semicircle as a whole by virtue of Theorem 2. So

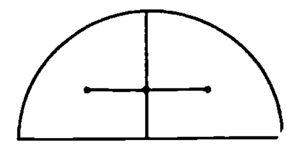


Figure 6-11

$$\bar{y} = \frac{3(36) + (3 + \frac{12}{3\pi})(9\pi/2)}{36 + 9\pi/2} = \frac{28 + 3\pi}{8 + \pi} \approx 3.359.$$

An interesting relation between centroids and solids of revolution is known as the Theorem of Pappus.

Theorem 3. If a plane region is rotated about a line in its plane not intersecting it, the volume of the resulting solid is equal to the area of the region times the circumference of the circle described by its centroid.

Proof. Take the axis of rotation as the y-axis. Then (Figure 6-12)

$$\bar{x} = \frac{\int_a^b xh\,dx}{A},$$

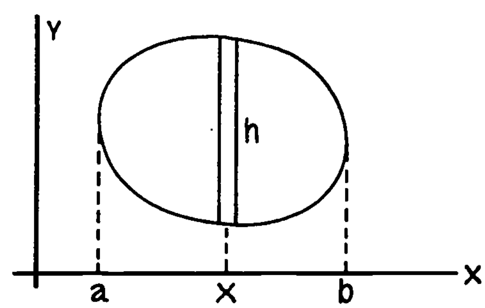


Figure 6-12

where A is the area of the region. Using the cylindrical shell method we have for the volume of the solid of revolution,

$$V = \int_a^b (2\pi x)h \, dx$$

$$= 2\pi \int_a^b xh \, dx$$

$$= (2\pi\bar{x})A,$$

as was to be proved.

The Theorem of Pappus can be used in either direction.

Example 6. (a) To find the centroid of a right triangle (Figure 6-13) observe that if rotated about the side AC we get a cone of radius a and altitude b. Hence

$$\frac{1}{3}\pi a^2 b = (2\pi\bar{x})\frac{1}{2}ab, \quad \text{or} \quad \bar{x} = a/3.$$

Similarly we find $\bar{y} = b/3$.

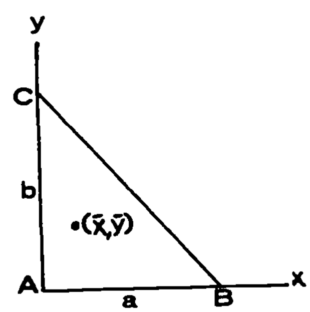


Figure 6-13

(b) The torus, obtained by rotating a circle of radius a about a line b units from the center, b > a, has by the Theorem of Pappus a volume of $(2\pi b)(\pi a^2) = 2\pi^2 a^2 b$, as was found in Chapter 3.

PROBLEMS

1. The density of air decreases with height roughly according to the formula $\rho = \rho_0 e^{-h/a}$, where ρ_0 is the density at the surface and $a \approx 6$ miles.

   (a) Where is the center of gravity of a cylindrical column of air 12 miles high?

   (b) Assuming that the formula holds for arbitrarily large h (it doesn't) where is the center of gravity of an infinitely high column?

2. Find the centroid of each of the following regions.

   (a) Bounded by $y = \sqrt{x}$, $y = 0$, $x = 1$.

   (b) Bounded by $y = ax^2$, $y = b$.

   (c) Bounded by the x-axis and the arch of $y = \sin x$ from $x = 0$ to $x = \pi$.

   (d) The infinite region in the fourth quadrant bounded by $x = 0$, $y = 0$, $y = \log x$.

   (e) The infinite region bounded by the x-axis and $y = \dfrac{1}{1 + x^2}$.

3. Why cannot the following be added to Problem 2?

   (f) The infinite region in the first quadrant bounded by the axes and $y = \dfrac{x}{1 + x^2}$.

4. Is the following statement true? "The centroid of the infinite region bounded by the x-axis and the curve $y = (x^2 + 1)^{-1/2}$ lies on the y-axis".

5. Find the centroid of the region bounded by the x-axis and the first arch of the cycloid.

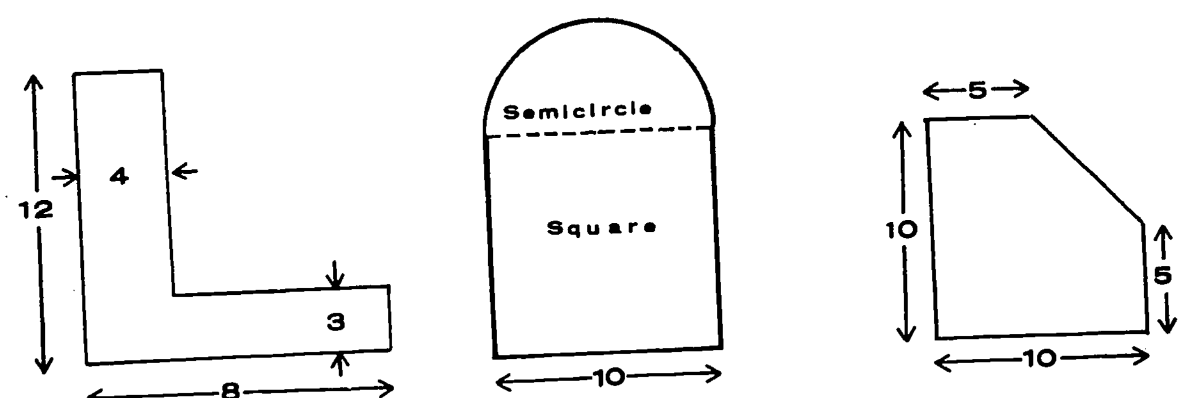
6. Find the centroid of a quadrant of an ellipse, using the parametric form

$$x = a \cos \theta, \quad y = b \sin \theta.$$

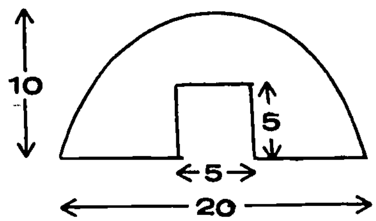
7. Do Problems 3 and 4 of Section I by using the Theorem of Pappus and the results of Problem 2(c) above.

8. Find the centroid of a quadrant of a circle, using the Theorem of Pappus and the formula for the volume of a sphere. Check with Example 4.

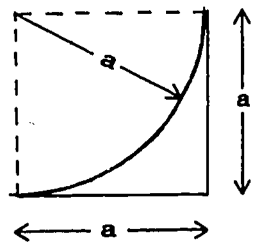
9. Find the centroid of each of the following regions.



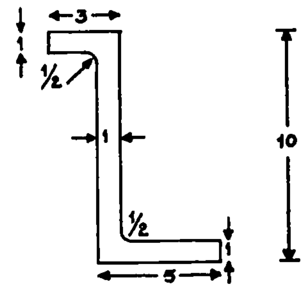
(a) (b) (c)



(d)



(e)



(f)

10. (a) Prove that if a region is divided into two parts its centroid is on the line segment joining the centroids of the two parts.

(b) Prove that if a region is divided into three parts its centroid is in the triangle (possibly degenerate) formed by the centroids of the three parts.

(c) Formulate and prove a statement like the above for four points; for n points.

11. Given a triangle A_1, A_2, A_3 and masses m_1, m_2, m_3 at the vertices. As in (b) above, the center of mass lies in the triangle. Prove conversely that for any point P in or on the triangle there are non-negative masses m_1, m_2, m_3 whose center of mass is P. Prove also that m_1, m_2, m_3 are unique to within a common factor. The numbers (m_1, m_2, m_3) are called <u>barycentric coordinates</u> of P.

12. (a) For the triangle (0,0), (0,1), (2,0) find a set of barycentric coordinates of each of the points (1/2,1/2), (1,0), (0,1), (x,y).

(b) Find the xy-coordinates of the points with barycentric coordinates (1,1,1), (1,2,3), (0,1,1), (1,0,0), (a,b,c).

Chapter 13

INFINITE SERIES

1. Taylor Series.

At various points in our study of calculus we have found Taylor's Theorem (Section 6-4) a useful tool in computation or in deriving properties of functions. The time has now come to apply this theorem systematically. First we give a restatement and a proof.

Theorem 1. If $f, f', f'', \ldots, f^{(n+1)}$ are unicon on an interval containing a and x, then

$$(1)\ f(x) = f(a) + f'(a)(x - a) + \frac{1}{2!} f''(a)(x - a)^2 + \ldots + \frac{1}{n!} f^{(n)}(a)(x - a)^n + \frac{1}{n!} \int_a^x f^{(n+1)}(t)(x - t)^n dt.$$

Proof. The method is just successive integration by parts, differentiating f at each step and picking the constant of integration properly. We start with

$$(2)\ f(x) - f(a) = \int_a^x f'(t)dt,$$

the second form of the Fundamental Theorem. Taking

$$u = f'(t), \qquad dv = dt,$$

$$du = f''(t)dt, \qquad v = t - x = -(x - t),$$

we get

$$f(x) = f(a) - f'(t)(x - t)\Big|_a^x + \int_a^x f''(t)(x - t)dt$$

$$= f(a) + f'(a)(x - a) + \int_a^x f''(t)(x - t)dt.$$

Now we take

$$u = f''(t), \qquad dv = (x - t)dt,$$

$$du = f'''(t)dt, \qquad v = -\frac{1}{2}(x - t)^2,$$

to get

$$f(x) = f(a) + f'(a)(x - a)$$

$$- \frac{1}{2}f''(t)(x - t)^2\Big|_a^x + \frac{1}{2}\int_a^x f'''(t)(x - t)^2 dt$$

$$= f(a) + f'(a)(x - a) + \frac{1}{2}f''(a)(x - a)^2$$

$$+ \frac{1}{2}\int_a^x f'''(t)(x - t)^2 dt.$$

It is easy to see that a continuation of this process will give the desired result. (The more cautious mathematician may wish to give a proof by mathematical induction).

Notice that we have not implied that $x > a$, for (2) holds even if $x < a$.

It is customary to define $f^{(0)} = f$, and $0! = 1$. With these conventions we can write (1) in the form

$$(3) \qquad f(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(a)(x - a)^k + R_n(x),$$

where

$$(4) \qquad R_n(x) = \frac{1}{n!} \int_a^x f^{(n+1)}(t)(x - t)^n dt.$$

The first part of the right-hand side of (3) is called the _Taylor expansion of f, about a, to_ n + 1 _terms_. $R_n(x)$ is the _remainder after this expansion_ (or _after_ n + 1 _terms_, or _after the_ (n + 1)st _term_). The Taylor expansion about 0 is also known as a _Maclaurin_ expansion.

You may have noticed that the expression for $R_n(x)$ given in (4) is not the same as the one given in the earlier statement of Taylor's Theorem. This other form of $R_n(x)$, known as the derivative form, is easily derived from the integral form by use of the mean value theorem for integrals (Section 12-5). Since $(x - t)^n$ does not change sign in the interval [a,x], and since $f^{(n+1)}$ was assumed unicon, and hence continuous, the hypotheses of the mean value theorem hold and we have

$$R_n(x) = \frac{1}{n!} f^{(n+1)}(\xi) \int_a^x (x - t)^n dt$$

$$= \frac{1}{n!} f^{(n+1)}(\xi) \left[- \frac{(x - t)^{n+1}}{n + 1} \right]_a^x$$

or

(5) $$R_n(x) = \frac{1}{(n + 1)!} (x - a)^{n+1} f^{(n+1)}(\xi),$$

with ξ between a and x.

Example 1. For the expansion of log x about 1 we note that

$$f(x) = \log x,$$
$$f'(x) = 1/x,$$
$$f''(x) = -1/x^2,$$
$$f'''(x) = 2/x^3,$$
$$\cdots\cdots\cdots$$
$$f^{(k)}(x) = (-1)^{k-1}(k - 1)!/x^k.$$

It follows that

(6) $$\log x = 0 + (x - 1) - \frac{1}{2}(x - 1)^2 + \frac{1}{3}(x - 1)^3 - \frac{1}{4}(x - 1)^4 + \ldots + (-1)^{n-1} \frac{1}{n}(x - 1)^n + R_n(x$$

with

(7) $$R_n(x) = (-1)^n \int_1^x \frac{(t - x)^n}{t^{n+1}}\, dt$$

or

(8) $$R_n(x) = (-1)^n \frac{1}{n + 1}(x - 1)^{n+1} \frac{1}{\xi^{n+1}}, \quad \xi \text{ between } 1 \text{ and } x.$$

Since log x and all its derivatives are defined in the interval $(0,\infty)$ the above expansion, with remainder, is good for any $x > 0$.

From (8) we can easily get a bound for $|R_n(x)|$. If $x \geq 1$ then $1/\xi^{n+1}$ is maximum when $\xi = 1$; if $x < 1$ then $1/\xi^{n+1}$ is maximum when $\xi = x$. Hence

$$|R_n(x)| \leq \frac{1}{n + 1} |x - 1|^{n+1} \times \begin{cases} 1 & \text{if } x \geq 1 \\ 1/x^{n+1} & \text{if } x < 1 \end{cases}$$

$$= \frac{1}{n + 1} \times \begin{cases} (x - 1)^{n+1} & \text{if } x \geq 1 \\ (\frac{1}{x} - 1)^{n+1} & \text{if } x < 1. \end{cases}$$

That Taylor's Theorem is not the only way to get a Taylor expansion is shown by the following example, which is typical of methods we shall develop later.

Example 2. The Maclaurin expansion of $f(x) = \frac{1}{1 - x}$ can be obtained by dividing 1 by 1 - x in increasing powers of x. We get the algebraic identity

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

We leave it to the reader (Problem 1) to show that the first n + 1 terms on the right-hand side constitute the terms of a Taylor expansion. Once this is done it follows that the remaining term is the remainder. That is,

$$R_n(x) = \frac{x^{n+1}}{1 - x} = \frac{1}{n!} \int_0^x f^{(n+1)}(t)(x - t)^n dt$$

$$= (n + 1) \int_0^x \frac{(x - t)^n}{(1 - t)^{n+2}} dt.$$

This equation can be checked by evaluating the integral (see Problem 2).

If, in (9), we change n to n - 1 and replace x by -t we get

$$\frac{1}{1 + t} = 1 - t + t^2 - t^3 + \dots + (-1)^{n-1}t^{n-1} + (-1)^n \frac{t^n}{1 + t}.$$

Now integrate each side of this identity from 0 to x. This gives

(10) $$\log(1 + x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \ldots + (-1)^{n-1}\frac{1}{n}x^n + R_n(x),$$

where

(11) $$R_n(x) = (-1)^n \int_0^x \frac{t^n}{1 + t}\,dt.$$

This is the Maclaurin expansion of $\log(1 + x)$.

Finally, in (10) put $1 + x = y$. We get

(12) $$\log y = (y - 1) - \frac{1}{2}(y - 1)^2 + \frac{1}{3}(y - 1)^3 - \ldots$$

$$+ (-1)^{n-1}\frac{1}{n}(y - 1)^n + S_n(y),$$

with

$$S_n(y) = (-1)^n \int_0^{y-1} \frac{t^n}{1 + t}\,dt.$$

The expansion (12) is obviously the same as (6), and so $S_n(y)$ must be equal to the $R_n(x)$ in (7). These examples illustrate the many forms in which the remainder terms may appear.

If <u>all</u> derivatives of f exist and are unicon we have the possibility of letting $n \to \infty$ in (1). That this can be done in some cases is evident from our earlier use of the Maclaurin expansion of e^x. We have

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!} + R_n(x).$$

$$R_n(x) = \frac{x^{n+1}}{(n+1)!} e^{\xi}.$$

For $x > 0$, for example, $e^{\xi} \leq e^x$, and so

$$|R_n(x)| \leq \frac{x^{n+1}}{(n+1)!} e^x.$$

Now if $N \geq 2x$, then for $n > N$,

$$\frac{x^{n+1}}{(n+1)!} e^x = e^x \frac{x^N}{N!} \frac{x}{N+1} \frac{x}{N+2} \cdots \frac{x}{n} \frac{x}{n+1}$$

$$< e^x \frac{x^N}{N!} \frac{1}{2^{n-N+1}}$$

As $n \to \infty$ the last quantity $\to 0$, and hence

(12) $$e^x = \lim_{n\to\infty} \left(1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!}\right).$$

We write (12) in the forms

$$e^x = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \ldots$$

$$= \sum_{n=0}^{\infty} \frac{x^n}{n!} ,$$

and call it the Maclaurin series of e^x. In general, if f and all its derivatives exist at a, the expression

$$\sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(a)(x - a)^n$$

is the _Taylor series of_ f(x) _at_ a. The _Maclaurin series_ is simply the case a = 0. A Taylor series of a function may or may not converge for a given value of x, and if it does converge its limit may or may not be the value of the function for that value of x. This is why we have to pay so much attention to the remainder in the finite Taylor expansion.

Example 3. Consider the remainder in the Maclaurin expansion of log(1 + x) given by (11), for the case x > 0. We have

$$|R_n(x)| = \int_0^x \frac{t^n}{1 + t}\, dt.$$

Since t varies from 0 to x, 1 + t lies between 1 and 1 + x. Hence

$$\int_0^x \frac{t^n}{1 + x}\, dt \le |R_n(x)| \le \int_0^x \frac{t^n}{1}\, dt,$$

or, evaluating the integrals,

$$\frac{1}{1 + x} \frac{x^{n+1}}{n + 1} \le |R_n(x)| \le \frac{x^{n+1}}{n + 1}.$$

$$\text{As } n \to \infty, \; x^{n+1} \to \begin{cases} 0 \text{ if } x < 1, \\ 1 \text{ if } x = 1, \\ \infty \text{ if } x > 1. \end{cases}$$

Hence the Maclaurin series converges to the value of the function if and only if $x \leq 1$. We leave it to the reader to show that for $x < 0$ the series converges to the function if and only if $x > -1$. Hence the <u>interval of convergence</u> of the Maclaurin series of log (1 + x) is $-1 < x \leq 1$.

For easy reference we list here some important Maclaurin series and their intervals of convergence. They are derived either in the text or in problems.

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots, \quad -\infty < x < \infty.$$

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots,$$

$$-\infty < x < \infty.$$

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots, \quad -\infty < x < \infty.$$

$$\log(1 + x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots,$$

$$-1 < x \leq 1.$$

$$\arctan x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n + 1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots,$$

$$-1 \leq x \leq 1.$$

$$(1 + x)^a = \sum_{n=0}^{\infty} \frac{a(a - 1) \ldots (a - n + 1)}{n!} x^n$$

$$= 1 + ax + \frac{a(a - 1)}{2!} x^2 + \frac{a(a - 1)(a - 2)}{3!} x^3$$

$$+ \cdots, \quad -1 < x \leq 1.$$

For further study of Taylor series, and in particular their relation to the operations of differentiation and integration, we must first consider infinite series in general. This is the subject of the next section.

PROBLEMS

1. Show that (9) is actually the Maclaurin expansion of

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

2. Show that $(n+1)\int_0^x \frac{(x-t)^n}{(1-t)^{n+2}}\,dt = \frac{x^{n+1}}{1-x}$.

[Hint. Write the integral as $\int_0^x \left(\frac{x-t}{1-t}\right)^n \frac{1}{(1-t)^2}\,dt$

and make a substitution.]

3. Show that for $x < 0$, $\int_0^x \frac{t^n}{1+t}\,dt$ converges to 0 as $n \to \infty$

if and only if $x > -1$.

4. Find the maximum possible error in approximating each of the following functions as indicated.

(a) Log x by 4 terms of the expansion about 1, for $|x - 1| \leq .2$.

(b) $x^{1/3}$ by 3 terms of the expansion about 8, for $|x - 8| \leq 2$.

(c) sin x by 3 non-zero terms of the expansion about 0, for $|x| \leq \pi/4$.

(d) $\cos x$ by 4 terms of the expansion about $\pi/3$, for $|x - \pi/3| \leq .1$.

5. (a) Find the Maclaurin series for $\sin x$ and prove that it converges for all x.

(b) Do the same for $\cos x$.

6. (a) Find a form of the remainder for $\log(1 - x)$ similar to the one given in (11) for $\log(1 + x)$.

(b) By subtracting, show that for $0 \leq x < 1$,

$$\log\frac{1 + x}{1 - x} = 2\left[x + \frac{x^3}{3} + \frac{x^5}{5} + \ldots + \frac{x^{2m-1}}{2m - 1}\right] + T_m,$$

where $T_m = \int_0^x \frac{2t^{2m}}{1 - t^2}dt$.

(c) Since the integrand in T_m is always positive, by the Mean Value Theorem for integrals we have

$$T_m = \frac{2\xi^{2m}}{1 - \xi^2}, \quad 0 \leq \xi \leq x.$$

Show that the right hand side is an increasing function of ξ and hence that

$$T_m \leq \frac{2x^{2m}}{1 - x^2} .$$

(d) For what value of x is $\frac{1 + x}{1 - x} = 3$? How many terms of the series in (b) would be needed to compute log 3 to 3D? Do you know a faster way to compute log 3?

7. In each of the following find the first three non-zero terms of the expansion. Find an expression for the general term if this can be done easily.

(a) e^{-2x} about 0.

(b) sin x about $\pi/4$.

(c) e^x about a.

(d) $5x^4 - 2x^2 + x + 2$ about 0; about -1.

(e) tan x about 0.

(f) $\sqrt{x}$ about 9.

(g) log x about 4.

(h) $\sqrt{1 + x} + \sqrt{1 - x}$ about 0.

(i) $e^{\sin x}$ about 0.

(j) arcsin x about 0.

8. We wish to approximate $\int_0^1 e^{x^2} dx$ to 5D.

(a) Show that for $0 \le y \le 1$,

$$e^y = 1 + y + \frac{y^2}{2!} + \ldots + \frac{y^n}{n!} + R_n,$$

where $0 \le R_n \le \frac{y^{n+1}}{(n + 1)!} e.$

(b) Putting $y = x^2$ in the above and integrating, show that

$$\int_0^1 e^{x^2} dx = 1 + \frac{1}{3} + \frac{1}{5 \times 2!} + \ldots + \frac{1}{(2n + 1)n!} + T_n,$$

where $0 \le T_n \le \frac{e}{(2n + 3)(n + 1)!}$.

(c) Show that 8 terms of the series (n = 7) is enough to give the required accuracy.

9. Use the method of the previous problem to calculate

(a) $\int_0^1 \frac{\sin x}{x} dx$ to 3D.

(b) $\int_1^2 x^{-1}e^{-x} dx$ to 2D. Compare with Example 2 of Section 11-9. Which method do you prefer, and why?

10. To see that a convergent Taylor series of a function may have little connection with the function values consider the function

$$f(x) = \begin{cases} 0 & \text{if } x = 0, \\ e^{-1/x^2} & \text{if } x \neq 0. \end{cases}$$

(a) Show that for any $n > 0$,

$$\lim_{x \to 0} \frac{e^{-1/x^2}}{x^n} = 0.$$

[Hint. Put $y = \frac{1}{x^2}$ and take limits as $y \to \infty$.]

(b) Using

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

show that $f'(0) = 0$.

(c) Compute $f'(x)$. Using

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

show that $f''(0) = 0$.

(d) Show how one can conclude that $f^{(n)}(0) = 0$ for $n = 3,4,\ldots$.

(e) What is the Maclaurin series of $f(x)$? For what values of x does it converge? For what values of x does it converge to $f(x)$?

11. Use the integral form of the remainder in the Maclaurin expansion of e^{-x} to do Example 5 of Section 11-7.

2. <u>Convergence of series</u>.

Basically, an infinite series is just an infinite sequence written in a special way. Let $a_1, a_2, a_3, \ldots$ be any sequence and define another sequence $S_1, S_2, S_3, \ldots$ by

$$S_1 = a_1,$$
$$S_2 = a_1 + a_2,$$
$$S_3 = a_1 + a_2 + a_3,$$

and in general

$$S_n = \sum_{k=1}^{n} a_k.$$

The S's can also be defined recursively by

$$S_1 = a_1$$
$$S_{n+1} = S_n + a_{n+1}, \qquad n = 1, 2, \ldots;$$

this is a convenient way of computing them in a flow chart or a computer program. The sequence $S_1, S_2, S_3, \ldots$ is called the <u>sequence of partial sums of the infinite series</u>

$$a_1 + a_2 + a_3 + \ldots \qquad \text{or} \qquad \sum_{n=1}^{\infty} a_n.$$

<u>Example 1</u>. (a) The series $\sum_{n=1}^{\infty} n$ has the partial sums

$$1, 3, 6, 10, \ldots, \frac{n(n+1)}{2}, \ldots .$$

(b) The series $\sum_{n=1}^{\infty} 2^{-n}$ has the partial sums

$$\frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \frac{15}{16}, \ldots, \frac{2^n - 1}{2^n}, \ldots .$$

(c) The series $\sum_{n=1}^{\infty} (-1)^{n-1}$ has the partial sums

$$1, 0, 1, 0, 1, 0, \ldots .$$

An infinite series is said to <u>converge</u> if the sequence of partial sums converges. If

$$\lim_{n\to\infty} S_n = L$$

we write

$$\sum_{n=1}^{\infty} a_n = L$$

and say that L is the <u>sum</u> of the series.

In Example 1, series (b) converges to the sum 1, while series (a) and (c) diverge, i.e., do not converge.

Since series are so closely related to sequences many of the properties of sequences developed in Chapter 2 can, with suitable changes, be applied to series. Some of the results of this process are stated in Problem 1. We prove here some other theorems whose proofs are not quite so direct.

Theorem 1. If $\sum_{n=1}^{\infty} a_n$ converges then $\lim_{n\to\infty} a_n = 0$.

Proof. Let $\sum_{n=1}^{\infty} a_n = A$ and let $S_n = \sum_{k=1}^{n} a_k$ be the n-th partial sum. Since the sequence $S_1, S_2, \ldots$ converges to A, given any $\varepsilon > 0$ there is an N such that

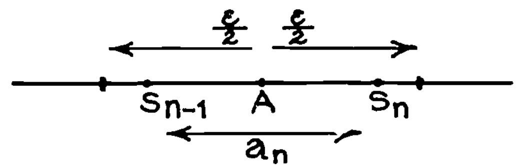


Figure 2-1

$$|S_n - A| < \frac{\varepsilon}{2} \quad \text{whenever} \quad n > N.$$

Then

$$|a_n| = |S_n - S_{n-1}| = |(S_n - A) - (S_{n-1} - A)|$$

$$\leq |S_n - A| + |S_{n-1} - A| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

whenever $n > N + 1$. Hence $\lim_{n\to\infty} a_n = 0$.

Briefly, the proof says that since the S_n's get close together when n is large their differences, the a_n's, get small.

The converse of this theorem is not true. That is, if $\lim_{n\to\infty} a_n = 0$ then $\sum_{n=1}^{\infty} a_n$ does not necessarily converge. An example is the so-called harmonic series, $\sum_{n=1}^{\infty} \frac{1}{n}$

Theorem 2. The harmonic series diverges.

Proof. Compare the two series,

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{9} + \frac{1}{10} + \dots ,$$

$$\frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{16} + \frac{1}{16} + \dots ,$$

where, in the second series, there are 2^{j-1} terms of the form $\frac{1}{2^j}$. For the partial sums, T_n, of the second series we have

$$T_1 = \frac{1}{2}, \quad T_2 = \frac{2}{2}, \quad T_4 = \frac{3}{2}, \quad T_8 = \frac{4}{2}, \quad \dots, \quad T_{2^n} = \frac{n+1}{2}, \quad \dots .$$

Each term of the first series is at least as large as the corresponding term of the second series, and so for the partial sums of the harmonic series, $S_{2^n} \geq \frac{n+1}{2}$. Since these increase without bound the series cannot converge.

This example illustrates that Theorem 1 cannot be used to prove convergence of a series. It can be used to prove divergence, for an equivalent statement of the theorem is: If the sequence $a_1, a_2, a_3, \dots$ either diverges or converges to a limit not zero then $\sum_{n=1}^{\infty} a_n$ diverges. Thus we can see at once that Examples 1 (a) and (c) diverge since their sequences of terms diverge.

There is, however, a useful special case in which we can prove convergence.

<u>Theorem 3</u>. Let $a_1, a_2, a_3, \ldots$ be a decreasing sequence of positive numbers with limit 0. Then $\sum_{n=1}^{\infty} (-1)^{n-1} a_n$ converges.

<u>Proof</u>. Consider the odd partial sums S_{2n+1} and the even ones S_{2n}. We have

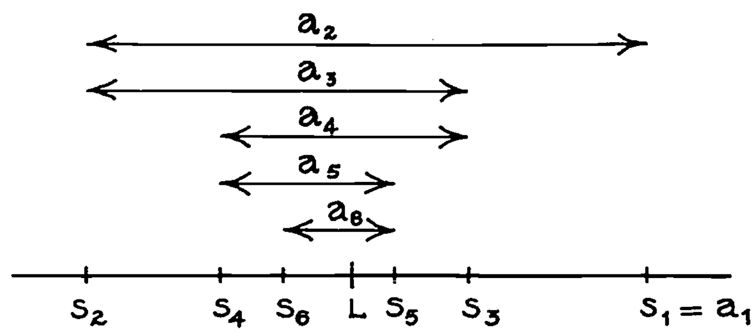


Figure 2-2

$$S_{2(n+1)+1} = S_{2n+1} - a_{2n+2} + a_{2n+3} \leq S_{2n+1}$$

since $a_{2n+3} \leq a_{2n+2}$. Thus the odd partial sums form a decreasing sequence. This is shown clearly in Figure 2-2. Similarly, the even partial sums form an increasing sequence, since

$$S_{2(n+1)} = S_{2n} + a_{2n+1} - a_{2n+2} \geq S_{2n}$$

because $a_{2n+1} \geq a_{2n+2}$. To apply the Completeness Axiom (Section 2-9) to these two sequences we need only to show that

$$\lim_{n\to\infty} (S_{2n+1} - S_{2n}) = 0.$$

But since $S_{2n+1} - S_{2n} = a_{2n+1}$ this is one of the hypotheses of the theorem. Hence the two sequences

$$S_1, S_3, S_5, \ldots \quad \text{and} \quad S_2, S_4, S_6, \ldots$$

have the same limit L. It follows that the sequence $S_1, S_2, S_3, \ldots$ has limit L; that is, $\sum_{n=1}^{\infty} (-1)^{n-1} a_n$ converges to the sum L.

As an example of the application of this theorem we see that the <u>alternating</u> harmonic series,

$$\sum_{n=1}^{\infty} (-1)^{n-1}/n = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots ,$$

does converge. (See, also, Example 3 of Section 1).

Before stating a useful corollary of Theorem 3 we give a definition. If the series $\sum_{n=1}^{\infty} a_n = L$ then the <u>remainder after n terms</u> of the series is

$$R_n = L - S_n = \sum_{k=n+1}^{\infty} a_k.$$

If we approximate the sum L by a partial sum S_n, $|R_n|$ is the error of the approximation. Hence we are interested in bounds on R_n or on $|R_n|$. For series of the type of Theorem 3 we have this convenient result.

<u>Corollary 1</u>. For the series in Theorem 3, R_N lies between 0 and the $(N + 1)$st term.

Proof. If N is even, N = 2n, we have

$$S_{2n} \leq L \leq S_{2n+1},$$

and so

$$0 \leq L - S_{2n} = R_{2n} \leq S_{2n+1} - S_{2n} = a_{2n+1}.$$

If N = 2n + 1, then

$$S_{2n+1} \geq L \geq S_{2n+2} ,$$

$$0 \geq L - S_{2n+1} = R_{2n+1} \geq S_{2n+2} - S_{2n+1} = -a_{2n+2}.$$

This corollary is well illustrated by Figure 2-2. Since L lies somewhere between S_5 and S_6 it is evident that S_5-L, which is $-R_5$, is $\leq S_5 - S_6$, which is $-a_6$.

In this figure we know only that L lies between S_5 and S_6. If we want to make an estimate of L it is natural to take the value half-way between these two, that is, $S_5 - \frac{1}{2}a_6$. Then the true value of L can differ from this by at most $\frac{1}{2}a_6$. In general we have the following:

Corollary 2. If $T_n = S_n + \frac{1}{2}(-1)^n a_{n+1}$, then

$$|L - T_n| \leq \frac{1}{2}a_{n+1}$$

Proof. If n is even we have

$$S_n \leq L \leq S_{n+1} = S_n + a_{n+1},$$

$$S_n - \frac{1}{2}a_{n+1} \leq L - \frac{1}{2}a_{n+1} \leq S_n + \frac{1}{2}a_{n+1},$$

$$-\frac{1}{2}a_{n+1} \leq L - S_n - \frac{1}{2}a_{n+1} \leq \frac{1}{2}a_{n+1}.$$

That is,

$$|L - T_n| \leq \frac{1}{2}a_{n+1}.$$

A similar procedure works if n is odd.

A series whose terms alternate in sign is called an alternating series. It is evident that Theorem 3 and its corollary hold for any alternating series for which the absolute values of the terms approach zero monotonically. That monotonicity is essential both for proof of convergence and for the bounds of Corollary I is shown in Problem 2.

Example 2. The Maclaurin series for e^x,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

is an alternating series if $x < 0$. The remainder $R_n(x)$ therefore lies between 0 and $x^{n+1}/(n + 1)!$, and we need not bother with the complicated remainder terms in integral or derivative forms. For instance,

$$e^{-1} = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + - \ldots + \frac{1}{10!} + R_{10}$$

$$\approx 0.367879464 + R_{10}$$

with $0 \geq R_{10} \geq -\frac{1}{11!} \approx -2.51 \times 10^{-8}$. Hence if we use Corollary 2 and take

$$e^{-1} \approx 0.367879451$$

the absolute value of the error is less than 1.3×10^{-8}.

Computer programming of such a computation is relatively simple. Figure 2-3 shows a typical flow diagram for computing $\sum_{n=1}^{\infty} a_n$ when the signs of the a_n alternate. As in Figure 9-3 of Chapter 11, MAX is a bound on

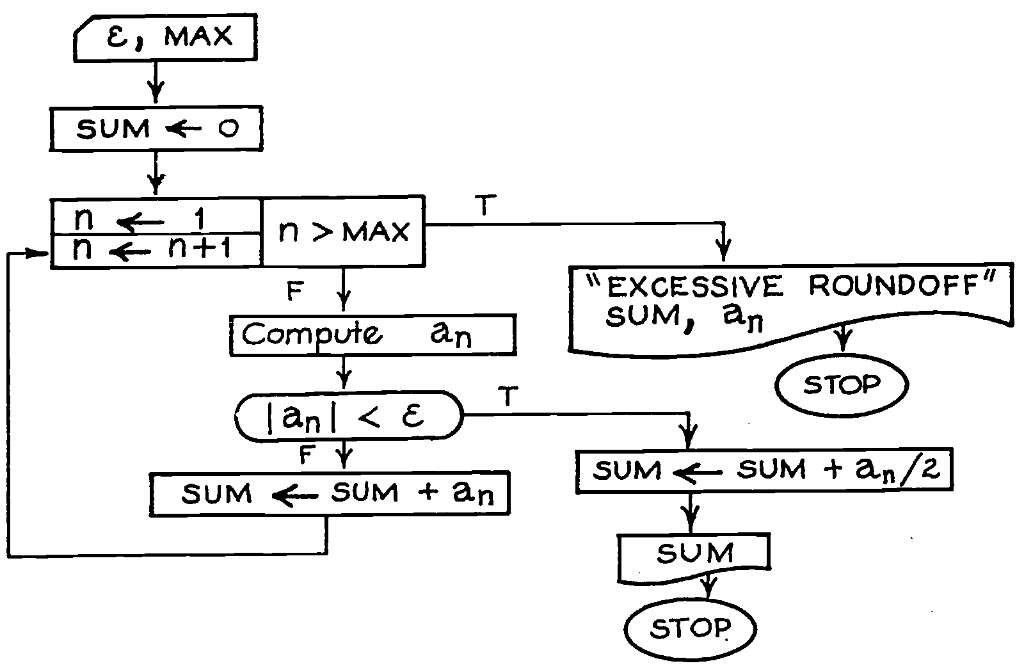


Figure 2-3

the number of terms that can be added without exceeding a roundoff error of $\varepsilon/2$.

The box

Compute a_n

will often contain a recursion formula. For Example 2, for instance, it would be

$a \leftarrow -a/n$

,

and the box

$SUM \leftarrow 0$

would be replaced by

$SUM \leftarrow 1$
$a \leftarrow 1$

to give the proper start to the recursion process.

We note here a property of series similar to the one for sequences stated on page 159. It is that the convergence or divergence of a series is not affected by removal, addition, or change of value of any <u>finite</u> number of terms. This can affect the <u>sum</u> of the series, if convergent, but the question of convergence or divergence depends only on what happens for $n > N$ and we can choose N so large that all the changes occur for $n < N$. Thus the series

$$5 + 4 + 3 - 2 + 1 + 0 + 0 - 10 + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \dots$$

behaves essentially like the alternating harmonic series and Theorem 3 and Corollary 1 can be applied to it beginning with the ninth term. We often cover such possibilities by appropriate use of the phrase "for sufficiently large n."

Another trivial but sometimes convenient modification that we can make is in numbering the terms. In the above series, for example, it might be best to start the subscript n with the value -5, so that the general term after the eighth can be written as $(-1)^{n-1}/n$. In such cases, when the initial value of n is not 1, we generally still call a_n the n-th term, speaking, if necessary, of the "zeroth term", the "minus second term", etc.

PROBLEMS

1. Prove the following properties of series by interpreting them in terms of sequences of partial sums.

(a) If $\sum_{n=1}^{\infty} a_n = A$ and $\sum_{n=1}^{\infty} b_n = B$, then

$$\sum_{n=1}^{\infty} (a_n \pm b_n) = A \pm B.$$

(b) If $k \neq 0$, then $\sum_{n=1}^{\infty} ka_n$ converges if and only if $\sum_{n=1}^{\infty} a_n$ converges; if $\sum_{n=1}^{\infty} a_n = A$ then $\sum_{n=1}^{\infty} ka_n = kA$.

(c) If $\sum_{n=1}^{\infty} a_n = A$ and $\sum_{n=1}^{\infty} b_n = B$, and if $a_n \leq b_n$ for all n, then $A \leq B$.

2. (a) Define $a_1, a_2, a_3, \ldots$ by

$$a_{2n-1} = \frac{1}{n}, \quad a_{2n} = -10^{-n}, \quad n = 1, 2, \ldots .$$

Show that $\sum_{n=1}^{\infty} a_n$ is an alternating series whose terms approach 0, but which diverges.

(b) Define $a_1, a_2, a_3, \ldots$ by

$$a_n = 3(-1/2)^n \quad \text{if} \quad n \text{ is not divisible by } 3.$$

$$a_n = 10(-1/2)^n \quad \text{if} \quad n \text{ is divisible by } 3.$$

Show that $\sum_{n=1}^{\infty} a_n$ is a convergent alternating series for which the conclusions of Corollary 1 are not true.

3. What conclusions, if any, can be from the following facts by means of Theorem 1?

(a) $\sum_{n=1}^{\infty} \frac{n^3}{3^n}$ converge

(b) $\sum_{n=1}^{\infty} (1 - \frac{\log n}{n})^n$ diverges.

(c) $\lim_{n\to\infty} \frac{\sin n}{n} = 0.$

(d) The sequence $\frac{1}{\sin 1}, \frac{1}{\sin 2}, \ldots, \frac{1}{\sin n}, \ldots$ does not have the limit 0.

4. How many terms would be needed to get 5D accuracy in summing each of the following alternating series?

(a) $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$.

(b) $1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \ldots$.

(c) $\sum_{n=0}^{\infty} \frac{x^n}{n!}$, $x = -.5$ and $x = -2$.

(d) $x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots$, $x = .8$.

5. Write a flow chart to compute the n^{th} term, $n \geq 2$, of

$$f(x) = 1 - \frac{x^2}{2\cdot 7} + \frac{x^4}{2\cdot 4\cdot 7\cdot 11} - \ldots + \frac{(-1)^n\, x^{2n}}{(2\cdot 4\cdots 2n)(7\cdot 11\cdots (4n+3))} + \ldots$$

first without using a recursive formula and then using a formula that computes a_n from a_{n-1}. Which way is the most efficient? Remember this as you work the next two problems.

6. (a) Modify Figure 2-3 to compute e^x, for $x < 0$, from its Maclaurin series.

(b) Write a program from your flow diagram and test it by computing e^{-1} and e^{-2} to 5D.

(c) Make the necessary modifications in your program and print a table of e^{-x} to 5D for $x = 0(.1)5$.

7. Carry out parts (a) and (b) of Problem 6 for sin x and cos x. For part (c) make a 5D table of sin x and cos x for x = 0(1)45 degrees.

8. Use the method of Problem 8, Section 1, but modified by the use of Corollary 2, to approximate the following integrals.

(a) $\int_0^1 \cos x^2 \, dx$ to 2D.

(b) $\int_0^{.5} \sqrt{x}\, e^{-x} \, dx$ to 3D.

9. Problem 4(a) shows that it is impractical to try to sum the alternating harmonic series directly. For such slowly converging series there are methods of "accelerating" the rate of convergence. If one plots several partial sums, as in Figure 2-2, he soon notices that the midpoints of successive segments (S_n, S_{n+1}) are much closer to the limit than the ends of the segments. Let $T_n = (S_n + S_{n+1})/2$ be these midpoints.

(a) Use the Squeeze Theorem of Section 2-7 to prove that $\lim_{n\to\infty} T_n = \lim_{n\to\infty} S_n$.

(b) If we define

$$b_1 = T_1, \quad b_n = (-1)^{n-1}(T_n - T_{n-1}), \quad n = 2,3,\ldots$$

then T_n is the n-th partial sum of $\sum_{n=1}^{\infty}(-1)^{n-1}b_n$. Show that

$$b_1 = (2a_1 - a_2)/2, \quad b_n = (a_n - a_{n+1})/2,$$

and so $\sum_{n=1}^{\infty}(-1)^{n-1}b_n$ is also an alternating series with the same sum as $\sum_{n=1}^{\infty}(-1)^{n-1}a_n$.

(c) For $a_n = \frac{1}{n}$ show that $b_1 = \frac{3}{4}$, $b_n = \frac{1}{2n(n+1)}$, $n > 1$.

For $a_1 = \frac{3}{4}$, $a_n = \frac{1}{2n(n+1)}$, $n > 1$, show that

$b_1 = \frac{17}{24}$, $b_n = \frac{2}{4n(n+1)(n+2)}$, $n > 1$.

For $a_1 = \frac{17}{24}$, $a_n = \frac{2}{4n(n+1)(n+2)}$, $n > 1$,

find b_n.

(d) Sum the alternating harmonic series correct to 5D.

95

3. Tests for Convergence.

For series that do not satisfy the conditions of Theorem 3 of the last section the question of convergence or divergence is more difficult to answer. Of great importance in this connection is the following extension of the Completeness Axiom.

Theorem 1. A bounded monotone sequence is convergent.

Note that a monotone sequence is automatically bounded on one side by its first term, e.g. any increasing sequence is bounded below. A proof of the theorem is given at the end of this section.

Corollary 1. An infinite series of positive terms converges if and only if the partial sums are bounded.

The proof is left to the reader.

The way this Corollary is most often used is illustrated in the following example.

Example 1. We know that the series

(1) $$1 + \frac{1}{2} + \frac{1}{4} + \ldots + \frac{1}{2^n} + \ldots$$

converges to the sum 2. How about the series

(2) $$1 + \frac{1}{1 \times 2} + \frac{1}{2 \times 4} + \frac{1}{3 \times 8} + \ldots + \frac{1}{n2^n} + \ldots ?$$

Each term of (2) is less than or equal to the corresponding term of (1), and so the partial sums of (2) are less than or equal to the partial sums of (1). By Corollary 1 the partial sums of (1) are bounded, since (1) converges, and so the partial sums of (2) are bounded. Hence, again by Corollary 1, (2) converges.

Knowing that (2) converges we can speak of its remainders. By the same argument, the remainders of (2) are less than or equal to the remainders of (1), that is, $R_n \leq 2^{-n}$. This gives us some idea of how many terms to take to get the sum to a given accuracy.

Example 2. Compare the harmonic series,

(3) $$1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} + \ldots$$

and

(4) $$0 + \frac{2}{3} + \frac{3}{8} + \frac{4}{15} + \ldots + \frac{n}{n^2 - 1} + \ldots \quad .$$

Beginning with the second term, each term of (4) is greater than the corresponding term of (3), and hence, except for an additive constant, the partial sums of (4) are greater than the partial sums of (3). Since (3) diverges, a double application of Corollary I, as in the preceding example, shows that (4) also diverges.

The general method used in these examples can be stated as follows:

<u>Comparison Test I</u>. Let $a_n \geq 0$, $b_n \geq 0$, $b_n \geq a_n$, for all sufficiently large n.

(a) If $\sum b_n$ converges then $\sum a_n$ converges, and if S_n and R_n are their respective remainders after n terms then $0 \leq R_n \leq S_n$.

(b) If $\sum a_n$ diverges then $\sum b_n$ diverges.

For brevity we have written merely $\sum$ instead of $\sum_{n=1}^{\infty}$. We shall do this whenever there is no danger of the meaning being mistaken.

To apply this test we need some standard series for comparison purposes. So far we have only one useful divergent series, the harmonic series, and one simple convergent series, the geometric series

$$1 + r + r^2 + \ldots + r^n + \ldots ,$$

for which $R_n = \frac{r^{n+1}}{1 - r}$, (see Example 2 in Section 1) and which therefore converges when $|r| < 1$. We anticipate a result of the next section with the statement:

$$\sum_{n=1}^{\infty} n^{-p} \text{ converges if and only if } p > 1.$$

This series is called a p-series.

Example 3. Test $\sum \frac{\sqrt{n + 2}}{n^2 - 10}$ for convergence. For large n the constants 2 and 10 are insignificant compared with n and n^2, and so the term is approximately $\frac{\sqrt{n}}{n^2}$. This suggests comparison with the p-series for $p = 3/2$, and also suggests that we look for convergence. However our term is obviously greater than $n^{-3/2}$, and it does not seem that we can apply Part (a) of the test. The way out of this dilemma is to multiply the p-series by a suitable constant; this does not affect its convergence. So we ask: Is

$$\frac{\sqrt{n + 2}}{n^2 - 10} \leq c \frac{\sqrt{n}}{n^2}$$

for a suitable c? Rewriting the inequality as

$$\frac{\sqrt{n + 2}}{\sqrt{n}} \leq c \frac{n^2 - 10}{n^2}$$

or

$$\sqrt{1 + \frac{2}{n}} \leq c\left(1 - \frac{10}{n^2}\right)$$

it is evident that this is true for $c = 2$ and, say, $n > 100$. Hence our series converges.

The above tricks of comparing the approximate sizes of corresponding terms for large n, and of multiplying the comparison series by a constant, are very useful. They are combined in the following statement:

Comparison Test 2. If, for all sufficiently large n, $a_n > 0$, $b_n > 0$, and $\frac{a_n}{b_n} \begin{cases} \le M_1 \\ \ge M_2 > 0 \end{cases}$ then if $\sum b_n \begin{cases} \text{converges} \\ \text{diverges} \end{cases}$ so does $\sum a_n$.

Note that if both conditions are satisfied, i.e. $0 < M_2 \le \frac{a_n}{b_n} \le M_1$, then $\sum a_n$ and $\sum b_n$ either both converge or both diverge.

Proof. If $\sum b_n$ converges so does $\sum M_1 b_n$. Hence if $\frac{a_n}{b_n} \le M_1$, i.e. $a_n \le M_1 b_n$, $\sum a_n$ converges by Test 1. The other half of the theorem is proved similarly.

The following special case of Test 2 is often useful. It will apply, for example, to Example 2.

Comparison Test 3. If, for all suffficiently large n, $a_n > 0$ and $b_n > 0$, and if $\lim_{n\to\infty} \frac{a_n}{b_n} = c \le \infty$, then if

$$\begin{Bmatrix} c < \infty \\ c > 0 \end{Bmatrix} \text{ and } \sum b_n \begin{Bmatrix} \text{converges} \\ \text{diverges} \end{Bmatrix} \text{ so does } \sum a_n.$$

The proof of this from Test 2 is left to the reader.

Example 4. Test $\sum_{n=1}^{\infty} n^5(.9)^n$ for convergence. We know that $\sum_{n=1}^{\infty}(.9)^n$ converges but our terms are greater than those of this comparison series so the tests tell us nothing. But we also know from the limit computations of Section 10-4 that exponential terms dominate powers. So we take the convergent series $\sum(.99)^n$ as a comparison. Then

$$\lim_{n\to\infty} \frac{n^5(.9)^n}{(.99)^n} = \lim_{n\to\infty} \frac{n^5}{(1.1)^n} = 0$$

by Problem 5(b) of Section 10-4. Hence our series converges.

The comparison tests, derived from Corollary 1, apply only to series whose terms are ultimately positive. For other series we can get some information by considering the series whose terms are the absolute values of those of the given series. If this series converges we say the given