# Undergraduate Texts in Mathematics

Editors S. Axler F.W. Gehring K.A.Ribet

Springer Science+Business Media, LLC

Introduction to Linear Algebra 1997, ISBN 0-387-96205-0

Calculus of Several Variables, Third Edition 1987, ISBN 0-387-96405-3

**Basic Mathematics** 1988, ISBN 0-387-96787-7

Math! Encounters with High School Students 1985, ISBN 0-387-96129-1

Geometry: A High School Course (with Gene Morrow), Second Edition 1988, ISBN 0-387-96654-4

# Other Books by Lang Published by Springer-Verlag

Linear Algebra • Undergraduate Algebra • Real and Functional Analysis • Undergraduate Analysis • Fundamentals of Differential Geometry • Introduction to Arakelov Theory • Riemann-Roch Algebra (with William Fulton) • Complex Multiplication • Introduction to Modular Forms • Modular Units (with Daniel Kubert) • Fundamentals of Diophantine Geometry • Elliptic Functions • Number Theory III • Survey on Diophantine Geometry • Cyclotomic Fields I and II •  $SL_2(\mathbf{R})$  • Abelian Varieties • Introduction to Algebraic and Abelian Functions • Algebraic Number Theory • Introduction to Complex Hyperbolic Spaces • Elliptic Curves: Diophantine Analysis • A First Course in Calculus • The Beauty of Doing Mathematics • THE FILE • Introduction to Diophantine Approximations • CHALLENGES • Differential and Riemannian Manifolds • Math Talks for Undergraduates • Collected Papers I-V • Spherical Inversion on SLn ( $\mathbf{R}$ ) • Algebra Serge Lang

# **Short Calculus**

The Original Edition of "A First Course in Calculus"

With 30 Illustrations



Serge Lang Department of Mathematics Yale University New Haven, CT 06520 USA

Editorial Board

S. Axler Mathematics Department San Francisco State University San Francisco, CA 94132 USA F.W. Gehring Mathematics Department East Hall University of Michigan Ann Arbor, MI 48109 USA K.A. Ribet Mathematics Department University of California, Berkeley Berkeley, CA 94720-3840 USA

Mathematics Subject Classification (2000): 26-01, 26A06

Library of Congress Cataloging-in-Publication Data Lang, Serge, 1927– Short calculus : the original edition of "A First Course in Calculus" / Serge Lang. p. cm. — (Undergraduate texts in mathematics) Includes bibliographical references and index. ISBN 978-0-387-95327-4 ISBN 978-1-4613-0077-9 (eBook) DOI 10.1007/978-1-4613-0077-9 1. Calculus I. Title. II. Series. QA303.2.L.36 2001 515—dc21 2001041076

Printed on acid-free paper.

First printing entitled *A First Course in Calculus* published by Addison-Wesley Publishing Co., Inc., 1964.

© 2002 Springer Science+Business Media New York

Originally published by Springer-Verlag New York, Inc. in 2002

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher Springer Science+Business Media, LLC, except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computersoftware, or by similar or dissimilar methodology now known or hereafter developed is forbidden. The use of general descriptive names, trade names, trademarks, etc., in this publication, even if the former are not especially identified, is not to be taken as a sign that such names, as understood by the Trade Marks and Merchandise Marks Act, may accordingly be used freely by anyone.

Production managed by Yong-Soon Hwang; manufacturing supervised by Joe Quatela.

987654321

ISBN 978-0-387-95327-4

SPIN 10842616

# Foreword

The First Course in Calculus went through five editions since the early sixties. Sociological and educational conditions have evolved in various ways during four decades. Hence it has been found worth while to make the original edition again available. It is also worth while repeating here most of the foreword which I wrote almost forty years ago.

The purpose of a first course in Calculus is to teach the student the basic notions of derivative and integral, and the basic techniques and applications which accompany them.

At present in the United States, this material is covered mostly during the first year of college. Ideally, the material should be taught to students who are approximately sixteen years of age, and belongs properly in the secondary schools. (I have talked with several students of that age, and find them perfectly able to understand what it is all about.)

Irrespective of when it is taught, I believe that the presentation remains more or less invariant. The very talented student, with an obvious aptitude for mathematics, will rapidly require a course in functions of one real variable, more or less as it is understood by professional mathematicians. This book is not primarily addressed to such students (although I hope they will be able to acquire from it a good introduction at an early age).

I have not written this course in the style I would use for an advanced monograph, on sophisticated topics. One writes an advanced monograph for oneself, because one wants to give permanent form to one's vision of some beautiful part of mathematics, not otherwise accessible, somewhat in the manner of a composer setting down his symphony in musical notation.

This book is written for the student, to provide an immediate, and pleasant, access to the subject. I hope that I have struck a proper compromise between dwelling too much on special details, and not giving enough technical exercises, necessary to acquire the desired familiarity with the subject. In any case, certain routine habits of sophisticated mathematicians are unsuitable for a first course.

This does not mean that so-called rigour has to be abandoned. The logical development of the mathematics of this course from the most basic axioms proceeds through the following stages: Set theory Integers (whole numbers) Rational numbers (fractions) Numbers (i.e. real numbers) Limits Derivatives and forward.

No one in his right mind suggests that one should begin a course with set theory. It happens that the most satisfactory place to jump into the subject is between limits and derivatives. In other words, any student is ready to accept as intuitively obvious the notions of numbers and limits and their basic properties. For some reason, there is a fashion which holds that the best place to enter the subject logically is between numbers and limits. Experience shows that the students do *not* have the proper psychological background to accept this, and resist it tremendously. Of course, there is still another fashion, which is to omit proofs completely. This does not teach mathematics, and puts students at a serious disadvantage for subsequent courses, and the understanding of what goes on.

In fact, it turns out that one can have the best of all these ideas. The arguments which show how the properties of limits can be reduced to those of numbers form a self-contained whole. Logically, it belongs *before* the subject matter of our course. Nevertheless, we have inserted it as an appendix. If any students feel the need for it, they need but read it and visualize it as Chapter 0. In that case, everything that follows is as rigorous as any mathematician would wish it (so far as objects which receive an analytic definition are concerned). Not one word need be changed in any proof. I hope this takes care once and for all of possible controversies concerning so-called rigour.

Some objects receive a geometric definition, and there are applications to physical concepts. In that case, it is of course necessary to insert one step to bridge the physical notion and its mathematical counterpart. The major instances of this are the functions sine and cosine, and the area, as an integral.

For sine and cosine, we rely on the notions of plane geometry. If one accepts standard theorems concerning plane figures then our proofs satisfy the above-mentioned standards.

For the integral, we first give a geometric argument. We then show, using the usual Riemann sums, how this geometric argument has a perfect counterpart when we require the rules of the game to reduce all definitions and proofs to numbers. This should satisfy everybody. Furthermore, the theory of the integral is so presented that only its existence depends either

#### FOREWORD

on a geometric argument or a slightly involved theoretical investigation (upper and lower sums). According to the level of ability of a class, the teacher may therefore dose the theory according to ad hoc judgement.

It is not generally recognized that some of the major difficulties in teaching mathematics are analogous to those in teaching a foreign language. (The secondary schools are responsible for this. Proper training in the secondary schools could entirely eliminate this difficulty.) Consequently, I have made great efforts to carry the student verbally, so to say, in using proper mathematical language. Some proofs are omitted. For instance, they would be of the following type. In the theory of maxima and minima, or increasing and decreasing functions, we carry out in full just one of the cases. The other is left as an exercise. The changes needed in the proof are slight, amounting mainly to the insertion of an occasional minus sign, but they force students to understand the situation and train them in writing clearly. This is very valuable. Aside from that, such an omission allows the teacher to put greater emphasis on certain topics, if necessary, by carrying out the other proof. As in learning languages, repetition is one of the fundamental tools, and a certain amount of mechanical learning, as distinguished from logical thinking, is both healthy and necessary.

I have made no great innovations in the exposition of calculus. Since the subject was discovered some 300 years ago, it was out of the question. Rather, I have omitted some specialized topics which no longer belong in the curriculum. Stirling's formula is included only for reference, and can be skipped, or used to provide exercises. Taylor's formula is proved with the integral form of the remainder, which is then properly estimated. The proof with integration by parts is more natural than the other (differentiating some complicated expression pulled out of nowhere), and is the one which generalizes to the higher dimensional case. I have placed integration after differentiation, because otherwise one has no technique available to evaluate integrals. But on the whole, everything is fairly standard.

I have cut down the amount of analytic geometry to what is both necessary and sufficient for a general first course in this type of mathematics. For some applications, more is required, but these applications are fairly specialized. For instance, if one needs the special properties concerning the focus of a parabola in a course on optics, then that is the place to present them, not in a general course which is to serve mathematicians, physicists, chemists, biologists, and engineers, to mention but a few. What is important is that the basic idea of representing a graph by a figure in the plane should be thoroughly understood, together with basic examples. The more abstruse properties of ellipses, parabolas, and hyperbolas should be skipped.

As for the question: why republish a forty year old edition? I answer:

Because for various reasons, a need exists for a short, straightforward and clear introduction to the subject. Adding various topics may be useful in some respects, and adding more exercises also, but such additions may also clutter up the book, especially for students with no or weak background.

To conclude, if I may be allowed another personal note here, I learned how to teach the present course from Artin, the year I wrote my Doctor's thesis. I could not have had a better introduction to the subject.

> Serge Lang New Haven, 2002

# **Contents**

#### CHAPTER I

## Numbers and Functions

1.	Integers, r	ational	nu	mbe	$\mathbf{rs}$	and	rea	l nu	ımł	bers					1
2.	Inequalitie	es.				•									4
3.	Functions			•					•					•	8
4.	Powers						•				•				11

## CHAPTER II

# **Graphs and Curves**

Coordinates		•			•	•				•					•		14
Graphs .																	17
The straight lin	ne.					•						•					<b>22</b>
Distance betwe	een tv	vo j	poin	$\mathbf{ts}$													26
Curves and equ	uatior	ıs	•		•	•											27
The circle .						•											28
The parabola.	Char	iges	s of	coo	rdi	nate	es										31
The hyperbola																	33
	Coordinates Graphs . The straight li Distance betwee Curves and eq The circle . The parabola. The hyperbola	Coordinates Graphs The straight line . Distance between tw Curves and equation The circle The parabola. Char The hyperbola .	Coordinates Graphs The straight line Distance between two p Curves and equations The circle The parabola. Changes The hyperbola	Coordinates Graphs The straight line Distance between two poin Curves and equations . The circle The parabola. Changes of The hyperbola	Coordinates	Coordinates	Coordinates	Coordinates	CoordinatesGraphsThe straight lineDistance between two pointsCurves and equationsThe circleThe parabola.Changes of coordinates.The hyperbola	CoordinatesGraphsThe straight lineDistance between two pointsCurves and equationsThe circleThe parabola.Changes of coordinatesThe hyperbola	Coordinates						

#### CHAPTER III

#### The Derivative

The slope of a cur	ve	•			•	•								•		•	35
The derivative .		•	•				•		•					•			39
Limits			•						•				•				44
Powers																	<b>48</b>
Sums, products, an	nd	quo	tie	$\mathbf{nts}$					•				•				51
The chain rule .																	57
Rate of change .			•	•		•		•	•	•	•	•	•	•		•	63
	The slope of a cur The derivative Limits . Powers Sums, products, an The chain rule Rate of change	The slope of a curve The derivative . Limits Powers Sums, products, and The chain rule . Rate of change .	The slope of a curve . The derivative Limits Powers Sums, products, and quo The chain rule Rate of change	The slope of a curve The derivative Limits Powers Sums, products, and quotien The chain rule Rate of change	The slope of a curve The derivative	The slope of a curve.The derivative.Limits.Powers.Sums, products, and quotientsThe chain rule.Rate of change.	The slope of a curve										

#### CHAPTER IV

#### Sine and Cosine

1.	The sine and	cosine	fun	ction	s.	•											65
2.	The graphs	• •	•			•	•	•	•	•	•	•	•	•	•	•	69

#### CONTENTS

3.	Addition formula					•						71
4.	The derivatives	•										74
5.	Two basic limits		•			•					•	76

# CHAPTER V

#### The Mean Value Theorem

1.	The maximum and minimum theorem.						<b>79</b>
2.	Existence of maxima and minima						85
3.	The mean value theorem						87
4.	Increasing and decreasing functions .				•		89

#### CHAPTER VI

## **Sketching Curves**

1.	Behavior as $x$ becom	nes	very	r lai	rge		•	•	•		•		•	94
2.	Curve sketching .													97
3.	Polar coordinates													102
4.	Parametric curves	•	•	•	•	•			•	•	•	•		104

# CHAPTER VII

#### **Inverse Functions**

1.	Definition of inverse functions	•						107
2.	Derivative of inverse functions	5			•			109
3.	The arcsine							111
4.	The arctangent							113

#### CHAPTER VIII

#### **Exponents and Logarithms**

1.	The logarithm .	•			•				•			118
2.	The exponential func	tio	n					•				123
3.	The general exponent	tial	fur	ncti	on		•			•		126
4.	Order of magnitude											128
5.	Some applications											132

#### CHAPTER IX

#### Integration

1.	The indefinite integral				•							134
2.	Continuous functions											137
3.	Area											138
4.	Upper and lower sums	•				•	•	•	•			142

#### CONTENTS

5.	The fundamental theorem							146
6.	The basic properties							150

#### CHAPTER X

## **Properties of the Integral**

1.	Further connection	with	$\mathbf{the}$	der	iva	tive	е	•	•	•	•			153
2.	Sums	•												154
3.	Inequalities							•					•	157
4.	Improper integrals										•			158

#### CHAPTER XI

# **Techniques of Integration**

1.	Substitution				•							163
2.	Integration by parts .											166
3.	Trigonometric integrals											169
4.	Partial fractions	•	•			•	•	•	•	•		173

#### CHAPTER XII

#### Some Substantial Exercises

1.	An estimate for $(n!)^{1/n}$								178
2.	Stirling's formula .					•			179
3.	Wallis' product								181

## CHAPTER XIII

#### **Applications of Integration**

1.	Length of curves							183
2.	Area in polar coordinates							187
3.	Volumes of revolution .							188
4.	Work							190
5.	Moments							192

## CHAPTER XIV

## **Taylor's Formula**

1.	Taylor's formula				÷			•			195
2.	Estimate for the remainder							•		•	199
3.	Trigonometric functions					•					201
4.	Exponential function	•									203
5.	Logarithm	•						•		•	205
6.	The arctangent			•	•	•	•			<b>.</b> .	<b>206</b>
7.	The binomial expansion .	•			•						<b>208</b>

#### CONTENTS

# CHAPTER XV

#### Series

1.	Convergent series		•								•					211
2.	Series with positive tern	ns		•	•											213
3.	The integral test															216
4.	Absolute convergence															218
5.	Power series															220
6.	Differentiation and inte	grat	ion	of	pov	ver	seri	es	•	•	•	•	•	•	•	223
Ap	pendix 1. $\epsilon$ and $\delta$ .	•	•	•	•	•	•	•	•					•		227
1.	Least upper bound .															227
2.	Limits				•											<b>228</b>
3.	Points of accumulation			•												<b>234</b>
4.	Continuous functions	•	•	•	•	•	•	•	•	•	•	•	•	•	•	236
Ap	pendix 2. Physics and	M٤	ıth	ema	atic	:5	•		•			•	•	•	•	239
An	swers		•			•		•			•	•				<b>24</b> 1
Su	pplementary Exercises	•	•	•		•	•	•				•				250
Inc	lex			•								•				260

# CHAPTER I

# **Numbers and Functions**

In starting the study of any sort of mathematics, we cannot prove everything. Every time that we introduce a new concept, we must define it in terms of a concept whose meaning is already known to us, and it is impossible to keep going backwards defining forever. Thus we must choose our starting place, what we assume to be known, and what we are willing to explain and prove in terms of these assumptions.

At the beginning of this chapter, we shall describe most of the things which we assume known for this course. Actually, this involves very little. Roughly speaking, we assume that you know about numbers, addition, subtraction, multiplication, and division (by numbers other than 0). We shall recall the properties of inequalities (when a number is greater than another). On a few occasions we shall take for granted certain properties of numbers which might not have occurred to you before and which will always be made precise. Proofs of these properties will be supplied in the appendix for those of you who are interested.

#### §1. Integers, rational numbers and real numbers

The most common numbers are the numbers  $1, 2, 3, \ldots$  which are called *positive integers*.

The numbers -1, -2, -3, -4, ... are called *negative integers*. When we want to speak of the positive integers together with the negative integers and 0, we call them simply *integers*. Thus the integers are 0, 1, -1, 2, -2, 3, -3, ....

The sum and product of two integers are again integers.

In addition to the integers we have fractions, like  $\frac{3}{4}, \frac{5}{7}, -\frac{1}{8}, -\frac{101}{27}, \frac{8}{16}, \ldots$ , which may be positive or negative, and which can be written as quotients m/n, where m, n are integers and n is not equal to 0. Such fractions are called rational numbers. Every integer m is a rational number, because it can be written as m/1, but of course it is not true that every rational number is an integer. We observe that the sum and product of two rational numbers are again rational numbers. If a/b and m/n are two rational numbers (a, b, m, n being integers and b, n unequal to 0), then

their sum and product are given by the following formulas, which you know from elementary school:

$$\frac{a}{b} \frac{m}{n} = \frac{am}{bn},$$
$$\frac{a}{b} + \frac{m}{n} = \frac{an + bm}{bn}.$$

In this second formula, we have simply put the two fractions over the common denominator bn.

We can represent the integers and rational numbers geometrically on a straight line. We first select a unit length. The integers are multiples of this unit, and the rational numbers are fractional parts of this unit. We have drawn a few rational numbers on the line below.



Observe that the negative integers and rational numbers occur to the left of zero.

Finally, we have the numbers which can be represented by infinite decimals, like  $\sqrt{2} = 1.414...$  or  $\pi = 3.14159...$ , and which will be called *real numbers* or simply *numbers*.

Geometrically, the numbers are represented as the collection of all points on the above straight line, not only those which are a rational part of the unit length or a multiple of it.

We note that the sum and product of two numbers are numbers. If a is a number unequal to zero, then there is a unique number b such that ab = ba = 1, and we write

$$b = \frac{1}{a}$$
 or  $b = a^{-1}$ .

We say that b is the *inverse* of a, or "a inverse". We emphasize that the expression

# 1/0 or $0^{-1}$ is not defined.

In other words, we cannot divide by zero, and we do not attribute any meaning to the symbols 1/0 or  $0^{-1}$ .

However, if a is a number then the product  $0 \cdot a$  is defined and is equal to 0. The product of any number and 0 is 0. Furthermore, if b is any number unequal to 0, then 0/b is defined and equal to 0. It can also be written  $0 \cdot (1/b)$ .

If a is a rational number  $\neq 0$ , then 1/a is also a rational number. Indeed, if we can write a = m/n, with integers m, n both different from 0, then

$$\frac{1}{a} = \frac{n}{m}$$

is also a rational number.

Not all numbers are rational numbers. For instance,  $\sqrt{2}$  is not a rational number, and we shall now prove this fact.

We recall that the *even numbers* are the integers  $\pm 2, \pm 4, \pm 6, \pm 8, \ldots$ , which can be written in the form 2n for some integer n. An *odd* number is an integer like  $\pm 1, \pm 3, \pm 5, \pm 7, \ldots$ , which can be written in the form 2n + 1 for some integer n. Thus  $6 = 2 \cdot 3$  is even (we select n = 3) and

$$11 = 2 \cdot 5 + 1$$

is odd (we select n = 5).

We observe that the square of an even number is even. Indeed, if n is an integer and 2n is an even number, then

$$(2n)^2 = 4n^2$$

is an even number, which can be written  $2(2n^2)$ , the product of 2 and the integer  $2n^2$ .

The square of an odd number is odd. To prove this, let 2n + 1 be an odd number (*n* being an integer). Then its square is

$$(2n + 1)^2 = 4n^2 + 4n + 1$$
  
= 2(2n<sup>2</sup> + 2n) + 1

Since  $2n^2 + 2n$  is an integer, we have written the square of our odd number in the form 2m + 1 for some integer *m*, and thus have shown that our square is odd.

We are now ready to prove that the square root of 2 is not a rational number. Suppose it were. This would mean that we can find a rational number a such that  $a^2 = 2$ . We can write

$$a=rac{m}{n}$$
 ,

where m, n are integers, and neither m nor n is 0. Furthermore, we can assume that not both m, n are even because we can put the fraction m/n in lowest form and cancel as many powers of 2 dividing both m and n as possible. Thus we can assume that at least one of the integers m or n is odd.

From our assumption that  $a^2 = 2$  we get  $(m/n)^2 = 2$  or

$$\frac{m^2}{n^2} = 2$$

Multiplying both sides of this equation by  $n^2$  yields

$$m^2 = 2n^2$$

and the right-hand side is even. By what we saw above, this means that m is even and we can therefore write m = 2k for some integer k. Substituting, we obtain

$$(2k)^2 = 2n^2$$

or  $4k^2 = 2n^2$ . We cancel 2 and get  $2k^2 = n^2$ . This means that  $n^2$  is even, and consequently, by what we saw above, that n is even. Thus we have reached the conclusion that both m, n are even, which contradicts the fact that we put our fraction in lowest form. We can therefore conclude that there was no fraction m/n whose square is 2.

It is usually very difficult to determine whether a given number is a rational number or not. For instance, the fact that  $\pi$  is not rational was discovered only at the end of the 18th century.

#### §2. Inequalities

Aside from addition, multiplication, subtraction and division (by numbers other than 0), we shall now discuss another important feature of the real numbers.

We have the *positive numbers*, represented geometrically on the straight line by those numbers unequal to 0 and lying to the right of 0. If a is a positive number, we write a > 0. You have no doubt already worked with positive numbers, and with inequalities. The next two properties are the most basic ones, concerning positivity.

POS 1. If a, b are positive, so is the product ab and the sum a + b.

POS 2. If a is a number, then either a is positive, or a = 0, or -a is positive, and these possibilities are mutually exclusive.

If a number is not positive and not 0, then we say that this number is *negative*. By POS 2, if a is negative then -a is positive.

Although you know already that the number 1 is positive, it can in fact be *proved* from our two properties. It may interest you to see the proof, which runs as follows and is very simple. By POS 2, we know that either 1 or -1 is positive. If 1 is not positive, then -1 is positive. By POS 1, it must then follow that (-1)(-1) is positive. But this product is equal to 1. Consequently, it must be 1 which is positive, and not -1. Using property POS 1, we could now conclude that 1 + 1 = 2 is positive, that 2 + 1 = 3 is positive, and so forth.

If a > 0 we shall also say that a is greater than 0. If we wish to say that a is positive or equal to 0, we write

 $a \ge 0$ 

and read this "a greater than or equal to zero".

Given two numbers a, b we shall say that a is greater than b and write a > b if a - b > 0. We write a < 0 if -a > 0 and a < b if b > a. Thus 3 > 2 because 3 - 2 > 0.

We shall write  $a \ge b$  when we want to say that a is greater than or equal to b. Thus  $3 \ge 2$  and  $3 \ge 3$  are both true inequalities.

Using only our two properties POS 1 and POS 2 we shall now prove all the usual rules concerning inequalities. You probably know these already, but proving them systematically will both sharpen your wits and etch these rules more profoundly in your mind.

In what follows, let a, b, c be numbers.

Rule 1. If a > b and b > c then a > c. Rule 2. If a > b and c > 0 then ac > bc. Rule 3. If a > b and c < 0 then ac < bc.

Rule 2 expresses the fact that an inequality which is multiplied by a positive number is *preserved*. Rule 3 tells us that if we multiply both sides of an inequality by a negative number, then the inequality gets *reversed*. For instance, we have the inequality

1 < 3.

Since 2 > 0 we also have  $2 \cdot 1 < 2 \cdot 3$ . But -2 is negative, and if we multiply both sides by -2 we get

-2 > -6.

In the geometric representation of the real numbers on the line, -2 lies to the right of -6. This gives us the geometric representation of the fact that -2 is greater than -6.

To prove Rule 1, suppose that a > b and b > c. By definition, this means that (a - b) > 0 and (b - c) > 0. Using property POS 1, we conclude that

$$a-b+b-c>0,$$

and canceling b gives us (a - c) > 0. By definition, this means a > c, as was to be shown.

To prove Rule 2, suppose that a > b and c > 0. By definition,

a-b>0.

Hence using the property of POS 1 concerning the product of positive numbers, we conclude that

$$(a-b)c>0.$$

The left-hand side of this inequality is none other than ac - bc, which is therefore >0. Again by definition, this gives us

We leave the proof of Rule 3 as an exercise.

If a is a number, then we define the *absolute value* of a to be:

$$a \text{ itself if } a \text{ is } \geq 0$$
.  
 $-a \text{ if } a \text{ is } < 0$ .

In the second case, when a is negative, then -a is positive. Thus the absolute value of a number is always a positive number, or 0. For instance, the absolute value of 3 is 3 itself. The absolute value of -3 is -(-3) = 3. The absolute value of  $-\frac{1}{2}$  is  $\frac{1}{2}$ . The absolute value of  $\sqrt{2}$  is  $\sqrt{2}$  and the absolute value of  $-\sqrt{2}$  is  $\sqrt{2}$ . It is customary to denote the absolute value of a number by two bars beside the number. Thus the absolute value of a number a is written |a|. For instance, |3| = 3 and |-3| = 3 also. We have by definition |0| = 0.

Let a be a number >0. Then there exists a number whose square is a. This is one of the facts which we take for granted about numbers. If  $b^2 = a$  then we observe that

$$(-b)^2 = b^2$$

is also equal to a. Thus either b or -b is positive. We agree to denote by  $\sqrt{a}$  the positive square root and call it simply the square root of a. Thus  $\sqrt{4}$  is equal to 2 and not -2, even though  $(-2)^2 = 4$ . This is the most practical convention about the use of the  $\sqrt{-3}$  sign that we can make. Of course, the square root of 0 is 0 itself. A negative number does not have a square root.

THEOREM 1. If a is a number, then  $|a|^2 = a^2$  and

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

*Proof.* If a is positive then |a| = a and our first assertion is clear. If a is negative, then |a| = -a and

$$(-a)^2 = a^2,$$

INEQUALITIES

so we again get  $|a|^2 = a^2$ . When a = 0 our first assertion simply means 0 = 0. Finally, taking the (positive) square root, we get

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

THEOREM 2. If a, b are numbers, then

$$|ab| = |a| |b|.$$

This theorem expresses the fact that the absolute value of a product is the product of the absolute values. We shall leave the proof as an exercise. As an example, we see that

$$|-6| = |(-3) \cdot 2| = |-3| |2| = 3 \cdot 2 = 6.$$

There is one final inequality which is extremely important.

THEOREM 3. If a, b are two numbers, then

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

*Proof.* We first observe that either ab is positive, or it is negative, or it is 0. In any case, we have

$$ab \leq |ab| = |a| |b|.$$

Hence, multiplying both sides by 2, we obtain the inequality

$$2ab \leq 2 |a| |b|.$$
  
$$(a+b)^2 = a^2 + 2ab + b^2$$
  
$$\leq a^2 + 2 |a| |b| + b^2$$
  
$$= (|a| + |b|)^2.$$

We can take the square root of both sides and use Theorem 1 to conclude that

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

thereby proving our theorem.

From this we get

You will find plenty of exercises below to give you practice with inequalities. We shall work out some numerical examples to show you the way.

Example 1. Determine the numbers satisfying the equality

$$|x+1| = 2.$$

This equality means that either x + 1 = 2 or -(x + 1) = 2, because the absolute value of x + 1 is either (x + 1) itself or -(x + 1). In the first case, solving for x gives us x = 1, and in the second case, we get -x - 1 = 2 or x = -3. Thus the answer is x = 1 or x = -3.

Example 2. Determine all intervals of numbers satisfying the inequality

$$|x+1| > 2.$$

We distinguish cases.

Case 1.  $x + 1 \ge 0$ . Then |x + 1| = x + 1, and in this case, we must find those x such that  $x + 1 \ge 0$  and x + 1 > 2. The inequality x + 1 > 2 implies the inequality  $x + 1 \ge 0$ . Hence all x such that x + 1 > 2 will do, i.e. all x such that x > 1.

Case 2. x + 1 < 0. Then |x + 1| = -(x + 1), and in this case we must find those numbers x such that x + 1 < 0 and -x - 1 > 2. These inequalities are equivalent with the inequalities x < -1 and x < -3. The set of numbers x satisfying these inequalities is precisely the set of numbers x satisfying x < -3.

Putting our two cases together, we find that the required numbers x are those such that x > 1 or x < -3.

#### Exercises

Determine all intervals of numbers x satisfying the following inequalities.

1. |x| < 32.  $|2x+1| \leq 1$ 3.  $|x^2-2| \leq 1$ 4. |x-5| < |x+1|5. (x+1)(x-2) < 06. (x-1)(x+1) > 07. (x-5)(x+5) < 08.  $x(x+1) \leq 0$ 9.  $x^2(x-1) \geq 0$ 10.  $(x-5)^2(x+10) \leq 0$ 11.  $(x-5)^4(x+10) \leq 0$ 12.  $(2x+1)^6(x-1) \geq 0$ 13.  $(4x+7)^{20}(2x+8) < 0$ 

Prove the following inequalities for all numbers x, y.

14.  $|x + y| \ge |x| - |y|$ 15.  $|x - y| \ge |x| - |y|$ 16.  $|x - y| \le |x| + |y|$ 

#### §3. Functions

A *function* (of numbers) is a rule which to any given number associates another number.

It is customary to denote a function by some letter, just as a letter "x" denotes a number. Thus if we denote a given function by f, and x is a

FUNCTIONS

number, then we denote by f(x) the number associated with x by the function. This of course does not mean "f times x". There is no multiplication involved here. The symbols f(x) are read "f of x".

For example, the rule could be "square the number". For this function, we associate the number  $x^2$  to the number x. If f is the function "square the number", then  $f(x) = x^2$ . In particular, the square of 2 is 4 and hence f(2) = 4. The square of 7 is 49 and f(7) = 49. The square of  $\sqrt{2}$  is 2, and hence  $f(\sqrt{2}) = 2$ . The square of (x + 1) is  $x^2 + 2x + 1$  and thus  $f(x + 1) = x^2 + 2x + 1$ . If h is any number,

$$f(x+h) = x^2 + 2xh + h^2.$$

To take another example, let g be the function "add 1 to the number". Then to each number x we associate the number x + 1. Therefore g(x) = x + 1 and g(1) = 2. Also, g(2) = 3, g(3) = 4,  $g(\sqrt{2}) = \sqrt{2} + 1$ , and g(x + 1) = x + 2 for any number x.

We can view the absolute value as a function, defined by the rule: Given any number a, we associate the number a itself if  $a \ge 0$ , and we associate the number -a if a < 0. Let F denote the absolute value function. Then F(x) = |x| for any number x. We have in particular F(2) = 2, and F(-2) = 2 also. The absolute value is not defined by means of a formula like  $x^2$  or x + 1. We give you another example of such a function which is not defined by a formula.

We consider the function G described by the following rule:

G(x) = 0 if x is a rational number.

G(x) = 1 if x is not a rational number.

Then in particular,  $G(2) = G(\frac{2}{3}) = G(-\frac{3}{4}) = 0$  but

 $G(\sqrt{2}) = 1.$ 

You must be aware that you can construct a function just by prescribing arbitrarily the rule associating a number to a given one.

If f is a function and x a number, then f(x) is called the *value* of the function at x. Thus if f is the function "square the number", the value of f at 2 is 4 and the value of f at  $\frac{1}{2}$  is  $\frac{1}{4}$ .

In order to describe a function, we need simply to give its value at any number x. Thus we usually speak of a function f(x), meaning by that the function f whose value at x is f(x). For instance, we would say "Let f(x) be the function  $x^3 + 5$ " instead of saying "Let f be the function cube the number and add 5". The advantages of speaking and writing in this way are obvious.

[I, §3]

We would also like to be able to define a function for some numbers and leave it undefined for others. For instance we would like to say that  $\sqrt{x}$  is a function (the square root function, whose value at a number x is the square root of that number), but we observe that a negative number does not have a square root. Hence it is desirable to make the notion of function somewhat more precise by stating explicitly for what numbers it is defined. For instance, the square root function is defined only for numbers  $\geq 0$ . This function is denoted by  $\sqrt{x}$ . The value  $\sqrt{x}$  is the unique number  $\geq 0$  whose square is x.

Let us give another example of a function which is not defined for all numbers. The function f(x) = 1/x is defined only for numbers  $\neq 0$ . For this particular function, we have f(1) = 1,  $f(2) = \frac{1}{2}$ ,  $f(\frac{1}{2}) = 2$  and

$$f(\sqrt{2}) = \frac{1}{\sqrt{2}} \cdot$$

One final word before we pass to the exercises: There is no magic reason why we should always use the letter x to describe a function f(x). Thus instead of speaking of the function f(x) = 1/x we could just as well say f(y) = 1/y or f(q) = 1/q. Unfortunately, the most neutral way of writing would be f(blank) = 1/blank, and this is really not convenient.

#### Exercises

1. Let f(x) = 1/x. What is  $f(\frac{3}{4}), f(-\frac{2}{3})$ ?

2. Let f(x) = 1/x again. What is f(2x + 1) (for any number x such that  $x \neq -\frac{1}{2}$ )?

- 3. Let g(x) = |x| x. What is g(1), g(-1), g(-54)?
- 4. Let  $f(y) = 2y y^2$ . What is f(z), f(w)?

5. For what numbers could you define a function f(x) by the formula

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

What is the value of this function for x = 5?

6. For what numbers could you define a function f(x) by the formula  $f(x) = \sqrt[3]{x}$  (cube root of x)? What is f(27)?

 7. Let f(x) = x/|x|, defined for  $x \neq 0$ . What is:

 a. f(1) b. f(2) c. f(-3) d.  $f(-\frac{4}{3})$  

 8. Let f(x) = x + |x|. What is:
 a.  $f(\frac{1}{2})$  b. f(2) c. f(-4) d. f(-5) 

 9. Let  $f(x) = 2x + x^2 - 5$ . What is:
 a. f(1) b. f(-1) c. f(x + 1) 

10. For what numbers could you define a function f(x) by the formula  $f(x) = \sqrt[4]{x}$  (fourth root of x)? What is f(16)?

POWERS

#### §4. Powers

In this section we just summarize some elementary arithmetic.

Let *n* be an integer  $\geq 1$  and let *a* be any number. Then  $a^n$  is the product of *a* with itself *n* times. For example, let a = 3. If n = 2, then  $a^2 = 9$ . If n = 3, then  $a^3 = 27$ . Thus we obtain a function which is called the *n*-th power. If *f* denotes this function, then  $f(x) = x^n$ .

We recall the rule

$$x^{m+n} = x^m x^n$$

for any number x and integers  $m, n \ge 1$ .

Again, let *n* be an integer  $\geq 1$ , and let *a* be a positive number. We define  $a^{1/n}$  to be the unique positive number *b* such that  $b^n = a$ . (That there exists such a unique number *b* is taken for granted as part of the properties of numbers.) We get a function called the *n*-th root. Thus if *f* is the 4-th root, then f(16) = 2 and f(81) = 3.

The *n*-th root function can also be defined at 0, the *n*-th root of 0 being 0 itself.

Question: If n is an odd integer like 1, 3, 5, 7, ..., can you define an n-th root function for all numbers?

If a, b are two numbers  $\geq 0$  and n is an integer  $\geq 1$  then

$$(ab)^{1/n} = a^{1/n}b^{1/n}.$$

There is another useful and elementary rule. Let m, n be integers  $\geq 1$  and let a be a number  $\geq 0$ . We define  $a^{m/n}$  to be  $(a^{1/n})^m$  which is also equal to  $(a^m)^{1/n}$ . This allows us to define fractional powers, and gives us a function  $f(x) = x^{m/n}$ 

defined for  $x \ge 0$ .

We now come to powers with negative numbers or 0. We want to define  $x^a$  when a is a negative rational number or 0 and x > 0. We want the fundamental rule

$$x^{a+b} = x^a x^b$$

to be true. This means that we must define  $x^0$  to be 1. For instance, since

$$2^3 = 2^{3+0} = 2^3 2^0,$$

we see from this example that the only way in which this equation holds is if  $2^0 = 1$ . Similarly, in general, if the relation

$$x^a = x^{a+0} = x^a x^0$$

is true, then  $x^0$  must be equal to 1.

Suppose finally that a is a positive rational number, and let x be a number >0. We define  $x^{-a}$  to be

$$\frac{1}{x^a}$$
.

Thus

$$2^{-3} = \frac{1}{2^3} = \frac{1}{8}$$
, and  $4^{-2/3} = 1/4^{2/3}$ .

We observe that in this special case,

$$(4^{-2/3})(4^{2/3}) = 4^0 = 1.$$

In general,  $x^{a}x^{-a} = x^{0} = 1$ .

We are tempted to define  $x^a$  even when a is not a rational number. This is more subtle. For instance, it is absolutely meaningless to say that

 $2^{\sqrt{2}}$ 

is the product of 2 square root of 2 times itself. The problem of defining  $2^a$  (or  $x^a$ ) when a is not rational will be postponed to a later chapter. Until that chapter, when we deal with such a power, we shall assume that there is a function, written  $x^a$ , described as we have done above for rational numbers, and satisfying the fundamental relation

$$x^{a+b} = x^a x^b$$
$$x^0 = 1.$$

*Example.* We have a function  $f(x) = x^{\sqrt{2}}$  defined for all x > 0. It is actually hard to describe its values for special numbers, like  $2^{\sqrt{2}}$ . It was unknown for a very long time whether  $2^{\sqrt{2}}$  is a rational number or not. The solution (*it is not*) was found only in 1927 by the mathematician Gelfond, who became famous for solving a problem that was known to be very hard.

**Warning.** Do not confuse a function like  $x^2$  and a function like  $2^x$ . Given a number c > 0, we can view  $c^x$  as a function defined for all x. (It will be discussed in detail in Chapter VIII.) This function is called an *exponential function*. Thus  $2^x$  and  $10^x$  are exponential functions. We shall select a number

$$e = 2.718...$$

and the exponential function  $e^x$  as having special properties which make it better than any other exponential function. The meaning of our use of the word "better" will be explained in Chapter VIII.

#### POWERS

#### EXERCISES

Find  $a^x$  and  $x^a$  for the following values of x and a.

1. $a = 2$ and $x = 3$	2. $a = 5$ and $x = -1$
3. $a = \frac{1}{2}$ and $x = 4$	4. $a = \frac{1}{3}$ and $x = 2$
5. $a = -\frac{1}{2}$ and $x = 4$	6. $a = 3$ and $x = 2$
7. $a = -3$ and $x = -1$	8. $a = -2$ and $x = -2$
9. $a = -1$ and $x = -4$	10. $a = -\frac{1}{2}$ and $x = 9$

11. Determine whether  $2^{\sqrt{2}} + 3^{\sqrt{3}}$  is a rational number. (This is actually a major research problem whose answer is not known today. Later in this course we shall deal with numbers e and  $\pi$ . Although it is known that neither e nor  $\pi$  is rational, it is not known whether  $e\pi$  or  $e + \pi$  is rational.)

# CHAPTER II

# **Graphs and Curves**

The ideas contained in this chapter allow us to translate certain statements backwards and forwards between the language of numbers and the language of geometry.

It is extremely basic for what follows, because we can use our geometric intuition to help us solve problems concerning numbers and functions, and conversely, we can use theorems concerning numbers and functions to yield results about geometry.

#### **§1.** Coordinates

Once a unit length is selected, we can represent numbers as points on a line. We shall now extend this procedure to the plane, and to pairs of numbers.

We visualize a horizontal line and a vertical line intersecting at an origin O.



These lines will be called *coordinate axes* or simply *axes*.

We select a unit length and cut the horizontal line into segments of lengths 1, 2, 3,  $\ldots$  to the left and to the right, and do the same to the vertical line, but up and down, as indicated in the next figure.

On the vertical line we visualize the points going below 0 as corresponding to the negative integers, just as we visualized points on the left of the

14



horizontal line as corresponding to negative integers. We follow the same idea as that used in grading a thermometer, where the numbers below zero are regarded as negative.

We can now cut the plane into squares whose sides have length 1.



We can describe each point where two lines intersect by a pair of integers. Suppose that we are given a pair of integers like (1, 2). We go to the right of the origin 1 unit and vertically up 2 units to get the point (1, 2) which has been indicated above. We have also indicated the point (3, 4). The diagram is just like a map.

Furthermore, we could also use negative numbers. For instance to describe the point (-3, -2) we go to the left of the origin 3 units and vertically downwards 2 units.

There is actually no reason why we should limit ourselves to points which are described by integers. For instance we can also have the point  $(\frac{1}{2}, -1)$  and the point  $(-\sqrt{2}, 3)$  as on the next figure. We have not drawn all the squares on the plane. We have drawn only the relevant lines to find our two points.



In general, if we take any point P in the plane and draw the perpendicular lines to the horizontal axis and to the vertical axis, we obtain two numbers x, y as in the figure below.



The perpendicular line from P to the horizontal axis determines a number x which is negative in the figure because it lies to the left of the origin. The number y determined by the perpendicular from P to the vertical axis is positive because it lies above the origin. The two numbers x, y are called the *coordinates* of the point P, and we can write P = (x, y).

Every pair of numbers (x, y) determines a point of the plane. We find the point by going a distance x from the origin 0 in the horizontal direction and then a distance y in the vertical direction. If x is positive we go to the right of 0. If x is negative, we go to the left of 0. If y is positive we go vertically upwards, and if y is negative we go vertically downwards.

The coordinates of the origin are (0, 0).

We usually call the horizontal axis the *x*-axis and the vertical axis the *y*-axis. If a point P is described by two numbers, say (5, -10), it is customary to call the first number its *x*-coordinate and the second number its *y*-coordinate. Thus 5 is the *x*-coordinate, and -10 the *y*-coordinate of our point.

Of course, we could use other letters besides x and y, for instance t and s, or u and v.

[II, §2]

GRAPHS

Our two axes separate the plane into four quadrants which are numbered as indicated in the figure:



If (x, y) is a point in the first quadrant, then both x and y are >0. If (x, y) is a point in the fourth quadrant, then x > 0 but y < 0.

#### Exercises

- 1. Plot the following points: (-1, 1), (0, 5), (-5, -2), (1, 0).
- 2. Plot the following points:  $(\frac{1}{2}, 3), (-\frac{1}{3}, -\frac{1}{2}), (\frac{4}{3}, -2), (-\frac{1}{4}, -\frac{1}{2}).$

3. Let (x, y) be the coordinates of a point in the second quadrant. Is x positive or negative? Is y positive or negative?

4. Let (x, y) be the coordinates of a point in the third quadrant. Is x positive or negative? Is y positive or negative?

- 5. Plot the following points: (1.2, -2.3), (1.7, 3).
- 6. Plot the following points:  $(-2.5, \frac{1}{3}), (-3.5, \frac{5}{4}).$
- 7. Plot the following points: (1.5, -1), (-1.5, -1).

#### §2. Graphs

Let f be a function. We define the graph of f to be the collection of all pairs of numbers (x, f(x)) whose first coordinate is any number for which f is defined and whose second coordinate is the value of the function at the first coordinate.

For example, the graph of the function  $f(x) = x^2$  consists of all pairs (x, y) such that  $y = x^2$ . In other words, it is the collection of all pairs  $(x, x^2)$ , like (1, 1), (2, 4), (-1, 1), (-3, 9), etc.

Since each pair of numbers corresponds to a point on the plane (once a system of axes and a unit length have been selected), we can view the graph of f as a collection of points in the plane. The graph of the function

 $f(x) = x^2$  has been drawn in the figure below, together with the points which we gave above as examples.



To determine the graph, we plot a lot of points making a table giving the x- and y-coordinates.

x	f(x)	x	f(x)
1	1	-1	1
<b>2</b>	4	-2	4
3	9	-3	9
$\frac{1}{2}$	$\frac{1}{4}$	$-\frac{1}{2}$	$\frac{1}{4}$

At this stage of the game there is no other way for you to determine the graph of a function other than this trial and error method. Later, we shall develop techniques which give you greater efficiency in doing it.

We shall now give several examples of graphs of functions which occur very frequently in the sequel.

*Example* 1. Consider the function f(x) = x. The points on its graph are of type (x, x). The first coordinate must be equal to the second. Thus  $f(1) = 1, f(-\sqrt{2}) = -\sqrt{2}$ , etc. The graph looks like this:



*Example 2.* Let f(x) = -x. Its graph looks like this:



Observe that the graphs of the preceding two functions are straight lines. We shall study the general case of a straight line later.

*Example 3.* Let f(x) = |x|. When  $x \ge 0$ , we know that f(x) = x. When  $x \le 0$ , we know that f(x) = -x. Hence the graph of |x| is obtained by combining the preceding two, and looks like this:



All values of y are  $\geq 0$ , whether x is positive or negative.

*Example* 4. There is an even simpler type of function than the ones we have just looked at, namely the constant functions. For instance, we can define a function f such that f(x) = 2 for all numbers x. The rule is to associate the number 2 to any number x. It is a very simple rule, and the graph of this function is a horizontal line, intersecting the vertical axis at the point (0, 2).



If we took the function f(x) = -1, then the graph would be a horizontal line intersecting the vertical axis at the point (0, -1).

In general, let c be a fixed number. The graph of the function f(x) = c is the horizontal line intersecting the vertical axis at the point (0, c). The function f(x) = c is called a *constant* function.

*Example 5.* The last of our examples is the function f(x) = 1/x (defined for  $x \neq 0$ ). By plotting a few points of the graph, you will see that it looks like this:



For instance, you can plot the following points:

x	1/x	x	1/x
1	1	1	1
<b>2</b>	$\frac{1}{2}$	-2	$-\frac{1}{2}$
3	$\frac{1}{3}$	-3	$-\frac{1}{3}$
$\frac{1}{2}$	2	$-\frac{1}{2}$	-2
$\frac{1}{3}$	3	$-\frac{1}{3}$	<b>3</b>

As x becomes very large positive, 1/x becomes very small. As x approaches 0 from the right, 1/x becomes very large. A similar phenomenon occurs when x approaches 0 from the left; then x is negative and 1/x is negative also. Hence in that case, 1/x is very large negative.

In trying to determine how the graph of a function looks, you can already watch out for the following:

The points at which the graph intersects the two coordinate axes.

What happens when x becomes very large positive and very large negative.

On the whole, however, in working out the exercises, your main technique is just to plot a lot of points until it becomes clear to you what the graph looks like.

#### Exercises

Sketch the graphs of the following functions and plot at least three points on each graph.

1. $x + 1$	2. $2x$	3. $3x$
4. 4 <i>x</i>	5. $2x + 1$	6. $5x + \frac{1}{2}$
7. $\frac{x}{2} + 3$	8. $-3x + 2$	9. $2x^2 - 1$
10. $-3x^2 + 1$	11. $x^3$	12. $x^4$
13. $\sqrt{x}$	14. $x^{-1/2}$	15. $2x$
16. $x + 1$	17. $ x  + x$	18. $ x  + 2x$
19. $- x $	20. $- x  + x$	21. $\frac{1}{x+2}$
22. $\frac{1}{x-2}$	23. $\frac{1}{x+3}$	24. $\frac{1}{x-3}$
25. $\frac{2}{x-2}$	26. $\frac{2}{x+2}$	27. $\frac{2}{x}$
28. $\frac{-2}{x+5}$	29. $\frac{3}{x+1}$	30. $\frac{x}{ x }$

(In Exercises 13, 14, and 21 through 30, the functions are not defined for all values of x.)

31. Sketch the graph of the function f(x) such that: f(x) = 0 if  $x \leq 0$ .

f(x) = 0 if  $x \equiv 0$ . f(x) = 1 if x > 0.

32. Sketch the graph of the function f(x) such that:

f(x) = x if x < 0. f(0) = 2.

 $f(x) = x \quad \text{if} \quad x > 0.$ 

- 33. Sketch the graph of the function f(x) such that:  $f(x) = x^2$  if x < 0. f(x) = x if  $x \ge 0$ .
- 34. Sketch the graph of the function f(x) such that:

f(x) = |x| + x if  $-1 \le x \le 1$ .

f(x) = 3 if x > 1. (f(x) is not defined for other values of x.)

35. Sketch the graph of the function f(x) such that:  $f(x) = x^3$  if  $x \leq 0$ . f(x) = 1 if 0 < x < 2.  $f(x) = x^2$  if x > 2

$$J(x) = x^2 \quad \text{if} \quad x \leq 2.$$

36. Sketch the graph of the function f(x) such that:

f(x) = x if  $0 < x \le 1$ . f(x) = x - 1 if  $1 < x \le 2$ .

f(x) = x - 2 if  $2 < x \le 3$ . f(x) = x - 3 if  $3 < x \le 4$ .

(We leave f(x) undefined for other values of x, but try to define it yourself in such a way as to preserve the symmetry of the graph.)

#### §3. The straight line

One of the most basic types of functions is the type whose graph represents a straight line. We have already seen that the graph of the function f(x) = x is a straight line. If we take f(x) = 2x, then the line slants up much more steeply, and even more so for f(x) = 3x. The graph of the function f(x) = 10,000x would look almost vertical. In general, let *a* be a positive number  $\neq 0$ . Then the graph of the function

$$f(x) = ax$$

represents a straight line. The point (2, 2a) lies on the line because f(2) = 2a. The point  $(\sqrt{2}, \sqrt{2}a)$  also lies on the line, and if c is any number, the point (c, ca) lies on the line. The (x, y) coordinates of these points are obtained by making a similarity transformation, starting with the coordinates (1, a) and multiplying them by some number c.

We can visualize this procedure by means of similar triangles. In the figure below, we have a straight line. If we select a point (x, y) on the line and drop the perpendicular from this point to the x-axis, we obtain a right triangle.



If x is the length of the base of the smaller triangle in the figure, and y its height, and if cx is the length of the base of the bigger triangle, then cy is the height of the bigger triangle: The smaller triangle is similar to the bigger one.

If a is a number <0, then the graph of the function f(x) = ax is also a straight line, which slants to the left. For instance, the graphs of f(x) = -x or f(x) = -2x.

Let a, b be two numbers. The graph of the function g(x) = ax + b is also a straight line, which is parallel to the line determined by the function f(x) = ax. In order to convince you of this, we observe the following. When x = 0 we see that g(x) = b. Let y' = y - b. The equation y' = ax is of the type discussed above. If we have a point (x, y') on the straight line y' = ax, then we get a point (x, y' + b) on the straight line y = ax + b, by simply adding b to the second coordinate. This means that the graph of the function g(x) = ax + b is the straight line parallel to the line determined by the function f(x) = ax and passing through the point (0, b).

Example 1. Let g(x) = 2x + 1. When x = 0, then g(x) = 1. The graph looks like this:

(0, 1)





We shall frequently speak of a function f(x) = ax + b as a straight line (although of course, it is its graph which is a straight line).

The number a which is the coefficient of x is called the *slope* of the line. It determines how much the line is slanted. As we have already seen in many examples, when the slope is positive, the line is slanted to the right, and when the slope is negative, the line is slanted to the left. The relationship y = ax + b is also called the *equation* of the line. It gives us the relation between the x- and y-coordinates of a point on the line.

Let f(x) = ax + b be a straight line, and let  $(x_1, y_1)$  and  $(x_2, y_2)$  be two points of the line. It is easy to find the slope of the line in terms of the coordinates of these two points. By definition, we know that

$$y_1 = ax_1 + b$$
$$y_2 = ax_2 + b.$$

and

[II, §3]

Subtracting, we get

$$y_2 - y_1 = ax_2 - ax_1 = a(x_2 - x_1).$$

Consequently, if the two points are distinct,  $x_2 \neq x_1$ , then we can divide by  $x_2 - x_1$  and obtain

$$a = \frac{y_2 - y_1}{x_2 - x_1}$$

This formula gives us the slope in terms of the coordinates of two distinct points on the line.

*Example 3.* Look at the line f(x) = 2x + 5. Letting x = 1, we have f(x) = 7 and letting x = -1, we get f(x) = 3. Thus the points (1, 7) and (-1, 3) are on the line. The slope is 2, and is equal to

$$\frac{7-3}{1-(-1)}$$

as it should be.

Geometrically, our quotient

$$\frac{y_2 - y_1}{x_2 - x_1}$$

is simply the ratio of the vertical side and horizontal side of the triangle in the next diagram:



Conversely, given two points in the plane, it is easy to determine the equation of the line passing through them.

*Example* 4. Let (1, 2) and (2, -1) be the two points. What is the slope of the line between them? What is the equation of the line?

We first find the slope. It must be the quotient

$$\frac{y_2 - y_1}{x_2 - x_1}$$

[II, §3]
[II, §3]

which is equal to

$$\frac{-1-2}{2-1} = -3.$$

Thus we know that our line is given by the equation

y = -3x + b

for some number b. Furthermore, we know also that the line must pass through the point (1, 2). If f(x) = -3x + b, then we must have f(1) = 2. From this we can solve for b, namely

$$2 = -3 \cdot 1 + b$$

yields b = 2 + 3 = 5. Thus the equation of the line is

$$f(x) = -3x + 5.$$

Observe that it does not matter which point we call  $(x_1, y_1)$  and which we call  $(x_2, y_2)$ . We would get the same answer for the slope.

Knowing two points on a line, we first determine the slope and then solve for the constant b, using the coordinates of one of the points.

We can also determine the equation of a line provided we know the slope and one point.

*Example* 5. Find the equation of the line having slope -7 and passing through the point (-1, 2).

The equation must be of type

$$y = -7x + b$$

for some number b. Furthermore, when x = -1, we must have y = 2. Thus

$$2 = (-7)(-1) + b$$

and b = -5. Hence the equation of the line is

$$y = -7x - 5.$$

Finally, we should mention vertical lines. These cannot be represented by equations of type y = ax + b. Suppose that we have a vertical line intersecting the x-axis at the point (2, 0). The y-coordinate of any point on the line can be arbitrary. Thus the equation of the line is simple x = 2. In general, the equation of the vertical line intersecting the x-axis at the point (c, 0) is x = c.

#### Exercises

Sketch the graphs of the following lines:

1. 
$$y = -2x + 5$$
 2.  $y = 5x - 3$ 

 3.  $y = \frac{x}{2} + 7$ 
 4.  $y = -\frac{x}{3} + 1$ 

What is the equation of the line passing through the following points?

5. (-1, 1) and (2, -7)6.  $(3, \frac{1}{2})$  and (4, -1)7.  $(\sqrt{2}, -1)$  and  $(\sqrt{2}, 1)$ 8. (-3, -5) and  $(\sqrt{3}, 4)$ 

What is the equation of the line having the given slope and passing through the given point?

9. slope 4 and point $(1, 1)$	10. slope $-2$ and point $(\frac{1}{2}, 1)$
11. slope $-\frac{1}{2}$ and point $(\sqrt{2}, 3)$	12. slope $\sqrt{3}$ and point (-1, 5)

Sketch the graphs of the following lines:

13.	x =	5	14. $x$	= -1	15. x	; =	-3
16.	y =	-4	17. y	= 2	18. y	· =	0.

### §4. Distance between two points

Let  $(x_1, y_1)$  and  $(x_2, y_2)$  be two points in the plane, for instance as in the following diagrams.



We can then make up a right triangle. By the Pythagoras theorem, the length of the line segment joining our two points can be determined from the lengths of the two sides. The square of the bottom side is  $(x_2 - x_1)^2$ , which is also equal to  $(x_1 - x_2)^2$ . This is clear in part (a) of the figure. It is also true in part (b) (convince yourself by working out an example, as for instance in Example 2 below). The length of the vertical side is  $(y_2 - y_1)^2$ , which is equal to  $(y_1 - y_2)^2$ . If L denotes the length of the line segment, then

$$L^{2} = (x_{1} - x_{2})^{2} + (y_{1} - y_{2})^{2}$$

[II, §4]

[II, §5]

and consequently,

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

*Example* 1. Let the two points be (1, 2) and (1, 3). Then the length of the line segment between them is

$$\sqrt{(1-1)^2 + (3-2)^2} = 1.$$

The length L is also called the *distance* between the two points.

*Example 2.* Find the distance between the points (-1, 5) and (4, -3).

The distance is

$$\sqrt{(4 - (-1))^2 + (-3 - 5)^2} = \sqrt{89}.$$

(You should plot these points, and convince yourself that the minus signs do not affect the validity of our formula for the length of the line segment between the two points.)

#### Exercises

Find the distance between the following points:

- 1. The points (-3, -5) and (1, 4)
- 2. The points (1, 1) and (0, 2)
- 3. The points (-1, 4) and (3, -2)
- 4. The points (1, -1) and (-1, 2)
- 5. The points  $(\frac{1}{2}, 2)$  and (1, 1)

6. Find the coordinates of the fourth corner of a rectangle, three of whose corners are (-1, 2), (4, 2), (-1, -3).

7. What are the lengths of the sides of the rectangle in Exercise 6?

### **§5.** Curves and equations

Let F(x, y) be an expression involving a pair of numbers (x, y). Let c be a number. We consider the equation

$$F(x, y) = c$$

The collection of points (a, b) in the plane satisfying this equation, i.e. such that

$$F(a, b) = c,$$

is called the *graph* of the equation. This graph is also known as a curve, and we will usually not make a distinction between the equation

$$F(x,y) = c$$

and the curve which represents the equation.

For example,

$$x + y = 2$$

is the equation of a straight line, and its graph is the straight line. We shall study below important examples of equations which arise frequently.

If f is a function, then we can form the expression y - f(x), and the graph of the equation

$$y-f(x)=0$$

is none other than the graph of the function f as we discussed it in §2.

You should observe that there are equations of type

$$F(x, y) = c$$

which are not obtained from a function y = f(x). For instance, the equation  $x^2 + y^2 = 1$  is such an equation.

We shall now study important examples of graphs of equations

$$F(x, y) = 0$$
 or  $F(x, y) = c$ .

# §6. The circle

The function  $F(x, y) = x^2 + y^2$  has a simple geometric interpretation. It is the square of the distance of the point (x, y) from the origin (0, 0). Thus the points (x, y) satisfying the equation

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

are simply those points whose distance from the origin is 1. It is the circle of radius 1, with center at the origin.

Similarly, the points (x, y) satisfying the equation

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

are those points whose distance from the origin is 2. They constitute the circle of radius 2. In general, if c is any number >0, then the graph of the

[II, §6]

THE CIRCLE

equation

 $x^2 + y^2 = c^2$ 

is the circle of radius c, with center at the origin.

We have already remarked that the equation

 $x^2 + y^2 = 1$ 

or  $x^2 + y^2 - 1 = 0$  is not of the type y - f(x) = 0, i.e. does not come from a function y = f(x). However, we can write our equation in the form

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

For any value of x between -1 and +1, we can solve for y and get

$$y = \sqrt{1 - x^2}$$
 or  $y = -\sqrt{1 - x^2}$ 

If  $x \neq 1$  or  $x \neq -1$ , then we get two values of y for each value of x. Geometrically, these two values correspond to the points indicated on the following diagram.



There is a function, defined for  $-1 \leq x \leq 1$ , such that

$$f(x)=\sqrt{1-x^2},$$

and the graph of this function is the upper half of our circle. Similarly, there is another function

$$g(x) = -\sqrt{1-x^2},$$

also defined for  $-1 \leq x \leq 1$ , whose graph is the lower half of the circle. Neither of these functions is defined for other values of x.

We now ask for the equation of the circle whose center is (1, 2) and whose radius has length 3. It consists of the points (x, y) whose distance from (1, 2) is 3. These are the points satisfying the equation

$$(x-1)^2 + (y-2)^2 = 9.$$





The graph of this equation has been drawn above.

To pick another example, we wish to determine those points at a distance 2 from the point (-1, -3). They are the points (x, y) satisfying the equation

$$(x - (-1))^2 + (y - (-3))^2 = 4$$

or, in other words,

$$(x+1)^2 + (y+3)^2 = 4.$$

(Observe carefully the cancellation of minus signs!) Thus the graph of this equation is the circle of radius 2 and center (-1, -3).

In general, let a, b be two numbers and r a number >0. Then the circle of radius r and center (a, b) is the graph of the equation

$$(x-a)^2 + (y-b)^2 = r^2$$

In our last example, we have

$$r = 2, \quad a = -1, \quad b = -3.$$

*Example* 1. Suppose that we have a quadratic expression like

$$x^{2} + 2x$$
.

We can complete the square and write it as

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

Similarly, given the expression  $y^2 - 3y$ , we can write it

$$y^2 - 3y = (y - \frac{3}{2})^2 - \frac{9}{4}.$$

Given an equation

$$x^2 + y^2 + 2x - 3y - 5 = 0$$

we can use the trick of completing the square to see what its graph looks

like. Our equation can be written in the form

$$(x+1)^2 + (y-\frac{3}{2})^2 = 5 + \frac{9}{4} + 1 = \frac{33}{4}.$$

Thus our equation is a circle of center  $(-1, \frac{3}{2})$  and radius  $\sqrt{33/4}$ .

#### EXERCISES

Sketch the graph of the following equations:

1. $(x-2)^2 + (y+1)^2 = 25$	2. $x^2 + (y - 1)^2 = 9$
3. $(x+1)^2 + y^2 = 3$	$4. \ \frac{x^2}{9} + \frac{y^2}{16} = 1$
5. $\frac{x^2}{4} + \frac{y^2}{9} = 1$	$6. \ \frac{x^2}{5} + \frac{y^2}{16} = 1$
7. $\frac{x^2}{4} + \frac{y^2}{25} = 1$	8. $\frac{(x-1)^2}{9} + \frac{(y+2)^2}{16} = 1$
9. $4x^2 + 25y^2 = 100$	10. $\frac{(x+1)^2}{3} + \frac{(y+2)^2}{4} = 1$
11. $25x^2 + 16y^2 = 400$	12. $(x-1)^2 + \frac{(y+3)^2}{4} = 1$

(In Exercises 4 through 12, the graph of the equation is called an *ellipse*. It is a stretched-out circle. Investigate for yourself the effect of changing the coefficients of  $x^2$  and  $y^2$  in these equations.)

### §7. The parabola. Changes of coordinates

We have seen what the graph of the equation  $y = x^2$  looks like. Suppose that we graph the equation  $y = (x - 1)^2$ . We shall find that it looks exactly the same, but as if the origin were placed at the point (1, 0).

Similarly, the curve  $y - 2 = (x - 4)^2$  looks again like  $y = x^2$  except that the whole curve has been moved as if the origin were the point (4, 2). The graphs of these equations have been drawn on the next diagram.

We can formalize these remarks as follows. Suppose that in our given coordinate system we pick a point (a, b) as a new origin. We let new coordinates be x' = x - a and y' = y - b. Thus when x = a we have x' = 0 and when y = b we have y' = 0. If we have a curve

$$y' = x'^2$$

in the new coordinate system whose origin is at the point (a, b), then it gives rise to the equation

$$(y-b) = (x-a)^2$$

[II, §7]



in terms of the old coordinate system. This type of curve is known as a *parabola*.

We can apply the same technique of completing the square that we did for the circle.

*Example.* What is the graph of the equation

$$2y - x^2 - 4x + 6 = 0?$$

Completing the square, we can write

$$x^2 + 4x = (x+2)^2 - 4.$$

 $2y = (x+2)^2 - 10$ 

Thus our equation can be rewritten

or

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

We choose a new coordinate system

$$x' = x + 2$$
 and  $y' = y + 5$ 

so that our equation becomes

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

or  $y' = \frac{1}{2}x'^2$ . This is a function whose graph you already know, and whose sketch we leave to you.

Finally, we remark that if we have an equation

$$x - y^2 = 0$$

or  $x = y^2$ , then we get a parabola which is tilted horizontally. (Draw the

graph yourself.) We can then apply the technique of changing the coordinate system to see what the graph of a more general equation is like, for instance the graph of

$$x - y^2 + 2y + 5 = 0.$$

#### Exercises

Sketch the graph of the following equations:

1. 
$$y = -x + 2$$
  
3.  $x - 4y^2 = 0$   
2.  $y = 2x^2 + x - 3$   
4.  $x - y^2 + y + 1 = 0$ 

Complete the square in the following equations and change the coordinate system to put them into the form

$$x'^2 + y'^2 = r^2$$

or

 $y' = cx'^2$ 

or

$$x' = cy'^2$$

with a suitable constant c.

5. $x^2 + y^2 - 4x + 2y - 20 = 0$	$6. \ x^2 + y^2 - 2y - 8 = 0$
7. $x^2 + y^2 + 2x - 2 = 0$	$8. \ y - 2x^2 - x + 3 = 0$
9. $y - x^2 - 4x - 5 = 0$	$10. \ y - x^2 + 2x + 3 = 0$
$11. \ x^2 + y^2 + 2x - 4y = -3$	12. $x^2 + y^2 - 4x - 2y = -3$
$13. \ x - 2y^2 - y + 3 = 0$	14. $x - y^2 - 4y = 5$

# §8. The hyperbola

We have already seen what the graph of the equation

$$xy = 1$$

looks like. It is of course the same as the graph of the function f(x) = 1/x (defined for  $x \neq 0$ ). If we pick a coordinate system whose origin is at the point (a, b), the equation

$$y-b=\frac{1}{x-a}$$

is known as a hyperbola. In terms of the new coordinate system x' = x - aand y' = y - b, our hyperbola has the old type of equation

$$x'y'=1.$$

1

If we are given an equation like

$$xy - 2x + 3y + 4 = 5,$$

we can factor the left-hand side and rewrite the equation as

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

or

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

In terms of the coordinate system x' = x + 3 and y' = y - 2, we get the equation

$$x'y'=-5.$$

The graph of this equation has been drawn on the following diagram.



EXERCISES

Sketch the graphs of the following curves:

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

 3.  $xy - 4 = 0$ 
 4.  $y = \frac{2}{1-x}$ 

 5.  $y = \frac{1}{x+1}$ 
 6.  $(x+2)(y-1) = 1$ 

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

 9.  $y = \frac{1}{x-2} + 4$ 
 10.  $y = \frac{1}{x+1} - 2$ 

 11.  $y = \frac{4x-7}{x-2}$ 
 12.  $y = \frac{-2x-1}{x+1}$ 

 13.  $y = \frac{x+1}{x-1}$ 
 14.  $y = \frac{x-1}{x+1}$ 

# CHAPTER III

# The Derivative

The two fundamental notions of this course are those of the derivative and the integral. We take up the first one in this chapter.

The derivative will give us the slope of a curve at a point. It has also applications to physics, where it can be interpreted as the rate of change.

We shall develop some basic techniques which will allow you to compute the derivative in all the standard situations which you are likely to encounter in practice.

# §1. The slope of a curve

Consider a curve, and take a point P on the curve. We wish to define the notions of slope of the curve at that point, and tangent line to the curve at that point. Sometimes the statement is made that the tangent to the curve at the point is the line which touches the curve only at that point. This is pure nonsense, as the subsequent pictures will convince you.



In Figs. 1, 2, and 3, we look at the tangent line to the curve at the point P. In Fig. 1 the line cuts the curve at the other point Q. In Fig. 2 the line is also tangent to the curve at the point Q. In Fig. 3 the curve is supposed to be very flat near the point P, and the horizontal line cuts the curve at P, but we would like to say that it is tangent to the curve at P if the curve is very flat. The vertical line cuts the curve only at P, but is not tangent.



FIGURE 3

Observe also that you cannot get out of the difficulties by trying to distinguish a line "cutting" the curve, or "touching the curve", or by saying that the line should lie on one side of the curve (cf. Fig. 1).

We therefore have to give up the idea of touching the curve only at one point, and look for another idea.

We have to face two problems. One of them is to give the correct geometric idea which allows us to define the tangent to the curve, and the other is to test whether this idea allows us to compute effectively this tangent line when the curve is given by a simple equation with numerical coefficients. It is a remarkable thing that our solution of the first problem will in fact give us a solution to the second.

In the first chapter, we have seen that knowing the slope of a straight line and one point on the straight line allows us to determine the equation of the line. We shall therefore define the slope of a curve at a point and then get its tangent afterward by using the method of Chapter II.

Our examples show us that to define the slope of the curve at P, we should not consider what happens at a point Q which is far removed from P. Rather, it is what happens near P which is important.

Let us therefore take any point Q on the given curve y = f(x), and assume that  $Q \neq P$ . Then the two points P, Q determine a straight line with a certain slope which depends on P, Q and which we shall write as S(P,Q). Suppose that the point Q approaches the point P on the curve (but stays distinct from P). Then, as Q comes near P, the slope S(P,Q)of the line passing through P and Q should approach the (unknown) slope of the (unknown) tangent line to the curve at P. In the following diagram, we have drawn the tangent line to the curve at P and two lines between Pand another point on the curve close to P (Fig. 4). The point  $Q_2$  is closer to P on the curve and so the slope of the line between P and  $Q_1$ .

If the limit of the slope S(P, Q) exists as Q approaches P, then it should be regarded as the slope of the curve itself at P. This is the basic idea



FIGURE 4

behind our definition of the slope of the curve at P. We take it as a definition, perhaps the most important definition in this book. To repeat:

Given a curve y = f(x), let P be a point on the curve. The *slope* of the curve at P is the limit of the slope of lines between P and another point Q on the curve, as Q approaches P.

The idea of defining the slope in this manner was discovered in the seventeenth century by Newton and Leibnitz. We shall see that this definition allows us to determine the slope effectively in practice.

First we observe that when y = ax + b is a straight line, then the slope of the line between any two distinct points on the curve is always the same, and is the slope of the line as we defined it in the preceding chapter.

Let us now look at the next simplest example,

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

We wish to determine the slope of this curve at the point (1, 1).

We look at a point near (1, 1), for instance a point whose x-coordinate is 1.1. Then  $f(1.1) = (1.1)^2 = 1.21$ . Thus the point (1.1, 1.21) lies on the curve. The slope of the line between two points  $(x_1, y_1)$  and  $(x_2, y_2)$  is

$$\frac{y_2-y_1}{x_2-x_1}$$

Therefore the slope of the line between (1, 1) and (1.1, 1.21) is

$$\frac{1.21 - 1}{1.1 - 1} = \frac{0.21}{0.1} = 2.1.$$

In general, the x-coordinate of a point near (1, 1) can be written 1 + h, where h is some small number, positive or negative, but  $h \neq 0$ . We have

$$f(1 + h) = (1 + h)^2 = 1 + 2h + h^2.$$

Thus the point  $(1 + h, 1 + 2h + h^2)$  lies on the curve. When h is positive, the line between our two points would look like that in Fig. 5.



FIGURE 5

When h is negative, then 1 + h is smaller than 1 and the line would look like this:



FIGURE 6

For instance, h could be -0.1 and 1 + h = 0.9.

The slope of the line between our two points is therefore the quotient

$$rac{(1+2h+h^2)-1}{(1+h)-1},$$
 $rac{2h+h^2}{h}=2+h.$ 

which is equal to

When the point whose x-coordinate is 
$$1 + h$$
 approaches our point  $(1, 1)$ , the number h approaches 0. As h approaches 0, the slope of the line between our two points approaches 2, which is therefore the slope of the curve at the point  $(1, 1)$  by definition.

You will appreciate how simple the computation turns out to be, and how easy it was to get this slope!

Let us take another example. We wish to find the slope of the same curve  $(fx) = x^2$  at the point (-2, 4). Again we take a nearby point whose

x-coordinate is -2 + h for small  $h \neq 0$ . The y-coordinate of this nearby point is

$$f(-2+h) = (-2+h)^2 = 4 - 4h + h^2.$$

The slope of the line between the two points is therefore

$$\frac{4-4h+h^2-4}{-2+h-(-2)}=\frac{-4h+h^2}{h}=-4+h.$$

As h approaches 0, the nearby point approaches the point (-2, 4) and we see that the slope approaches -4.

#### Exercises

Find the slopes of the following curves at the indicated points:

- 1.  $y = 2x^2$  at the point (1, 2)
- 2.  $y = x^2 + 1$  at the point (-1, 2)
- 3. y = 2x 7 at the point (2, -3)
- 4.  $y = x^3$  at the point  $(\frac{1}{2}, \frac{1}{8})$
- 5. y = 1/x at the point  $(2, \frac{1}{2})$
- 6.  $y = x^2 + 2x$  at the point (-1, -1)

### §2. The derivative

We continue to consider the function  $y = x^2$ . Instead of picking a definite numerical value for the x-coordinate of a point, we could work at an arbitrary point on the curve. Its coordinates are then  $(x, x^2)$ . We write the x-coordinate of a point nearby as x + h for some small number h, positive or negative, but  $h \neq 0$ . The y-coordinate of this nearby point is

$$(x+h)^2 = x^2 + 2xh + h^2.$$

Hence the slope of the line between them is

$$\frac{(x+h)^2 - x^2}{(x+h) - x} = \frac{x^2 + 2xh + h^2 - x^2}{x+h - x}$$
$$= \frac{2xh + h^2}{h}$$
$$= 2x + h.$$

As h approaches 0, 2x + h approaches 2x. Consequently, the slope of the curve  $y = x^2$  at an arbitrary point (x, y) is 2x. In particular, when x = 1 the slope is 2 and when x = -2 the slope is -4, as we found out

before by the explicit computation using the special x-coordinates 1 and -2.

This time, however, we have found out a general formula giving us the slope for any point on the curve. Thus when x = 3 the slope is 6 and when x = -10 the slope is -20.

The example we have just worked out gives us the procedure for treating more general functions.

Given a function f(x), we form the quotient

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

This quotient is the slope of the line between the points

$$(x, f(x))$$
 and  $(x+h, f(x+h))$ .

We shall call it the Newton quotient. If it approaches a limit as h approaches 0, then this limit is called the *derivative* of f at x, and we say that f is *dif*ferentiable at x. The limit will be written in an abbreviated fashion,

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

The derivative will be written f'(x), and we thus have

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

The derivative may thus be viewed as a function f', which is defined at all numbers x such that the Newton quotient approaches a limit as h tends to 0.

We say that f is differentiable if it is differentiable at all points for which it is defined. For instance the function  $f(x) = x^2$  is differentiable and its derivative is 2x.

It will also be convenient to use another notation for the derivative, namely

$$f'(x) = \frac{df}{dx}$$

(or df/dx). Thus the two expressions f'(x) and df/dx mean the same thing. We emphasize however that in the expression df/dx we do not multiply f or x by d or divide df by dx. The expression is to be read as a whole. We shall find out later that the expression, under certain circumstances, behaves as if we were dividing, and it is for this reason that we adopt this classical way of writing the derivative. We work out some examples before giving you exercises on this section.

*Example* 1. Let f(x) = 2x + 1. Find the derivative f'(x).

We form the Newton quotient. We have f(x + h) = 2(x + h) + 1. Thus

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

As h approaches 0 (which we write also  $h \to 0$ ), this number is equal to 2 and hence the limit is 2. Thus

$$f'(x) = 2$$

for all values of x. The derivative is constant.

*Example 2.* Find the slope of the graph of the function  $f(x) = 2x^2$  at the point whose x-coordinate is 3.

We may just as well find the slope at an arbitrary point on the graph. It is the derivative f'(x). We have

$$f(x + h) = 2(x + h)^2 = 2(x^2 + 2xh + h^2).$$

The Newton quotient is

$$\frac{f(x+h) - f(x)}{h} = \frac{2(x^2 + 2xh + h^2) - 2x^2}{h}$$
$$= \frac{4xh + 2h^2}{h}$$
$$= 4x + 2h.$$

As  $h \to 0$  the limit is 4x. Hence f'(x) = 4x. At the point x = 3 we get f'(3) = 12, which is the desired slope.

*Example 3.* Find the equation of the tangent line to the curve  $y = 2x^2$  at the point whose x-coordinate is -2.

In the preceding example we have computed the general formula for the slope of the tangent line. It is

$$f'(x) = 4x.$$

At the point x = -2 the slope is therefore -8. The tangent line has an equation

$$y = -8x + b$$

for some number b. The y-coordinate of our point is  $2(-2)^2 = 8$ . Hence

we must have

8 = -8(-2) + bb = -8.

and solving for b yields

Thus the equation of the tangent line is

$$y = -8x - 8.$$

In defining the Newton quotient, we can take h positive or negative. It is sometimes convenient when taking the limit to look only at values of hwhich are positive. In this manner we get what is called the *right derivative*. If in taking the limit of the Newton quotient we took only negative values for h, we would get the *left derivative*.

*Example* 4. Let f(x) = |x|. Find its right derivative and its left derivative when x = 0.

The right derivative is the limit

$$\lim_{\substack{h\to 0\\h>0}}\frac{f(0+h)-f(0)}{h}$$

When h > 0, we have f(0 + h) = f(h) = h, and f(0) = 0. Thus

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

The limit as  $h \to 0$  and h > 0 is therefore 1.

The left derivative is the limit

$$\lim_{\substack{h\to 0\\h<0}}\frac{f(0+h)-f(0)}{h}$$

When h < 0 we have

$$f(0+h) = f(h) = -h.$$

Hence

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

The limit as  $h \to 0$  and h < 0 is therefore -1.

We see that the right derivative at 0 is 1 and the left derivative is -1. They are not equal. We would of course expect this from the graph of our function f(x) = |x|, which looks like that in Fig. 7.



FIGURE 7

Both the right derivative of f and the left derivative of f exist but they are not equal.

We could rephrase our definition of the derivative and say that the derivative of a function f(x) is defined when the right derivative and the left derivative exist and they are equal, in which case this common value is simply called the *derivative*.

*Example* 5. Let f(x) be equal to x if  $0 < x \leq 1$  and x - 1 if  $1 < x \leq 2$ . We do not define f for other values of x. Then the graph of f looks like this:



FIGURE 8

The left derivative of f at 1 exists and is equal to 1, but the right derivative of f at 1 does not exist. We leave the verification of the first assertion to you. To verify the second assertion, we must see whether the limit

$$\lim_{\substack{h \to 0 \\ h > 0}} \frac{f(1+h) - f(1)}{h}$$

exists. Since 1 + h > 1 we have

$$f(1+h) = 1+h-1 = h.$$

Also f(1) = 1. Thus the Newton quotient is

$$\frac{f(1+h) - f(1)}{h} = \frac{h-1}{h} = 1 - \frac{1}{h}$$

As h approaches 0 the quotient 1/h has no limit since it becomes arbitrarily large. Thus the Newton quotient has no limit for h > 0 and the function does not have a right derivative when x = 1.

#### THE DERIVATIVE

#### Exercises

Find the derivatives of the following functions:

1. $x^2 + 1$	2. $x^{3}$
3. $2x^3$	4. $3x^2$
5. $x^2 - 5$	6. $2x^2 + x$
7. $2x^2 - 3x$	8. $\frac{1}{2}x^3 + 2x$
9. $\frac{1}{x+1}$	10. $\frac{2}{x+1}$

11. In exercises 1 through 10, find the slope of the graph at the point whose x-coordinate is 2, and find the equation of the tangent line at that point.

12. Let f(x) be defined as follows:

f(x) = -x if  $x \le 0$  f(x) = 2 if x > 0. Find f'(x) when x = -1. Find the right and left derivatives of f at x = 0, if they exist.

13. Let f(x) = |x| + x. Does f'(0) exist? Does f'(x) exist for values of x other than 0?

14. Let f(x) = 0 if  $x \le 1$  and f(x) = x if x > 1. Sketch the graph. Find the right and left derivatives of f when x = 1. Find f'(x) for all other values of x.

15. Determine whether the following functions have a derivative at 0, and if so, what is the derivative.

(a) f(x) = x|x| (b)  $f(x) = x^2|x|$  (c)  $f(x) = x^3|x|$ 

# §3. Limits

In defining the slope of a curve at a point, or the derivative, we used the notion of limit, which we regarded as intuitively clear. It is indeed. You can see in the Appendix how one may define limits using only properties of numbers, but we do not worry about this here. However, we shall make a list of the properties of limits which will be used in the sequel, just to be sure of what we assume about them, and also to give you a technique for computing limits.

First, we note that if F is a constant function, F(x) = c for all x, then

$$\lim_{h\to 0} F(h) = c$$

is the constant itself. If F(h) = h, then

$$\lim_{h\to 0} F(h) = 0$$

The next properties relate limits with addition, subtraction, multiplication, division, and inequalities. [III, §3]

LIMITS

Suppose that we have two functions F(x) and G(x) which are defined for the same numbers. Then we can form the sum of the two functions F + G, whose value at a point x is F(x) + G(x). Thus when  $F(x) = x^4$ and  $G(x) = 5x^{3/2}$  we have

$$F(x) + G(x) = x^4 + 5x^{3/2}$$

The value F(x) + G(x) is also written (F + G)(x). The first property of limits concerns the sum of two functions.

Property 1. Suppose that we have two functions F and G defined for small values of h, and assume that the limits

$$\lim_{h\to 0} F(h) \quad \text{and} \quad \lim_{h\to 0} G(h)$$

exist. Then

$$\lim_{h \to 0} \left[ F(h) + G(h) \right]$$

exists and

$$\lim_{h\to 0} (F+G)(h) = \lim_{h\to 0} F(h) + \lim_{h\to 0} G(h)$$

In other words the limit of a sum is equal to the sum of the limits.

A similar statement holds for the difference F-G.

After the sum we discuss the product. Suppose we have two functions F and G defined for the same numbers. Then we can form their product FG whose value at a number x is

$$(FG)(x) = F(x)G(x).$$

For instance if  $F(x) = 2x^2 - 2^x$  and  $G(x) = x^2 + 5x$ , then the product is

$$(FG)(x) = (2x^2 - 2^x)(x^2 + 5x).$$

Property 2. Let F, G be two functions defined for small values of h, and assume that

$$\lim_{h \to 0} F(h) \quad \text{and} \quad \lim_{h \to 0} G(h)$$

exist. Then the limit of the product exists and we have

$$\lim_{h \to 0} (FG)(h) = \lim_{h \to 0} [F(h)G(h)]$$
$$= \lim_{h \to 0} F(h) \cdot \lim_{h \to 0} G(h)$$

In words, we can say that the product of the limits is equal to the limit of the product.

As a special case, suppose that F(x) is the constant function F(x) = c. Then we can form the function cG, product of the constant by G, and we have

$$\lim_{h\to 0} cG(h) = c \cdot \lim_{h\to 0} G(h).$$

Thirdly, we come to quotients. Let F, G be as before, but assume that  $G(x) \neq 0$  for any x. Then we can form the quotient function F/G whose value at x is

$$\frac{F}{G}(x) = \frac{F(x)}{G(x)}$$
.

Property 3. Assume that the limits

$$\lim_{h \to 0} F(h) \quad \text{and} \quad \lim_{h \to 0} G(h)$$

exist, and that

$$\lim_{h\to 0} G(h) \neq 0.$$

Then the limit of the quotient exists and we have

$$\lim_{h\to 0}\frac{F(h)}{G(h)}=\frac{\lim F(h)}{\lim G(h)}\cdot$$

In words, the quotient of the limits is equal to the limit of the quotient.

As we have done above, we shall sometimes omit writing  $h \rightarrow 0$  for the sake of simplicity.

Property 4. Let F, G be two functions defined for small values of h, and assume that  $G(h) \leq F(h)$ . Assume also that

$$\lim_{h\to 0} F(h) \quad \text{and} \quad \lim_{h\to 0} G(h)$$

exist. Then

$$\lim_{h\to 0} G(h) \leq \lim_{h\to 0} F(h).$$

Property 5. Let the assumptions be as in Property 4, and in addition, assume that

$$\lim_{h\to 0} G(h) = \lim_{h\to 0} F(h).$$

Let E be another function defined for the same numbers as F, G such that

$$G(h) \leq E(h) \leq F(h)$$

[III, §3]

LIMITS

for all small values of h. Then

$$\lim_{h\to 0} E(h)$$

exists and is equal to the limits of F and G.

Property 5 is known as the squeezing process. You will find many applications of it in the sequel.

Example 1. Find the limit

$$\lim_{h \to 0} \frac{2xh+3}{x^2-4h}$$

when  $x \neq 0$ .

The numerator of our quotient approaches 3 when  $h \to 0$  and the denominator approaches  $x^2$ . Thus the quotient approaches  $3/x^2$ . We can justify these steps more formally by applying our three properties. For instance:

$$\lim_{h \to 0} (2xh + 3) = \lim_{h \to 0} (2xh) + \lim_{h \to 0} 3$$
$$= \lim_{h \to 0} (2x) \lim_{h \to 0} (h) + \lim_{h \to 0} 3$$
$$= 2x \cdot 0 + 3$$
$$= 3.$$

For the denominator, we have

$$\lim (x^2 - 4h) = \lim x^2 + \lim (-4h)$$
  
=  $x^2 + \lim (-4) \lim (h)$   
=  $x^2 + (-4) \cdot 0$   
=  $x^2$ .

Using the rule for the quotient, we get  $3/x^2$ .

The properties of limits which we have stated above will allow you to compute limits in determining derivatives. We illustrate this by an example.

Example 2. Let f(x) = 1/x (defined for  $x \neq 0$ ). Find the derivative df/dx.

The Newton quotient is

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

We put everything over a common denominator (x + h)xh. The Newton

quotient is equal to

$$\frac{x-(x+h)}{(x+h)xh} = \frac{-h}{(x+h)xh} = \frac{-1}{(x+h)x}.$$

Thus we have to determine the limit of a quotient as h approaches 0. Using the property of a product, we have

$$\lim (x+h)x = \lim (x+h) \lim x$$
$$= x^2.$$

Using the property of quotients, we see that the Newton quotient for the function 1/x approaches  $-1/x^2$ . Thus

$$\frac{df}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \frac{-1}{x^2}$$

#### Exercises

Find the derivatives of the following functions, justifying the steps in taking limits by means of the first three properties:

1.  $f(x) = 2x^{2} + 3x$ 2.  $f(x) = \frac{1}{2x + 1}$ 3.  $f(x) = \frac{x}{x + 1}$ 5.  $f(x) = \frac{x}{2x - 1}$ 7.  $f(x) = x^{4}$ 9.  $f(x) = 2x^{3}$ 2.  $f(x) = \frac{1}{2x + 1}$ 4. f(x) = x(x + 1)6.  $f(x) = 3x^{3}$ 8.  $f(x) = x^{5}$ 10.  $f(x) = \frac{1}{2}x^{3} + x$ 

# §4. Powers

We have seen that the derivative of the function  $x^2$  is 2x. Let us consider the function  $f(x) = x^3$  and find its derivative. We have

$$f(x + h) = (x + h)^3 = x^3 + 3x^2h + 3xh^2 + h^3.$$

Hence the Newton quotient is

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

Using the properties of limits of sums and products, we see that  $3x^2$ 

[III, §4]

POWERS

remains equal to itself as h approaches 0, that 3xh and  $h^3$  both approach 0. Hence

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = 3x^2.$$

This suggests that in general, whenever  $f(x) = x^n$  for some positive integer *n*, the derivative f'(x) should be  $nx^{n-1}$ . This is indeed the case, and we state it as a theorem.

THEOREM 1. Let n be an integer  $\geq 1$  and let  $f(x) = x^n$ . Then

$$\frac{df}{dx} = nx^{n-1}.$$

*Proof.* We have

$$f(x + h) = (x + h)^n = (x + h)(x + h) \cdots (x + h)_n$$

the product being taken n times. Selecting x from each factor gives us a term  $x^n$ . If we take x from all but one factor and h from the remaining factor, we get  $hx^{n-1}$  taken n times. This gives us a term  $nx^{n-1}h$ . All other terms will involve selecting h from at least two factors, and the corresponding term will be divisible by  $h^2$ . Thus we get

$$f(x + h) = (x + h)^n = x^n + nx^{n-1}h + h^2g(x, h)$$

where g(x, h) is simply some expression involving powers of x and h with numerical coefficients which it is unnecessary for us to determine. However, using the rules for limits of sums and products we can conclude that

$$\lim_{h\to 0}g(x,h)$$

will be some number which it is unnecessary for us to determine.

The Newton quotient is therefore

$$\frac{f(x+h) - f(x)}{h} = \frac{x^n + nx^{n-1}h + h^2g(x,h) - x^n}{h}$$

We can cancel  $x^n$  and are left with

$$\frac{nx^{n-1}h + h^2g(x,h)}{h}$$

We can now divide numerator and denominator by h, thereby giving us

$$nx^{n-1} + hg(x, h).$$

THE DERIVATIVE

As h approaches 0, the term  $nx^{n-1}$  remains unchanged. The limit of h as h tends to 0 is 0, and hence by the product rule, the term hg(x, h) approaches 0 when h tends to 0. Thus finally

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=nx^{n-1},$$

which proves our theorem.

THEOREM 2. Let a be any number and let  $f(x) = x^a$  (defined for x > 0). Then f(x) has a derivative, which is

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

It would not be difficult to prove Theorem 2 when a is a negative integer. It is best however to wait until we have a rule giving us the derivative of a quotient before doing it. We could also give a proof when a is a rational number. However, we shall prove the general result in a later chapter, and thus we prefer to wait until then, when we have more techniques available.

Examples. If 
$$f(x) = x^{10}$$
 then  $f'(x) = 10x^9$ .  
If  $f(x) = x^{3/2}$  (for  $x > 0$ ) then  $f'(x) = \frac{3}{2}x^{1/2}$ .  
If  $f(x) = x^{-5/4}$  then  $f'(x) = -\frac{5}{4}x^{-9/4}$ .  
If  $f(x) = x^{\sqrt{2}}$  then  $f'(x) = \sqrt{2} x^{\sqrt{2}-1}$ .

Note especially the special case when f(x) = x. Then f'(x) = 1.

#### Exercises

- 1. Write out the expansion of  $(x + h)^4$  in terms of powers of x and h.
- 2. Find the derivative of the function  $x^4$  directly, using the Newton quotient.
- 3. What are the derivatives of the following functions? (a)  $x^{2/3}$  (b)  $x^{-3/2}$  (c)  $x^{7/6}$

4. What is the equation of the tangent line to the curve  $y = x^9$  at the point (1, 1)?

5. What is the slope of the curve  $y = x^{2/3}$  at the point (8, 4)? What is the equation of the tangent line at that point?

6. Give the slope and equation of the tangent line to the curve  $y = x^{-3/4}$  at the point whose x-coordinate is 16.

7. Give the slope and equation of the tangent line to the curve  $y = \sqrt{x}$  at the point whose x-coordinate is 3.

8. Give the derivatives of the following functions at the indicated points:

(a)  $f(x) = x^{1/4}$  at x = 5(b)  $f(x) = x^{-1/4}$  at x = 7(c)  $f(x) = x^{\sqrt{2}}$  at x = 10(d)  $f(x) = x^{\pi}$  at x = 7

# §5. Sums, products, and quotients

In this section we shall derive several rules which allow you to find the derivatives for sums, products, and quotients of functions when you know the derivative of each factor.

Before stating and proving these rules, we make one remark concerning the derivative.

Let f(x) be a function having a derivative f'(x). Since the quotient

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

approaches a limit as h approaches 0, and since

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

using the rules for sums and products of limits, we conclude that

$$\lim_{h\to 0} f(x+h) = f(x),$$

and that f(x + h) - f(x) approaches 0 as h approaches 0.

Of course, we can never substitute h = 0 in our quotient, because then it becomes 0/0, which is meaningless. Geometrically, letting h = 0amounts to taking the two points on the curve equal to each other. It is then impossible to have a unique straight line through one point. Our procedure of taking the limit of the Newton quotient is meaningful only if  $h \neq 0$ .

Let c be a number and f(x) a function which has a derivative f'(x) for all values of x for which it is defined. We can multiply f by the constant c to get another function cf whose value at x is cf(x).

The derivative of cf is then given by the formula

$$(cf)'(x) = c \cdot f'(x);$$

in other words, the derivative of a constant times a function is the constant times the derivative of the function.

To prove this rule, we use the definition of derivative. The Newton quotient for the function cf is

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

Let us take the limit as h approaches 0. Then c remains fixed, and

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

approaches f'(x). According to the rule for the product of limits, we see that our Newton quotient approaches cf'(x), as was to be proved.

For example, let  $f(x) = 3x^2$ . Then f'(x) = 6x. If  $f(x) = 17x^{1/2}$ , then  $f'(x) = \frac{17}{2}x^{-1/2}$ . If  $f(x) = 10x^a$ , then  $f'(x) = 10ax^{a-1}$ . Next we look at the sum of two functions.

Let f(x) and g(x) be two functions which have derivatives f'(x) and g'(x), respectively. Then the sum f(x) + q(x) has a derivative, and

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

The derivative of a sum is equal to the sum of the derivatives.

To prove this, we have by definition

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

Therefore the Newton quotient for f + q is

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

Collecting terms and separating the fraction, we see that this expression is equal to

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

Taking the limit as h approaches 0 and using the rule for the limit of a sum, we see that this last sum approaches f'(x) + g'(x) as h approaches 0. This proves what we wanted.

For example, the derivative of the function  $x^3 + x^2$  is  $3x^2 + 2x$ . The derivative of the function  $4x^{1/2} + 5x^{-10}$  is

$$2x^{-1/2} - 50x^{-11}.$$

Carried away by our enthusiasm at determining so easily the derivative of functions built up from others by means of constants and sums, we might [III, §5]

now be tempted to state the rule that the derivative of a product is the product of the derivatives. Unfortunately, this is false. To see that the rule is false, we look at an example.

Let f(x) = x and  $g(x) = x^2$ . Then f'(x) = 1 and g'(x) = 2x. Therefore f'(x)g'(x) = 2x. However, the derivative of the product  $(fg)(x) = x^3$ is  $3x^2$ , which is certainly not equal to 2x. Thus the product of the derivatives is not equal to the derivative of the product.

Through trial and error the correct rule was discovered. It can be stated as follows:

Let f(x) and g(x) be two functions having derivatives f'(x) and g'(x). Then the product function f(x)g(x) has a derivative, which is given by the formula

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

In words, the derivative of the product is equal to the derivative of the first times the second, plus the first times the derivative of the second.

The proof is not very much more difficult than the proofs we have already encountered. By definition, we have

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

Consequently the Newton quotient for the product function fg is

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

At this point, it looks a little hopeless to transform this quotient in such a way that we see easily what limit it approaches as h tends to 0. But we rewrite our quotient by inserting

$$-f(x)g(x+h) + f(x)g(x+h)$$

in the numerator. This certainly does not change the value of our quotient, which now looks like

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

We can split this fraction into a sum of two fractions:

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

THE DERIVATIVE

We can factor g(x + h) in the first term, and f(x) in the second term, to obtain

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

The situation is now well under control. As h tends to 0, g(x + h) tends to g(x), and the two quotients in the expression we have just written tend to f'(x) and g'(x) respectively. Thus the Newton quotient for fg tends to

f'(x)g(x) + f(x)g'(x),

thereby proving our assertion.

To illustrate the rules for products, let us find the derivative of

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

Applying the rule, we see that it is equal to

$$1 \cdot (3x^2) + (x+1)6x$$

Similarly, let  $f(x) = 2x^5 + 5x^4$  and  $g(x) = 2x^{1/2} + x^{-1}$ . Then the derivative of f(x)g(x) is

$$(10x^4 + 20x^3)(2x^{1/2} + x^{-1}) + (2x^5 + 5x^4)\left(x^{-1/2} - \frac{1}{x^2}\right),$$

which you may and should leave just like that without attempting to simplify the expression.

The last rule of this section concerns the derivative of a quotient. We begin with a special case.

Let g(x) be a function having a derivative g'(x), and such that  $g(x) \neq 0$ . Then the derivative of the quotient 1/g(x) exists, and is equal to

$$\frac{-1}{g(x)^2}\,g'(x).$$

To prove this, we look at the Newton quotient

$$\frac{1}{\frac{g(x+h)}{h}} - \frac{1}{\frac{g(x)}{h}}$$

which is equal to

$$\frac{g(x)-g(x+h)}{g(x+h)g(x)h}=-\frac{1}{g(x+h)g(x)}\frac{g(x+h)-g(x)}{h}$$

[III, §5]

Letting h approach 0 we see immediately that our expression approaches

$$\frac{-1}{g(x)^2}g'(x)$$

as desired.

The general case of the rule for quotients can now be easily stated and proved.

Let f(x) and g(x) be two functions having derivatives f'(x) and g'(x) respectively, and such that  $g(x) \neq 0$ . Then the derivative of the quotient f(x)/g(x) exists, and is equal to

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

Putting this into words yields: The bottom times the derivative of the top, minus the top times the derivative of the bottom, over the bottom squared (which you can memorize like a poem).

To prove this rule, we write our quotient in the form

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

and use the rule for the derivative of a product, together with the special case we have just proved. We obtain

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

Putting this expression over the common denominator  $g(x)^2$  yields

$$\frac{g(x)f'(x) - f(x)g'(x)}{g(x)^2},$$

which is the desired derivative.

We work out some examples.

Let  $f(x) = x^2 + 1$  and  $g(x) = 3x^4 - 2x$ . Then the derivative of f(x)/g(x) is

$$\frac{(3x^4-2x)2x-(x^2+1)(12x^3-2)}{(3x^4-2x)^2}.$$

Still another: The derivative of 2x/(x+4) is

$$\frac{(x+4)\cdot 2-2x\cdot 1}{(x+4)^2}\cdot$$

THE DERIVATIVE

For future reference, we write the various rules which we have proved into the df/dx notation. The first one:

$$\frac{d(cf)}{dx} = c \cdot \frac{df}{dx} \cdot$$

The sum:

$$rac{d(f+g)}{dx} = rac{df}{dx} + rac{dg}{dx} \cdot$$

The product:

$$\frac{d(fg)}{dx} = \frac{df}{dx}g(x) + f(x)\frac{dg}{dx}.$$

The quotient:

$$\frac{d(f/g)}{dx} = \frac{g \frac{df}{dx} - f \frac{dg}{dx}}{g(x)^2} \cdot$$

#### Exercises

Find the derivatives of the following functions:

To break the monotony of the letter x, let us use another.

15. 
$$f(t) = \frac{t^2 + 2t - 1}{(t+1)(t-1)}$$
 16.  $\frac{t^{-5/4}}{t^2 + t - 1}$ 

17. What is the slope of the curve

$$y = \frac{t}{t+5}$$

at the point t = 2? What is the equation of the tangent line at this point?

18. What is the slope of the curve

$$y = \frac{t^2}{t^2 + 1}$$

at t = 1? What is the equation of the tangent line?

# §6. The chain rule

We know how to build up new functions from old ones by means of sums, products, and quotients. There is one other important way of building up new functions. We shall first give examples of this new way.

Consider the function  $(x + 2)^{10}$ . We can say that this function is made up of the 10-th power function, and the function x + 2. Namely, given a number x, we first add 2 to it, and then take the 10-th power. Let g(x) =x + 2 and let f be the 10-th power function. Then we can take the value of f at x + 2, namely

$$f(x+2) = (x+2)^{10}$$

and we can also write it as

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

Another example: Consider the function  $(3x^4 - 1)^{1/2}$ . If we let  $g(x) = 3x^4 - 1$  and f be the square root function, then

$$f(g(x)) = \sqrt{3x^4 - 1} = (3x^4 - 1)^{1/2}.$$

In order not to get confused by the letter x, which cannot serve us any more in all contexts, we use another letter to denote the values of g. Thus we may write  $f(u) = u^{1/2}$ .

Similarly, let f(u) be the function u + 5 and g(x) = 2x. Then

$$f(g(x)) = f(2x) = 2x + 5.$$

One more example of the same type: Let

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

and

$$g(x) = x^{10}.$$

Then

$$f(g(x)) = \frac{1}{x^{10} + 2}$$

In order to give you sufficient practice with many types of functions, we now mention several of them whose definitions will be given later. These will be sin and cos (which we read sine and cosine), log (which we read logarithm or simply log), and the exponential function exp. We shall select a special number e (whose value is approximately 2.718...), such that the function exp is given by

$$\exp\left(x\right) = e^x.$$

We now see how we make new functions out of these.

Let  $f(u) = \sin u$  and  $g(x) = x^2$ . Then

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

Let  $f(u) = e^u$  and  $g(x) = \cos x$ . Then

$$f(g(x)) = e^{\cos x}.$$

Let  $f(v) = \log v$  and  $g(t) = t^3 - 1$ . Then

$$f(g(t)) = \log (t^3 - 1).$$

Let  $g(w) = w^{10}$  and  $f(z) = \log z + \sin z$ . Then

$$g(f(z)) = (\log z + \sin z)^{10}$$

You should practice with part (a) of the exercises, in order to assimilate properly the terminology and mechanisms of these combined functions.

Whenever we have two functions f and g such that f is defined for all numbers which are values of g, then we can build a new function denoted by  $f \circ g$  whose value at a number x is

$$(f \circ g)(x) = f(g(x)).$$

The rule defining this new function is: Take the number x, find the number g(x), and then take the value of f at g(x). This is the value of  $f \circ g$  at x. The function  $f \circ g$  is called the *composite function* of f and g. We say that g is the *inner* function and that f is the *outer* function.

It is important to keep in mind that we can compose two functions only when the outer function is defined at all values of the inner function. For instance, let  $f(u) = u^{1/2}$  and  $g(x) = -x^2$ . Then we cannot form the composite function  $f \circ g$  because f is defined only for positive numbers (or 0) and the values of g are all negative, or 0. Thus  $(-x^2)^{1/2}$  does not make sense.

However, for the moment you are asked to learn the mechanism of composite functions just the way you learned the multiplication table, in order to acquire efficient conditioned reflexes when you meet composite functions. Hence for the drills given by the exercises at the end of the section, you should forget for a while the meaning of the symbols and operate with them formally, just to learn the formal rules properly.

We come to the problem of taking the derivative of a composite function.

We start with an example. Suppose we want to find the derivative of the function  $(x + 1)^{10}$ . The Newton quotient would be a very long expression,

THE CHAIN RULE

which it would be essentially hopeless to disentangle by brute force, the way we have up to now. It is therefore a pleasant surprise that there will be an easy way of finding the derivative. We tell you the answer right away: The derivative of this function is  $10(x + 1)^9$ . This looks very much related to the derivative of powers.

Before proving and stating the general theorem, we give you other examples. The derivative of  $(x^2 + 2x)^{3/2}$  is  $\frac{3}{2}(x^2 + 2x)^{1/2}(2x + 2)$ . Observe carefully the extra term 2x + 2, which is the derivative of the expression  $x^2 + 2x$ .

The derivative of  $(x^2 + x)^{10}$  is  $10(x^2 + x)^9(2x + 1)$ . Observe again the presence of the term 2x + 1, which is the derivative of  $x^2 + x$ .

Can you guess the general rule from the preceding assertions? The general rule was also discovered by trial and error, but we profit from three centuries of experience, and thus we are able to state it and prove it very simply, as follows.

Let f and g be two functions having derivatives, and such that f is defined at all numbers which are values of g. Then the composite function  $f \circ g$  has a derivative, given by the formula

$$(f \circ g)'(x) = f'(g(x))g'(x).$$

This can be expressed in words by saying that we take the *derivative of the* outer function times the derivative of the inner function (or the derivative of what is inside).

The preceding assertion is known as the *chain rule*, and we shall now prove it.

We must consider the Newton quotient of the composite function  $f \circ g$ . By definition, it is

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

Put u = g(x), and let

$$k = g(x+h) - g(x).$$

Then k depends on h, and tends to 0 as h approaches 0. Our Newton quotient is equal to

$$\frac{f(u+k)-f(u)}{h}$$

Suppose that k is unequal to 0 for all small values of h. Then we can multiply and divide this quotient by k, and obtain

$$\frac{f(u+k)-f(u)}{k}\frac{k}{h} = \frac{f(u+k)-f(u)}{k}\frac{g(x+h)-g(x)}{h}$$

THE DERIVATIVE

If we let h approach 0 and use the rule for the limit of a product, we see that our Newton quotient approaches

f'(u)g'(x),

and this would prove our chain rule, under the assumption that k is not 0.

It does not happen very often that k = 0 for arbitrarily small values of h, but when it does happen, the preceding argument breaks down. For those of you who are interested, we shall show you how the argument can be slightly modified so as to be valid in all cases. The uninterested reader can just skip it.

We go back to the definition of the derivative of f. Given a number u such that f(u) is defined, we know that

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

Therefore, the limit of the expression

$$\varphi(k) = \frac{f(u+k) - f(u)}{k} - f'(u)$$

as k approaches 0 is equal to 0. In symbols:

$$\lim_{k\to 0}\varphi(k)=0.$$

Multiplying by k we obtain

$$k\varphi(k) = f(u+k) - f(u) - kf'(u)$$
$$f(u+k) - f(u) = k \cdot f'(u) + k \cdot \varphi(k).$$

So far, this is valid only when k is not 0. But if we define  $\varphi(0)$  to be 0, then we note that the relationship we have just derived is still valid when k = 0 because k does not appear in the denominator. Substituting k = 0just yields

$$f(u) - f(u) = 0,$$

which is certainly true.

Now let u = g(x) and let k = g(x + h) - g(x). As h approaches zero, so does k.

The Newton quotient for the composite function  $f \circ g$  is

$$\frac{f(g(x+h))-f(g(x))}{h}=\frac{f(u+k)-f(u)}{h},$$
which, by the expression we have just derived, is equal to

$$rac{k \cdot f'(u) + k \cdot arphi(k)}{h}$$
 ,

or substituting the value for k, is equal to

$$rac{g(x+h)\,-\,g(x)}{h}f^{\prime}(u)\,+\,rac{g(x+h)\,-\,g(x)}{h}\,arphi(k).$$

Taking the limit as h approaches 0, we see that the first term approaches g'(x)f'(u). So far as the second term is concerned, taking the limit, we get

$$\lim_{h\to 0}\frac{g(x+h)-g(x)}{h}\varphi(k)=g'(x)\cdot 0=0$$

because the limit of  $\varphi(k)$  as h or k goes to 0 is 0. This proves that the Newton quotient of  $f \circ g$  approaches

and concludes the proof of the chain rule in general.

We are in a position to see the reason for the notation dg/dx. The chain rule in this notation can be expressed by the formula

$$\frac{d(f \circ g)}{dx} = \frac{df}{du} \frac{du}{dx}$$

if u = g(x) is a function of x. Thus the derivative behaves as if we could cancel the du. As long as we have proved this result, there is nothing wrong with working like a machine in computing derivatives of composite functions, and we shall give you several examples before the exercises.

Let  $f(u) = u^{10}$  and  $u = g(x) = x^2 + 1$ . Then  $f'(u) = 10u^9$  and g'(x) = 2x. Thus

$$\frac{d(f \circ g)}{dx} = 10u^9 \cdot 2x = 10(x^2 + 1)^9 2x.$$

Let  $f(u) = 2u^{1/2}$  and g(x) = 5x + 1. Then  $f'(u) = u^{-1/2}$  and g'(x) = 5. Thus

$$\frac{d(f \circ g)}{dx} = (5x+1)^{-1/2} \cdot 5.$$

(Pay attention to the constant 5, which is the derivative of 5x + 1. You are very likely to forget it.)

In order to give you more extensive drilling than would be afforded by the functions we have considered, like powers, we summarize the derivatives of the elementary functions which are to be considered later.

$$\frac{d (\sin x)}{dx} = \cos x.$$
$$\frac{d (\cos x)}{dx} = -\sin x.$$

 $\frac{d(e^x)}{dx} = e^x$  (yes,  $e^x$ , the same as the function!).

$$\frac{d\ (\log x)}{dx} = \frac{1}{x} \cdot$$

In view of these, and the chain rule, we see that the derivative of  $(\sin x)^7$  is 7  $(\sin x)^6 (\cos x)$ . (Here again we emphasize the appearance of  $\cos x$ , which is the derivative of what's inside our composite function.)

The derivative of 
$$(\log x)^{1/2}$$
 is  $\frac{1}{2} (\log x)^{-1/2} \cdot \frac{1}{x}$ .

The derivative of  $e^{\sin x}$  is  $e^{\sin x} (\cos x)$ .

The derivative of  $\cos(2x^2)$  is  $-\sin(2x^2) \cdot 4x$ . (The 4x is the derivative of  $2x^2$ .)

## Exercises

(a) In each case, find two functions f(u) and g(x) such that the indicated function is of type f(g(x)).

(b) Find the derivative of the indicated function.

1.	$(x+1)^{8}$	2.	$(2x - 5)^{1/2}$
3.	$(\sin x)^3$	4.	$(\log x)^5$
5.	$\sin 2x$	6.	$\log{(x^2+1)}$
7.	$e^{\cos x}$	8.	$\log\left(e^{x}+\sin x\right)$
9.	$\sin\left(\log x + \frac{1}{x}\right)$	10.	$\frac{x+1}{\sin 2x}$
11.	$(2x^2+3)^3$	12.	$\cos(\sin 5x)$
13.	$\log(\cos 2x)$	14.	$\sin\left[\left(2x+5\right)^2\right]$
15.	$\sin\left[\cos\left(x+1\right)\right]$	16.	$\sin(e^x)$
17.	$\frac{1}{(3x-1)^4}$	18.	$\frac{1}{(4x)^3}$
19.	$\frac{1}{(\sin 2x)^2}$	20.	$\frac{1}{(\cos 2x)^2}$

(Note: Do not attempt to simplify your answers.)

# §7. Rate of change

The derivative has an interesting physical interpretation, which was very closely connected with it in its historical development, and is worth mentioning.

Suppose that a particle moves along some straight line a certain distance depending on time t. Then the distance s is a function of t, which we write s = f(t).

For two values of the time,  $t_1$ , and  $t_2$ , the quotient

$$\frac{f(t_2) - f(t_1)}{t_2 - t_1}$$

can be regarded as a sort of average speed of the particle. At a given time  $t_0$ , it is therefore reasonable to regard the limit

$$\lim_{t\to t_0}\frac{f(t)-f(t_0)}{t-t_0}$$

as the rate of change of s with respect to t. This is none other than the derivative f'(t).

For instance if the particle is an object dropping under the influence of gravity, then experimental data show that

$$s = \frac{1}{2}gt^2,$$

where g is the gravitational constant. In that case,

$$\frac{ds}{dt} = gt$$

is its speed.

The rate of change of the speed is the acceleration. In the case of gravity, we take the derivative of the speed and we get simply the constant g.

In general, given a function y = f(x), the derivative f'(x) is interpreted as the rate of change of f(x). Thus f' is also a function. If it turns out to be also differentiable (this being usually the case), then its derivative is called the *second derivative* of f and is denoted by f''(x). For instance, the first derivative of  $(x^2 + 1)^2$  is  $2(x^2 + 1)2x = 4x^3 + 4x$ , and the second derivative f''(x) is  $12x^2 + 4$ .

### Exercises

Find the second derivatives of the following functions:

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

There is no reason to stop at the second derivative, and one can of course con-

- tinue with the third, fourth, etc.
  - 3. Find the 80-th derivative of  $x^7 + 5x 1$ .
  - 4. Find the 7-th derivative of  $x^7 + 5x 1$ .
  - 5. Find the third derivative of  $x^2 + 1$ .
  - 6. Find the third derivative of  $x^3 + 2x 5$ .

Find the derivatives of the following functions:

7.	$\frac{1}{\sin 3x}$	8.	$(\sin x)(\cos x)$
9.	$(x^2+1)e^x$	10.	$(x^3+2x)(\sin 3x)$
11.	$\frac{1}{\sin x + \cos x}$	12.	$\frac{\sin 2x}{e^x}$
13.	$\frac{\log x}{x^2+3}$	14.	$\frac{x+1}{\cos 2x}$
15.	$(2x - 3)(e^x + x)$	16.	$(x^3 - 1)(e^{3x} + 5x)$
17.	$\frac{x^3+1}{x-1}$	18.	$\frac{x^2-1}{2x+3}$
19.	$(x^{4/3} - e^x)(2x + 1)$	20.	$(\sin 3x)(x^{1/4}-1)$
21.	$\sin\left(x^2+5x\right)$	22.	$e^{3x^2+8}$
23.	$\frac{1}{\log\left(x^4+1\right)}$	24.	$\frac{1}{\log\left(x^{1/2}+2x\right)}$
25.	$\frac{\Delta x}{e^x}$	26.	Relax.

27. A particle is moving so that at time t, the distance traveled is given by  $s(t) = t^3 - 2t + 1$ . At what time is the acceleration equal to 0?

28. A cube is expanding in such a way that its edge is changing at a rate of 5 in./sec. When its edge is 4 in. long, find the rate of change of its volume.

29. A sphere is increasing so that its radius increases at the rate of 1 in./sec. How fast is its volume changing when its radius is 3 in.? (The volume of a sphere is  $4\pi r^3/3$ .)

30. What is the rate of change of the area of a circle with respect to its radius, diameter, circumference?

31. A point moves along the graph of  $y = 1/(x^2 + 4)$  so that its x-coordinate changes at the rate of 3 units per second. What is the rate of change of its y-coordinate when x = 2?

# CHAPTER IV

# **Sine and Cosine**

From the sine of an angle and the cosine of an angle, we shall define functions of numbers, and determine their derivatives.

It is convenient to recall all the facts about trigonometry which we need in the sequel, especially the formula giving us the sine and cosine of the sum of two angles. Thus our treatment of the trigonometric functions is self-contained—you do not need to know anything about sine and cosine before starting to read this chapter. However, most of the proofs of statements in §1 come from plane geometry and will be left to you.

# §1. The sine and cosine functions

Suppose that we have given coordinate axes, and a certain angle, as shown on the figure.



We select a point (x, y) (not the origin) on the line determining our angle A. We let  $r = \sqrt{x^2 + y^2}$ . Then r is the distance from (0, 0) to the point (x, y). We define

sine 
$$A = \frac{y}{r} = \frac{y}{\sqrt{x^2 + y^2}}$$
  
cosine  $A = \frac{x}{r} = \frac{x}{\sqrt{x^2 + y^2}}$ 

If we select another point  $(x_1, y_1)$  on the line determining our angle A and use its coordinates to get the sine and cosine, then we shall obtain the same values as with (x, y). Indeed, there is a positive number c such that

$$x_1 = cx \quad \text{and} \quad y_1 = cy.$$

Thus

$$rac{y_1}{\sqrt{x_1^2+x_1^2}}=rac{cy}{\sqrt{c^2x^2+c^2y^2}}$$

We can factor c from the denominator, and then cancel c in both the numerator and denominator to get

$$\frac{y}{\sqrt{x^2+y^2}}$$

In this way we see that sine A does not depend on the choice of coordinates (x, y).

The geometric interpretation of the above argument simply states that the triangles in the following diagram are similar.



The angle A can go all the way around. For instance, we could have an angle determined by a point (x, y) in the second or third quadrant.



When the angle A is in the first quadrant, then its sine and cosine are positive because both coordinates x, y are positive. When the angle A is in the second quadrant, its sine is positive because y is positive, but its cosine is negative because x is negative.

When A is in the third quadrant, sine A is negative and cosine A is negative also.

In order to define the sine of a *number*, we select a unit for measuring angles. We let  $\pi$  be the area of the circle of radius 1. We then choose a unit angle such that the flat angle is equal to  $\pi$  times the unit angle. (See the following figures.) The right angle has measure  $\pi/2$ . The full angle going once around is then  $2\pi$ .

The unit of measurement for which the flat angle is  $\pi$  is called the *radian*. Thus the right angle has  $\pi/2$  radians.

66



There is another current unit of measurement for which the flat angle is 180. This unit is called the *degree*. Thus the flat angle has 180 degrees, and the right angle has 90 degrees. We also have

> 360 degrees =  $2\pi$  radians 60 degrees =  $\pi/3$  radians 45 degrees =  $\pi/4$  radians 30 degrees =  $\pi/6$  radians.

We make a table of the sines and cosines of these angles.

Angle	Sine	Cosine
$\pi/6$	1/2	$\sqrt{3}/2$
$\pi/4$	$1/\sqrt{2}$	$1/\sqrt{2}$
$\pi/3$	$\sqrt{3}/2$	1/2
$\pi/2$	1	0
$\pi$	0	-1
$2\pi$	0	1

Unless otherwise specified, we always use the radian measure, and our table is given for this measure.

The values of this table are easily determined, using properties of similar triangles and plane geometry. For instance, we get the sine of the angle  $\pi/4$  radians from a right triangle with two equal sides:



We can determine the sine of  $\pi/4$  by means of the point (1, 1). Then  $r = \sqrt{2}$  and sine  $\pi/4$  radians is  $1/\sqrt{2}$ . Similarly for the cosine.

The following is an important rule relating the sine and cosine.

THEOREM 1. For any angle A we have

cosine 
$$A = sine\left(A + \frac{\pi}{2}\right)$$

or

sine 
$$A = \operatorname{cosine}\left(A - \frac{\pi}{2}\right)$$

*Proof.* You can use theorems of plane geometry to prove our theorem. We leave this to you.

We wish to define a function of numbers, which will also be called the *sine*. The rule is:

For any number x we associate to it the number which is the sine of x radians.

This function is denoted by sin x and is defined for all x. Thus sin  $\pi = 0$ ,  $\sin \pi/2 = 1$ ,  $\sin 2\pi = 0$ ,  $\sin 0 = 0$ .

Similarly, we have the *cosine function*, which is defined for all numbers x by the rule:

 $\cos x$  is the number which is the cosine of the angle x radians.

Thus  $\cos 0 = 1$  and  $\cos \pi = -1$ . We can also define the tangent function,  $\tan x$ , which is the quotient

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

and is defined for all numbers x such that  $\cos x \neq 0$ . These are the numbers x which are unequal to

$$\frac{\pi}{2}$$
,  $\frac{3\pi}{2}$ ,  $\frac{5\pi}{2}$ , ...;

in general  $x \neq (2n+1)\pi/2$  for some integer n.

If we had used the measure of angles in degrees we would obtain *another* sine function which is not equal to the sine function which we defined in terms of radians. Suppose we call this other sine function sin<sup>\*</sup>. Then

$$\sin^*\left(180\right) = \sin\,\pi,$$

and in general

$$\sin^* (180x) = \sin \pi x$$

for any number x. Thus

$$\sin^* x = \sin\left(\frac{\pi}{180}x\right)$$

is the formula relating our two sine functions. It will become clear later why we always pick the radian measure instead of any other. [IV, §2]

THE GRAPHS

At present we have no means of computing values for the sine and cosine other than the very special cases listed above (and similar ones, based on simple symmetries of right triangles). It will be only in Chapter XIV that we shall develop a method which will allow us to find sin x and cos x for any value of x, up to any degree of accuracy that you wish.

#### Exercises

Find the following values of the sin function and cos function:

 1.  $\sin 3\pi/4$  2.  $\sin 2\pi/6$  

 3.  $\sin \frac{7\pi}{12}$  4.  $\sin \left(\pi - \frac{\pi}{6}\right)$  

 5.  $\cos \left(\pi + \frac{\pi}{6}\right)$  6.  $\cos \left(\pi + \frac{2\pi}{6}\right)$  

 7.  $\cos \left(2\pi - \frac{\pi}{6}\right)$  8.  $\cos \frac{5\pi}{4}$ 

Find the following values:

9.  $\tan \frac{\pi}{4}$ 10.  $\tan \frac{2\pi}{6}$ 11.  $\tan \frac{5\pi}{4}$ 12.  $\tan \left(2\pi - \frac{\pi}{4}\right)$ 

13. Prove by plane geometry that  $\sin(\pi - x) = \sin x$ .

14. Prove by plane geometry that  $\cos(\pi - x) = -\cos x$ .

15. Prove by plane geometry that  $\sin (2\pi - x) = -\sin x$ .

16. Prove by plane geometry that  $\sin(-x) = -\sin x$ .

17. Prove by plane geometry that  $\cos(-x) = \cos x$ .

18. Let a be a given number. Determine all numbers x such that  $\sin x = \sin a$ . (You may suppose that  $0 \le a < 2\pi$ , and distinguish the cases  $a = \pi/2$ ,  $a = -\pi/2$  and  $a \ne \pm \pi/2$ .)

## §2. The graphs

We wish to sketch the graph of the sine function.

We know that  $\sin 0 = 0$ . As x goes from 0 to  $\pi/2$ , the sine of x increases until x reaches  $\pi/2$ , at which point the sine is equal to 1.

As x ranges from  $\pi/2$  to  $\pi$ , the sine decreases until it becomes sin  $\pi = 0$ .

As x ranges from  $\pi$  to  $3\pi/2$  the sine becomes negative, but otherwise behaves in a similar way to the first quadrant, until it reaches  $\sin 3\pi/2 = -1$ .

Finally, as x goes from  $3\pi/2$  to  $2\pi$ , the sine of x goes from -1 to 0, and we are ready to start all over again.

The graph looks like this:



The graph of the cosine will look like that of the sine, but it starts with  $\cos 0 = 1$ . In the next picture the scale used on the vertical axis is different from that on the horizontal axis so that we will have more room for the arches of the graph.



If we go once around by  $2\pi$ , both the sine and cosine take on the same values, in other words

$$\sin (x + 2\pi) = \sin x$$
$$\cos (x + 2\pi) = \cos x$$

for all x. This holds whether x is positive or negative, and the same would be true if we took  $x - 2\pi$  instead of  $x + 2\pi$ .

You might legitimately ask why one arch of the sine (or cosine) curve looks the way you have drawn it, and not the following way:



In the next section, we shall find the slope of the curve  $y = \sin x$ . It is equal to  $\cos x$ . Thus when x = 0, the slope is  $\cos 0 = 1$ . Furthermore, when  $x = \pi/2$ , we have  $\cos \pi/2 = 0$  and hence the slope is 0. This means that the curve becomes horizontal, and cannot have a peak the way we have drawn it above.

At present we have no means for computing more values of  $\sin x$  and  $\cos x$ . However, using the few that we know and the derivative, we can convince ourselves that the graphs look as we have drawn them.

#### EXERCISES

- 1. Draw the graph of  $\tan x$ .
- 2. Let sec  $x = 1/\cos x$  be defined when  $\cos x \neq 0$ . Draw the graph of sec x.
- 3. Let  $\cot x = 1/\tan x$ . Draw the graph of  $\cot x$ . (Sec and  $\cot are$  abbreviations for the secant and  $\cot angent$ .)
- 4. Draw the graph of the function  $\sin(1/x)$ .
- 5. Draw the graph of the function  $x \sin(1/x)$ .
- 6. Draw the graph of the function  $x^2 \sin(1/x)$ . (In Exercises 4, 5, 6 the function is not defined for x = 0.)

7. Let  $f(x) = x^2 \sin(1/x)$  when  $x \neq 0$  and let f(0) = 0. Using the Newton quotient, show that f has a derivative at 0 and that f'(0) = 0.

8. Let  $f(x) = x \sin(1/x)$  when  $x \neq 0$  and let f(0) = 0. Show that f does not have a derivative at 0. (Look at small values of x like

$$h = \frac{2}{n\pi}$$

*n* being a large integer. Try n = 1, 2, 3, 4, etc. and see what happens to the values of f(x) and the Newton quotient

$$\frac{f(h) - f(0)}{h} .)$$

## §3. Addition formula

In this section we shall state and prove the most important formulas about sine and cosine.

To begin with, using the Pythagoras theorem, we observe that

$$(\sin x)^2 + (\cos x)^2 = 1$$

for all x. Indeed, if we have an angle A and we determine its sine and cosine from the right triangle, as in the following figure,



then we have

$$a^2 + b^2 = r^2.$$

Dividing by  $r^2$  yields

$$\left(\frac{a}{r}\right)^2 + \left(\frac{b}{r}\right)^2 = 1.$$

The same argument works when A is greater than  $\pi/2$ , by means of a triangle like this one:



In both cases, we have sine A = a/r and cosine A = b/r, so that we have the relation

 $(sine A)^2 + (cosine A)^2 = 1.$ 

It is customary to write the square of the sine and cosine as  $\sin^2 A$  and  $\cos^2 A$ .

Our main result is the addition formula.

THEOREM 2. For any angles A and B, we have

 $\sin (A + B) = \sin A \cos B + \cos A \sin B$  $\cos (A + B) = \cos A \cos B - \sin A \sin B.$ 

Proof. We shall prove the second formula first.

We consider two angles A, B and their sum:



We take two points P, Q as indicated, at a distance 1 from the origin O. We shall now compute the distance from P to Q, using two different coordinate systems.

First, we take a coordinate system as usual:



Then the coordinates of P are (1, 0) and those of Q are

 $(\cos{(A+B)},\sin{(A+B)}).$ 

The square of the distance between P and Q is

$$\sin^2 (A + B) + (\cos (A + B) - 1)^2$$
,

which is equal to

$$-2\cos\left(A+B\right)+2$$

Next we place the coordinate system as follows:



Then the coordinates of P become

$$(\cos A, \sin (-A)) = (\cos A, -\sin A).$$

Those of Q are simply  $(\cos B, \sin B)$ . The square of the distance between P and Q is

$$(\sin B + \sin A)^2 + (\cos B - \cos A)^2,$$

which is equal to

$$\sin^2 B + 2 \sin B \sin A + \sin^2 A + \cos^2 B - 2 \cos B \cos A + \cos^2 A$$
  
= 2 + 2 \sin A \sin B - 2 \cos A \cos B.

If we set the squares of the two distances equal to each other, we get our formula.

The addition formula for the sine can be obtained by the following device:

$$\sin (A + B) = \cos \left(A + B - \frac{\pi}{2}\right)$$
$$= \cos A \cos \left(B - \frac{\pi}{2}\right) - \sin A \sin \left(B - \frac{\pi}{2}\right)$$
$$= \cos A \sin B + \sin A \sin \left(\frac{\pi}{2} - B\right)$$
$$= \cos A \sin B + \sin A \cos B,$$

thereby proving our formula.

# §4. The derivatives

We shall prove:

**THEOREM** 3. The functions  $\sin x$  and  $\cos x$  have derivatives and

$$\frac{d (\sin x)}{dx} = \cos x$$
$$\frac{d (\cos x)}{dx} = -\sin x$$

*Proof.* We shall first determine the derivative of  $\sin x$ . We have to look at the Newton quotient of  $\sin x$ . It is

$$\frac{\sin\left(x+h\right)-\sin x}{h}$$

Using the addition formula to expand sin (x + h), we see that the Newton quotient is equal to

$$\frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \cdot$$

We put together the two terms involving  $\sin x$ :

$$\frac{\cos x \sin h + \sin x (\cos h - 1)}{h}$$

and separate our quotient into a sum of two terms:

$$\cos x \frac{\sin h}{h} + \sin x \frac{\cos h - 1}{h} \cdot$$

We now face the problem of finding the limit of  $\frac{\sin h}{h}$  and  $\frac{\cos h - 1}{h}$  as h approaches 0. This is a somewhat more difficult problem than those we encountered previously. We cannot tell right away what these limits will be. In the next section, we shall prove that

$$\lim_{h \to 0} \frac{\sin h}{h} = 1 \quad \text{and} \quad \lim_{h \to 0} \frac{\cos h - 1}{h} = 0.$$

Once we know these limits, then we see immediately that the first term approaches  $\cos x$  and the second term approaches

$$(\sin x) \cdot 0 = 0$$

Hence

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

[IV, §4]

THE DERIVATIVES

This proves that

$$\frac{d\,(\sin\,x)}{dx} = \,\cos\,x.$$

To find the derivative of  $\cos x$ , we could proceed in the same way, and we would encounter the same limits. However, there is a trick which avoids this.

We know that  $\cos x = \sin\left(x + \frac{\pi}{2}\right)$ . Let  $u = x + \frac{\pi}{2}$  and use the chain rule. We get

$$\frac{d\ (\cos x)}{dx} = \frac{d\ (\sin u)}{du}\ \frac{du}{dx}$$

However, du/dx = 1. Hence

$$\frac{d(\cos x)}{dx} = \cos u = \cos \left(x + \frac{\pi}{2}\right) = -\sin x,$$

thereby proving our theorem.

*Remark.* It is not true that the derivative of the function  $\sin^* x$  is  $\cos^* x$ . Using the chain rule, find out what its derivative is. The reason for using the radian measure of angles is to get a function  $\sin x$  whose derivative is  $\cos x$ .

## Exercises

1. What is the derivative of  $\tan x$ ?

Find the derivative of the following functions:

2.	$\sin(3x)$	3.	$\cos(5x)$
4.	$\sin\left(4x^2+x\right)$	5.	$\tan (x^3 - 5)$
6.	$\tan (x^4 - x^3)$	7.	$\tan(\sin x)$
8.	$\sin(\tan x)$	9.	$\cos(\tan x)$
~			

10. What is the slope of the curve  $y = \sin x$  at the point whose x-coordinate is  $\pi$ ?

Find the slope of the following curves at the indicated point (we just give the *x*-coordinate of the point):

11.  $y = \cos (3x)$  at  $x = \pi/3$ 12.  $y = \sin x$  at  $x = \pi/6$ 13.  $y = \sin x + \cos x$  at  $x = 3\pi/4$ 14.  $y = \tan x$  at  $x = -\pi/4$ 15.  $y = \frac{1}{\sin x}$  at  $x = -\pi/6$ 

## §5. Two basic limits

We shall first prove that

$$\lim_{h \to 0} \frac{\sin h}{h} = 1$$

Both the numerator and the denominator approach 0 as h approaches 0, and we get no information by trying some cancellation procedure, the way we did it for powers.

Let us assume first that h is positive, and look at the following diagram.

We take a circle of radius 1 and an angle of h radians. Let s be the altitude of the small triangle, and t that of the big triangle. Then

$$\sin h = \frac{s}{1} = s$$

and

$$\tan h = \frac{t}{1} = t = \frac{\sin h}{\cos h}.$$

We see that:

area of triangle OAB < area of sector OCB < area of triangle OCD.

The base OA of the small triangle is equal to  $\cos h$  and its altitude is  $\sin h$ .

The base OC of the big triangle is equal to 1. Its altitude is

$$t=\frac{\sin h}{\cos h}$$

The area of each triangle is  $\frac{1}{2}$  the base times the altitude.

The area of the sector is the fraction  $h/2\pi$  of the area of the circle, which is  $\pi$ . Hence the area of the sector is h/2. Thus we obtain:

$$\frac{1}{2}\cos h\sin h < \frac{1}{2}h < \frac{1}{2}\frac{\sin h}{\cos h}.$$

We multiply everywhere by 2 and get

$$\cos h \sin h < h < \frac{\sin h}{\cos h}$$



[IV, §5]

There are really two inequalities here. The first one is

$$\cos h \sin h < h.$$

Since h is positive, we can divide it by h and then divide by  $\cos h$ , which is also positive. This yields

$$\frac{\sin h}{h} < \frac{1}{\cos h} \cdot$$

The second inequality is

$$h < \frac{\sin h}{\cos h}$$
.

We multiply it by  $\cos h$  and divide it by h to get

$$\cos h < \frac{\sin h}{h}$$
.

Putting our two inequalities together, we get

$$\cos h < \frac{\sin h}{h} < \frac{1}{\cos h}.$$

The game is almost won. Letting h approach 0, we see that  $(\sin h)/h$  is squeezed between two quantities which approach 1. Hence it must approach 1 also, and our proof is complete.

We still have to consider the limit

$$\lim_{h\to 0}\frac{\cos h-1}{h}$$

and show that it is 0. We have

$$\frac{\cos h - 1}{h} = \frac{(\cos h - 1)(\cos h + 1)}{h(\cos h + 1)}$$
$$= \frac{\cos^2 h - 1}{h(\cos h + 1)}$$
$$= \frac{-\sin^2 h}{h(\cos h + 1)}.$$

We can write this last expression in the form

$$-rac{\sin h}{h}(\sin h)rac{1}{\cos h+1}\cdot$$

Using the property concerning the product of limits, we have a product of

three factors. The first is

$$-\frac{\sin h}{h}$$

and approaches -1 as h approaches 0.

The second is  $\sin h$  and approaches 0 as h approaches 0.

The third is

$$\frac{1}{\cos h + 1}$$

and its limit is  $\frac{1}{2}$  as h approaches 0.

Therefore the limit of the product is 0, and everything is proved!

We still have one thing to take care of. We computed our limit when h > 0. Suppose that h < 0. We can write

$$h = -k$$

with k > 0. Then

$$\frac{\sin\left(-k\right)}{-k} = \frac{-\sin k}{-k} = \frac{\sin k}{k} \cdot$$

As h tends to 0, so does k. Hence we are reduced to our previous limit because k > 0. A similar remark applies to our other limit involving  $\cos h$ .

## EXERCISES

Find the following limits, as h approaches 0.

1. 
$$\frac{\sin 2h}{h} \left[ \text{Hint: Put } k = 2h. \text{ Then } \frac{\sin 2h}{h} = 2\frac{\sin k}{k} \cdot \right]$$
  
2. 
$$\frac{\sin 3h}{h}$$
  
3. 
$$\frac{\sin h}{3h}$$
  
4. 
$$\frac{\tan h}{\sin h}$$
  
5. 
$$\frac{\cos 2h}{1 + \sin h}$$
  
6. 
$$\frac{\sin h^2}{h}$$
  
7. 
$$\frac{\sin 2h^2}{3h}$$
  
8. 
$$\frac{\sin h^3}{h^3}$$
  
9. 
$$\frac{\sin 2h^3}{h^3}$$
  
10. 
$$\frac{h \sin h}{\sin 2h^2}$$
  
11. 
$$\frac{(\sin h) (\sin 2h)}{(\sin 3h)h}$$

# CHAPTER V

# The Mean Value Theorem

Given a curve, y = f(x), we shall use the derivative to give us information about the curve. For instance, we shall find the maximum and minimum of the graph, and regions where the curve is increasing or decreasing. We shall use the mean value theorem, which is basic in the theory of derivatives.

# §1. The maximum and minimum theorem

Let f be a differentiable function. A point c such that f'(c) = 0 is called a *critical point* of the function. The derivative being zero means that the slope of the tangent line is 0 and thus that the tangent line itself is horizontal. We have drawn three examples of this phenomenon.



The third example is that of a function like  $f(x) = x^3$ . We have  $f'(x) = 3x^2$  and hence when x = 0, f'(0) = 0.

The other two examples are those of a maximum and a minimum, respectively, if we look at the graph of the function only near our point c. We shall now formalize these notions.

Let a, b be two numbers with a < b. We shall repeatedly deal with the interval of numbers between a and b. Sometimes we want to include the end points a and b, and sometimes we do not. We need some terminology to distinguish between these various types of intervals, and the standard terminology is as follows:

The collection of numbers x such that a < x < b is called the *open* interval between a and b.

The collection of numbers x such that  $a \leq x \leq b$  is called the *closed interval* between a and b. We denote this closed interval by the symbols [a, b]. (A single point will also be called a closed interval.)

If we wish to include only one end point, we shall say that the interval is *half closed*. We have of course two half-closed intervals, namely the one consisting of the numbers x with  $a \leq x < b$ , and the other one consisting of the numbers x with  $a < x \leq b$ .

Sometimes, if a is a number, we call the collection of numbers x > a (or x < a) an open interval. The context will always make this clear.

Let f be a function, and c a number at which f is defined. We shall say that c is a maximum of the function if

$$f(c) \geq f(x)$$

for all numbers x at which f is defined. If we have only  $f(c) \ge f(x)$  for all numbers x in some interval, then we say that c is a maximum of the function *in that interval*.

Example 1. Let  $f(x) = \sin x$ . Then  $\pi/2$  is a maximum for f because  $f(\pi/2) = 1$  and  $\sin x \leq 1$  for all values of x. Note that  $-3\pi/2$  is also a maximum for  $\sin x$ .

*Example 2.* Let f(x) = 2x, and view f as a function defined only on the interval

 $0 \leq x \leq 2.$ 

Then 2 is a maximum for the function in this interval because f(2) = 4 and  $f(x) \leq 4$  for all x in the interval.

*Example 3.* Let f(x) = 1/x. We know that f is not defined for x = 0. This function has no maximum. It becomes arbitrarily large when x comes close to 0 and x > 0.

We have illustrated our three examples in Figs. 4, 5, and 6.

In the next theorem, we shall prove that under certain circumstances, the derivative of a function at a maximum is 0.



THEOREM 1. Let f be a function which is defined and differentiable on the open interval a < x < b. Let c be a number in the interval which is a maximum for the function. (In other words,  $f(c) \ge f(x)$  for all x in the interval.) Then f'(c) = 0.

*Proof.* If we take small values of h (positive or negative), the number c + h will lie in the interval.



Let us first take h positive (see Fig. 7). We must have

$$f(c) \geq f(c+h)$$

no matter what h is (provided h is small). Therefore  $f(c + h) - f(c) \leq 0$ . Since h > 0, the Newton quotient

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

is  $\leq 0$ . Hence the limit is  $\leq 0$ , or in symbols:

$$\lim_{\substack{h\to 0\\h>0}}\frac{f(c+h)-f(c)}{h} \leq 0.$$

Now take h negative, say h = -k with k > 0. Then

$$f(c - k) - f(c) \leq 0, \quad f(c) - f(c - k) \geq 0$$

and the quotient is

$$\frac{f(c-k)-f(c)}{-k} = \frac{f(c)-f(c-k)}{k}$$

Thus the Newton quotient is  $\geq 0$ . Taking the limit as h (or k) approaches 0, we see that

$$\lim_{\substack{h\to 0\\h<0}}\frac{f(c+h)-f(c)}{h} \ge 0.$$

The only way in which our two limits can be equal is that they should both be 0. Therefore f'(c) = 0.

We can interpret our arguments geometrically by saying that the line between our two points slants to the left when we take h > 0 and slants to the right when we take h < 0. As h approaches 0, both lines must approach the tangent line to the curve. The only way this is possible is for the tangent line at the point whose x-coordinate is c to be horizontal. This means that its slope is 0, i.e. f'(c) = 0.

Everything we have done with a maximum could have been done with a minimum.

Let f be a function. We say that a number c is a minimum for f if  $f(c) \leq f(x)$  for all x at which the function is defined.

Theorem 1 remains true when we replace the word "maximum" by the word "minimum". It will be a good exercise for you to prove Theorem 1 for the minimum. When we refer to Theorem 1 we shall use it in both cases.

We illustrate various minima with the graphs of certain functions.

In Fig. 8 the function has a minimum. In Fig. 9 the minimum is at the end point of the interval. In Figs. 3 and 6 the function has no minimum.



FIGURE 9

[V, §1]

In the following picture, the point  $c_1$  looks like a maximum and the point  $c_2$  looks like a minimum, provided we stay close to these points, and don't look at what happens to the curve farther away.



FIGURE 10

There is a name for such points.

We shall say that a point c is a *local minimum* of the function f if there exists an interval

$$a_1 < c < b_1$$

such that  $f(c) \leq f(x)$  for all numbers x with  $a_1 \leq x \leq b_1$ .

Similarly, we define the notion of *local maximum*. (Do it yourself.) In Fig. 10, the point  $c_3$  is a local maximum,  $c_4$  is a local minimum, and  $c_5$  is a local maximum.

The actual maximum and minimum occur at the end points.

In practice, a function usually has only a finite number of critical points, and it is easy to find all points c such that f'(c) = 0. One can then determine by inspection which of these are maxima, which are minima, and which are neither.

*Example* 1. Find the critical points of the function  $f(x) = x^3 - 1$ . We have  $f'(x) = 3x^2$ . Hence there is only one critical point, namely x = 0.

Example 2. Find the critical points of the function

$$y = x^3 - 2x + 1.$$

The derivative is  $3x^2 - 2$ . It is equal to 0 precisely when

$$x^2 = \frac{2}{3},$$

which means  $x = \sqrt{2/3}$  or  $-\sqrt{2/3}$ . These are the critical points.

*Example 3.* Find the local maximum and minimum of the function  $f(x) = x^3 - 2x + 1$ .

The local maximum and minimum must be a critical point, hence we have only two possibilities, which we found in Example 2. These are  $x = \sqrt{2/3}$  and  $x = -\sqrt{2/3}$ . We make a small table of values of our function:



The number  $\sqrt{2/3}$  is between 0 and 1. Since f(1) = 0 and f(0) = 1, it follows that  $\sqrt{2/3}$  must be a local minimum.

Similarly,  $-\sqrt{2/3}$  is between 0 and -1. But f(0) = 1 and f(-1) = 2. Hence  $-\sqrt{2/3}$  is a local maximum. The sketch of the graph looks like this:



FIGURE 11

### Exercises

Find the critical points of the following functions:

1. $x^2 - 2x + 5$	2. $2x^2 - 3x - 1$
3. $3x^2 - x + 1$	4. $-x^2 + 2x + 2$
5. $-2x^2 + 3x - 1$	6. $x^3 + 2$
7. $x^3 - 3x$	8. $\sin x + \cos x$
9. $\cos x$	10. $\sin x$

11. A box with open top is to be made with a square base and a constant surface C. Determine the sides of the box if the volume is to be a maximum.

12. A container in the shape of a cylinder with open top is to have a fixed surface area C. Find the radius of its base and its height if it is to have maximum volume.

13. Do the above two problems when the box and container are closed at the top.

(The area of a circle of radius x is  $\pi x^2$  and its length is  $2\pi x$ . The volume of a cylinder whose base has radius x and of height y is  $\pi x^2 y$ .)

# §2. Existence of maxima and minima

It is important to know when a function has a maximum and a minimum. We shall describe a condition under which it does.

Up to now we have defined derivatives by means of a Newton quotient taken with h positive and negative. However, in Chapter II, we also described the right derivative and the left derivative.

Let f(x) be a function defined on an interval

$$a \leq x \leq b.$$

(We assume throughout that a < b.) We shall say that it is differentiable in this interval if it is differentiable in the interval

$$a < x < b$$
,

and if, in addition, it has a right derivative at a, and a left derivative at b. Thus we assume that the limits

$$\lim_{\substack{h \to 0 \\ h > 0}} \frac{f(a+h) - f(a)}{h}$$

and

$$\lim_{\substack{h\to 0\\h<0}}\frac{f(b+h)-f(b)}{h}$$

exist. These limits will be denoted by f'(a) and f'(b) just as with the ordinary derivative.

If we have a point x in the interval which is not equal to the end points, then f'(x) has the usual meaning.

Since the function may not be defined outside the interval, the quotient

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

would not be defined when h < 0, and similarly the quotient

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

would not be defined when h > 0.

We need some criterion to know when a function has a maximum and a minimum in an interval.

A function f is said to be *continuous* if for every x such that f is defined, we have

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

If f is differentiable, then it must be continuous.

In this chapter, we shall deal mostly with functions defined on an interval with end points a, b such that a < b, and these functions will be differentiable in the open interval, and continuous at the end points.

THEOREM 2. Let f(x) be a function which is continuous on the closed interval  $a \leq x \leq b$ . Then f has a maximum and also has a minimum in this interval.

This means that there is a point  $c_1$  in the interval such that  $f(c_1) \ge f(x)$  for all x in the interval, and there is a point  $c_2$  such that  $f(c_2) \le f(x)$  for all x in the interval.

If you look back to Fig. 3 and Fig. 6 of §1, you will see graphs of functions which have no maximum or minimum. The reason for this is that the functions are not defined over closed intervals. In the case of a hyperbola, as in Fig. 6, we could define the function at 0 in an arbitrary way, for instance let f(0) = 997. We still would not get a maximum or a minimum.

We shall take Theorem 2 for granted, without proof. If you are interested in seeing a proof, you can look in the appendix. The proof must be carried out by using special properties of numbers.

Combining Theorems 1 and 2, we obtain:

THEOREM 3. Let a, b be two numbers, a < b. Let f be a function which is continuous over the closed interval

 $a \leq x \leq b$ 

and differentiable on the open interval a < x < b. Assume that

$$f(a) = f(b) = 0.$$

Then there exists a point c such that

a < c < b

and such that f'(c) = 0.

*Proof.* If the function is constant in the interval, then its derivative is 0 and any point in the open interval a < x < b will do.

If the function is not constant, then there exists some point in the interval where the function is not 0, and this point cannot be one of the end points a or b. Suppose that some value of our function is positive. By Theorem 2, the function has a maximum. Let c be this maximum. Then f(c) must be greater than 0, and c cannot be either one of the end points.

Consequently

a < c < b.

By Theorem 1, we must have f'(c) = 0. This proves our theorem in case the function is positive somewhere in the interval.

If the function is negative for some number in the interval, then we use Theorem 2 to get a minimum, and we argue in a similar way, using Theorem 1 (applied to a minimum). (Write out the argument in full as an exercise.)

The following picture illustrates our Theorem 3.



FIGURE 12

## §3. The mean value theorem

Let f(x) be a function which is differentiable in the closed interval

 $a \leq x \leq b.$ 

We continue to assume throughout that a < b. This time we do not assume, as in Theorem 3, that f(a) = f(b) = 0. We shall prove that there exists a point c between a and b such that the slope of the tangent line at (c, f(c)) is the same as the slope of the line between the end points of our graph. In other words, the tangent line is parallel to the line passing through the end points of our graph.



FIGURE 13

The slope of the line between the two end points is

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

because the coordinates of the end points are (a, f(a)) and (b, f(b)) respectively. Thus we have to find a point c such that

$$f'(c) = \frac{f(b) - f(a)}{b - a} \cdot$$

THEOREM 4. Let a < b as before. Let f be a function which is continuous in the closed interval  $a \leq x \leq b$ , and differentiable in the interval a < x < b. Then there exists a point c such that a < c < b and

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

*Proof.* The equation of the line between the two end points is

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

Indeed, the slope

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

is the coefficient of x. When x = a, y = f(a). Hence we have written down the equation of the line having the given slope and passing through a given point. When x = b, we note that y = f(b).

We now consider geometrically the difference between f(x) and the straight line. In other words, we consider the function

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

Then

$$g(a) = f(a) - f(a) = 0$$

g(b) = f(b) - f(b) = 0

and also.

We can therefore apply Theorem 3 to the function g(x). We know that there is a point c between a and b, and not equal to a or b, such that

$$g'(c) = 0.$$

But

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

[V, §4]

Consequently

$$0 = g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}$$

This gives us the desired value for f'(c).

The difference between f(x) and the straight line becomes 0 at the end points. This is the geometric idea which allows us to apply our Theorem 3.

*Example* 1. Let a = 1 and b = 2. Let  $f(x) = x^2$ . Find a point c as in Theorem 4.

We have f'(x) = 2x, and

$$\frac{f(b) - f(a)}{b - a} = \frac{2^2 - 1^2}{2 - 1} = \frac{3}{1} = 3.$$

We have to solve the equation 2x = 3. We get x = 3/2. Thus c = 3/2 is a point such that f'(c) has the required value.

*Example 2.* Let  $f(x) = x^3 + 2x$ , and let a = -1, b = 2. Find a number c as in Theorem 4.

We have  $f'(x) = 3x^2 + 2$ , and

$$\frac{f(b) - f(a)}{b - a} = \frac{12 - (-3)}{2 - (-1)} = \frac{15}{3} = 5.$$

We must solve  $3x^2 + 2 = 5$  and get x = 1 or x = -1. We take c = 1.

## Exercises

Find a number c as in the mean value theorem for each one of the following functions:

$1. f(x) = x^3,$	$1 \leq x \leq 3$
2. $f(x) = (x - 1)^3$ ,	$-1 \leq x \leq 2$
$3. f(x) = x^3,$	$-1 \leq x \leq 3$
4. $f(x) = x^2 + 5x$ ,	$0 \leq x \leq 2$

# §4. Increasing and decreasing functions

Let f be a function defined on some interval (which may be open or closed). We shall say that f is *increasing* over this interval if

$$f(x_1) \leq f(x_2)$$

whenever  $x_1$  and  $x_2$  are two points of the interval such that  $x_1 \leq x_2$ .

Thus, if a number lies to the right of another, the value of the function at the larger number must be greater than or equal to the value of the function at the smaller number.

In the next figure, we have drawn the graph of an increasing function.



FIGURE 14

We say that a function defined on some interval is *decreasing* over this interval if

$$f(x_1) \ge f(x_2)$$

whenever  $x_1$  and  $x_2$  are two points of the interval such that  $x_1 \leq x_2$ .

Observe that a constant function (whose graph is horizontal) is both increasing and decreasing.

If we want to omit the equality sign in our definitions, we shall use the word *strictly* to qualify decreasing or increasing. Thus a function f is *strictly increasing* if

$$f(x_1) < f(x_2)$$

whenever  $x_1 < x_2$ , and f is strictly decreasing if

whenever 
$$x_1 < x_2$$
.

The mean value theorem gives us a good test to determine when a function is increasing or decreasing.

**THEOREM** 5. Let f be a function which is continuous in some interval, and differentiable in the interval (excluding the end points).

If f'(x) = 0 in the interval (excluding the end points), then f is constant.

If f'(x) > 0 in the interval (excluding the end points), then f is strictly increasing.

If f'(x) < 0 in the interval (excluding the end points), then f is strictly decreasing.

*Proof.* Let  $x_1$  and  $x_2$  be two points of the interval, and suppose  $x_1 < x_2$ . By the mean value theorem, there exists a point c such that  $x_1 < c < x_2$  and

$$f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

#### [V, §4]

The difference  $x_2 - x_1$  is positive, and we have

$$f(x_2) - f(x_1) = (x_2 - x_1)f'(c).$$

If the derivative of the function is 0 throughout the interval, excluding the end points, then the right-hand side of (1) is equal to 0, and hence  $f(x_1) = f(x_2)$ . This means that, given any two points of the interval, the values of the function are equal, and hence the function is constant.

If the derivative f'(x) is >0 for all x in the interval, excluding the end points, then f'(c) > 0 (because c is in the interval). Hence the product  $(x_2 - x_1)f'(c)$  is positive, and

$$f(x_1) < f(x_2).$$

This proves that the function is increasing.

We leave the proof of the last assertion to you, as an exercise.

*Example* 1. Determine the regions of increase and decrease for the function  $f(x) = x^2 + 5$ .

The derivative is f'(x) = 2x. It is >0 when x > 0 and hence f is strictly increasing when x > 0. The derivative is negative when x < 0 and hence f is strictly decreasing when x < 0. We leave the sketch of the graph to you.

*Example 2.* Determine the regions of increase and decrease for the function  $f(x) = x^3 - 2x + 1$ .

The derivative is  $3x^2 - 2$ . The condition

$$3x^2 - 2 > 0$$

is equivalent with  $3x^2 > 2$  or  $x^2 > 2/3$ . Thus when  $x > \sqrt{2/3}$  or  $-x > \sqrt{2/3}$ , we have  $x^2 > 2/3$ . The function is strictly increasing when  $x > \sqrt{2/3}$  and when  $x < -\sqrt{2/3}$ .

The condition

f'(x) < 0

is equivalent to  $3x^2 - 2 < 0$  or  $x^2 < 2/3$ . Thus the function is decreasing when

$$-\sqrt{2/3} < x < \sqrt{2/3}$$

*Example 3.* Prove that  $\sin x \leq x$  for  $x \geq 0$ .

We let  $f(x) = x - \sin x$ . Then f(0) = 0. Furthermore,

$$f'(x) = 1 - \cos x.$$

Since  $\cos x \leq 1$  for all x, it follows that  $f'(x) \geq 0$  for all x. Hence f(x) is an increasing function. Hence  $f(x) \geq 0$  for all  $x \geq 0$  and therefore  $\sin x \leq x$ .

We shall now give a criterion which will allow us to decide when a function is positive or negative throughout an interval. For this, we need what is known as the *intermediate value theorem*.

THEOREM 6. Let  $x_1$  and  $x_2$  be two numbers, with  $x_1 \leq x_2$ . Let f be a function which is defined on the interval

$$x_1 \leq x \leq x_2$$

and assume that f is continuous. Let  $y_1 = f(x_1)$  and  $y_2 = f(x_2)$ . Let  $y_3$  be a number between  $y_1$  and  $y_2$ . Then there exists a number  $x_3$  between  $x_1$  and  $x_2$  such that  $f(x_3) = y_3$ .

The proof of Theorem 6 is given in the appendix. It is the type of proof which reduces the theorem to properties of numbers, and which we omit.

As a consequence of Theorem 6, we see that a function which is continuous in an interval, is positive at some point of the interval, and is not equal to 0 at any point of the interval, must be positive throughout the interval. Indeed, if it were negative at some point of the interval, the intermediate value theorem would ensure that there is a point in the interval at which it is equal to 0. For our purposes, we therefore have the following criterion.

THEOREM 7. Let f be a function having a derivative f' at every point of an interval, and assume that f' is continuous. If f' is positive at some point of the interval, and  $f'(x) \neq 0$  for any x in the interval except possibly at the end points, then f is strictly increasing in the interval.

In practice, we can use Theorem 7 by checking that f' is not equal to 0 at any point. This is easier than checking that f'(x) > 0 at every point of an interval.

Needless to say, an analogous statement holds when f' is assumed to be negative throughout the interval. State it yourself as an exercise.

Before closing this chapter, we emphasize the first assertion of Theorem 5. We had already seen that the derivative of a constant is 0. We now have proved the converse. It gives us the following very important result:

**THEOREM 8.** Let f(x) and g(x) be two functions which are differentiable in some interval and assume that

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

[V, §4]

for all x in the interval. Then there is a constant C such that

f(x) = g(x) + C

for all x in the interval.

*Proof.* Let  $\varphi(x) = f(x) - g(x)$  be the difference of our two functions. Then

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

Hence  $\varphi(x)$  is constant, i.e.  $\varphi(x) = C$  for some number C.

## Exercises

Determine when the following functions are increasing and decreasing.

- 1.  $f(x) = x^3 + 1$ 2. f(x) = (x 1)(x 2)(x 3)3.  $f(x) = x^2 x + 5$ 4.  $f(x) = \sin x + \cos x$ 5.  $f(x) = \sin 2x$  ( $0 \le x \le 2\pi$ )6.  $f(x) = x^4 3x^2 + 1$ 7.  $f(x) = x^3 + x 2$ 8.  $f(x) = -x^3 + 2x + 1$ 9.  $f(x) = 2x^3 + 5$ 10.  $f(x) = 5x^2 + 1$
- 11. Consider only values of  $x \ge 0$ , and let

$$f_1(x) = x - \sin x \qquad f_2(x) = -1 + \frac{x^2}{2} + \cos x$$

$$f_3(x) = -x + \frac{x^3}{3 \cdot 2} + \sin x \qquad f_4(x) = 1 - \frac{x^2}{2} + \frac{x^4}{4 \cdot 3 \cdot 2} - \cos x$$

$$f_5(x) = x - \frac{x^3}{3 \cdot 2} + \frac{x^5}{5 \cdot 4 \cdot 3 \cdot 2} - \sin x$$

(a) Determine whether  $f_1(x)$  is increasing or decreasing. Using the value of  $f_1(x)$  at 0, show that  $\sin x \leq x$ .

(b) Determine which of the other functions are increasing or decreasing. Using the value of each function at 0, prove the following inequalities:

$$x - \frac{x^3}{3 \cdot 2} \le \sin x \le x - \frac{x^3}{3 \cdot 2} + \frac{x^5}{5 \cdot 4 \cdot 3 \cdot 2}$$
$$1 - \frac{x^2}{2} \le \cos x \le 1 - \frac{x^2}{2} + \frac{x^4}{4 \cdot 3 \cdot 2}$$

(c) Show how the above procedure can be continued to get further inequalities for  $\sin x$  and  $\cos x$ . Give the general formula.

12. Assume that there is a function f(x) such that  $f(x) \neq 0$  for any x, and f'(x) = f(x). Let g(x) be any function such that g'(x) = g(x). Show that there is a constant C such that g(x) = Cf(x). [Hint: Differentiate the quotient g/f.]

# CHAPTER VI

# **Sketching Curves**

We have developed enough techniques to be able to sketch curves and graphs of functions much more efficiently than before. We shall investigate systematically the behavior of a curve, and the mean value theorem will play a fundamental role.

We shall especially look for the following aspects of the curve:

- 1. Intersections with the coordinate axes.
- 2. Critical points.
- 3. Regions of increase.
- 4. Regions of decrease.
- 5. Maxima and minima (including the local ones).
- 6. Behavior as x becomes very large positive and very large negative.
- 7. Values of x near which y becomes very large positive or very large negative.

These seven pieces of information will be quite sufficient to give us a fairly accurate idea of what the graph looks like.

We shall also introduce a new way of describing points of the plane and functions, namely polar coordinates. These are especially useful in connection with the trigonometric functions.

## §1. Behavior as x becomes very large

Suppose we have a function f defined for all sufficiently large numbers. Then we get substantial information concerning our function by investigating how it behaves as x becomes large.

For instance, sin x oscillates between -1 and +1.

However, polynomials don't oscillate. When  $f(x) = x^2$ , as x becomes large positive, so does  $x^2$ . Similarly with the function  $x^3$ , or  $x^4$  (etc.).

Example 1. Consider a polynomial

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

We can write it in the form

$$x^3\left(1+\frac{2}{x^2}-\frac{1}{x^3}\right).$$
94

© Springer Science+Business Media New York 2002

When x becomes very large, the expression

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

approaches 1. In particular, given a small number  $\delta > 0$ , we have, for all x sufficiently large, the inequality

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

Therefore f(x) satisfies the inequality

$$x^{3}(1 - \delta) < f(x) < x^{3}(1 + \delta).$$

This tells us that f(x) behaves very much like  $x^3$  when x is very large.

A similar argument can be applied to any polynomial.

Example 2. Consider a quotient of polynomials like

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

Dividing numerator and denominator by  $x^3$ , we get

$$Q(x) = \frac{1 + \frac{2}{x^2} - \frac{1}{x^3}}{2 - \frac{1}{x^2} + \frac{1}{x^3}}$$

As x becomes very large, the numerator approaches 1 and the denominator approaches 2. Thus our fraction approaches  $\frac{1}{2}$ .

Example 3. Consider the quotient

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

Does it approach a limit as x becomes very large?

If we divide numerator and denominator by  $x^3$ , then we see that our quotient can be written

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

As x becomes very large, the numerator approaches 0 and the denominator approaches 1. Consequently the quotient approaches 0.

Example 4. Consider the quotient

$$Q(x) = \frac{x^3 - 1}{x^2 + 5}$$

and determine what happens when x becomes very large.

We divide numerator and denominator by  $x^2$ . This gives

$$Q(x) = \frac{x - \frac{1}{x^2}}{1 + \frac{5}{x^2}}.$$

As x becomes very large, the numerator is approximately equal to x and the denominator approaches 1. Thus the quotient is approximately equal to x.

These four examples are typical of what happens when we deal with quotients of polynomials.

Instead of saying "when x becomes very large", we shall also say "as x approaches infinity", or even better, "as x approaches plus infinity". Thus in Example 2 we would write

$$\lim_{x\to\infty}Q(x)=\frac{1}{2}.$$

In Example 3 we would write

$$\lim_{x\to\infty}Q(x)=0.$$

In Example 4, we would write

$$\lim_{x\to\infty}Q(x) = \infty.$$

We emphasize that this way of writing is an *abbreviation*, a shorthand, for the sentences we wrote in our various examples. There is NO NUMBER WHICH IS CALLED INFINITY. The symbol  $\infty$  will not be used except in the context we have just described.

Of course, we could also investigate what happens when x becomes very large *negative*, or, as we shall also say, when x approaches minus infinity, which we write  $-\infty$ .

In Example 2, we see that as x becomes very large negative, our quotient Q(x) still approaches  $\frac{1}{2}$ , because a fraction like  $1/x^3$  becomes very small. (For instance 1/-10,000 = -1/10,000 is very small negative.) Thus we would write

$$\lim_{x \to -\infty} Q(x) = \frac{1}{2}.$$
#### Exercises

Find the limits of the following quotients Q(x) as x becomes very large positive or negative. In other words, find

$$\lim_{x \to \infty} Q(x) \quad \text{and} \quad \lim_{x \to -\infty} Q(x).$$
1.  $\frac{2x^3 - x}{x^4 - 1}$ 
2.  $\frac{\sin x}{x}$ 
3.  $\frac{\cos x}{x}$ 
4.  $\frac{x^2 + 1}{\pi x^2 - 1}$ 
5.  $\frac{\sin 4x}{x^3}$ 
6.  $\frac{5x^4 - x^3 + 3x + 2}{x^3 - 1}$ 
7.  $\frac{-x^2 + 1}{x + 5}$ 
8.  $\frac{2x^4 - 1}{-4x^4 + x^2}$ 
9.  $\frac{2x^4 - 1}{-4x^3 + x^2}$ 
10.  $\frac{2x^4 - 1}{-4x^5 + x^2}$ 

Describe the behavior of the following polynomials as x becomes very large positive and very large negative.

 11.  $x^3 - x + 1$  12.  $-x^3 - x + 1$  

 13.  $x^4 + 3x^3 + 2$  14.  $-x^4 + 3x^3 + 2$  

 15.  $2x^5 + x^2 - 100$  16.  $-3x^5 + x + 1000$  

 17.  $10x^6 - x^4$  18.  $-3x^6 + x^3 + 1$ 

19. A function f(x) which can be expressed as follows:

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0,$$

where n is a positive integer and the  $a_n, a_{n-1}, \ldots, a_0$  are numbers, is called a polynomial. If  $a_n \neq 0$ , then n is called the *degree* of the polynomial. Describe the behavior of f(x) as x becomes very large positive or negative, n is odd or even, and  $a_n > 0$  or  $a_n < 0$ . (You will have eight cases to consider.)

# §2. Curve sketching

We shall put together all the information we have gathered up to now to get an accurate picture of the graph of a function. We deal systematically with the seven properties stated in the introduction, and our discussion will take the form of working out examples.

Example 1. Sketch the graph of the curve

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

and determine the seven properties stated in the introduction.

1. When x = 0, we have f(x) = -1. When x = 1, f(x) = 0.

2. The derivative is

$$f'(x) = \frac{2}{(x+1)^2}.$$

(You can compute it using the quotient rule.) It is never 0, and therefore the function has no critical points.

3. The denominator is a square and hence is always positive. Thus f'(x) > 0 for all x. The function is increasing for all x. Of course, the function is not defined for x = -1 and neither is the derivative. Thus it would be more accurate to say that the function is increasing in the region

x < -1

and is increasing in the region x > -1.

4. There is no region of decrease.

5. Since the derivative is never 0, there is no relative maximum or minimum.

6. As x becomes very large positive, our function approaches 1 (using the method of the preceding section). As x becomes very large negative, our function also approaches 1.

Finally, there is one more useful piece of information which we can look into, when f(x) itself becomes very large positive or negative:

7. As x approaches -1, the denominator approaches 0 and the numerator approaches -2. If x approaches -1 from the right, then the denominator is positive, and the numerator is negative. Hence the fraction

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

is negative, and is very large negative.



98

[VI, §2]

CURVE SKETCHING

99

If x approaches -1 from the left, then x - 1 is negative, but x + 1 is negative also. Hence f(x) is positive and very large, since the denominator is small when x is close to -1. Putting all this information together, we see that the graph looks like that in the preceding figure.

We have drawn the two lines x = -1 and y = 1, as these play an important role when x approaches -1 and when x becomes very large, positive or negative.

Example 2. Sketch the graph of the curve

$$y = -x^3 + 3x - 5.$$

1. When x = 0, we have y = -5.

2. The derivative is

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

It is 0 when  $3x^2 = 3$ , which is equivalent to saying that  $x^2 = 1$ , or  $x = \pm 1$ . These are the critical points.

3. The derivative is positive when  $-3x^2 + 3 > 0$ , which amounts to saying that

$$3x^2 < 3$$
 or  $x^2 < 1$ .

This is equivalent to the condition

-1 < x < 1,

which is therefore a region of increase.

4. When  $-3x^2 + 3 < 0$ , the function decreases. This is the region given by the inequality

 $3x^2 > 3$ 

or  $x^2 > 1$ . Thus when

x > 1 or x < -1,

the function decreases.

5. Since the function decreases when x < -1 and increases when x > -1 (and is close to -1), we conclude that the point -1 is a local minimum. Also, f(-1) = -7.

Similarly, the point 1 is a relative maximum and f(1) = -3.

6. As x becomes very large positive,  $x^3$  is very large positive and  $-x^3$  is very large negative. Hence our function becomes very large negative, as we see if we put it in the form

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

Similarly, as x becomes very large negative, our function becomes very large positive.

Putting all this information together, we see that the graph looks like this:



### Exercises

Sketch the following curves, indicating all the information stated in the introduction:

1.  $y = \frac{x^2 + 2}{x - 2}$ 2.  $y = \frac{x-3}{x^2+1}$ 3.  $y = \frac{x+1}{x^2+1}$ 4.  $y = \sin^2 x$ 6.  $y = \frac{x^2 - 1}{x}$ 5.  $y = \cos^2 x$ 8.  $y = \frac{x^3 + 1}{x + 1}$ 7.  $y = \tan^2 x$ 10.  $y = \frac{2x^2 - 1}{r^2 - 2}$ 9.  $y = x^4 - 2x^3 + 1$ 11.  $y = \frac{2x - 3}{3x + 1}$ 12.  $y = x^4 + 4x$ 13.  $y = x^5 + x$ 14.  $y = x^6 + 6x$ 15.  $y = x^7 + x$ 16.  $y = x^8 + x$ 

17. Which of the following polynomials have a minimum (for all x)?

(a)  $x^6 - x + 2$ (b)  $x^5 - x + 2$ (c)  $-x^6 - x + 2$ (d)  $-x^5 - x + 2$ (e)  $x^6 + x + 2$ (f)  $x^5 + x + 2$ 

18. Which of the polynomials in Exercise 17 have a maximum (for all x)? 19. Sketch the curves of Exercise 17.

Sketch the following curves:

20.  $x^3 + x - 1$ 21.  $x^3 - x - 1$ 22.  $-x^3 + 2x + 5$ 23.  $-2x^3 + x + 2$ 24.  $x^3 - x^2 + 1$ 25.  $x^3 + x$ 

26. Let a, b, c, d be four distinct numbers. What would the following curves look like? You may assume a < b < c < d.

- (i) (x a)(x b)
- (ii) (x a)(x b)(x c)
- (iii) (x a)(x b)(x c)(x d)

### REMARK

In the classical study of functions f(x), one sometimes asks for points such that f''(x) = 0, f''(x) > 0, and f''(x) < 0. If we view the second derivative of f as the rate of change of the slope of our curve, in a region where f is increasing, the positivity of the second derivative means that the rate of change of the slope is positive. This can be interpreted geometrically by saying that the curve is bending upwards. If on the other hand the second derivative is negative, the curve is bending downwards. The following two pictures illustrate this:



A point where a curve changes its behavior from bending upwards to bending downwards (or vice versa) is called an *inflection point*. If the curve is the graph of a function f(x), then f''(x) must be 0 at such a point. The following picture illustrates this:



SKETCHING CURVES

The determination of inflection points of course gives us more information about our curves than we listed above. However, I believe that this additional piece of information is not worth spending much time on, especially since the above-mentioned seven pieces of information already tend to be slightly oppressive. We shall therefore not go deeper into the question, leaving it as an optional topic of study for anyone who has a strong urge to look into it. The reader will find the discussion in the supplementary exercises for Chapter V.

## §3. Polar coordinates

Instead of describing a point in the plane by its coordinates with respect to two perpendicular axes, we can also describe it as follows. We draw a line between the point and a given origin. The angle which this line makes with the horizontal axis and the distance between the point and the origin determine our point. Thus the point is described by a pair of numbers  $(r, \theta)$ , which are called its *polar coordinates*.



If we have our usual axes and x, y are the ordinary coordinates of our point, then we see that

$$rac{x}{r} = \cos heta$$
 and  $rac{y}{r} = \sin heta,$   
 $x = r \cos heta$  and  $y = r \sin heta.$ 

whence

This allows us to change from polar coordinates to ordinary coordinates.

It is to be understood that r is always supposed to be  $\geq 0$ . In terms of the ordinary coordinates, we have

$$r = \sqrt{x^2 + y^2}.$$

*Example* 1. Find polar coordinates of the point whose ordinary coordinates are  $(1, \sqrt{3})$ .

The polar coordinates are  $(2, \pi/3)$ .

We observe that we may have several polar coordinates corresponding to the same point. The point whose polar coordinates are  $(r, \theta + 2\pi)$  is the same as the point  $(r, \theta)$ . Thus in our example above,  $(2, \pi/6 + 2\pi)$  would also be polar coordinates for our point. In practice, we usually use the value for the angle which lies between 0 and  $2\pi$ .

Let f be a function whose values are  $\geq 0$ . If we set  $r = f(\theta)$ , then the set of points  $(\theta, f(\theta))$  is the graph of the function in polar coordinates. We can also view  $r = f(\theta)$  as the equation of a curve.

*Example 2.* Sketch the graph of the function  $r = \sin \theta$  for  $0 \leq \theta \leq \pi$ .

First we make a table of values, as indicated. As  $\theta$  ranges from 0 to  $\pi/2$ , sin  $\theta$  increases until it reaches 1. As  $\theta$  goes from  $\pi/2$  to  $\pi$ , the sine decreases back to 0. Hence the graph looks like this:



*Example* 3. The equation of the circle of radius 3 and center at the origin in polar coordinates is simply

$$r = 3$$

### Exercises

- 1. Plot the following points in polar coordinates: (a)  $(2, \pi/4)$  (b)  $(3, \pi/6)$  (c)  $(1, -\pi/4)$  (d)  $(2, -3\pi/6)$
- 2. Same directions as in Exercise 1.

(a) (1, 1) (b) (4, -3)

(These are polar coordinates. Just show approximately the angle represented by the given coordinates.)

3. Find polar coordinates for the following points given in the usual x- and y-coordinates:

(a) (1, 1) (b) (-1, -1) (c)  $(3, 3\sqrt{3})$  (d) (-1, 0)

Sketch the graphs of the following curves given in polar coordinates:

4. r = 55.  $r = \sin 2\theta$ 6.  $r = \theta$ 7.  $r = \sin \theta + \cos \theta$ 8.  $r = \frac{2}{2 - \cos \theta}$ 9.  $r = \sin 3\theta$ 10.  $r^2 = 2a^2 \cos 2\theta$  (a > 0) 11.  $r = 1 + \cos \theta$ 12.  $r = a \cos \theta$  (a > 0) 13.  $r = a \sin \theta$  (a > 0)

### §4. Parametric curves

There is one other way in which we can describe a curve. Suppose that we look at a point which moves in the plane. Its coordinates can be given as a function of time t. Thus, when we give two functions of t, say

$$x = f(t), \qquad y = g(t),$$

we may view these as describing a point moving along a curve.

For example, if we let  $x = \cos t$  and  $y = \sin t$ , then our point moves along a circle, counterclockwise, with uniform speed.

When (x, y) is described by two functions of t as above, we say that we have a *parametrization of the curve* in terms of the *parameter t*.

*Example 1.* Sketch the curve  $x = t^2$ ,  $y = t^3$ .

We can make a table of values as usual. We also investigate when x and y are increasing or decreasing functions of t. For instance, taking the derivative, we get

 $\frac{dx}{dt} = 2t$ 

and

$$\frac{dy}{dt} = 3t^2.$$

Thus x increases when t > 0 and decreases when t < 0. The y-coordinate is increasing since  $t^2 > 0$  (unless t = 0). Furthermore, the x-coordinate is always positive (unless t = 0). Thus the graph looks like this:



Example 2. We look at the circle

$$x = \cos \theta, \qquad y = \sin \theta.$$

We shall find another parametrization. Consider the circle of radius 1 and the point (x, y) on it:



Let

 $t = \frac{y}{x+1} \qquad (x \neq -1).$ 

The equation of the circle is

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

Then we can interpret t geometrically to be  $\tan \frac{\theta}{2}$ .

From our expression for t we obtain

$$y = (x + 1)t$$
 and  $y^2 = (x + 1)^2 t^2$ .

On the other hand,

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

In the two expressions for  $y^2$ , we cancel (x + 1)  $(x \neq -1)$  and obtain

$$1 - x = (x + 1)t^2$$
.

From this we solve for x, and get

$$x = \frac{1-t^2}{1+t^2}$$

From the formula y = t(x + 1), we get

$$y = \frac{2t}{1+t^2}$$

Thus, for  $x \neq -1$ , we can get t from x and y, but conversely we can recover x, y from values of t, which can be given arbitrarily. This allows us to find points on the circle explicitly simply by giving t arbitrary values, which may be selected to be rational numbers. Points (x, y) such that x and y are rational numbers are called *rational points*. The above parametrization shows us how to get all of them (except when x = -1).

#### Exercises

Sketch the following curves given in parametric form:

1. x = t + 1, y = 3t + 42.  $x = 1 + t^2$ , y = 3 - t3.  $x = 1 - t^2$ , y = t  $(-1 \le t \le 1)$ 4.  $x = t - \sin t$ ,  $y = 1 - \cos t$ 5.  $x = 2(t - \sin t)$ ,  $y = 2(1 - \cos t)$ 

One can also give a parametrization of curves in polar coordinates. Sketch the following curves in polar coordinates:

6.  $r = t^3$ ,  $\theta = \pi t^2$  7. r = t,  $\theta = t^2$ 

8. Find at least five rational points on the circle of radius 1. (It is very difficult in general to find rational points on curves. It is rather remarkable that we did it on the circle.)

9. Using the method above, find an analogous parametrization of the circle

$$x^2 + y^2 = 9.$$

10. Prove that on any curve

 $x^n + y^n = 1$ 

 $(n = \text{positive integer} \geq 3)$ , there is only a finite number of rational points. (If you solve this, you will be instantly world-famous among mathematicians. It is in fact believed that the only rational points are those for which x or y = 0. This is Fermat's problem, and has been verified for many values of n. It is also unknown whether there exist infinitely many rational points on a curve of type

$$y^2 = f(x),$$

where f is a polynomial of degree  $\geq 5$ , having distinct roots, for instance

$$y^2 = (x - 1)(x - 2)(x - 3)(x - 4)(x - 5)$$

This is a special case of a more general conjecture made by Mordell some forty years ago, and unproved to this day.)

# CHAPTER VII

# **Inverse Functions**

Suppose that we have a function, for instance

$$y = 3x - 5.$$

Then we can solve for x in terms of y, namely

$$x = \frac{1}{3}(y+5).$$

Thus x can be expressed as a function of y.

Although we are able to solve by means of an explicit formula, there are interesting cases where x can be expressed as a function of y, but without such an explicit formula. In this chapter, we shall investigate such cases.

### **§1.** Definition of inverse functions

Let y = f(x) be a function, defined for all x in some interval. If, for each value  $y_1$  of y, there exists exactly one value  $x_1$  of x in the interval such that  $f(x_1) = y_1$ , then we can define an *inverse function* 

$$x = g(y)$$

by the rule: Given a number y, we associate with it the unique number x in the interval such that f(x) = y.

Our inverse function is defined only at those numbers which are values of f. We have the fundamental relation f(g(y)) = y and g(f(x)) = x.

For example, consider the function  $y = x^2$ , which we view as being defined only for  $x \ge 0$ . Every positive number (or 0) can be written uniquely as the square of a positive number (or 0). Hence we can define the inverse function, which will also be defined for  $y \ge 0$ , but not for y < 0.

The next theorem gives us a good criterion when the inverse function is defined.

THEOREM 1. Let f(x) be a function which is strictly increasing. Then the inverse function exists.

107

*Proof.* This is practically obvious: Given a number  $y_1$  and a number  $x_1$  such that  $f(x_1) = y_1$ , there cannot be another number  $x_2$  such that  $f(x_2) = y_1$  unless  $x_2 = x_1$ , because if  $x_2 \neq x_1$ , then either  $x_2 > x_1$ , in which case  $f(x_2) > f(x_1)$ , or  $x_2 < x_1$ , in which case  $f(x_2) < f(x_1)$ .

Since the positivity of the derivative gives us a good test when a function is strictly increasing, we are able to define inverse functions whenever the function is differentiable and its derivative is positive.

As usual, what we have said above applies as well to functions which are strictly decreasing, and whose derivatives are negative.

Using the intermediate value theorem, we now conclude:

THEOREM 2. Let f be a continuous function on the closed interval  $a \leq x \leq b$  and assume that f is strictly increasing. Let  $f(a) = \alpha$  and  $f(b) = \beta$ . Then the inverse function is defined on the closed interval  $[\alpha, \beta]$ .

*Proof.* Given any number  $\gamma$  between  $\alpha$  and  $\beta$ , there exists a number c between a and b such that  $f(c) = \gamma$ , by the intermediate value theorem. Our assertion now follows from Theorem 1.

If we let g be this inverse function, then  $g(\alpha) = a$  and  $g(\beta) = b$ . Furthermore, the inverse function is characterized by the relation

$$f(x) = y$$
 if and only if  $x = g(y)$ .

*Example 1.* Let  $f(x) = x^3 - 2x + 1$ , viewed as a function on the interval  $x > \sqrt{2/3}$ . Can you define the inverse function? For what numbers? If g is the inverse function, what is g(0)? What is g(5)?

Since  $f'(x) = 3x^2 - 2$ , the derivative is positive when  $x > \sqrt{2/3}$ . Hence our function is strictly increasing and the inverse function is defined.

We know (from the techniques of the previous chapter) that  $x = \sqrt{2/3}$  is a minimum of f, and  $f(\sqrt{2/3}) = (2/3)^{3/2} - 2(2/3)^{1/2} + 1$ . The inverse function is defined therefore for  $y > f(\sqrt{2/3})$ .

Since f(1) = 0, we get g(0) = 1. Since f(2) = 5, we get g(5) = 2.

Please note that we do not give an explicit formula for our inverse function.

*Example* 2. Let  $f(x) = x^n$  (*n* being a positive integer). We view *f* as defined only for numbers x > 0. Since f'(x) is  $nx^{n-1}$ , the function is strictly increasing. Hence the inverse function exists. This inverse function *g* is in fact what we mean by the *n*-th root. In particular, we have proved that every positive number has an *n*-th root.

In all the exercises of the previous chapter you determined intervals over which certain functions increase and decrease. You can now define inverse functions for such intervals. In most cases, you cannot write down an explicit formula for such inverse functions.

# §2. Derivative of inverse functions

We shall state below a theorem which allows us to determine the derivative of an inverse function when we know the derivative of the given function.

THEOREM 3. Let a, b be two numbers, a < b. Let f be a function which is differentiable on the interval a < x < b and such that its derivative f'(x) is >0 for all x in this open interval. Then the inverse function x = g(y) exists, and we have

$$g'(y) = rac{1}{f'(x)}$$

*Proof.* We are supposed to investigate the Newton quotient

$$\frac{g(y+k)-g(y)}{k}\cdot$$

The following picture illustrates the situation:



By the intermediate value theorem, every number of the form y + k with small values of k can be written as a value of f. We let x = g(y) and we let h = g(y + k) - g(y). Then

$$x = g(y)$$
 and  $g(y + k) = x + h$ .

Furthermore, y + k = f(x + h) and hence

$$k = f(x+h) - f(x).$$

The Newton quotient for g can therefore be written

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

and we see that it is the reciprocal of the Newton quotient for f, namely

$$rac{1}{rac{f(x+h)-f(x)}{h}}\cdot$$

As h approaches 0, we know that k approaches 0. Conversely, as k approaches 0, we know that there exists exactly one value of h such that f(x + h) = y + k, because the inverse function is defined. Consequently, the corresponding value of h must also approach 0.

If we now take the limit of the reciprocal of the Newton quotient of f, as h (or k) approaches 0, we get

$$\frac{1}{f'(x)}$$
.

By definition, this is the derivative g'(y) and our theorem is proved.

*Example 1.* Let  $f(x) = x^3 - 2x + 1$ , viewed as a function on the interval  $x > \sqrt{2/3}$ . Let g be the inverse function. What is g'(0)? What is g'(5)?

Since  $f'(x) = 3x^2 - 2$  we know that

$$g'(y) = \frac{1}{f'(x)}$$

whenever y = f(x) or x = g(y). But

$$f(1) = 0$$

and

$$g(0) = 1$$

Therefore

$$g'(0) = \frac{1}{f'(1)} = 1.$$

Similarly, f(2) = 5 and g(5) = 2. Hence

$$g'(5) = rac{1}{f'(2)} = rac{1}{10},$$

because f'(2) = 10.

Please note that the derivative g'(y) is given in terms of the inverse function. We don't have a formula in terms of y.

The theorem giving us the derivative of the inverse function could also be expressed by saying that

$$\frac{dx}{dy} = \frac{1}{dy/dx} \cdot$$

Here also, the derivative behaves as *if* we were taking a quotient. Thus the notation is very suggestive and we can use it from now on without thinking, because we proved a theorem justifying it.

THE ARCSINE

#### Exercises

In the exercises of Chapter V, §4, restrict f to an interval so that the inverse function is defined in an interval containing the indicated point, and find the derivative of the inverse function at that point. (Let g denote this inverse function in every case.)

1. g'(2)2. g'(6)3. g'(7)4. g'(-1)5.  $g'(\sqrt{3}/2)$ 6. g'(-1)7. g'(0)8. g'(2)9. g'(21)10. g'(11)

# §3. The arcsine

It is impossible to define an inverse function for the function  $y = \sin x$ because to each value of y there correspond infinitely many values of x (differing by  $2\pi$ ). However, if we restrict our attention to special intervals, we can define the inverse function.

We restrict the sine function to the interval

$$-\frac{\pi}{2} \leq x \leq \frac{\pi}{2}$$

The derivative of  $\sin x$  is  $\cos x$  and in that interval, we have

$$0 < \cos x$$

except when  $x = \pi/2$  or  $x = -\pi/2$  in which case the cosine is 0.

Therefore, in the interval

$$-\frac{\pi}{2} \leq x \leq \frac{\pi}{2}$$

the function is strictly increasing by Theorem 2 of Chapter V, §4. The inverse function exists, and is called the *arcsine*.

Let  $f(x) = \sin x$ , and  $x = \arcsin y$ , the inverse function. Since f(0) = 0we have  $\arcsin 0 = 0$ . Furthermore, since  $\sin (-\pi/2) = -1$  and  $\sin (\pi/2) = 1$ , we know that the inverse function is defined over the interval going from -1 to +1, i.e. for

$$-1 \leq y \leq 1.$$

In words, we can say loosely that  $\arcsin x$  is the angle whose sine is x. (We throw in the word loosely because, strictly speaking,  $\arcsin x$  is a number, and not an angle, and also because we mean the angle between  $-\pi/2$  and  $\pi/2$ .)

The derivative of  $\sin x$  is positive for

$$-\pi/2 < x < \pi/2$$

Since the derivative of the inverse function x = g(y) is 1/f'(x), the deriva-

tive of  $\arcsin y$  is also positive, in the interval

-1 < y < 1.

Therefore the inverse function is strictly increasing in that interval. Its graph looks like this:



At this point, we cannot call the horizontal axis simultaneously the xand y-axis. Thus for the moment, the horizontal axis as drawn above is the y-axis and the vertical axis is the x-axis.

According to the general rule for the derivative of inverse functions, we know that when

$$y = \sin x$$
 and  $x = \arcsin y$ 

the derivative is

$$\frac{dx}{dy} = \frac{1}{\cos x}$$

When x is very close to  $\pi/2$ , we know that  $\cos x$  is close to 0. Therefore the derivative is very large. Hence the curve is almost vertical. Similarly, when x is close to  $-\pi/2$  and y is close to -1, the curve is almost vertical, as drawn.

Finally, it turns out that we can express our derivative explicitly as a function of y. Indeed, we have the relation

$$\sin^2 x + \cos^2 x = 1,$$
$$\cos^2 x = 1 - \sin^2 x.$$

whence

In the interval between 
$$-\pi/2$$
 and  $\pi/2$ , the cosine is  $\geq 0$ . Hence we can take the square root, and we get

$$\cos x = \sqrt{1 - \sin^2 x}$$

in that interval. Since  $y = \sin x$ , we can write our derivative in the form

$$\frac{dx}{dy} = \frac{1}{\sqrt{1 - y^2}}$$

which is expressed entirely in terms of y.

Having now obtained all the information we want concerning the arcsine, we shift back our letters to the usual ones. We state the main properties as a theorem.

**THEOREM** 4. View the sine function as defined on the interval

 $[-\pi/2, \pi/2].$ 

Then the inverse function is defined on the interval [-1, 1]. Call it q(x) =arcsin x. Then g is differentiable in the open interval -1 < x < 1, and

$$g'(x) = \frac{1}{\sqrt{1-x^2}}$$

### Exercises

1. Define the inverse function accosine, viewing the cosine only on the interval  $[0, \pi]$ .

2. What is the derivative of arccosine?

3. Let  $g(x) = \arcsin x$ . What is  $g'(\frac{1}{2})$ ? What is  $g'(1/\sqrt{2})$ ? What is  $g(\frac{1}{2})$ ? What is  $g(1/\sqrt{2})$ ?

4. Let  $g(x) = \arccos x$ . What is  $g'(\frac{1}{2})$ ? What is  $g'(1/\sqrt{2})$ ? What is  $g(\frac{1}{2})$ ? What is  $q(1/\sqrt{2})$ ?

5. Let sec  $x = 1/\cos x$ . Define the inverse function of the secant over a suitable interval and obtain a formula for the derivative of this inverse function.

Find the following numbers.

6.	$\arcsin (\sin 3\pi/2)$	7.	$\arcsin (\sin 2\pi)$
8.	$\arccos(\cos 3\pi/2)$	9.	$\arccos(\cos -\pi/2)$

10. arcsin (sin  $-3\pi/4$ )

# §4. The arctangent

Let  $f(x) = \tan x$  and view this function as defined over the interval

$$-\frac{\pi}{2} < x < \frac{\pi}{2} \cdot$$

As x goes from  $-\pi/2$  to  $\pi/2$ , the tangent goes from very large negative

[VII, §4]

values to very large positive values. As x approaches  $\pi/2$ , the tangent has in fact arbitrarily large positive values, and similarly when x approaches  $-\pi/2$ , the tangent has arbitrarily large negative values.

We recall that the graph of the tangent looks like this:



The derivative of  $\tan x$  is

$$\frac{d\,(\tan\,x)}{dx}=\,1\,+\,\tan^2x.$$

Hence the derivative is always positive, and our function is strictly increasing. Therefore the inverse function is defined for all numbers. We call it the *arctangent*. Its derivative is also positive (because it is the reciprocal of the derivative of the tangent) and hence the arctan is strictly increasing also.

The graph looks like this:



Let  $x = g(y) = \arctan y$ . Then

$$g'(y) = \frac{1}{1 + \tan^2 x}$$

so that

$$g'(y) = \frac{1}{1+y^2}.$$

Here again we are able to get an explicit formula for the derivative of the inverse function.

As with the arcsine, when dealing simultaneously with the function and its inverse function, we have to keep our letters x, y separate. However, we now summarize the properties of the arctan in terms of our usual notation.

**THEOREM 5.** The inverse function of the tangent is defined for all numbers. Call it the arctangent. Then it has a derivative, and that derivative is given by the relation

$$\frac{d (\arctan x)}{dx} = \frac{1}{1+x^2} \cdot$$

As x becomes very large positive, arctan x approaches  $\pi/2$ . As x becomes very large negative, arctan x approaches  $-\pi/2$ . The arctangent is strictly increasing for all x.

In words, we can say that  $\arctan x$  is the angle whose tangent is x, between  $-\pi/2$  and  $\pi/2$ .

### Exercises

1. Let g be the arctan function. What is g(1)? What is  $g(1/\sqrt{3})$ ? What is g(-1)? What is  $g(\sqrt{3})$ ?

2. Let g be the arctan function. What is g'(1)? What is  $g'(1/\sqrt{3})$ ? What is g'(-1)? What is  $g'(\sqrt{3})$ ?

3. Suppose you were to define an inverse function for the tangent in the interval  $\pi/2 < x < 3\pi/2$ . What would be the derivative of this inverse function?

- (a)  $\arctan(\tan 3\pi/4)$ ? (b)  $\arctan$
- (c)  $\arctan(\tan 5\pi/6)$ ?
- (b)  $\arctan(\tan 2\pi)$ ?
- (d)  $\arctan(\tan -\frac{5\pi}{6})$ ?

# CHAPTER VIII

# **Exponents and Logarithms**

We remember that we had trouble at the very beginning with the function  $2^x$  (or  $3^x$ , or  $10^x$ ). It was intuitively very plausible that there should be such functions, satisfying the fundamental equation

$$2^{x+y} = 2^x 2^y$$

for all numbers x, y, and  $2^0 = 1$ , but we had difficulties in saying what we meant by  $2^{\sqrt{2}}$  (or  $2^{\pi}$ ).

It is the purpose of this chapter to give a systematic treatment of this function, and others like it.

We shall see that its inverse function is defined for positive numbers. It is called the log (or rather the log to the base 2). Thus  $y = 2^x$  if and only if  $x = \log_2 y$ . For instance,

 $3 = \log_2 8$ 

and

 $8 = 2^{3}$ 

are two ways of saying the same thing.

Let us assume for the moment that we can make sense of the function  $2^x$ , and let us see how we could find its derivative.

We form the Newton quotient. It is

$$\frac{2^{x+h}-2^x}{h}$$

Using the fundamental equation we see that this quotient is equal to

$$\frac{2^{x}2^{h}-2^{x}}{h}=2^{x}\frac{2^{h}-1}{h}\cdot$$

As h approaches 0,  $2^x$  remains fixed, but it is very difficult to see what happens to

$$\frac{2^h-1}{h} \cdot \frac{116}{116}$$

© Springer Science+Business Media New York 2002

It is not at all clear that this quotient approaches a limit. Roughly speaking, we meet a difficulty which is analogous to the one we met when we tried to find the derivative of  $\sin x$ . However, in the present situation, a direct approach would lead to much greater difficulties than those which we met when we discussed

$$\lim_{h\to 0}\frac{\sin h}{h}.$$

It is in fact true that

$$\lim_{h \to 0} \frac{2^h - 1}{h}$$

exists. We see that it does not depend on x. It depends only on 2.

If we tried to take the derivative of  $10^x$ , we would end up with the problem of determining the limit

$$\lim_{h\to 0}\frac{10^h-1}{h},$$

which is also independent of x.

There seems to be no reason for selecting 2, or 10, or any other number a in investigating the function  $a^x$ . However, we shall see that there exists a number called e such that

$$\lim_{h \to 0} \frac{e^h - 1}{h}$$

is equal to 1. This is perfectly marvelous, because if we then form the Newton quotient for  $e^x$ , we get

$$rac{e^{x+h}-e^x}{h}=\,e^x\cdotrac{e^h-1}{h}\cdot$$

Therefore its limit as h approaches 0 is  $e^x$  and hence

$$\frac{d(e^x)}{dx} = e^x.$$

We shall find out eventually how to compute e. Its value is 2.718.... In Chapter XIV you will have the technique to compute e to any degree of accuracy you wish.

The sales talk which precedes must now be set aside in your mind. We start from scratch, and in order to develop the theory in the easiest way, it is best to start with the log, and not with the exponential function. We shall then have no difficulties in determining all the limits which arise.

## §1. The logarithm

We define a function  $\log x$  to be the area under the curve 1/x between 1 and x if  $x \ge 1$ , and the negative of the area of the curve 1/x between 1 and x if 0 < x < 1. In particular,  $\log 1 = 0$ .



The shaded portion of our picture represents the area under the curve between 1 and x and we have taken an instance where x > 1.

If x < 1 and x > 0, we would have the following picture:



If 0 < x < 1, we have said that  $\log x$  is equal to the negative of the area. Thus  $\log x < 0$  if 0 < x < 1 and  $\log x > 0$  if x > 1.

(In defining the log, we have appealed to our geometric intuition of area, just as we did when we defined the sine and cosine. In the next chapter, we shall see how one can avoid this appeal to geometry, and give a purely analytic definition.)

The fundamental fact concerning the logarithm is the following.

**THEOREM** 1. The function  $\log x$  is differentiable, and

$$\frac{d\ (\log x)}{dx} = \frac{1}{x} \cdot$$

Proof. We form the Newton quotient

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

and have to prove that it approaches 1/x as a limit when h approaches 0.

Let us take h > 0 for the moment. Then  $\log (x + h) - \log x$  is the area under the curve between x and x + h. Since the curve 1/x is decreasing, this area satisfies the following inequalities:

$$h\frac{1}{x+h} < \log (x+h) - \log x < h\frac{1}{x}$$

Indeed, 1/x is the height of the big rectangle as drawn on the next figure, and 1/(x + h) is the height of the small rectangle. Since h is the base of the rectangle, and since the area under the curve 1/x between x and x + h is in between the two rectangles, we see that it satisfies our inequalities.



We divide both sides of our inequalities by the positive number h. Then the inequalities are preserved, and we get

$$\frac{1}{x+h} < \frac{\log\left(x+h\right) - \log x}{h} < \frac{1}{x}$$

As h approaches 0, our Newton quotient is squeezed between 1/(x + h) and 1/x and consequently approaches 1/x. This proves our theorem in case h > 0.

When h < 0 we use an entirely similar argument, which we leave as an exercise. (You have to pay attention to the sign of the log. Also when you divide an inequality by h and h < 0 then the inequality gets reversed. However, you will again see that the Newton quotient is squeezed between 1/x and 1/(x + h).)

From now on, we do not need our geometric intuition any more. The notion of area was used only to give us the existence of a function whose derivative is 1/x, and whose value at 1 is 0. All the arguments which follow depend only on the fact that we have such a function.

Let, therefore,  $\log x$  be a function defined for x > 0, such that  $\log 1 = 0$ , and such that

$$\frac{d\ (\log x)}{dx} = \frac{1}{x} \cdot$$

If g(x) is another such function, then it differs from log x by a constant C, i.e.

$$g(x) = \log x + C$$

for all x > 0. (Use Theorem 8 of Chapter V, §4.) Setting x = 1 gives g(1) = C. Since we assume that g(1) = 0 we see that  $g(x) = \log x$ . Thus there is only one function having the properties mentioned above.

THEOREM 2. If a, b are two numbers > 0, then

$$\log (ab) = \log a + \log b$$

*Proof.* Consider the function  $f(x) = \log (ax)$ , defined for all x > 0. We can take its derivative by the chain rule, and we obtain

$$rac{df}{dx} = rac{1}{ax} \cdot a = rac{1}{x} \cdot$$

 $\left(\operatorname{Put} u = ax \text{ and remember that } \frac{df}{dx} = \frac{df}{du} \frac{du}{dx} \cdot \right)$  Therefore, our functions f(x) and  $\log x$  have the same derivative. Consequently, they differ by a constant:

$$\log(ax) = \log x + C.$$

This is true for all x > 0. In particular, it is true for x = 1. This yields

$$\log a = C$$

and determines the constant C as being equal to  $\log a$ .

But the above relation being true for all x > 0, it is also true if we set x = b. In that case, we obtain

$$\log (ab) = \log b + \log a,$$

thereby proving our theorem.

Please appreciate the elegance and efficiency of the arguments!

THEOREM 3. The function  $\log x$  is strictly increasing for x > 0. It takes on arbitrarily large positive and negative values.

[VIII, §1]

#### THE LOGARITHM

*Proof.* Since the derivative is 1/x, which is positive for all x > 0, our function is strictly increasing. Since  $\log 1 = 0$ , we conclude for instance that  $\log 2 > 0$ .

Using Theorem 2 we now see that

 $log 4 = log (2 \cdot 2) = 2 log 2,$   $log 8 = log (4 \cdot 2) = log 4 + log 2 = 3 log 2,$  $log 16 = log (8 \cdot 2) = log 8 + log 2 = 4 log 2,$ 

and so on. In general,

 $\log 2^n = n \log 2$ 

for any positive integer n. As n becomes very large,  $n \log 2$  also becomes very large.

Concerning the negative values, observe that

$$0 = \log 1 = \log \left(2 \cdot \frac{1}{2}\right) = \log 2 + \log \left(\frac{1}{2}\right)$$

Therefore

$$\log\left(\frac{1}{2}\right) = -\log 2,$$

and arguing as before,

$$\log\left(\frac{1}{2^n}\right) = -n\log 2.$$

For large positive integers n, the number  $-n \log 2$  is very large negative.

These same arguments can be applied to prove the following theorem.

THEOREM 4. If n is an integer, positive or negative, and a is a number >0, then

$$\log\left(a^{n}\right) = n\log a.$$

*Proof.* We proceed stepwise, assuming first that n is positive. Then

$$\log (a^2) = \log a + \log a = 2 \log a.$$

We continue with  $a^3$ ,  $a^4$ , etc. We leave the arguments as an exercise.

Suppose that n is negative, say n = -m with m positive. Then

$$0 = \log 1 = \log (a^m \cdot a^{-m}) = \log (a^m) + \log (a^{-m}).$$

Therefore

$$\log (a^n) = \log (a^{-m}) = -m \log (a) = n \log a,$$

thereby proving our theorem.

By the intermediate value theorem, the function log takes on *all* values. Its graph looks like that in the next figure.



*Remark.* We shall sometimes consider composite functions of the type  $\log(f(x))$ . Since the log is not defined for numbers <0, this expression is defined only for numbers x such that f(x) > 0. This is to be understood whenever we write such an expression.

Thus when we write  $\log (x - 2)$ , this is defined only when x - 2 > 0, in other words x > 2. When we write  $\log (\sin x)$ , this is meaningful only when  $\sin x > 0$ . It is not defined when  $\sin x \leq 0$ .

#### Exercises

1. What is the tangent line to the curve  $y = \log x$  at the point whose x-coordinate is (a) 2, (b) 5, (c)  $\frac{1}{2}$ .

2. What is the equation of the tangent line of the curve  $y = \log (x^2 + 1)$  at the point whose x-coordinate is (a) -1, (b) 2, (c) -3.

+3))

3. Find the derivatives of the following functions:

(a)	$\log(\sin x)$	(b)	$\sin (\log (2x))$
(c)	$\log (x^2 + 5)$	(d)	$\frac{\log 2x}{\sin x}$

4. What is the equation of the tangent line of the curve  $y = \log (x + 1)$  at the point whose x-coordinate is 3?

5. What is the equation of the tangent line of the curve  $y = \log (2x - 5)$  at the point whose x-coordinate is 4?

6. Prove that  $\log x < x$  for all x > 1. [Hint: Let  $f(x) = x - \log x$ , find f(1) and show that f is strictly increasing.]

7. Let h be a positive number. Compare the area under the curve 1/x between 1 and 1 + h with the area of suitable rectangles to show that

$$\frac{h}{1+h} < \log\left(1+h\right) < h.$$

8. Prove that

$$\lim_{h \to 0} \frac{1}{h} \log (1+h) = 1.$$

[VIII, §2]

9. Prove that for every positive integer n, we have

$$\frac{1}{n+1} < \log\left(1 + \frac{1}{n}\right) < \frac{1}{n}$$

10. Let

$$a_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} - \log n$$

for each integer  $n \ge 2$ . Show that  $a_{n+1} < a_n$ .

11. Let 
$$b_n = a_n - \frac{1}{n}$$
. Show that  $b_{n+1} > b_n$ .

12. The sequence of positive numbers  $a_n$  decreases, and the sequence of positive numbers  $b_n$  increases. Since  $b_n - a_n$  becomes arbitrarily small as n becomes large, it follows that there is a unique number C such that

 $b_n < C < a_n$ 

for all positive integers n. This number is called *Euler's constant*. Prove that this constant is not a rational number. (Undying fame awaits you if you can do this exercise.)

## §2. The exponential function

We can apply the theory of the inverse function. Since the function log is strictly increasing, the inverse function is defined and we call it exp. Since log takes on *all* values, the inverse function is defined for *all* numbers, positive or negative.

Since  $0 = \log 1$ , we have by definition

$$1 = \exp(0).$$

THEOREM 5. If z, w are two numbers, then

$$\exp(z+w) = \exp(z) \cdot \exp(w).$$

*Proof.* Let  $a = \exp z$  and  $b = \exp w$ . Then  $z = \log a$  and  $w = \log b$  by definition of the inverse function. By Theorem 2, we know that

$$z + w = \log (ab).$$

By the definition of the inverse function, this means that

$$\exp\left(z+w\right)=ab.$$

However,  $ab = \exp(z) \cdot \exp(w)$ . Our theorem is proved.

We define the number e to be exp (1). This is the same as saying that  $\log e = 1$  or exp (1) = e.

(With the geometric interpretation of the log as area under the curve 1/x, this means that e is the number such that the area between 1 and e is equal to 1.)

Using Theorem 5, we conclude that

$$\exp(2) = \exp(1 + 1) = \exp(1) \exp(1) = e^2.$$

Similarly,

$$\exp(3) = \exp(2 + 1) = \exp(2) \exp(1) = e^2 \cdot e = e^3.$$

Proceeding stepwise, we conclude that

$$\exp\left(n\right) = e^n$$

for every positive integer n.

If n is a negative integer, write n = -m where m is positive. Then

$$1 = \exp(0) = \exp(m - m) = \exp(m) \exp(-m).$$

Dividing by exp (m), which is  $e^m$ , we get

$$\exp\left(-m\right)=\frac{1}{e^{m}}\cdot$$

We see therefore that our exp function gives us the power function for positive and negative integers m.

THEOREM 6. The function exp is differentiable, and

$$\frac{d\ (\exp x)}{dx} = \exp x.$$

*Proof.* By the theory of the derivative of the inverse function, we know that it is differentiable. If  $y = \exp x$  and  $x = \log y$ , then the theory of the derivative of inverse functions gives us

$$\frac{dy}{dx} = \frac{1}{dx/dy} \cdot$$

But dx/dy = 1/y. Hence

$$\frac{dy}{dx} = \frac{1}{1/y} = y = \exp x,$$

thereby proving our theorem.

From now on, we agree to write  $e^x$  instead of exp (x). In view of Theorem 5, we have the rule

$$e^{z+w} = e^z e^w$$

for all numbers z, w, and  $e^0 = 1$ .

[VIII, §2]

The preceding theorem then looks like

$$\frac{d(e^x)}{dx} = e^x$$

By definition, the derivative of  $e^x$  is the limit of the Newton quotient

$$\frac{e^{x+h}-e^x}{h}=e^x\frac{e^h-1}{h}$$

as h approaches 0. Hence, at the very end of our theory, we now obtain in a very natural way the fact that

$$\lim_{h \to 0} \frac{e^h - 1}{h} = 1.$$

As we said in the introduction, this fact would have been very troublesome to obtain directly.

We are now in a position to see that the graph of  $e^x$  looks like this:



The function is always strictly increasing, and since e > 1,  $e^n$  becomes very large as n becomes very large. Hence so does  $e^x$  for any x.

Actually,  $e^x$  increases quite fast. We shall make this a little more precise in §4.

#### EXERCISES

1. What is the equation of the tangent line to the curve  $y = e^{2x}$  at the point whose x-coordinate is (a) 1, (b) -2, (c) 0.

2. What is the equation of the tangent line to the curve  $y = e^{x/2}$  at the point whose x-coordinate is (a) -4, (b) 1, (c) 0.

3. What is the equation of the tangent line to the curve  $y = xe^x$  at the point whose x-coordinate is 2?

- 4. Find the derivatives of the following functions:
  - (a)  $e^{\sin 3x}$  (b)  $\log (e^x + \sin x)$
  - (c)  $\sin(e^{x+2})$  (d)  $\sin(e^{4x-5})$

125

5. Sketch the curve  $y = e^{-1/x}$  (defined for  $x \neq 0$ ).

6. Sketch the curve  $y = e^{-1/x^2}$  (defined for  $x \neq 0$ ).

7. Let f(x) be a differentiable function over some interval satisfying the relation f'(x) = Kf(x) for some constant K. Show that there is a constant C such that  $f(x) = Ce^{Kx}$ .

8. Show that  $1 + x < e^x < \frac{1}{1 - x}$  for 0 < x < 1, and for -1 < x < 0.

# §3. The general exponential function

Let a be a positive number, and x any number. We define

Thus

$$a^{x} = \exp (x \log a) = e^{x \log a}.$$
$$a^{\sqrt{2}} = e^{\sqrt{2} \log a}.$$

Using the properties of the log and exp, it is easy to prove that

$$a^{x+y} = a^x a^y$$

for all numbers x and y and that  $a^0 = 1$ . Furthermore,  $(a^x)^y = a^{xy}$ .

We leave the proof as an exercise. However, we observe that when x is a positive integer n, then  $a^x$  is indeed the product of a with itself n times. For instance, take x = 2. Then

$$e^{2 \log a}$$

is equal to

 $e^{\log a + \log a} = (e^{\log a})^2 = a \cdot a,$ 

and similarly,

```
e^{3 \log a}
```

is equal to

```
e^{(\log a) + (\log a) + (\log a)}
```

which is equal to  $(e^{\log a})^3$ . Thus our definition of  $a^x$  as  $e^{x \log a}$  is consistent with our original notation of  $a^n$  when x is a positive integer n. This is the justification for the definition in the general case.

The moral of the story is that, by going all the way around the difficulties, and giving up on a frontal attack on the function  $a^x$ , we have recovered it at the end, together with all of the desired properties. For instance, we have its derivative:

THEOREM 7. The derivative of  $a^x$  is  $a^x (\log a)$ .

[VIII, §3]

*Proof.* We use the chain rule. Let  $u = (\log a)x$ . Then  $du/dx = \log a$  and  $a^x = e^u$ . Hence

$$\frac{d(a^x)}{dx} = a^x (\log a)$$

as desired.

In particular,

$$\frac{d(2^x)}{dx} = 2^x \log 2.$$

This result clarifies the mysterious limit

$$\lim_{h\to 0}\frac{2^h-1}{h}$$

which we encountered in the introduction. We are now able to see that this limit is  $\log 2$ , and it arises in a very natural way. We also see that when we take a = e we obtain the only exponential function which is equal to its own derivative. For any other choice of the constant a, we would get an extraneous constant appearing in the derivative.

As an application of our theory of the exponential function, we also can take care of the general power function (which we had left dangling in Chapter III).

THEOREM 8. Let c be any number, and let

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

be defined for x > 0. Then f'(x) exists and is equal to

$$f'(x) = cx^{c-1}.$$

Proof. By definition,

$$f(x) = e^{c \log x} = e^u$$

if we put  $u = c \log x$ . Then

$$\frac{du}{dx} = \frac{c}{x} \cdot$$

Using the chain rule, we see that

$$f'(x) = e^{u} \cdot \frac{du}{dx} = e^{c \log x} \cdot \frac{c}{x}$$
$$= x^{c} \cdot \frac{c}{x} = cx^{c-1}.$$

This proves our theorem.

When x, y are two numbers such that  $y = 2^x$ , it is customary to say that x is the log of y to the base 2. Similarly, if a is a number >0, and  $y = a^x = e^{x \log a}$ , we say that x is the log of y to the base a. When  $y = e^x$ , we simply say that  $x = \log y$ .

The log to the base a is sometimes written  $\log_a$ . It occurs very infrequently, and you might as well forget about it.

#### Exercises

- 1. What is the derivative of  $10^x$ ?  $7^x$ ?
- 2. What is the derivative of  $3^x$ ?  $\pi^x$ ?

3. What is the derivative of the function  $x^x$  (defined for x > 0)? [Hint:  $x^x = e^{x \log x}$ .]

4. What is the derivative of the function  $x^{(x^x)}$ ?

5. Referring to Exercise 8 of §1, what is

$$\lim_{\substack{h \to 0 \\ h > 0}} (1+h)^{1/h}?$$

6. What about when h < 0, does the limit come out the same?

7. Sketch the curves  $y = 2^x$  and  $y = 2^{-x}$ .

8. Find the equation of the tangent line to the curve  $y = x^x$  at the point x = 1.

9. Find the equation of the tangent line to each curve of Exercise 1 at x = 0.

10. Find the equation of the tangent line to each curve of Exercise 2 at x = 2.

# §4. Order of magnitude

The area under the curve 1/x between 1 and 2 is at least equal to 1. (Visualize a rectangle like that in the picture below.)



Since  $\log e = 1$  and  $\log 4 = 2 \log 2 \ge 1$ , it follows that e < 4. On the other hand, the area under our curve between 1 and 2 is less than 1 because 1/x is at most equal to 1 in that interval. Thus we have

2 < e < 4.

We shall learn how to compute e to any degree of accuracy in a later chapter.

In order to prove Theorem 10, we need an auxiliary statement.

THEOREM 9. Let a be a number >0. Then

$$\frac{(1+a)^n}{n}$$

becomes very large as n becomes very large.

Proof. We can write

$$(1+a)^n = 1 + na + \frac{n(n-1)}{2}a^2 + b,$$

where b is some number  $\geq 0$ . This is easily seen by expanding the product of (1 + a) with itself n times. Consequently, dividing by n, we obtain

$$\frac{(1+a)^n}{n} = \frac{1}{n} + a + \frac{n-1}{2}a^2 + \frac{b}{n},$$

and b/n is  $\geq 0$ . As *n* becomes large, we see that the term  $\frac{n-1}{2}a^2$  becomes

large. All the other terms are  $\geq 0$ . Hence we have proved our theorem.

We apply this to the case of  $e^n/n$ . We know that e can be written in the form 1 + a with a > 0. Hence we see that  $e^n/n$  becomes arbitrarily large as n increases indefinitely.

THEOREM 10. The function  $f(x) = e^x/x$  is strictly increasing for x > 1. As x becomes very large positive, so does f(x).

Proof. To verify the first assertion, take the derivative

$$f'(x) = \frac{xe^x - e^x}{x^2} = \frac{e^x}{x^2} (x - 1).$$

We know that  $e^x$  and  $x^2$  are >0 for all x > 0. Hence our derivative f'(x) is positive when x > 1, and so our function is strictly increasing.

When x is an integer n, we know that the function f(n) becomes arbitrarily large. It follows that it does also for all x when x becomes arbitrarily large. This proves our theorem.

Observe that we can conclude that

$$\frac{x}{e^x}$$
 and  $\frac{n}{e^n}$ 

approach 0 as x and n become large. In fact, for any number a > 0 we also observe that

$$\frac{n}{(1+a)^n}$$

approaches 0 as n becomes large.

COROLLARY 1. As x becomes arbitrarily large, the function  $x - \log x$  also becomes arbitrarily large.

*Proof.* The log of  $e^x/x$  is  $x - \log x$ . Our assertion follows from the properties of the log proved previously, namely when y becomes very large, so does log y.

COROLLARY 2. As x becomes very large, the quotient

$$\frac{x}{\log x}$$

also becomes very large.

*Proof.* Let  $y = \log x$ . Then  $x = e^y$  and our quotient is of the form

 $\frac{e^{y}}{y}$  .

We know that  $\log x$  becomes large when x becomes large. Thus our assertion follows from the theorem.

COROLLARY 3. As x becomes very large,  $x^{1/x}$  approaches 1 as a limit.

*Proof.* We have  $x^{1/x} = e^{(\log x)/x}$ . As x becomes very large,  $(\log x)/x$  approaches 0 and hence its exponent approaches 1.

This corollary is used frequently when we look at integers n instead of arbitrary numbers x. Thus

 $n^{1/n}$ 

approaches 1 as n becomes very large.

The next theorem is a refinement of some previous ones.

**THEOREM** 11. Let m be a positive integer. Then the function

$$\frac{e^x}{x^m}$$

is strictly increasing for x > m and becomes very large when x becomes very large.

130

[VIII, §4]

*Proof.* Let  $f(x) = e^x/x^m$ . Since

we have

$$x^m = e^{m \log x}$$

 $x = e^{\log x}$ .

Therefore

$$f(x) = e^{x - m \log x}.$$

We have

$$f'(x) = e^{x - m \log x} \left( 1 - \frac{m}{x} \right) \cdot$$

This is >0 when x > m and our first assertion is proved.

As to the second, to prove that  $e^{x}/x^{m}$  becomes very large when x does, it suffices to do it for its log. Taking the log we see that

$$\log f(x) = x - m (\log x),$$

which can also be written

$$(\log x)\left(\frac{x}{\log x} - m\right)$$
.

As x becomes large, so does  $\log x$ , and so does  $x/\log x$  by Corollary 2 above. Since m is fixed, our expression becomes very large, as desired.

### EXERCISES

1. Sketch the curve  $y = xe^x$ , giving the usual pieces of information from the beginning of Chapter VI. [*Hint:* The derivative is  $xe^x + e^x = e^x(x+1)$ . It is equal to 0 just when x = -1, which is therefore the only critical point. Take it from there.]

Sketch the graphs of the following functions. (In Exercises 6 through 10,  $x \neq 0$ .)

2. $xe^{-x}$	3. $xe^{-x^2}$	4. $x^2 e^{-x^2}$
5. $x^2 e^{-x}$	6. $e^{x}/x$	7. $e^{x}/x^{2}$
8. $e^{x}/x^{3}$	9. $e^{1/x}$	10. $xe^{1/x}$

11. Show that the equation  $e^x = ax$  has at least one solution for any number a except when  $0 \leq a < e$ .

12. Let f(x) be the function  $e^{-1/x^2}$  when  $x \neq 0$  and f(0) = 0. Show that f has a derivative at 0 and that f'(0) = 0.

13. Does f' have a derivative? If yes, what is it?

- 14. Does f have any further derivatives at 0?
- 15. Sketch the curve  $f(x) = x^x$ .

### [VIII, §5]

### **§5.** Some applications

It is worth while to mention briefly some applications of the exponential function to physics and chemistry.

It is known (from experimental data) that when a piece of radium is left to disintegrate, the rate of disintegration is proportional to the amount of radium left.

Suppose that at time t = 0 we have 10 grams of radium and let f(t) be the amount of radium left at time t. Then

$$\frac{df}{dt} = kf(t)$$

for some constant k. We take k negative since the physical interpretation is that the amount of substance decreases.

If we take the derivative of the quotient

$$\frac{f(t)}{e^{kt}}$$

and use the rule for the derivative of a quotient, we find

$$\frac{e^{kt}f'(t) - f(t)ke^{kt}}{e^{2kt}},$$

and this is equal to 0 (using our hypothesis concerning f'(t)). Hence there is a constant C such that

$$f(t) = C e^{kt}.$$

Let t = 0. Then f(0) = C. Thus C = 10, if we assumed that we started with 10 grams. In general, C is interpreted as the amount of initial substance when t = 0.

Similarly, consider a chemical reaction. It is frequently the case that the rate of the reaction is proportional to the quantity of reacting substance present. If f(t) denotes the amount of substance left after time t, then

$$\frac{df}{dt} = Kf(t)$$

for some constant K (determined experimentally in each case). We are therefore in a similar situation as before.

#### Exercises

1. Let  $f(t) = 10e^{Kt}$  for some constant K. Suppose that you know that f(1/2) = 2. Find K.
2. Let  $f(t) = Ce^{2t}$ . Suppose that you know f(2) = 5. Determine the constant C.

3. One gram of radium is left to disintegrate. After one million years, there is 0.1 gram left. What is the formula giving the rate of disintegration?

4. A certain chemical substance reacts in such a way that the rate of reaction is equal to the quantity of substance present. After one hour, there are 20 grams of substance left. How much substance was there at the beginning?

5. A radioactive substance disintegrates proportionally to the amount of substance present at a given time, say

$$f(t) = Ce^{Kt}.$$

At what time will there be exactly half the original amount left?

6. Suppose K = -4 in the preceding exercise. At what time will there be one-third of the substance left?

7. If bacteria increase in number at a rate proportional to the number present, how long will it take before 1,000,000 bacteria increase to 10,000,000 if it takes 12 minutes to increase to 2,000,000?

8. A substance decomposes at a rate proportional to the amount present. At the end of 3 minutes, 10 percent of the original substance has decomposed. When will half the original amount have decomposed?

## CHAPTER IX

# Integration

In this chapter, we solve, more or less simultaneously, the following problems:

(1) Given a function f(x), find a function F(x) such that

$$F'(x) = f(x).$$

This is the inverse of differentiation, and is called integration.

(2) Given a function f(x) which is  $\geq 0$ , give a definition of the area under the curve y = f(x) which does not appeal to geometric intuition.

Actually, in this chapter, we give the ideas behind the solutions of our two problems. The techniques which allow us to compute effectively when specific data are given will be postponed to the next chapter.

In carrying out (2) we shall follow an idea of Archimedes. It is to approximate the function f by horizontal functions, and the area under f by the sum of little rectangles.

The slightly theoretical sections 5 and 6 should be omitted for any class especially allergic to pure theory. The geometric argument involving area should suffice to justify the definite integral, and a mild sales talk on the integral as a limit of sums of small rectangles would be sufficient for the physical applications. Our axiomatization of the fundamental theorem allows the greatest flexibility concerning the extent to which these sections should be carried out in detail.

## **§1.** The indefinite integral

Let f(x) be a function defined over some interval. If F(x) is a function defined over the same interval and such that

$$F'(x) = f(x),$$

then we say that F is an *indefinite integral* of f. If G(x) is another indefinite integral of f, then G'(x) = f(x) also. Hence the derivative of the difference F - G is 0:

$$(F - G)'(x) = F'(x) - G'(x) = f(x) - f(x) = 0.$$

© Springer Science+Business Media New York 2002

[IX, §1]

Consequently, by Theorem 3 of Chapter V, 4, there is a constant C such that

$$F(x) = G(x) + C$$

for all x in the interval.

*Example* 1. An indefinite integral for  $\cos x$  would be  $\sin x$ . But  $\sin x + 5$  is also an indefinite integral for  $\cos x$ .

Example 2. log x is an indefinite integral for 1/x. So is log x + 10 or log  $x - \pi$ .

In the next chapter, we shall develop techniques for finding indefinite integrals. Here, we merely observe that every time we prove a formula for a derivative, it has an analogue for the integral.

It is customary to denote an indefinite integral of a function f by

$$\int f$$
 or  $\int f(x) dx$ 

In this second notation, the dx is meaningless by itself. It is only the full expression  $\int f(x) dx$  which is meaningful. When we study the method of substitution in the next chapter, we shall get further confirmation for the practicality of our notation.

We shall now make a table of some indefinite integrals, using the information which we have obtained about derivatives.

Let n be an integer,  $n \neq -1$ . Then we have

If 
$$n = -1$$
, then  
$$\int x^n dx = \frac{x^{n+1}}{n+1} \cdot \int \frac{1}{x} dx = \log x.$$

(This is true only in the interval x > 0.)

In the interval x > 0 we also have

$$\int x^c \, dx = \frac{x^{c+1}}{c+1}$$

for any number  $c \neq -1$ .

The following indefinite integrals are valid for all x.

$$\int \cos x \, dx = \sin x \qquad \int \sin x \, dx = -\cos x$$
$$\int e^x \, dx = e^x \qquad \int \frac{1}{1+x^2} \, dx = \arctan x$$

135

Finally, for -1 < x < 1, we have

$$\int \frac{1}{\sqrt{1-x^2}} \, dx = \arcsin x.$$

In practice, one frequently omits mentioning over what interval the various functions we deal with are defined. However, in any specific problem, one has to keep it in mind. For instance, if we write

$$\int x^{-1/3} \, dx = \frac{3}{2} \cdot x^{2/3},$$

this is valid for x > 0 and is also valid for x < 0. But 0 cannot be in any interval of definition of our functions. Thus we could have

$$\int x^{-1/3} dx = \frac{3}{2} \cdot x^{2/3} + 5$$
$$\int x^{-1/3} dx = \frac{3}{2} \cdot x^{2/3} - 2$$

when x < 0 and

when x > 0.

We agree throughout that indefinite integrals are defined only over intervals. Thus in considering the function 1/x, we have to consider *separately* the cases x > 0 and x < 0. For x > 0, we have already remarked that  $\log x$  is an indefinite integral. It turns out that for the interval x < 0 we can also find an indefinite integral, and in fact we have for x < 0,

$$\int \frac{1}{x} dx = \log \left( -x \right).$$

Observe that when x < 0, -x is positive, and thus  $\log(-x)$  is meaningful. The fact that the derivative of  $\log(-x)$  is equal to 1/x is true by the chain rule.

For x < 0, any other indefinite integral is given by

 $\log\left(-x\right)+C,$ 

where C is a constant.

It is sometimes stated that in all cases,

$$\int \frac{1}{x} dx = \log |x| + C.$$

With our conventions, we do not attribute any meaning to this, because our functions are not defined over intervals (the missing point 0 prevents this). In any case, the formula would be *false*. Indeed, for x < 0 we have

$$\int \frac{1}{x} dx = \log |x| + C_1,$$

136

and for x > 0 we have

$$\int \frac{1}{x} dx = \log |x| + C_2$$

However, the two constants need not be equal, and hence we cannot write

$$\int \frac{1}{x} dx = \log |x| + C$$

in all cases.

We prefer to stick to our convention that integrals are defined only over intervals. When we deal with the log, it is to be understood that we deal only with the case x > 0.

#### Exercises

Find indefinite integrals for the following functions:

1. 
$$\sin 2x$$
 2.  $\cos 3x$ 

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

(In these last two problems, specify the intervals over which you find an indefinite integral.)

## §2. Continuous functions

Let f(x) be a function. We shall say that f is continuous if

$$\lim_{h \to 0} f(x+h) = f(x)$$

for all x at which the function is defined.

It is understood that in taking the limit, only values of h for which f(x + h) is defined are considered. For instance, if f is defined on an interval

 $a \leq x \leq b$ 

(assuming a < b), then we would say that f is continuous at a if

$$\lim_{\substack{h \to 0 \\ h > 0}} f(a+h) = f(a).$$

We cannot take h < 0, since the function would not be defined for a + h if h < 0.

Geometrically speaking, a function is continuous if there is no break in its graph. All differentiable functions are continuous. We have already

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

has a limit, then the numerator f(x + h) - f(x) must approach 0, because

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

The following are graphs of functions which are not continuous.



In Fig. 1, we have the graph of a function like

$$f(x) = -1 \quad \text{if} \quad x \leq 0$$
  
$$f(x) = 1 \quad \text{if} \quad x > 0.$$

We see that

$$f(a+h)=f(h)=1$$

for all h > 0. Hence

$$\lim_{\substack{h\to 0\\h>0}} f(a+h) = 1,$$

which is unequal to f(0).

A similar phenomenon occurs in Fig. 2 where there is a break. (Cf. Example 5 of Chapter III, §2.)

## §3. Area

Let a < b be two numbers, and let f(x) be a continuous function defined on the interval  $a \leq x \leq b$ .

We wish to find a function F(x) which is differentiable in this interval, and such that

$$F'(x) = f(x).$$

In this section, we appeal to our geometric intuition concerning area. We assume that  $f(x) \ge 0$  for all x in the interval, and we define geometrically the function F(x) by saying that it is the numerical measure of the area under the curve between a and x.

The following picture illustrates this.



We thus have F(a) = 0. The area between a and a is 0.

**THEOREM** 1. The function F(x) is differentiable, and its derivative is f(x).

*Proof.* Since we defined F geometrically, we shall have to argue geometrically.

We have to consider the Newton quotient

$$\frac{F(x+h)-F(x)}{h}.$$

Suppose first that x is unequal to the end point b, and also suppose that we consider only values of h > 0.

Then F(x + h) - F(x) is the area between x and x + h. A magnified picture may look like this.



The shaded area represents F(x + h) - F(x).

We let c be a point in the closed interval [x, x + h] which is a maximum for our function f in that small interval. We let d be a point in the same closed interval which is a minimum for f in that interval. Thus

$$f(d) \leq f(t) \leq f(c)$$
$$x \leq t \leq x + h.$$

for all t satisfying

(We are forced to use another letter, t, since x is already being used.)

The area under the curve between x and x + h is bigger than the area of the small rectangle in the figure above, i.e. the rectangle having base h and height f(d).

INTEGRATION

The area under the curve between x and x + h is smaller than the area of the big rectangle, i.e. the rectangle having base h and height f(c).

This gives us

 $h \cdot f(d) \leq F(x+h) - F(x) \leq h \cdot f(c).$ 

Dividing by the positive number h yields

$$f(d) \leq \frac{F(x+h) - F(x)}{h} \leq f(c).$$

Since c, d are between x and x + h, as h approaches 0 both f(c) and f(d) approach f(x). Hence the Newton quotient for F is squeezed between two numbers which approach f(x). It must therefore approach f(x) itself, and we have proved Theorem 1, when h > 0.

The proof is essentially the same as the proof which we used to get the derivative of log x. The only difference in the present case is that we pick a maximum and a minimum without being able to give an explicit value for it, the way we could for the function 1/x. Otherwise, there is no difference in the arguments.

If x = b, we look at negative values for h. The argument in that case is entirely similar to the one we have written down in detail, and we find again that the Newton quotient of F is squeezed between f(c) and f(d). We leave it as an exercise.

Suppose that we are able to guess at a function G(x) whose derivative is f(x). Then we know that there is a constant C such that

$$F(x) = G(x) + C.$$
  
Let  $x = a$ . We get  
$$0 = F(a) = G(a) + C.$$

This shows that C = -G(a). Hence letting x = b yields

$$F(b) = G(b) - G(a).$$

Thus the area under the curve between a and b is G(b) - G(a). This is very useful to know in practice, because we can usually guess the function G.

If we deal with a continuous function f which may be negative in the interval [a, b], then we could still use our notion of area to find the function F(x). However, in those portions where the function is negative, we have to take F to be *minus* the area under the curve. We illustrate this by the following picture. In this case, F(x) would be the area between  $a_1$  and  $a_2$ , minus the area between  $a_2$  and x (for the point x indicated in the picture). The argument that F'(x) = f(x) goes through in the same way.



Thus by using our geometric intuition, we have found a function F(x) whose derivative is f(x).

*Example 1.* Find the area under the curve  $y = x^2$  between x = 1 and x = 2.

Let  $f(x) = x^2$ . If  $G(x) = x^3/3$  then G'(x) = f(x). Hence the area under the curve between 1 and 2 is

$$G(2) - G(1) = \frac{2^3}{3} - \frac{1^3}{3} = \frac{7}{3}$$

Example 2. Find the area under one arch of the function  $\sin x$ .

We have to find the area under the curve between 0 and  $\pi$ . Let  $G(x) = -\cos x$ . Then  $G'(x) = \sin x$ . Hence the area is

$$G(\pi) - G(0) = -\cos \pi - (-\cos 0)$$
  
= -(-1) + 1  
= 2.

Note how remarkable this is. The arch of the sine curve going from 0 to  $\pi$  seems to be a very irrational curve, and yet the area turns out to be the integer 2!

#### Exercises

Find the area under the given curves between the given bounds.

1.  $y = x^{3}$  between x = 1 and x = 5. 2. y = x between x = 0 and x = 2. 3.  $y = \cos x$ , one arch. 4. y = 1/x between x = 1 and x = 2. 5. y = 1/x between x = 1 and x = 3. 6.  $y = x^{4}$  between x = -1 and x = 1. 7.  $y = e^{x}$  between x = 0 and x = 1. INTEGRATION

### §4. Upper and lower sums

To show the existence of the integral, we still use the idea of approximating our curves by constant functions.

Let a, b be two numbers, with  $a \leq b$ . Let f be a continuous function in the interval  $a \leq x \leq b$ .

By a partition of the interval [a, b] we mean a sequence of numbers

$$a = x_0 \leq x_1 \leq x_2 \leq \cdots \leq x_n = b$$

between a and b, such that  $x_i \leq x_{i+1}$  (i = 0, 1, ..., n-1). For instance, we could take just two numbers,

$$x_0 = a$$
 and  $x_1 = b$ .

This will be called the trivial partition.

A partition divides our interval in a lot of smaller intervals  $[x_i, x_{i+1}]$ .

$$a = x_0 \quad x_1 \quad x_2 \quad x_3 \quad \cdots \quad x_{n-1} \quad x_n = b$$

Given any number between a and b, in addition to  $x_0, \ldots, x_n$ , we can add it to the partition to get a new partition having one more small interval. If we add enough intermediate numbers to the partition, then the intervals can be made arbitrarily small.

Let f be a function defined on the interval

$$a \leq x \leq b$$

and continuous. If  $c_i$  is a point between  $x_i$  and  $x_{i+1}$ , then we form the sum

$$f(c_0)(x_1 - x_0) + f(c_1)(x_2 - x_1) + \cdots + f(c_{n-1})(x_n - x_{n-1}).$$

Such a sum will be called a *Riemann sum*. Each value  $f(c_i)$  can be viewed as the height of a rectangle, and each  $(x_{i+1} - x_i)$  can be viewed as the length of the base.

Let  $s_i$  be a point between  $x_i$  and  $x_{i+1}$  such that f has a maximum in this small interval  $[x_i, x_{i+1}]$  at  $s_i$ . In other words,

$$f(x) \leq f(s_i)$$

for all x between  $x_i$  and  $x_{i+1}$ . The rectangles then look like those in the next figure. In the picture,  $s_0$  happens to be equal to  $x_1$ ,  $s_2 = x_2$ ,  $s_3 = x_4$ .

The main idea which we are going to carry out is that, as we make the intervals of our partitions smaller and smaller, the sum of the areas of the



rectangles will approach a limit, and this limit can be used to define the area under the curve.

Observe however that when f(x) becomes negative, the value  $f(s_i)$  may be negative. Thus the corresponding rectangle gives a negative contribution

$$f(s_i)(x_{i+1} - x_i)$$

to the sum.

If P is the partition given by the numbers

$$x_0 \leq x_1 \leq x_2 \leq \cdots \leq x_n,$$

then the sum

$$f(s_0)(x_1 - x_0) + f(s_1)(x_2 - x_1) + \cdots + f(s_{n-1})(x_n - x_{n-1})$$

will be called the *upper sum* associated with the function f, and the partition P of the interval [a, b]. We shall denote it by the symbols

 $U^{b}_{a}(P,f).$ 

Also, it is tiresome to write the sum by repeating each term, and so we shall use the abbreviation

$$\sum_{i=0}^{n-1} f(s_i)(x_{i+1} - x_i)$$

to mean the sum from 0 to n - 1 of the terms  $f(s_i)(x_{i+1} - x_i)$ . Thus, by definition,

$$U_a^b(P,f) = \sum_{i=0}^{n-1} f(s_i)(x_{i+1} - x_i).$$

Instead of taking a maximum  $s_i$  in the interval  $[x_i, x_{i+1}]$  we could have taken a minimum. Let  $t_i$  be a point in this interval, such that

$$f(t_i) \leq f(x)$$

for all x in the small interval  $[x_i, x_{i+1}]$ . We call the sum

$$f(t_0)(x_1 - x_0) + f(t_1)(x_2 - x_1) + \cdots + f(t_{n-1})(x_n - x_{n-1})$$

[IX, §4]

INTEGRATION

the lower sum associated with the function f, and the partition P of the interval [a, b]. The lower sum will be denoted by

 $L^b_a(P,f).$ 

With our convention concerning sums, we can therefore write

$$L_a^b(P,f) = \sum_{i=0}^{n-1} f(t_i)(x_{i+1} - x_i).$$

On the next picture, we have drawn a typical term of the sum.



For all numbers x in the interval  $[x_i, x_{i+1}]$  we have

$$f(t_i) \leq f(x) \leq f(s_i).$$

Since  $x_{i+1} - x_i$  is  $\geq 0$ , it follows that each term of the lower sum is less than or equal to each term of the upper sum. Therefore

$$L^b_a(P,f) \leq U^b_a(P,f).$$

Furthermore, any Riemann sum taken with points  $c_i$  (which are not necessarily maxima or minima) is between the lower and upper sum.

What happens to our sums when we add a new point to a partition? We shall see that the lower sum increases and the upper sum decreases.

THEOREM 2. Let f be a continuous function on the interval [a, b]. Let  $P = (x_0, \ldots, x_n)$  be a partition of [a, b]. Let  $\overline{x}$  be any number in the interval, and let Q be the partition obtained from P by adding  $\overline{x}$  to  $(x_0, \ldots, x_n)$ . Then

$$L^b_a(P,f) \leq L^b_a(Q,f) \leq U^b_a(Q,f) \leq U^b_a(P,f).$$

*Proof.* Let us look at the lower sums, for example. Suppose that our number  $\overline{x}$  is between  $x_j$  and  $x_{j+1}$ :

$$x_j \leq \overline{x} \leq x_{j+1}.$$

When we form the lower sum for P, it will be the same as the lower sum

[IX, §4]

for Q except that the term

 $f(t_j)(x_{j+1} - x_j)$ 

will now be replaced by two terms. If u is a minimum for f in the interval between  $x_j$  and  $\overline{x}$ , and v is a minimum for f in the interval between  $\overline{x}$  and  $x_{j+1}$ , then these two terms are

$$f(u)(\overline{x} - x_j) + f(v)(x_{j+1} - \overline{x}).$$

We can write  $f(t_j)(x_{j+1} - x_j)$  in the form

$$f(t_j)(x_{j+1} - x_j) = f(t_j)(\bar{x} - x_j) + f(t_j)(x_{j+1} - \bar{x}).$$

Since  $f(t_j)$  is less than or equal to f(u) or f(v) (because  $t_j$  was a minimum in the whole interval between  $x_j$  and  $x_{j+1}$ ), it follows that

$$f(t_j)(x_{j+1} - x_j) \leq f(u)(\bar{x} - x_j) + f(v)(x_{j+1} - \bar{x})$$

Thus when we replace the term in the sum for P by the two terms in the sum for Q, the value of the contribution of these two terms increases. Since all other terms are the same, our assertion is proved.

The assertion concerning the fact that the upper sum decreases is left as an exercise. The proof is very similar.

As a consequence of our theorem, we obtain:

COROLLARY. Every lower sum is less than or equal to every upper sum.

*Proof.* Let P and Q be two partitions. If we add to P all the points of Q and add to Q all the points of P, we obtain a partition R such that every point of P is a point of R and every point of Q is a point of R. Thus R is obtained by adding points to P and to Q. Consequently, we have the inequalities

$$L^b_a(P,f) \leq L^b_a(R,f) \leq U^b_a(R,f) \leq U^b_a(Q,f).$$

This proves our assertion.

It is now a very natural question to ask whether there is a *unique* point between the lower sums and the upper sums. In the next sections, we shall prove:

**THEOREM 3.** There exists a unique number which is greater or equal to every lower sum and less than or equal to every upper sum.

This number will be called the *definite integral* of f between a and b, and is denoted by

$$\int_{a}^{b} f$$

INTEGRATION

*Example.* Let  $f(x) = x^2$  and let the interval be [0, 1]. Write out the upper and lower sums for the partition consisting of  $(0, \frac{1}{2}, 1)$ .

The minimum of the function in the interval  $[0, \frac{1}{2}]$  is at 0, and f(0) = 0. The minimum of the function in the interval  $[\frac{1}{2}, 1]$  is at  $\frac{1}{2}$  and  $f(\frac{1}{2}) = \frac{1}{4}$ . Hence the lower sum is

$$f(0)(\frac{1}{2}-0) + f(\frac{1}{2})(1-\frac{1}{2}) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}.$$

The maximum of the function in the interval  $[0, \frac{1}{2}]$  is at  $\frac{1}{2}$  and the maximum of the function in the interval  $[\frac{1}{2}, 1]$  is at 1. Thus the upper sum is

$$f(\frac{1}{2})(\frac{1}{2}-0) + f(1)(1-\frac{1}{2}) = \frac{1}{8} + \frac{1}{2} = \frac{5}{8}$$

#### Exercises

Write out the lower and upper sums for the following functions and intervals. Use a partition such that the length of each small interval is  $\frac{1}{2}$ .

- 1.  $f(x) = x^2$  in the interval [1, 2].
- 2. f(x) = 1/x in the interval [1, 3].
- 3. f(x) = x in the interval [0, 2].
- 4. f(x) = 3 in the interval [0, 5].

5. Let f(x) = 1/x and let the interval be [1, 2]. Let n be a positive integer. Write out the upper and lower sum, using the partition such that the length of each small interval is 1/n.

6. Prove that

$$\frac{1}{n+1} + \frac{1}{n+2} + \dots + \frac{1}{n+n} \le \log 2 \le \frac{1}{n} + \frac{1}{n+1} + \dots + \frac{1}{2n-1}$$

7. Let  $f(x) = \log x$ . Let *n* be a positive integer. Write out the upper and lower sums, using the partition of the interval between 1 and *n* consisting of the integers from 1 to *n*, i.e. the partition (1, 2, ..., n).

## §5. The fundamental theorem

This section will contain a sketch of the proof of Theorem 3. We first need to discuss a special property of numbers.

Let S be a collection of numbers, with at least one number in the collection. (We also say that S is not empty.) An *upper bound* for S is a number B such that

$$x \leq B$$

for all x in the collection S. A least upper bound for S is an upper bound which is smallest among all upper bounds.

For instance, let S be the collection of numbers whose square is  $\leq 4$ . Then 5 is an upper bound for S and so is 3. However, 2 is the least upper bound.

Another example: Let S be the collection of numbers whose square is <2. Then  $\sqrt{2}$  is the least upper bound. (In this example, the least upper bound does not belong to our collection S.)

The property of numbers which we accept without proof is that every collection of numbers which is not empty and has an upper bound also has a least upper bound.

Similarly, we shall say that a number A is a *lower bound* for S if

 $A \leq x$ 

for all x in the collection S. A greatest lower bound for S is a lower bound which is largest among all lower bounds.

We shall also accept without proof that every non-empty collection of numbers which has a lower bound has a greatest lower bound.

For example, the collection of numbers 1,  $\frac{1}{2}$ ,  $\frac{1}{3}$ ,  $\frac{1}{4}$ , ... has a greatest lower bound, which is the number 0. Observe again that this greatest lower bound is not in the collection.

We return to our upper and lower sums.

Let f be as in §4. We consider the collection of numbers consisting of all lower sums

 $L_a^b(P,f)$ 

for all partitions P. This collection certainly has an upper bound (any upper sum will be an upper bound). We denote by

 $L_a^b(f)$ 

its least upper bound, and call it the *lower integral* of f, over the interval [a, b].

Every upper sum is an upper bound for the collection of lower sums. Therefore

$$L_a^b(f) \leq U_a^b(P, f)$$

for every partition P. Thus  $L_a^b(f)$  is a lower bound for the collection of upper sums. We denote by

 $U_a^b(f)$ 

the greatest lower bound of the collection of upper sums, and call it the

$$L^b_a(f) \leq U^b_a(f).$$

If x is any point in the interval, then we have the numbers

$$L_a^x(f)$$
 and  $U_a^x(f)$ .

Thus  $L_a^x(f)$  is the value at x of a function defined on the interval. Similarly,  $U_a^x(f)$  is the value at x of another function defined on the interval.

Our purpose is to prove that these functions are equal for all x. The way we shall do this is to prove that their derivative is equal to f. From this it will follow that there is a constant C such that

$$L_a^x(f) = U_a^x(f) + C,$$

for all x in the interval. This is true especially for x = a, so that

$$L^a_a(f) = U^a_a(f) + C.$$

Since both  $L_a^a(f)$  and  $U_a^a(f)$  are equal to 0, we get C = 0, and hence  $L_a^x(f) = U_a^x(f)$  for all x in the interval. Letting x = b gives us what we want.

The proof that the derivative of  $L_a^x(f)$  and  $U_a^x(f)$  is f(x) will depend on two properties, which we now state. These properties concern L and U. In order to be neutral, and speak of either L or U without committing ourselves, we shall use another letter, say I.

THEOREM 4. Let a, d be two numbers, with a < d. Let f be a continuous function on the interval [a, d]. Suppose that for each pair of numbers  $b \leq c$  in the interval we are able to associate a number denoted by  $I_b^c(f)$  satisfying the following properties:

Property 1. If M, m are two numbers such that

$$m \leq f(x) \leq M$$

for all x in the interval [b, c] then

$$m(c-b) \leq I_b^c(f) \leq M(c-b).$$

Property 2. We have

$$I_{a}^{b}(f) + I_{b}^{c}(f) = I_{a}^{c}(f).$$

Then the function  $I_a^x(f)$  is differentiable in the interval [a, d] and its derivative is f(x).

The fact that L and U satisfy our two properties will be proved in the next section. We shall now prove Theorem 4.

[IX, §5]

*Proof.* We have to form the Newton quotient

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

and see if it approaches a limit as  $h \to 0$ . (If x = a then it is to be understood that h > 0, and if x = b then h < 0. As usual, we prove that the function  $I_a^x(f)$  is right differentiable at a and left differentiable at b.)

Assume for the moment that h > 0. By Property 2, applied to the three numbers a, x, x + h, we conclude that our Newton quotient is equal to

$$\frac{I_a^x(f) + I_x^{x+h}(f) - I_a^x(f)}{h} = \frac{I_x^{x+h}(f)}{h}$$

This reduces our investigation of the Newton quotient to the interval between x and x + h.

Let s be a point between x and x + h such that f reaches a maximum in this small interval [x, x + h] and let t be a point in this interval such that f reaches a minimum.

We let m = f(t) and M = f(s) and apply Property 1 to the interval [x, x + h]. We obtain

 $f(t)(x + h - x) \leq I_x^{x+h}(f) \leq f(s)(x + h - x),$ 

which we can rewrite as

$$f(t) \cdot h \leq I_x^{x+h}(f) \leq f(s) \cdot h.$$

Dividing by the positive number h preserves the inequalities, and yields

$$f(t) \leq \frac{I_x^{x+h}(f)}{h} \leq f(s).$$

Since s, t lie between x and x + h, we must have (by continuity)

$$\lim_{h \to 0} f(s) = f(x)$$
$$\lim_{h \to 0} f(t) = f(x).$$

and

Thus our Newton quotient is squeezed between two numbers which approach 
$$f(x)$$
. It must therefore approach  $f(x)$ , and our theorem is proved when  $h > 0$ .

The argument when h < 0 is entirely similar. We omit it, except for the following remark concerning Property 2.

INTEGRATION

Suppose that we have two numbers b, c with c < b. We define

$$I_b^c(f) = -I_c^b(f),$$

whenever f is a continuous function on the interval [c, b]. Then it is easy to verify that Property 2 holds no matter what the numbers b, c are. For instance, suppose c < b. We show you how to prove the statement of Property 2 in that case.

By definition

$$I_b^c(f) = -I_c^b(f).$$

We know that

$$I_a^c(f) + I_c^b(f) = I_a^b(f)$$

if we use the ordinary case of Property 2 when the numbers are increasing. Substituting the value for  $I_b^c(f)$ , we find

$$I_a^c(f) - I_b^c(f) = I_a^b(f),$$

whence

 $I_a^c(f) = I_a^b(f) + I_b^c(f).$ 

## §6. The basic properties

We shall finally prove that L and U satisfy the basic properties.

Proof of Property 1. Suppose that we have two numbers m, M such that

 $m \leq f(x) \leq M$ 

for all x between b and c. Let P be the partition of the interval b, c consisting only of the end points. Then

$$m(c - b) \leq L_b^c(P, f) \leq L_b^c(f).$$

On the other hand,

$$L_b^c(f) \leq U_b^c(f) \leq M(c-b).$$

This proves our property.

Proof of Property 2. Let a, b, c be three numbers with  $a \leq b \leq c$ . Let f be a continuous function on the interval [a, c]. We shall prove:

$$L_a^b(f) + L_b^c(f) = L_a^c(f).$$

A similar statement holds for the upper integrals, namely

$$U_a^b(f) + U_b^c(f) = U_a^c(f).$$

*Proof.* We shall first prove that

$$L_a^b(f) + L_b^c(f) \leq L_a^c(f).$$

Since  $L_a^b(f)$  is the *least* upper bound of the lower sums built up with partitions, we can find such a lower sum which comes arbitrarily close to it. Thus, given any small number  $\epsilon > 0$ , we can find a partition P of the interval [a, b] such that

$$L_a^b(f) - \epsilon \leq L_a^b(P, f).$$

Similarly, we can find a partition Q of the interval [b, c] such that

$$L_b^c(f) - \epsilon \leq L_b^c(Q, f).$$

Adding, we get

$$L_a^b(f) + L_b^c(f) - 2\epsilon \leq L_a^b(P, f) + L_b^c(Q, f)$$

We can view P and Q together as giving a partition of the whole interval [a, c].

$$a \xrightarrow{P} b \xrightarrow{Q} c$$

We denote this partition simply by (P, Q). The two sums occurring on the right-hand side of our inequality are then equal to the lower sum built up from the partition (P, Q) over the whole interval [a, c]. Thus we can write

$$L_a^b(f) + L_b^c(f) - 2\epsilon \leq L_a^c((P,Q),f).$$

But any lower sum made up from a partition is smaller than the lower integral, which is an upper bound (even a least upper bound) of such lower sums. Thus our expression on the right satisfies the inequality

We obtain finally

$$L_a^b(f) + L_b^c(f) - 2\epsilon \leq L_a^c(f).$$

 $L_a^c((P,Q),f) \leq L_a^c(f).$ 

This is true for every  $\epsilon > 0$ . Let  $\epsilon$  approach 0. Then the left-hand side approaches

$$L_a^b(f) + L_b^c(f)$$

and our inequality follows.

We shall now prove the reverse inequality.

Let R be any partition of [a, c]. If we add b to R, then we get a partition of [a, c] which splits up into a partition of [a, b] and a partition of

[IX, §6]

[b, c]. Call these P and Q. Then

$$L_a^c(R,f) \leq L_a^c((P,Q),f) = L_a^b(P,f) + L_b^c(Q,f).$$

Since the lower integral is an upper bound for these lower sums, we know that the expression on the right is less than or equal to

$$L^b_a(f) + L^c_b(f),$$

which is therefore an upper bound for  $L_a^c(R, f)$ . The least upper bound  $L_a^c(f)$  is therefore less than or equal to our sum on the right, and the property is proved.

The proof that the upper integral U satisfies the two properties is entirely similar, and will be left as an exercise.

## CHAPTER X

# **Properties of the Integral**

This is a short chapter. It shows how the integral combines with addition and inequalities. There is no good formula for the integral of a product. The closest thing is integration by parts, which is postponed to the next chapter.

Connecting the integral with the derivative is what allows us to compute integrals. The fact that two functions having the same derivative differ by a constant is again exploited to the hilt.

## §1. Further connection with the derivative

Let f be a continuous function on some interval. Let a, b be two points of the interval such that  $a \leq b$ , and let F be a function which is differentiable on the interval and whose derivative is f.



Then we know that there is a constant C such that

$$\int_a^x f = F(x) + C$$

for all x in the interval. If we put x = a, we get

$$0=\int_a^a f=F(a)+C,$$

whence C = -F(a). We also have

$$\int_a^b f = F(b) + C.$$

From this we obtain

$$\int_a^b f = F(b) - F(a).$$

This is extremely useful in practice, because we can usually guess the function F, and once we have guessed it, we can then compute the integral by means of this relation.

S. Lang, *Short Calculus* © Springer Science+Business Media New York 2002

Furthermore, it is also practical to use the notation

 $F(x)\Big|_a^b$ 

instead of F(b) - F(a). Thus the integral

$$\int_0^{\pi} \sin x \, dx$$

is equal to

$$-\cos x\Big|_{0}^{\pi}$$

which is  $-\cos \pi - (-\cos 0) = 2$ .

As another example, suppose we want to find

$$\int_1^3 x^2 \, dx.$$

Let  $F(x) = x^3/3$ . Then  $F'(x) = x^2$ . Hence our integral is equal to

$$\left. \frac{x^3}{3} \right|_1^3 = \frac{27}{3} - \frac{1}{3} = \frac{26}{3} \cdot$$

Finally, we shall usually call the *indefinite integral* simply an *integral*, since the context makes clear what is meant. When we deal with a definite integral  $\int_a^b$ , the numbers a and b are sometimes called the *lower limit* and *upper limit*, respectively.

#### Exercises

Find the following integrals:

1. 
$$\int_{1}^{2} x^{5} dx$$
  
3.  $\int_{-\pi}^{\pi} \sin x \, dx$   
2.  $\int_{-1}^{1} x^{1/3} dx$   
4.  $\int_{0}^{\pi} \cos x \, dx$ 

## §2. Sums

Let f(x) and g(x) be two functions defined over some interval, and let F(x) and G(x) be (indefinite) integrals for f and g, respectively. Since the derivative of a sum is the sum of the derivatives, we see that F + G is an integral for f + g; in other words,

$$\int [f(x) + g(x)] \, dx = \int f(x) \, dx + \int g(x) \, dx.$$

 $[X, \S{2}]$ 

SUMS

Similarly, let c be a number. The derivative of cF(x) is cf(x). Hence

$$\int cf(x) \ dx = c \int f(x) \ dx.$$

A constant can be taken in and out of an integral.

*Example* 1. Find the integral of  $\sin x + 3x^4$ .

We have

$$\int (\sin x + 3x^4) \, dx = \int \sin x \, dx + \int 3x^4 \, dx$$
$$= -\cos x + 3x^5/5.$$

Any formula involving the indefinite integral yields a formula for the definite integral. Using the same notation as above, suppose we have to find

$$\int_a^b \left[f(x) + g(x)\right] \, dx.$$

We know that it is

$$F(x) + G(x)\Big|_a^b,$$

which is equal to

$$F(b) + G(b) - F(a) - G(a).$$

Thus we get the formula

$$\int_{a}^{b} [f(x) + g(x)] \, dx = \int_{a}^{b} f(x) \, dx + \int_{a}^{b} g(x) \, dx.$$

Similarly, for any constant c,

$$\int_a^b cf(x) \ dx = c \int_a^b f(x) \ dx.$$

Example 2. Find the integral

$$\int_0^{\pi} \left[\sin x + 3x^4\right] dx.$$

This (definite) integral is equal to

$$\begin{aligned} -\cos x + 3x^{5}/5\Big|_{0}^{\pi} &= -\cos \pi + 3\pi^{5}/5 - (-\cos 0 + 0) \\ &= 1 + 3\pi^{5}/5 + 1 \\ &= 2 + 3\pi^{5}/5. \end{aligned}$$

#### Exercises

Find the following integrals:

1. 
$$\int 4x^3 dx$$
  
2.  $\int (3x^4 - x^5) dx$   
3.  $\int (2\sin x + 3\cos x) dx$   
4.  $\int (3x^{2/3} + 5\cos x) dx$   
5.  $\int \left(5e^x + \frac{1}{x}\right) dx$   
6.  $\int_{-\pi}^{\pi} (\sin x + \cos x) dx$   
7.  $\int_{-1}^{1} 2x^5 dx$   
8.  $\int_{-1}^{2} e^x dx$   
9.  $\int_{-1}^{3} 4x^2 dx$ 

10. Find the area between the curves y = x and  $y = x^2$ . [Sketch the curve. If f(x) and g(x) are two continuous functions such that  $f(x) \ge g(x)$  on an interval [a, b], then the area between the two curves, from a to b, is

$$\int_a^b (f(x) - g(x)) \, dx.$$

In this problem, the curves intersect at x = 0 and x = 1.]



Hence the area is

$$\int_0^1 (x - x^2) \, dx = \frac{x^2}{2} - \frac{x^3}{3} \Big|_0^1 = \frac{1}{2} - \frac{1}{3}$$

11. Find the area between the curves y = x and  $y = x^3$ .

12. Find the area between the curves  $y = x^2$  and  $y = x^3$ .

13. Find the area between the curve y = (x - 1)(x - 2)(x - 3) and the x-axis. (Sketch the curve.)

14. Find the area between the curve y = (x + 1)(x - 1)(x + 2) and the x-axis.

15. Find the area between the curves  $y = \sin x$ ,  $y = \cos x$ , the y-axis, and the first point where these curves intersect for x > 0.

## **§3.** Inequalities

THEOREM 1. Let a, b be two numbers, with  $a \leq b$ . Let f, g be two continuous functions on the interval [a, b] and assume that  $f(x) \leq g(x)$  for all x in the interval. Then

$$\int_a^b f(x) \ dx \ \leq \ \int_a^b g(x) \ dx.$$

*Proof.* Since  $g(x) - f(x) \ge 0$ , we can use the basic Property 1 of Chapter IX, §5 (with m = 0) to conclude that

$$\int_a^b (g - f) \ge 0.$$

But

$$\int_a^b (g - f) = \int_a^b g - \int_a^b f.$$

Transposing the second integral on the right in our inequality, we obtain

$$\int_a^b g \ge \int_a^b f,$$

as desired.

Theorem 1 will be used mostly when g(x) = |f(x)|. Since a negative number is always  $\leq$  a positive number, we know that

 $f(x) \leq |f(x)|$ 

and

 $-f(x) \leq |f(x)|.$ 

THEOREM 2. Let a, b be two numbers, with  $a \leq b$ . Let f be a continuous function on the interval [a, b]. Then

$$\left|\int_{a}^{b}f(x) dx\right| \leq \int_{a}^{b}|f(x)| dx.$$

*Proof.* We simply let g(x) = |f(x)| in the preceding theorem. The absolute value of the integral on the left is equal to

$$\int_a^b f(x) \, dx \qquad \text{or} \qquad -\int_a^b f(x) \, dx.$$

We can apply Theorem 1 either to f(x) or -f(x) to get Theorem 2.

We make one other application of Theorem 2.

THEOREM 3. Let a, b be two numbers and f a continuous function on the closed interval between a and b. (We do not necessarily assume that a < b.) Let M be a number such that  $|f(x)| \leq M$  for all x in the interval. Then

$$\left|\int_{a}^{b} f(x) \ dx\right| \leq M \ |b - a|$$

*Proof.* If  $a \leq b$ , we can use Theorem 2 to get

$$\left|\int_a^b f(x) \, dx\right| \leq \int_a^b M \, dx = M \int_a^b dx = M(b - a).$$

If b < a, then

$$\int_a^b f = -\int_b^a f.$$

Taking the absolute value gives us the estimate M(a - b). Since a - b = |b - a| in case b < a, we have proved our theorem.

## §4. Improper integrals

We know that the area under curve 1/x between 1 and x is log x. Instead of taking x > 1, let us take 0 < x < 1. As x approaches 0, log x becomes very large negative. The integral

$$\int_x^1 \frac{1}{t} dt = \log t \Big|_x^1 = -\log x$$

is therefore very large positive. We can interpret this by saying that the area becomes very large.

However, it is remarkable that an entirely different situation will occur when we consider the area under the curve  $1/x^{1/2} = x^{-1/2}$ . We take x > 0, of course, and compute the integral

$$\int_{x}^{1} \frac{1}{t^{1/2}} dt = \frac{t^{1/2}}{1/2} \Big|_{x}^{1} = 2 - 2x^{1/2}.$$

As x approaches 0, this approaches 2, in spite of the fact that our curve  $y = x^{-1/2}$  gives rise to a chimney near the y-axis, and is not even defined for x = 0.

When that happens, we shall say that the integral

$$\int_0^1 t^{-1/2} \, dt$$

exists, even though the function is not defined at 0 and is not continuous in the closed interval [0, 1].

In general, suppose we have two numbers a, b with, say, a < b. Let f be a continuous function in the interval  $a < x \leq b$ . This means that

$$a+h \leq x \leq b.$$

We can then form our usual integral

$$\int_{a+h}^{b} f(x) \ dx.$$

If F is an indefinite integral for f over our interval, then the integral is equal to

$$F(b) - F(a+h)$$

If the limit

$$\lim_{h\to 0} F(a+h)$$

exists, then we say that the improper integral

$$\int_a^b f(x) \ dx$$

exists, and is equal to  $F(b) - \lim_{h \to 0} F(a+h)$ .

In our preceding examples, we can say that the improper integral

$$\int_0^1 \frac{1}{x} \, dx$$

does not exist, but that the improper integral

$$\int_0^1 x^{-1/2} \, dx$$

does exist. This second integral is equal to 2.

We make similar definitions when we deal with an interval  $a \leq x < b$ and a function f which is continuous on this interval. If the limit

$$\lim_{\substack{h\to 0\\h>0}}\int_a^{b-h}f(x)\ dx$$

exists, then we say that the improper integral exists, and it is equal to this limit.

Example 1. Show that the improper integral

$$\int_0^1 \frac{1}{x^2} \, dx$$

does not exist.

We consider

$$\int_{h}^{1} x^{-2} dx = \frac{x^{-1}}{-1} \Big|_{h}^{1} = -1 - \left(-\frac{1}{h}\right) = -1 + \frac{1}{h}.$$

This does not approach a limit as h approaches 0 and hence the improper integral does not exist.

There is another type of improper integral, dealing with large values.

Let a be a number and f a continuous function defined for  $x \ge a$ . Consider the integral

$$\int_a^B f(x) \ dx$$

for some number B > a. If F(x) is any indefinite integral of f, then our integral is equal to F(B) - F(a). If it approaches a limit as B becomes very large, then we define

$$\int_a^\infty f(x) \ dx \qquad \text{or} \qquad \int_a^\infty f$$

to be this limit, and say that the improper integral converges.

Thus  $\int_a^{\infty} f$  converges if

$$\lim_{B\to\infty}\int_a^B f$$

exists, and is equal to this limit. Otherwise, we say that the improper integral does not converge.

*Example 2.* Determine whether the improper integral  $\int_{1}^{\infty} \frac{1}{x} dx$  converges, and if it does, find its value.

We have, for a large number B,

$$\int_1^B \frac{1}{x} dx = \log B - \log 1 = \log B.$$

As B becomes large, so does log B, and hence the improper integral does not converge.

Let us look at the function  $1/x^2$ . Its graph looks like that in the next figure. At first sight, there seems to be no difference between this function and 1/x, except that  $1/x^2 < 1/x$  when x > 1. However, intuitively speaking, we shall find that  $1/x^2$  approaches 0 sufficiently faster than 1/x to guarantee that the area under the curve between 1 and B approaches a limit as B becomes large.



Example 3. Determine whether the improper integral

$$\int_1^\infty \frac{1}{x^2} \, dx$$

converges, and if it does, find its value.

For a large number B, we have

$$\int_{1}^{B} \frac{1}{x^{2}} dx = \frac{-1}{x} \Big|_{1}^{B} = -\frac{1}{B} + 1.$$

As B becomes large, 1/B approaches 0. Hence the limit as B becomes large exists and is equal to 1, which is the value of our integral. We thus have by definition

$$\int_1^\infty \frac{1}{x^2} dx = 1.$$

It is frequently possible to determine whether an improper integral converges without computing it, by comparing it with another which is known to converge. We give such a criterion in the following theorem.

THEOREM 4. Let f(x) and g(x) be two continuous functions defined for  $x \ge a$  and such that  $f(x) \ge 0$  and  $g(x) \ge 0$  for all  $x \ge a$ . Assume that  $f(x) \le g(x)$  and that the improper integral

$$\int_a^\infty g(x)\ dx$$

converges. Then so does the improper integral

$$\int_a^\infty f(x)\ dx.$$

(Intuitively speaking, we visualize the theorem as saying that if the area under the graph of g is finite, then the area under the graph of f is also finite because it is smaller.) *Proof.* Let B be a large number. Then according to the inequalities satisfied by the definite integral, we have

$$\int_a^B f \leq \int_a^B g.$$

As B becomes large, and increases, the integral on the right increases. But we know that it approaches a limit; call it L. This limit is a number, which is a bound for the integral of f between a and B, i.e.

$$\int_{a}^{B} f \leq L$$

As B increases, this integral of f also increases (the area under the graph increases because f is  $\geq 0$ ) but stays below L for all B. There must be a least upper bound for such integrals of f, and this least upper bound is the desired limit.

Example 4. Determine whether the improper integral

$$\int_1^\infty \frac{x}{x^3+1}\,dx$$

converges.

We don't try to evaluate this, but observe that

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

 $x^3 \leq x^3 + 1$ 

All the functions involved are  $\geq 0$  when  $x \geq 1$ . Using Example 3, we conclude that our improper integral converges.

#### Exercises

Determine whether the following improper integrals exist or not, and converge or not:

1. 
$$\int_{2}^{\infty} \frac{1}{x^{3/2}} dx$$
  
2.  $\int_{1}^{\infty} \frac{1}{x^{2/3}} dx$   
3.  $\int_{0}^{\infty} \frac{1}{1+x^{2}} dx$   
4.  $\int_{0}^{5} \frac{1}{5-x} dx$   
5.  $\int_{0}^{2} \frac{1}{x^{2}-2x} dx$   
6.  $\int_{1}^{\infty} e^{-x} dx$ 

, 7. Let B be a number >2. Find the area under the curve  $y = e^{-2x}$  between 2 and B. Does this area approach a limit when B becomes very large? If so, what limit?

for  $x \geq 1$ , whence

## CHAPTER XI

# **Techniques of Integration**

The purpose of this chapter is to teach you certain basic tricks to find indefinite integrals. It is of course easier to look up integral tables, but you should have a minimum of training in standard techniques.

## **§1.** Substitution

We shall formulate the analogue of the chain rule for integration.

Suppose that we have a function g(x) and another function f such that f(g(x)) is defined. (All these functions are supposed to be defined over suitable intervals.) We wish to evaluate an integral

$$\int f(g(x)) \frac{dg(x)}{dx} dx.$$

Let F(u) be an indefinite integral for f(u), so that

$$\int f(u) du = F(u)$$
, i.e.  $\frac{dF(u)}{du} = f(u)$ .

Then we assert that F(g(x)) is an integral for  $f(g(x)) \frac{dg}{dx}$ , or symbolically, that

$$\int f(g(x)) \frac{dg}{dx} dx = \int f(u) \, du.$$

This follows at once from the chain rule, because

$$\frac{dF}{dx} = \frac{dF}{du} \frac{du}{dx} = f(u) \frac{du}{dx} = f(g(x)) \frac{dg(x)}{dx}$$

Example 1. Find  $\int (x^2 + 1)^3 (2x) \, dx$ .

Put  $u = x^2 + 1$ . Then du/dx = 2x and our integral is in the form

$$\int f(u) \, \frac{du}{dx} \, dx,$$

the function f being  $f(u) = u^3$ .

163

Therefore our integral is equal to

$$\int f(u) \, du = \int u^3 \, du = \frac{u^4}{4} = \frac{(x^2 + 1)^4}{4} \, \cdot$$

We can check this by differentiating the expression on the right, using the chain rule. We get  $(x^2 + 1)^3(2x)$ , as desired.

Example 2. Find  $\int \sin(2x)(2) dx$ .

Put u = 2x. Then du/dx = 2. Hence our integral is in the form

$$\int \sin u \, du = -\cos u = -\cos (2x).$$

Observe that

$$\int \sin (2x) \, dx \neq -\cos (2x).$$

If we differentiate  $-\cos(2x)$ , we get  $\sin(2x) \cdot 2$ .

The integral in Example 2 could also be written

$$\int 2\sin\left(2x\right)\,dx.$$

It does not matter, of course, where we place the 2.

*Example 3.* Find  $\int \cos(3x) dx$ .

Let u = 3x. Then du/dx = 3. There is no extra 3 in our integral. However, we can take a constant in and out of an integral. Our integral is equal to

$$\frac{1}{3}\int 3\cos\left(3x\right)\,dx,$$

and this integral is in the form

$$\frac{1}{3}\int \cos u \, du.$$

Thus our integral is equal to  $\frac{1}{3} \sin u = \frac{1}{3} \sin (3x)$ .

It is convenient to use a purely formal notation which allows us to make a substitution u = g(x), as in the previous examples. Thus instead of writing

$$\frac{du}{dx} = 2x$$

in Example 1, we would write du = 2x dx. Similarly, in Example 2, we would write du = 2 dx, and in Example 3 we would write du = 3 dx. We do not attribute any meaning to this. It is merely a device of a type used in programming a computing machine. A machine does not think.

[XI, §1]

SUBSTITUTION

One simply adjusts certain electric circuits so that the machine performs a certain operation and comes out with the right answer. The fact that writing

$$du = \frac{du}{dx} dx$$

makes us come out with the right answer was *proved* once and for all when we established the relationship

$$\int f(g(x)) \frac{dg}{dx} dx = \int f(u) \, du.$$

The proof consisted in differentiating the answer and checking that it gives us the desired function.

Example 4. Find

$$\int (x^3 + x)^9 (3x^2 + 1) \, dx$$

Let

 $u=x^3+x.$ 

$$du = (3x^2 + 1) \, dx.$$

Hence our integral is of type  $\int f(u) du$  and is equal to

$$\int u^9 \, du = \frac{u^{10}}{10} = \frac{(x^3 + x)^{10}}{10} \cdot$$

We should also observe that the formula for integration by substitution applies to the definite integral. Suppose that x lies in an interval  $a \leq x \leq b$ . Then with our preceding notation, we have

$$\int_a^b f(g(x)) \frac{dg}{dx} dx = \int_{g(a)}^{g(b)} f(u) du.$$

This is essentially the definition of these symbols.

In Example 4, suppose that we consider the integral

$$\int_0^1 (x^3 + x)^9 (3x^2 + 1) \, dx,$$

with  $u = x^3 + x$ . When x = 0, u = 0, and when x = 1, u = 2. Thus our definite integral is equal to

$$\int_0^2 u^9 \, du = \frac{2^{10}}{10} \, \cdot \,$$

#### Exercises

Find the following integrals:

1.  $\int xe^{x^2} dx$ 2.  $\int x^3 e^{-x^4} dx$ 3.  $\int x^2(1+x^3) dx$ 4.  $\int \frac{\log x}{x} dx$ 5.  $\int \frac{1}{x (\log x)^n} dx$  (*n* = integer) 6.  $\int \frac{2x+1}{x^2+x+1} dx$ 7.  $\int \frac{x}{x+1} dx$ 8.  $\int \sin x \cos x dx$ 9.  $\int \sin^2 x \cos x dx$ 10.  $\int_0^{\pi} \sin^5 x \cos x dx$ 11.  $\int_0^{\pi} \cos^4 x \sin x dx$ 12.  $\int \frac{\sin x}{1+\cos^2 x} dx$ 13.  $\int \frac{\arctan x}{1+x^2} dx$ 14.  $\int_0^1 x^3 \sqrt{1-x^2} dx$ 

16. Find the area under the curve  $y = xe^{-x^2}$  between 0 and a number B > 0. Does this area approach a limit as B becomes very large? If so, what limit?

17. Find the area under the curve  $y = x^2 e^{-x^3}$  between 0 and a number B > 0. Does this area approach a limit as B becomes very large? If so, what limit?

In some integrals involving  $e^x$ , one can sometimes find the integral by the substitution  $u = e^x$ ,  $x = \log u$ , and dx = (1/u) du. You can combine this with the technique of §4 below to deal with the following integrals.

18. 
$$\int \sqrt{1+e^x} \, dx$$
  
19. 
$$\int \frac{1}{1+e^x} \, dx$$
  
20. 
$$\int \frac{1}{e^x+e^{-x}} \, dx$$
  
21. 
$$\int \frac{1}{\sqrt{e^x+1}} \, dx$$

## §2. Integration by parts

If f, g are two differentiable functions of x, then

$$\frac{d(fg)}{dx} = f(x)\frac{dg}{dx} + g(x)\frac{df}{dx}$$

Hence

$$f(x) \frac{dg}{dx} = \frac{d(fg)}{dx} - g(x) \frac{df}{dx}$$

[XI, §2]

INTEGRATION BY PARTS

Using the formula for the integral of a sum, which is the sum of the integrals, we obtain

$$\int f(x) \frac{dg}{dx} dx = f(x)g(x) - \int g(x) \frac{df}{dx} dx,$$

which is called the formula for integrating by parts.

If we let u = f(x) and v = g(x), then the formula can be abbreviated in our shorthand notation as follows:

$$\int u\,dv = uv - \int v\,du.$$

*Example* 1. Find the integral  $\int \log x \, dx$ .

Let  $u = \log x$  and v = x. Then du = (1/x) dx and dv = dx. Hence our integral is in the form  $\int u dv$  and is equal to

$$uv - \int v \, du = x \log x - \int 1 \, dx$$
$$= x \log x - x.$$

Example 2. Find  $\int e^x \sin x \, dx$ .

Let  $u = e^x$  and  $dv = \sin x \, dx$ . Then

 $du = e^x dx$  and  $v = -\cos x$ .

If we call our integral I, then

$$I = -e^{x} \cos x - \int -e^{x} \cos x$$
$$= -e^{x} \cos x + \int e^{x} \cos x \, dx.$$

This looks as if we were going around in circles. Don't lose heart. Rather, let  $t = e^x$  and  $dz = \cos x \, dx$ . Then

$$dt = e^x dx$$
 and  $z = \sin x$ .

The second integral becomes

$$\int t\,dz\,=\,e^x\sin\,x\,-\int\,e^x\sin\,x\,dx.$$

We have come back to our integral I but with a minus sign! Thus

$$I = e^x \sin x - e^x \cos x - I.$$

Hence

 $2I = e^x \sin x - e^x \cos x.$ 

and dividing by 2 gives us the value of I.

#### Exercises

Find the following integrals:

r

1. 
$$\int \arcsin x \, dx$$
  
2. 
$$\int \arctan x \, dx$$
  
3. 
$$\int e^{2x} \sin 3x \, dx$$
  
4. 
$$\int e^{-4x} \cos 2x \, dx$$
  
5. 
$$\int (\log x)^2 \, dx$$
  
6. 
$$\int (\log x)^3 \, dx$$
  
7. 
$$\int x^2 e^x \, dx$$
  
8. 
$$\int x^2 e^{-x} \, dx$$
  
9. 
$$\int x \sin x \, dx$$
  
10. 
$$\int x \cos x \, dx$$
  
11. 
$$\int x^2 \sin x \, dx$$
  
12. 
$$\int x^2 \cos x \, dx$$
  
13. 
$$\int x^3 \cos x^2 \, dx$$
  
14. 
$$\int x^5 \sqrt{1 - x^2} \, dx$$
  
15. 
$$\int x^2 \log x \, dx$$
  
16. 
$$\int x^3 \log x \, dx$$
  
17. 
$$\int x^2 (\log x)^2 \, dx$$
  
18. 
$$\int x^3 e^{-x^2} \, dx$$
  
19. 
$$\int \frac{x^7}{(1 - x^4)^2} \, dx$$
  
20. 
$$\int_{-\pi}^{\pi} x^2 \cos x \, dx$$

21. Let B be a number >0. Find the area under the curve  $y = xe^{-x}$  between 0 and B. Does this area approach a limit as B becomes very large?

22. Does the improper integral  $\int_{1}^{\infty} x^2 e^{-x} dx$  converge? 23. Does the improper integral  $\int_{1}^{\infty} x^3 e^{-x} dx$  converge? 24. Let B be a number >2. Find the area under the curve

$$y = \frac{1}{x \, (\log x)^2}$$

between 2 and B. Does this area approach a limit as B becomes very large? If so, what limit?
[XI, §3]

25. Does the improper integral

$$\int_3^\infty \frac{1}{x \; (\log x)^4} \, dx$$

converge? If yes, to what?

# §3. Trigonometric integrals

We shall investigate integrals involving sine and cosine. It will be useful to have the following formulas:

$$\sin^2 x = \frac{1 - \cos 2x}{2}$$
$$\cos^2 x = \frac{1 + \cos 2x}{2}$$

These are easily proved, using

$$\cos 2x = \cos^2 x - \sin^2 x$$

When we want to integrate  $\sin^2 x$ , we are thus reduced to

$$\int \frac{1}{2} dx - \frac{1}{2} \int \cos 2x \, dx = \frac{x}{2} - \frac{1}{4} \sin 2x.$$

There is a general way in which one can integrate  $\sin^n x$  for any positive integer n: integrating by parts. Let us take first an example.

*Example* 1. Find the integral  $\int \sin^3 x \, dx$ .

We write the integral in the form

$$\int \sin^2 x \sin x \, dx.$$

Let  $u = \sin^2 x$  and  $dv = \sin x \, dx$ . Then

$$du = 2 \sin x \cos x \, dx$$
 and  $v = -\cos x$ .

Thus our integral is equal to

$$- (\sin^2 x) (\cos x) - \int -\cos x (2 \sin x \cos x) dx$$
$$= -\sin^2 x \cos x + 2 \int \cos^2 x \sin x dx.$$

This last integral could then be determined by substitution, for instance

169

 $t = \cos x$  and  $dt = -\sin x \, dx$ . The last integral becomes  $-2\int t^2 \, dt$ , and hence

$$\int \sin^3 x \, dx = -\sin^2 x \cos x - \frac{2}{3} \cos^3 x.$$

To deal with an arbitrary positive integer n, we shall show how to reduce the integral  $\int \sin^n x \, dx$  to the integral  $\int \sin^{n-2} x \, dx$ . Proceeding stepwise downwards will give a method for getting the full answer.

**THEOREM** 1. For any positive integer n, we have

$$\int \sin^n x \, dx = -\frac{1}{n} \sin^{n-1} x \cos x + \frac{n-1}{n} \int \sin^{n-2} x \, dx.$$

*Proof.* We write the integral as

$$I_n = \int \sin^n x \, dx = \int \sin^{n-1} x \sin x \, dx.$$

Let  $u = \sin^{n-1} x$  and  $dv = \sin x \, dx$ . Then

$$du = (n - 1) \sin^{n-2} x \cos x \, dx$$
 and  $v = -\cos x$ .

Thus

$$I_n = -\sin^{n-1} x \cos x - \int -(n-1) \cos x \sin^{n-2} x \cos x \, dx$$
$$= -\sin^{n-1} x \cos x + (n-1) \int \sin^{n-2} x \cos^2 x \, dx.$$

We replace  $\cos^2 x$  by  $1 - \sin^2 x$  and get finally

$$I_n = -\sin^{n-1} x \cos x + (n-1)I_{n-2} - (n-1)I_n,$$

whence

$$nI_n = -\sin^{n-1} x \cos x + (n-1)I_{n-2}.$$

Dividing by n gives us our formula.

We leave the proof of the analogous formula for cosine as an exercise.

$$\int \cos^n x \, dx = \frac{1}{n} \cos^{n-1} x \sin x + \frac{n-1}{n} \int \cos^{n-2} x \, dx.$$

Integrals involving tangents can be done by a similar technique, because

$$\frac{d\tan x}{dx} = 1 + \tan^2 x.$$

These functions are less used than sine and cosine, and hence we don't write out the formulas, to lighten this printed page which would otherwise become oppressive.

[XI, §3]

It is also useful to remember the following trick:

$$\int \frac{1}{\cos x} dx = \int \sec x \, dx = \log (\sec x + \tan x).$$

This is done by substitution. We have

$$\frac{1}{\cos x} = \sec x = \frac{(\sec x)(\sec x + \tan x)}{\sec x + \tan x}$$

Let  $u = \sec x + \tan x$ . Then the integral is in the form

$$\int \frac{1}{u} \, du$$

(This is a good opportunity to emphasize that the formula we just obtained is valid on any interval such that  $\cos x \neq 0$  and  $\sec x + \tan x > 0$ . Otherwise the symbols are meaningless. Determine such an interval as an exercise.)

One can integrate mixed powers of sine and cosine by replacing  $\sin^2 x$  by  $1 - \cos^2 x$ , for instance.

*Example 2.* Find  $\int \sin^2 x \cos^2 x \, dx$ .

Replacing  $\sin^2 x$  by  $1 - \cos^2 x$ , we see that our integral is equal to

$$\int \cos^2 x \, dx \, - \int \cos^4 x \, dx$$

and we know how to find each one of these integrals.

When we meet an integral involving a square root, we can frequently get rid of the square root by making a trigonometric substitution.

Example 3. Find the area of a circle of radius 3.

The equation of the circle is

$$x^2 + y^2 = 9,$$

and the portion of the circle in the first quadrant is described by the function

$$y = \sqrt{3^2 - x^2}$$

One-fourth of the area is therefore given by the integral

$$\int_0^3 \sqrt{3^2 - x^2} \, dx.$$

[XI, §3]

Let  $x = 3 \sin t$ . Then  $dx = 3 \cos t dt$  and our integral becomes

$$\int_0^{\pi/2} \sqrt{3^2 - 3^2 \sin^2 t} \, 3 \cos t \, dt = \int_0^{\pi/2} 9 \cos^2 t \, dt = \frac{9\pi}{4}$$

(We see that  $1 - \sin^2 t = \cos^2 t$  and  $\sqrt{1 - \sin^2 t} = \cos t$  in the interval between 0 and  $\pi/4$ .) The total area of the circle is therefore  $9\pi$ .

### Exercises

Find the following integrals.

1. 
$$\int \sin^4 x \, dx$$
 2.  $\int \cos^3 x \, dx$  3.  $\int \sin^2 x \cos^3 x \, dx$ 

Find the area of the region enclosed by the following curves:

4. 
$$x^2 + \frac{y^2}{9} = 1.$$
 5.  $\frac{x^2}{4} + \frac{y^2}{16} = 1.$  6.  $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$ 

- 7. Find the area of a circle of radius r > 0.
- 8. For any two integers m, n prove the formulas:

$$\sin mx \sin nx = \frac{1}{2} [\cos (m-n)x - \cos (m+n)x] \\ \sin mx \cos nx = \frac{1}{2} [\sin (m+n)x + \sin (m-n)x] \\ \cos mx \cos nx = \frac{1}{2} [\cos (m+n)x + \cos (m-n)x]$$

9. Show that

$$\int_{-\pi}^{\pi} \sin 3x \cos 2x \, dx = 0.$$

10. Show that

$$\int_{-\pi}^{\pi}\cos 5x\,\cos 2x\,dx = 0.$$

11. Show in general that for any integers m, n we have

$$\int_{-\pi}^{\pi} \sin mx \cos nx \, dx = 0.$$

12. Show in general that

$$\int_{-\pi}^{\pi} \sin mx \sin nx \, dx = \begin{cases} 0 & \text{if } m \neq n, \\ \pi & \text{if } m = n. \end{cases}$$

13. Find  $\int \tan x \, dx$ .

Find the following integrals:

14. 
$$\int \frac{1}{\sqrt{9 - x^2}} dx$$
  
15. 
$$\int \frac{1}{\sqrt{3 - x^2}} dx$$
  
16. 
$$\int \frac{1}{\sqrt{2 - 4x^2}} dx$$
  
17. 
$$\int \frac{1}{\sqrt{a^2 - b^2 x^2}} dx$$

172

### §4. Partial fractions

Let f(x) and g(x) be two polynomials. We want to investigate the integral

$$\int \frac{f(x)}{g(x)} \, dx.$$

Using long division, one can reduce the problem to the case when the degree of f is less than the degree of g. The following example illustrates this reduction.

*Example 1.* Consider the two polynomials  $f(x) = x^3 - x + 1$  and  $g(x) = x^2 + 1$ . Dividing f by g (you should know how from high school) we obtain a quotient of x with remainder -2x + 1. Thus

$$x^{3} - x + 1 = (x^{2} + 1)x + (-2x + 1)x$$

Hence

$$rac{f(x)}{g(x)} = x^2 + 1 + rac{-2x+1}{x^2+1}$$

To find the integral of f(x)/g(x) we integrate  $x^2$  and 1, and the quotient on the right, which has the property that the degree of the numerator is less than the degree of the denominator.

From now on, we assume throughout that when we consider a quotient f(x)/g(x), the degree of f is less than the degree of g. Factoring out a constant if necessary, we also assume that g(x) can be written  $g(x) = x^d +$ lower terms. We shall begin by discussing special cases, and then describe afterwards how the general case can be reduced to these.

Case 1. If a is a number, find

$$\int \frac{1}{(x-a)^n} \, dx$$

if n is an integer  $\geq 1$ .

This is an old story. We know how to do it.

Case 2. If b is a number, find

$$\int \frac{1}{(x^2+b^2)^n}\,dx.$$

This is new. Using the substitution x = bz, dx = b dz reduces the integral to

$$\int \frac{1}{(x^2+1)^n} \, dx.$$

We shall now discuss how to find this integral. We use integration by parts when n > 1. (If n = 1, this is the arctangent.) Let the above integral be  $I_n$ . We shall start with  $I_{n-1}$  because our integration by parts will raise n instead of lowering n. We have

$$I_{n-1} = \int \frac{1}{(x^2+1)^{n-1}} \, dx.$$

Let  $u = \frac{1}{(x^2 + 1)^{n-1}}$  and dv = dx. Then

$$du = -(n-1) \frac{2x}{(x^2+1)^n} dx$$
 and  $v = x$ .

Thus

$$I_{n-1} = \frac{x}{(x^2+1)^{n-1}} + 2(n-1) \int \frac{x^2}{(x^2+1)^n} dx$$

We write  $x^2 = x^2 + 1 - 1$ . We obtain

$$I_{n-1} = \frac{x}{(x^2+1)^{n-1}} + 2(n-1)I_{n-1} - 2(n-1)I_n.$$

Therefore

$$2(n-1)I_n = \frac{x}{(x^2+1)^{n-1}} + (2n-3)I_{n-1},$$

whence

$$\int \frac{1}{(x^2+1)^n} dx = \frac{1}{2(n-1)} \frac{x}{(x^2+1)^{n-1}} + \frac{(2n-3)}{2(n-1)} \int \frac{1}{(x^2+1)^{n-1}} dx$$

Case 3. Find the integral

$$\int \frac{x}{(x^2+b^2)^n} \, dx.$$

This is an old story. We make the substitution  $u = x^2 + b^2$  and  $du = 2x \, dx$ . The integral is then equal to

$$\frac{1}{2}\int \frac{1}{u^n}\,du,$$

which you know how to do.

We shall now investigate a general quotient f(x)/g(x).

If one is given a polynomial of type  $x^2 + bx + c$ , then one completes the square. The polynomial can thus be written in the form

$$(x-\alpha)(x-\beta)$$
 or  $(x-\alpha)^2+\beta^2$ 

with suitable numbers  $\alpha$ ,  $\beta$ . For instance,

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

It can be shown that a polynomial g(x) can always be written as a product of terms of type

 $(x-\alpha)^n$  and  $[(x-\beta)^2+\gamma^2]^m$ ,

n, m being integers  $\geq 0$ . It can also be shown that a quotient f(x)/g(x) can be written as a sum of terms of the following type:

$$\frac{c_1}{x - \alpha} + \frac{c_2}{(x - \alpha)^2} + \dots + \frac{c_n}{(x - \alpha)^{n-1}} \\ + \frac{d_1 + e_1 x}{(x - \beta)^2 + \gamma^2} + \dots + \frac{d_m + e_m x}{[(x - \beta)^2 + \gamma^2]^m}$$

with suitable constants  $c_1, c_2, \ldots, d_1, d_2, \ldots, e_1, e_2, \ldots$ . The way one can determine these constants is to put the right-hand side over the common denominator g(x), equate the numerator f(x) with what is obtained on the right, and solve for the constants. We shall illustrate this by examples, and will not give the proofs of the preceding assertions, which are quite difficult.

Example 2. Express the quotient

$$\frac{1}{(x-2)(x-3)}$$

as a sum of the previous type.

This will be

$$rac{c_1}{x-2}+rac{c_2}{x-3}$$

Putting this over the common denominator, we get the numerator

$$c_1(x-3) + c_2(x-2) = (c_1 + c_2)x - 3c_1 - 2c_2,$$

which must be equal to 1. Thus we must have

$$c_1 + c_2 = 0$$
  
$$-3c_1 - 2c_2 = 1.$$

Solving for  $c_1$  and  $c_2$  gives  $c_2 = 1$  and  $c_1 = -1$ .

*Example 3.* Express the quotient

$$\frac{x+1}{(x-1)^2(x-2)}$$

as a sum of the previous type.

We want to find numbers  $c_1$ ,  $c_2$ ,  $c_3$  such that

$$\frac{x+1}{(x-1)^2(x-2)} = \frac{c_1}{x-1} + \frac{c_2}{(x-1)^2} + \frac{c_3}{x-2}$$

Putting the right-hand side over the common denominator  $(x-1)^2(x-2)$ , we get a numerator equal to

$$c_1(x-1)(x-2) + c_2(x-2) + c_3(x-1)^2$$
.

This can be rewritten as

$$(c_1 + c_3)x^2 + (-3c_1 + c_2 - 2c_3)x + 2c_1 - 2c_2 + c_3$$

and must be equal to x + 1. We equate the coefficients of  $x^2$ , x, and the constant terms. We get

$$c_1 + c_3 = 0$$
  
 $-3c_1 + c_2 - 2c_3 = 1$   
 $2c_1 - 2c_2 + c_3 = 1.$ 

This is a system of three linear equations in three unknowns, which you can solve to determine  $c_1$ ,  $c_2$ , and  $c_3$ . One finds  $c_1 = -3$ ,  $c_2 = -2$ ,  $c_3 = 3$ .

Example 4. Express the quotient

$$\frac{2x+5}{(x^2+1)^2(x-3)}$$

as a sum of the type described above.

We can find numbers  $c_1, c_2, \ldots$  such that the quotient is equal to

$$\frac{c_1+c_2x}{x^2+1}+\frac{c_3+c_4x}{(x^2+1)^2}+\frac{c_5}{x-3}$$

We put this over the common denominator  $(x^2 + 1)^2(x - 3)$ . The numerator is equal to

$$(c_1 + c_2 x)(x^2 + 1)(x - 3) + (c_3 + c_4 x)(x - 3) + c_5(x^2 + 1)^2$$

and must be equal to 2x + 5. If we equate the coefficients of  $x^4$ ,  $x^3$ ,  $x^2$ , x and

[XI, §4]

the constants, we get a system of five linear equations in five unknowns and it can be solved. It would be tedious to do it here.

We are now in a position to evaluate the integrals of the quotients in Examples 2, 3, and 4. We find:

In Example 2, the integral is

$$\int \frac{1}{(x-2)(x-3)} dx = -\log (x-2) + \log (x-3).$$

In Example 3, the integral is

$$\int \frac{x+1}{(x-1)^2(x-2)} \, dx = -3 \log \left(x-1\right) + \frac{2}{x-1} + 3 \log \left(x-2\right).$$

In Example 4, we find

$$\int \frac{2x+5}{(x^2+1)^2(x-3)} dx = c_1 \arctan x + \frac{1}{2}c_2 \log (x^2+1) + c_3 \int \frac{1}{(x^2+1)^2} dx + \frac{1}{2}c_4 \log (x^2+1) + c_5 \log (x-3).$$

The integral which we left standing is just that of Case 2. Find it explicitly as an exercise.

#### Exercises

- 1. Find the constants in Example 4.
- 2. Write out in full the integral

$$\int \frac{1}{(x^2+1)^2} \, dx.$$

Find the following integrals:

3. (a) 
$$\int \frac{1}{(x-3)(x+2)} dx$$
 (b)  $\int \frac{1}{(x+2)(x+1)} dx$   
4.  $\int \frac{x}{(x+1)(x+2)(x+3)} dx$  5.  $\int \frac{x+2}{x^2+x} dx$   
6.  $\int \frac{x}{(x+1)^2} dx$  7.  $\int \frac{x+1}{(x^2+9)^2} dx$   
8.  $\int \frac{4}{(x^2+16)^2} dx$  9.  $\int \frac{x}{(x+1)(x+2)^2} dx$   
10.  $\int \frac{1}{(x^2+1)^3} dx$  11.  $\int \frac{2x-3}{(x-1)(x+7)} dx$ 

# CHAPTER XII

# Some Substantial Exercises

We shall not use the contents of this section until the chapter on series, and even then we use only the estimate of  $(n!)^{1/n}$ . Thus this chapter may be skipped entirely. We include it mostly for reference, and to provide some good exercises for those interested.

# §1. An estimate for $(n!)^{1/n}$

Let n be a positive integer. We define n! (which we read n factorial) to be the product of the first n integers: Thus  $1 \cdot 2 \cdot 3 \cdots n$ . This is certainly less than  $n^n$  (the product of n with itself n times). We shall investigate to what extent it differs from  $n^n$ . (Not too much.)

In fact, what we shall prove first is that

$$n! = n^n e^{-n} d_n,$$

where  $d_n$  is a number such that  $d_n^{1/n}$  approaches 1 as *n* becomes large. This is a weaker statement than the result stated in the next theorem, whose proof is very simple and very easy to remember. It is a nice application of the lower and upper sum techniques.

THEOREM 1. Let n be a positive integer. Then

 $(n-1)! \leq n^n e^{-n} e \leq n!$ 

Proof. Exercise. Evaluate and compare the integral

$$\int_1^n \log x \, dx$$

with the upper and lower sum associated with the partition (1, 2, ..., n) of the interval [1, n]. Then exponentiate.

COROLLARY. As n becomes very large,

$$\frac{(n!)^{1/n}}{n} = \left[\frac{n!}{n^n}\right]^{1/n}$$

approaches 1/e.

178

© Springer Science+Business Media New York 2002

[XII, §2]

*Proof.* Take the n-th root of the right inequality in our theorem. We get

$$ne^{-1}e^{1/n} \leq (n!)^{1/n}.$$

Dividing by n yields

$$\frac{1}{e} e^{1/n} \leq \frac{(n!)^{1/n}}{n}$$

On the other hand, multiply both sides of the inequality

$$(n-1)! \leq n^n e^{-n} e$$

by n. We get  $n! \leq n^n e^{-n} e^{n}$ . Take an n-th root:

$$(n!)^{1/n} \leq n e^{-1} e^{1/n} n^{1/n}.$$

Dividing by n yields

$$\frac{(n!)^{1/n}}{n} \leq \frac{1}{e} e^{1/n} n^{1/n}$$

But we know that both  $n^{1/n}$  and  $e^{1/n}$  approach 1 as n becomes large. Thus our quotient is squeezed between two numbers approaching 1/e, and must therefore approach 1/e.

#### Exercises

1. Use the abbreviation lim to mean: limit as n becomes very large. Prove that  $n \to \infty$ 

(a)  $\lim_{n \to \infty} \left[ \frac{(3n)!}{n^{3n}} \right]^{1/n} = \frac{27}{e^3}$  (b)  $\lim_{n \to \infty} \left[ \frac{(3n)!}{n! n^{2n}} \right]^{1/n} = \frac{27}{e^2}$ 

2. Find the limit:

(a) 
$$\lim_{n \to \infty} \left[ \frac{(2n)!}{n^{2n}} \right]^{1/n}$$
 (b)  $\lim_{n \to \infty} \left[ \frac{(2n)!(5n)!}{n^{4n}(3n)!} \right]^{1/n}$ 

### §2. Stirling's formula

Using various refinements of the above method, one can prove the following theorem.

THEOREM 2. Let n be a positive integer. Then there is a number  $\theta$  between 0 and 1 such that

$$n! = \sqrt{2\pi n} n^n e^{-n} e^{\theta/12n}$$

We shall present another proof giving the main steps, and leave the details to you as exercises.

$$\varphi'(x) = \frac{x^2}{1-x^2}.$$

2. Let 
$$\psi(x) = \varphi(x) - \frac{x^3}{3(1-x^2)}$$
. Show that  
 $\psi'(x) = \frac{-2x^4}{3(1-x^2)^2}$ .

3. For 0 < x < 1, conclude that  $\varphi(x) > 0$  and  $\psi(x) < 0$ .

4. Deduce that for  $0 \leq x < 1$  we have

$$0 \leq \frac{1}{2} \log \frac{1+x}{1-x} - x \leq \frac{x^3}{3(1-x^2)}.$$
  
5. Let  $x = \frac{1}{2n+1}$ . Then  $\frac{1+x}{1-x} = \frac{n+1}{n}$  and

$$\frac{x^2}{3(1-x^2)} = \frac{1}{12(2n+1)(n^2+n)}$$

6. Conclude that

$$0 \leq \frac{1}{2} \log \frac{n+1}{n} - \frac{1}{2n+1} \leq \frac{1}{12(2n+1)(n^2+n)}$$
$$0 \leq (n+\frac{1}{2}) \log \frac{n+1}{n} - 1 \leq \frac{1}{12} \left(\frac{1}{n} - \frac{1}{n+1}\right).$$

7. Let

$$a_n = rac{n^{n+rac{1}{2}} e^{-n}}{n!}$$
 and  $b_n = a_n e^{1/12n}$ 

Then  $a_n \leq b_n$ . Show that

$$\frac{a_{n+1}}{a_n} \ge 1$$
 and  $\frac{b_{n+1}}{b_n} \le 1$ 

Thus the  $a_n$  are increasing and the  $b_n$  are decreasing. Hence there exists a unique number c such that

 $a_n \leq c \leq b_n$ 

for all n.

8. Conclude that

$$n! = c^{-1} n^{n + \frac{1}{2}} e^{-n} e^{\theta/12n}$$

for some number  $\theta$  between 0 and 1.

[XII, §2]

To get the value of the constant c, one has to use another argument, which will be described in the next section.

## §3. Wallis' product

Our first aim is to obtain the following limit, known as the Wallis product.

THEOREM 3. We have

$$\frac{\pi}{2} = \lim_{n \to \infty} \frac{2}{1} \frac{2}{3} \frac{4}{3} \frac{4}{5} \frac{6}{5} \frac{6}{7} \cdots \frac{2n}{2n-1} \frac{2n}{2n+1}.$$

Proof. The proof will again be presented as an exercise.

1. Using the recurrence formulas for the integrals of powers of the sine, prove that

$$\int_0^{\pi/2} \sin^{2n} x \, dx = \frac{2n-1}{2n} \frac{2n-3}{2n-2} \cdots \frac{1}{2} \frac{\pi}{2}$$
$$\int_0^{\pi/2} \sin^{2n+1} x \, dx = \frac{2n}{2n+1} \frac{2n-2}{2n-1} \cdots \frac{2}{3}$$

2. Using the fact that powers of the sine are decreasing as n = 1, 2, 3, ...and the first integral formula above, conclude that

$$1 \leq \frac{\int_0^{\pi/2} \sin^{2n-1} x \, dx}{\int_0^{\pi/2} \sin^{2n+1} x \, dx} \leq 1 + \frac{1}{2n}$$

3. Taking the ratio of the integrals of  $\sin^{2n} x$  and  $\sin^{2n+1} x$  between 0 and  $\pi/2$ , deduce Wallis' product.

COROLLARY. We have

$$\lim_{n \to \infty} \frac{(n!)^2 2^{2n}}{(2n)! n^{1/2}} = \pi^{1/2}.$$

Proof. Rewrite the Wallis product into the form

$$\frac{\pi}{2} = \lim_{n \to \infty} \frac{2^2 4^2 \cdots (2n-2)^2}{3^2 5^2 \cdots (2n-1)^2} 2n.$$

Take the square root and find the limit stated in the corollary.

Finally, show that the constant c in Stirling's formula is  $1/\sqrt{2\pi}$ , by arguing as follows. (Justify all the steps.)

$$c = \lim_{n \to \infty} \frac{(2n)^{2n + \frac{1}{2}e^{-2n}}}{(2n)!}$$
  
= 
$$\lim_{n \to \infty} \frac{(n!)^2 2^{2n} \sqrt{2}}{(2n)! n^{1/2}} \left[ \frac{n^{n + \frac{1}{2}e^{-n}}}{n!} \right]^2$$
  
= 
$$\sqrt{2\pi} \cdot c^2.$$

Thus  $c = 1/\sqrt{2\pi}$ .

# CHAPTER XIII

# **Applications of Integration**

Most of the applications are to physical concepts and show us how useful it is to have related differentiation and integration as limits of sum.

### §1. Length of curves

Let y = f(x) be a differentiable function over some interval [a, b] (with  $a \leq b$ ) and assume that its derivative f' is continuous. We wish to find a way to determine the length of the curve described by the graph. The main idea is to approximate the curve by small line segments and add these up.



Consequently, we consider a partition of our interval:

 $a = x_0 \leq x_1 \leq \cdots \leq x_n = b.$ 

For each  $x_i$  we have the point  $(x_i, f(x_i))$  on the curve y = f(x). We draw the line segments between two successive points. The length of such a segment is the length of the line between

 $(x_i, f(x_i))$  and  $(x_{i+1}, f(x_{i+1}))$ ,

and is equal to

$$\sqrt{(x_{i+1} - x_i)^2 + (f(x_{i+1}) - f(x_i))^2}$$

By the mean value theorem, we conclude that

$$f(x_{i+1}) - f(x_i) = (x_{i+1} - x_i)f'(c_i)$$
183

© Springer Science+Business Media New York 2002

for some number  $c_i$  between  $x_i$  and  $x_{i+1}$ . Using this, we see that the length of our line segment is

$$\sqrt{(x_{i+1}-x_i)^2+(x_{i+1}-x_i)^2f'(c_i)^2}$$

We can factor out  $(x_{i+1} - x_i)^2$  and we see that the sum of the length of these line segments is

$$\sum_{i=0}^{n-1} \sqrt{1+f'(c_i)^2} (x_{i+1}-x_i).$$

Let  $G(x) = \sqrt{1 + f'(x)^2}$ . Then G(x) is continuous, and we see that the sum we have just written down is

$$\sum_{i=0}^{n-1} G(c_i)(x_{i+1} - x_i).$$

This is precisely a Riemann sum used to find the integral. It is therefore very reasonable to *define* the length of our curve between a and b to be

$$\int_a^b \sqrt{1+f'(t)^2} \, dt.$$

(As an exercise, show that the least upper bound of the above sums is equal to the integral.)

**Parametric form.** We shall now see what happens when the curve is given in parametric form.

Suppose that our curve is given by

$$x = f(t), \qquad y = g(t),$$

with  $a \leq t \leq b$ , and assume that both f, g have continuous derivatives. As before, we cut our interval into a partition, say

 $a = t_0 \leq t_1 \leq \cdots \leq t_n = b.$ 

We then obtain points  $(f(t_i), g(t_i))$  on the curve, and the distance between two successive points is

$$\sqrt{(f(t_{i+1}) - f(t_i))^2 + (g(t_{i+1}) - g(t_i))^2}.$$

We use the mean value theorem for f and g. There are numbers  $c_i$  and  $d_i$  between  $t_i$  and  $t_{i+1}$  such that

$$f(t_{i+1}) - f(t_i) = f'(c_i)(t_{i+1} - t_i)$$
  
$$g(t_{i+1}) - g(t_i) = g'(d_i)(t_{i+1} - t_i).$$

Substituting these values and factoring out  $(t_{i+1} - t_i)$ , we see that the sum of the lengths of our line segments is equal to

$$\sum_{i=0}^{n-1} \sqrt{f'(c_i)^2 + g'(d_i)^2} (t_{i+1} - t_i).$$

Let

$$G(t) = \sqrt{f'(t)^2 + g'(t)^2}.$$

Then our sum is almost equal to

$$\sum_{i=0}^{n-1} G(c_i)(t_{i+1} - t_i),$$

which would be a Riemann sum for G. It is not, because it is not necessarily true that  $c_i = d_i$ . Nevertheless, what we have done makes it very reasonable to *define* the length of our curve (in parametric form) to be

$$\int_a^b \sqrt{f'(t)^2 + g'(t)^2} \, dt.$$

(A complete justification that this integral is a limit, in a suitable sense, of our sums would require some additional theory, which is irrelevant anyway since we just want to make it reasonable that the above integral should represent what we mean physically by length.)

Observe that when y = f(x) we can let t = x = g(t) and y = f(t). In that case, g'(t) = 1 and the formula for the length in parametric form is seen to be the same as the formula we obtained before for a curve y = f(x).

Example. Find the length of the curve

$$x = \cos t, \qquad y = \sin t$$

between t = 0 and  $t = \pi$ .

The length is the integral

$$\int_0^{\pi} \sqrt{(-\sin t)^2 + (\cos t)^2} \, dt.$$

In view of the relation  $(-\sin t)^2 = (\sin t)^2$  and a basic formula relating sine and cosine, we get

$$\int_0^{\pi} dt = \pi.$$

If we integrated between 0 and  $2\pi$  we would get  $2\pi$ . This is the length of the circle of radius 1.

**Polar coordinates.** Let us now find a formula for the length of curves given in polar coordinates. Say the curve is

$$r=f(\theta),$$

with  $\theta_1 \leq \theta \leq \theta_2$ . We know that

$$x = r \cos \theta = f(\theta) \cos \theta$$
$$y = r \sin \theta = f(\theta) \sin \theta.$$

This puts the curve in parametric form, just as in the preceding considerations. Consequently we can apply the definition as before, and we see that the length is

$$\int_{\theta_1}^{\theta_2} \sqrt{\left(\frac{dx}{d\theta}\right)^2 + \left(\frac{dy}{d\theta}\right)^2} \, d\theta$$

You can compute  $dx/d\theta$  and  $dy/d\theta$  using the rule for the derivative of a product. If you do this, you will find that many terms cancel, and that the integral is equal to

$$\int_{\theta_1}^{\theta_2} \sqrt{f(\theta)^2 + f'(\theta)^2} \ d\theta.$$

(The computation is very easy, and is good practice in simple identities involving sine and cosine. We leave it to you as an exercise. Anyhow, working it out will make you remember the formula better.)

#### Exercises

1. Carry out the preceding computation.

2. Find the length of a circle of radius r.

3. Find the length of the curve  $x = e^t \cos t$ ,  $y = e^t \sin t$  between t = 1 and t = 2.

4. Find the length of the curve  $x = \cos^3 t$ ,  $y = \sin^3 t$  (a) between t = 0 and  $t = \pi/4$ , and (b) between t = 0 and  $t = \pi$ .

Find the lengths of the following curves:

5.  $y = e^x$  between x = 0 and x = 1.

6.  $y = x^{3/2}$  between x = 1 and x = 3.

7.  $y = \frac{1}{2}(e^{x} + e^{-x})$  between x = -1 and x = 1.

8. Find the length of one loop of the curve  $r = 1 + \cos \theta$  (polar coordinates).

9. Same, with  $r = \cos \theta$ , between  $-\pi/2$  and  $\pi/2$ .

10. Find the length of the curve  $r = 2/\cos\theta$  between  $\theta = 0$  and  $\theta = \pi/3$ .

11. Find the length of the curve  $r = |\sin \theta|$  from  $\theta = 0$  to  $\theta = 2\pi$ .

12. Sketch the curve  $r = e^{\theta}$  (in polar coordinates), and also the curve  $r = e^{-\theta}$ .

13. Find the length of the curve  $r = e^{\theta}$  between  $\theta = 1$  and  $\theta = 2$ .

14. In general, give the length of the curve  $r = e^{\theta}$  between two values  $\theta_1$  and  $\theta_2$ .

### §2. Area in polar coordinates

Suppose we are given a function

$$r = f(\theta)$$

which is defined in some interval  $a \leq \theta \leq b$ . We assume that  $f(\theta) \geq 0$ and  $b \leq a + 2\pi$ .

We wish to find an integral expression for the area encompassed by the curve  $r = f(\theta)$  between the two bounds a and b.

Let us take a partition of [a, b], say

$$a = \theta_0 \leq \theta_1 \leq \cdots \leq \theta_n = b.$$

The picture between  $\theta_i$  and  $\theta_{i+1}$  might look like this:



We let  $s_i$  be a number between  $\theta_i$  and  $\theta_{i+1}$  such that  $f(s_i)$  is a maximum in that interval, and we let  $t_i$  be a number such that  $f(t_i)$  is a minimum in that interval. In the picture, we have drawn the circles (or rather the sectors) of radius  $f(s_i)$  and  $f(t_i)$ , respectively.

Then the area between  $\theta_i$ ,  $\theta_{i+1}$  and the curve lies between the two sectors. Denote it by  $A_i$ . The area of a sector having angle  $\theta_{i+1} - \theta_i$  and radius R is equal to the fraction

$$\frac{\theta_{i+1}-\theta_i}{2\pi}$$

[XIII, §3]

of the total area of the circle of radius R, namely  $\pi R^2$ . Hence we get the inequality

$$\frac{\theta_{i+1}-\theta_i}{2\pi}\pi f(t_i)^2 \leq A_i \leq \frac{\theta_{i+1}-\theta_i}{2\pi}\pi f(s_i)^2.$$

Let  $G(\theta) = \frac{1}{2}f(\theta)^2$ . We see that the sum of the small pieces of area  $A_i$  satisfies the inequalities

$$\sum_{i=0}^{n-1} G(t_i)(\theta_{i+1} - \theta_i) \leq \sum_{i=0}^{n-1} A_i \leq \sum_{i=0}^{n-1} G(s_i)(\theta_{i+1} - \theta_i).$$

Thus the desired area lies between the upper sum and lower sum associated with the partition. Thus it is reasonable to define the *area* to be

$$\int_{\theta_1}^{\theta_2} \frac{1}{2} f(\theta)^2 \, d\theta.$$

Example. Find the area bounded by one loop of the curve

$$r^2 = 2a^2 \cos 2\theta \qquad (a > 0).$$

Between  $-\frac{\pi}{4}$  and  $\frac{\pi}{4}$  the cosine is  $\geq 0$ . Thus we can write

$$r=\sqrt{2} a \sqrt{\cos 2\theta}$$
.

The area is therefore

$$\int_{-\pi/4}^{\pi/4} \frac{1}{2} 2a^2 \cos 2\theta \, d\theta = a^2$$

#### Exercises

Find the area enclosed by the following curves:

1.  $r = 2(1 + \cos \theta)$ 2.  $r^2 = a^2 \sin 2\theta$  (a > 0)3.  $r = 2a \cos \theta$ 4.  $r = \cos 3\theta, -\pi/6 \le \theta \le \pi/6$ 5.  $r = 1 + \sin \theta$ 6.  $r = 1 + \sin 2\theta$ 7.  $r = 2 + \cos \theta$ 8.  $r = 2 \cos 3\theta, -\pi/6 \le \theta \le \pi/6$ 

# §3. Volumes of revolution

Let y = f(x) be a continuous function of x on some interval  $a \leq x \leq b$ . Assume that  $f(x) \geq 0$  in this interval. If we revolve the curve y = f(x) around the x-axis, we obtain a solid, whose volume we wish to compute.



Take a partition of [a, b], say

$$a = x_0 \leq x_1 \leq \cdots \leq x_n = b$$

Let  $c_i$  be a minimum of f in the interval  $[x_i, x_{i+1}]$  and let  $d_i$  be a maximum of f in that interval. Then the solid of revolution in that small interval lies between a small cylinder and a big cylinder. The width of these cylinders is  $x_{i+1} - x_i$  and the radius is  $f(c_i)$  for the small cylinder and  $f(d_i)$  for the big one. Hence the volume of revolution, denoted by V, satisfies the inequalities

$$\sum_{i=0}^{n-1} \pi f(c_i)^2 (x_{i+1} - x_i) \leq V \leq \sum_{i=0}^{n-1} \pi f(d_i)^2 (x_{i+1} - x_i).$$

It is therefore reasonable to define this volume to be

$$\int_a^b \pi f(x)^2 \ dx.$$

*Example.* Compute the volume of the sphere of radius 1.

We take the function  $y = \sqrt{1 - x^2}$  between 0 and 1. If we rotate this curve around the x-axis, we shall get half the sphere. Its volume is therefore

$$\int_0^1 \pi (1 - x^2) \, dx = \frac{2}{3} \pi.$$

The volume of the full sphere is therefore  $\frac{4}{3}\pi$ .

#### Exercises

1. Find the volume of a sphere of radius r

Find the volumes of revolution of the following:

- 2.  $y = 1/\cos x$  between x = 0 and  $x = \pi/4$
- 3.  $y = \sin x$  between x = 0 and  $x = \pi/4$
- 4.  $y = \cos x$  between x = 0 and  $x = \pi/4$

5. The region between  $y = x^2$  and y = 5x

6.  $y = xe^{x/2}$  between x = 0 and x = 1

7.  $y = x^{1/2}e^{x/2}$  between x = 1 and x = 2

8.  $y = \log x$  between x = 1 and x = 2

9.  $y = \sqrt{1+x}$  between x = 1 and x = 5

10. Let B be a number >1. What is the volume of revolution of the curve  $y = e^{-x}$  between 1 and B? Does this volume approach a limit as B becomes large? If so, what limit?

11. Find the volume of a cone whose base has radius r, and of height h.

### §4. Work

Suppose a particle moves on a curve, and that the length of the curve is described by a variable u.

Let f(u) be a function. We interpret f as a force acting on the particle, in the direction of the curve. We want to find an integral expression for the work done by the force between two points on the curve.

Whatever our expression will turn out to be, it is reasonable that the work done should satisfy the following properties:

If a, b, c are three numbers, with  $a \leq b \leq c$ , then the work done between a and c is equal to the work done between a and b, plus the work done between b and c. If we denote the work done between a and b by  $W_a^b(f)$ , then we should have

$$W_a^c(f) = W_a^o(f) + W_b^c(f).$$

Furthermore, if we have a constant force M acting on the particle, it is reasonable to expect that the work done between a and b is

$$M(b - a).$$

Finally, if g is a stronger force than f, say  $f(u) \leq g(u)$ , on the interval [a, b], then we shall do more work with g than with f, meaning

$$W_a^b(f) \leq W_a^b(g).$$

In particular, if there are two constant forces m and M such that

$$m \leq f(u) \leq M$$

throughout the interval [a, b], then

$$m(b-a) \leq W_a^b(f) \leq M(b-a).$$

[XIII, §4]

But this condition, together with the first one expressed above, determines the integral uniquely! Hence there is only one reasonable way to associate a mathematical formula with work, compatible with physical requirements, and that is: The work done by the force f between a distance a and a distance b is

$$\int_a^b f(u) \ du.$$

If the particle or object happens to move along a straight line, say along the x-axis, then f is given as function of x, and our integral is simply

$$\int_a^b f(x) \ dx.$$

Furthermore, if the length of the curve u is given as a function of time t (as it is in practice, cf. §1) we see that the force becomes a function of t by the chain rule, namely f(u(t)). Thus between time  $t_1$  and  $t_2$  the work done is equal to

$$\int_{t_1}^{t_2} f(u(t)) \, \frac{du}{dt} \, dt.$$

This is the most practical expression for the work, since curves and forces are most frequently expressed as functions of time.

*Example.* Find the work done in stretching a spring from its unstretched position to a length of 10 cm. You may assume that the force needed to stretch the spring is proportional to the increase in length.

We visualize the spring as being horizontal, on the x-axis. Thus there is a constant K such that the force is given by

$$f(x) = Kx.$$

The work done is therefore

$$\int_0^{10} Kx \, dx = \frac{1}{2} K(100) \\ = 50 K.$$

Note. No exercises on work (and on moments in the next section) will be given. These belong properly in a physics course. However, the discussion concerning the uniqueness of our notion  $W_a^b(f)$  (and  $M_a^b(f)$  in the next section) does belong to this course.

#### **§5.** Moments

We wish to describe the notion of "moment" (with respect to the origin), which arises in physics. "Moment" should have the following properties.

Suppose that we have a mass K concentrated at a point b on the x-axis. Then its moment is Kb.

$$\begin{array}{c} K \\ \hline 0 \\ b \end{array}$$

Suppose that we have an interval [a, b], with a < b (which may be interpreted as a rod), and a constant density distribution C over the interval. Then the total mass is C(b - a). The moment of our constant distribution, which we denote by M, should then satisfy the inequalities

$$Ca(b - a) \leq M \leq Cb(b - a).$$

If we think of the total mass K = C(b - a) as being concentrated at the point b, then the right inequality states that the moment of our constant distribution is less than or equal to the moment of a total mass K placed at the point b. The left inequality states that the moment of our constant distribution is greater than or equal to that of a total mass K placed at the point a. If we assume  $0 \leq a < b$  then these two inequalities state that the moment should be larger, the farther our total mass is from the origin. We shall assume  $0 \leq a < b$  throughout our discussion.

Suppose that we have a variable density distribution over our interval [a, b]. This means that our distribution is represented by a function f(x) on the interval. Let us denote its "moment" by  $M_a^b(f)$ . The following two properties should then be satisfied.

Property 1. If f is continuous on the interval [a, b] and if there are two constants  $C_1, C_2$  both  $\geq 0$  such that

$$C_1 \leq f(x) \leq C_2$$

for all x in our interval, then the moment of f satisfies the inequalities

$$C_1a(b-a) \leq M_a^b(f) \leq C_2b(b-a).$$

(In other words, the moment of f should lie between the moments determined by the constant distributions  $C_1$  and  $C_2$ .)

Property 2. If f is continuous on the interval [a, c], with  $a \leq c$  and if b is a point such that  $a \leq b \leq c$ , then

$$M_a^c(f) = M_a^b(f) + M_b^c(f).$$

[XIII, §5]

MOMENTS

(In other words, the moment over the big interval is the sum of the moments over the first interval and the second interval.)

We shall now prove that there is one and only one way of obtaining a moment  $M_a^b(f)$  satisfying the above two properties, and that is the integral

$$\int_a^b x f(x) \ dx.$$

To begin with, we note that the integral satisfies Property 2. This is an old result.

As to Property 1, assume that f lies between two constants  $C_1$  and  $C_2$  as above. For the interval  $a \leq x \leq b$ , we get

$$C_1a \leq xf(x) \leq C_2b,$$

and hence by Theorem 1 of Chapter XI, §3 we conclude that

$$C_1a(b-a) \leq \int_a^b xf(x) dx \leq C_2b(b-a).$$

Thus the integral certainly satisfies our two properties.

We shall now prove conversely that any  $M_a^b(f)$  satisfying the two properties must be the integral written down above. This is done by our usual technique. We prove that the function  $M_a^x(f)$  has a derivative, and that this derivative is xf(x). Such a function must be the integral.

The Newton quotient is

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

Using Property 2, we see that the Newton quotient is equal to

$$\frac{M_x^{x+h}(f)}{h} \cdot$$

Let s be a maximum for f in the small interval between x and x + h and let t be a minimum for f in that small interval. If, say, h is positive, then by Property 1 we conclude that

$$f(t)x(x + h - x) \leq M_x^{x+h}(f) \leq f(s)(x + h)(x + h - x),$$

or in other words that

$$f(t)xh \leq M_x^{x+h}(f) \leq f(s)(x+h)h.$$

Dividing by h yields

$$xf(t) \leq \frac{M_x^{x+h}(f)}{h} \leq (x+h)f(s).$$

As h approaches 0, both f(t) and f(s) approach f(x) because s, t lie between x and x + h. The usual squeezing argument shows that the Newton quotient approaches xf(x), as was to be shown. If h is negative, a similar argument can be applied as usual.

Consider now a density distribution on the interval [a, b] represented by a continuous function f. We define the *total mass* (or simply *mass*) of this distribution over the interval to be the integral

$$\int_a^b f(x) \ dx.$$

Assume that  $f(x) \ge 0$  for all x in the interval, and that  $f(x) \ne 0$  for some point in the interval. Then the total mass is positive. We define the *center of gravity* of our distribution over the interval to be that point c such that the moment of our total mass concentrated at the point c is equal to the moment of f over the interval. This can be expressed by the relation

$$c \cdot \int_a^b f(x) \ dx = \int_a^b x f(x) \ dx.$$

In other words,

$$c = \frac{\int_a^b xf(x) \, dx}{\int_a^b f(x) \, dx} \cdot$$

If you study physics, you will recognize this as being the usual formula giving the center of gravity.

### CHAPTER XIV

# **Taylor's Formula**

We finally come to the point where we develop a method which allows us to compute the values of the elementary functions like sine, exp, and log. The method is to approximate these functions by polynomials, with an error term which is easily estimated. This error term will be given by an integral, and our first task is to estimate integrals. We then go through the elementary functions systematically, and derive the approximating polynomials.

You should review the estimates of Chapter X, §3, which will be used to estimate our error terms.

### §1. Taylor's formula

Let f be a function which is differentiable on some interval. We can then take its derivative f' on that interval. Suppose that this derivative is also differentiable. We need a notation for its derivative. We shall denote it by  $f^{(2)}$ . Similarly, if the derivative of the function  $f^{(2)}$  exists, we denote it by  $f^{(3)}$ , and so forth. In this system, the first derivative is denoted by  $f^{(1)}$ . (Of course, we can also write  $f^{(2)} = f''$ .)

In the d/dx notation, we also write

$$egin{aligned} f^{(2)}(x) &= rac{d^2 f}{dx^2}\,, \ f^{(3)}(x) &= rac{d^3 f}{dx^3}\,, \end{aligned}$$

and so forth.

Taylor's formula gives us a polynomial which approximates the function, in terms of the derivatives of the function. Since these derivatives are usually easy to compute, there is no difficulty in computing these polynomials.

For instance, if  $f(x) = \sin x$ , then  $f^{(1)}(x) = \cos x$ ,  $f^{(2)}(x) = -\sin x$ ,  $f^{(3)}(x) = -\cos x$ , and  $f^{(4)}(x) = \sin x$ . From there on, we start all over again.

In the case of  $e^x$ , it is even easier, namely  $f^{(n)}(x) = e^x$  for all positive integers n.

195

It is also customary to denote the function f itself by  $f^{(0)}$ . Thus  $f(x) = f^{(0)}(x)$ .

We need one more piece of notation before stating Taylor's formula. When we take successive derivatives of functions, the following numbers occur frequently:

1,  $2 \cdot 1$ ,  $3 \cdot 2 \cdot 1$ ,  $4 \cdot 3 \cdot 2 \cdot 1$ ,  $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$ , etc.

4!

5!

etc.

These numbers are denoted by

1!

2!

Thus

1! = 1	4! = 24
2! = 2	5! = 120
3! = 6	6! = 720

3!

When n is a positive integer, the symbol n! is read n factorial. Thus in general,

$$n! = n(n-1)(n-2)\cdots 2\cdot 1$$

is the product of the first n integers from 1 to n.

It is also convenient to agree that 0! = 1. This is the convention which makes certain formulas easiest to write.

We are now in a position to state Taylor's formula.

**THEOREM 1.** Let f be a function defined on a closed interval between two numbers a and b. Assume that the function has n derivatives on this interval, and that all of them are continuous functions. Then

$$f(b) = f(a) + \frac{f^{(1)}(a)}{1!} (b - a) + \frac{f^{(2)}(a)}{2!} (b - a)^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!} (b - a)^{n-1} + R_n,$$

where  $R_n$  (which is called the remainder term) is the integral

$$R_n = \int_a^b \frac{(b-t)^{n-1}}{(n-1)!} f^{(n)}(t) \, dt.$$

The most important case of Theorem 1 occurs when a = 0. In that case, the formula reads

$$f(b) = f(0) + \frac{f'(0)}{1!}b + \cdots + \frac{f^{(n-1)}(0)}{(n-1)!}b^{n-1} + R_n.$$

TAYLOR'S FORMULA

Furthermore, if x is any number between a and b, the same formula remains valid for this number x instead of b, simply by considering the interval between a and x instead of the interval between a and b. Thus the formula reads

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \cdots + \frac{f^{(n-1)}(0)}{(n-1)!}x^{n-1} + R_n,$$

where  $R_n$  is the integral

$$R_n = \int_0^x \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t) dt.$$

Each derivative f(0), f'(0), ...,  $f^{(n-1)}(0)$  is a number, and we see that the terms preceding  $R_n$  make up a polynomial in x. This is the approximating polynomial.

Of course, for the formula to be of any use, we have to estimate the remainder  $R_n$  and show that it becomes small when n becomes large, so that the polynomial does indeed approximate the function. We shall do this in the following sections for special functions. You may very well want to look at these sections before reading the proof, so as to familiarize yourselves with the symbols and nature of the theorem.

We shall now prove the theorem, for those of you who are interested in seeing the proof first. It is an application of integration by parts.

We proceed stepwise. We know that a function is the integral of its derivative. Thus when n = 1 we have

$$f(b) - f(a) = \int_a^b f'(t) dt.$$

Let u = f'(t) and dv = dt. Then du = f''(t) dt. We are tempted to put v = t. This is one case where we choose another indefinite integral, namely v = -(b - t), which differs from t by a constant. We still have dv = dt (the minus signs cancel!). Integrating by parts, we get

$$\int_{a}^{b} u \, dv = uv \Big|_{a}^{b} - \int_{a}^{b} v \, du$$
  
=  $-f'(t)(b-t) \Big|_{a}^{b} - \int_{a}^{b} -(b-t)f^{(2)}(t) \, dt$   
=  $f'(a)(b-a) + \int_{a}^{b} (b-t)f^{(2)}(t) \, dt.$ 

This is precisely the Taylor formula when n = 2.

We push it one step further, from 2 to 3. We rewrite the integral just obtained as

$$\int_{a}^{b} f^{(2)}(t)(b - t) dt.$$

Let  $u = f^{(2)}(t)$  and dv = (b - t) dt. Then

$$du = f^{(3)}(t) dt$$
 and  $v = \frac{-(b - t)^2}{2}$ .

Thus, integrating by parts, we find that our integral, which is of the form  $\int_a^b u \, dv$ , is equal to

$$uv\Big|_{a}^{b} - \int_{a}^{b} v \, du = -f^{(2)}(t) \, \frac{(b-t)^{2}}{2}\Big|_{a}^{b} - \int_{a}^{b} - \frac{(b-t)^{2}}{2} f^{(3)}(t) \, dt$$
$$= f^{(2)}(a) \, \frac{(b-a)^{2}}{2} + R_{3}.$$

Here,  $R_3$  is the desired remainder, and the term preceding it is just the proper term in the Taylor formula.

If you need it, do the next step yourself, from 3 to 4. We shall now show you how the general step goes, from n to n + 1.

Suppose that we have already obtained the first n-1 terms of the Taylor formula, with a remainder term

$$R_n = \int_a^b \frac{(b-t)^{n-1}}{(n-1)!} f^{(n)}(t) \, dt,$$

which we rewrite

$$R_n = \int_a^b f^{(n)}(t) \, \frac{(b-t)^{n-1}}{(n-1)!} \, dt.$$

Let  $u = f^{(n)}(t)$  and  $dv = \frac{(b-t)^{n-1}}{(n-1)!} dt$ . Then

$$du = f^{(n+1)}(t) dt$$
 and  $v = \frac{-(b-t)^n}{n!}$ .

(Observe how, when we integrate dv, we get n in the denominator, to climb from (n - 1)! to n!.)

Integrating by parts, we see that  $R_n$  is equal to

$$uv\Big|_{a}^{b} - \int_{a}^{b} v \, du = -f^{(n)}(t) \, \frac{(b-t)^{n}}{n!}\Big|_{a}^{b} - \int_{a}^{b} - \frac{(b-t)^{n}}{n!} f^{(n+1)}(t) \, dt$$
$$= f^{(n)}(a) \, \frac{(b-a)^{n}}{n!} + \int_{a}^{b} \frac{(b-t)^{n}}{n!} f^{(n+1)}(t) \, dt.$$

Thus we have split off one more term of the Taylor formula, and the new remainder is the desired  $R_{n+1}$ . This concludes the proof.

[XIV, §1]

### §2. Estimate for the remainder

THEOREM 2. In Taylor's formula of Theorem 1, there exists a number c in the interval [a, b] such that the remainder  $R_n$  is given by

$$R_n = \frac{f^{(n)}(c)(b-a)^n}{n!}.$$

If  $M_n$  is a number such that  $|f^{(n)}(x)| \leq M_n$  for all x in the interval, then

$$|R_n| \leq \frac{M_n |b-a|^n}{n!}$$

*Proof.* The second assertion follows at once from the first, taking the absolute value of the product of the n-th derivative and the length of the interval.

Let us prove the first assertion. Since  $f^{(n)}$  is continuous on the interval, there exists a point u in the interval such that  $f^{(n)}(u)$  is a maximum, and there exists a point v such that  $f^{(n)}(v)$  is a minimum for all values of  $f^{(n)}$ in our interval.

Let us assume that a < b. Then for any t in the interval, b - t is  $\geq 0$ , and hence

$$\frac{(b-t)^{n-1}}{(n-1)!}f^{(n)}(v) \leq \frac{(b-t)^{n-1}}{(n-1)!}f^{(n)}(t) \leq \frac{(b-t)^{n-1}}{(n-1)!}f^{(n)}(u).$$

Using Theorem 1 of Chapter X, §3, we conclude that similar inequalities hold when we take the integral. However,  $f^{(n)}(v)$  and  $f^{(n)}(u)$  are now fixed numbers which can be taken out of the integral sign. Consequently, we obtain

$$f^{(n)}(v)\int_{a}^{b}\frac{(b-t)^{n-1}}{(n-1)!}\,dt \leq R_{n} \leq f^{(n)}(u)\int_{a}^{b}\frac{(b-t)^{n-1}}{(n-1)!}\,dt.$$

We now perform the integration, which is very easy, and get

$$f^{(n)}(v) \frac{(b-a)^n}{n!} \leq R_n \leq f^{(n)}(u) \frac{(b-a)^n}{n!}$$

By the intermediate value theorem, the *n*-th derivative  $f^{(n)}(t)$  takes on all values between its minimum and maximum in the interval. Hence

$$f^{(n)}(t) \frac{(b-a)^n}{n!}$$

takes on all values between *its* minimum and maximum in the interval.

Hence there is some point c in the interval such that

$$R_n = f^{(n)}(c) \, rac{(b-a)^n}{n!}$$
 ,

which is what we wanted.

The proof in case b < a is similar, except that certain inequalities get reversed. We leave it as an exercise. (*Hint*: Interchange the limits of integration, taking the integral from b to a.)

The estimate of the remainder is particularly useful when b is close to a. In that case, let us rewrite Taylor's formula by setting b - a = h. We obtain:

**THEOREM** 3. Assumptions being as in Theorem 1, we have

$$f(a+h) = f(a) + f'(a)h + \dots + f^{(n-1)}(a) \frac{h^{n-1}}{(n-1)!} + R_n$$

with the estimate

$$|R_n| \leq M_n \frac{|h|^n}{n!},$$

where  $M_n$  is a bound for the absolute value of the n-th derivative of f between a and a + h.

In the following sections, we give several examples.

#### Exercises

(These exercises are mainly to show you how to recover some classical forms of the remainder. You may skip them without harm.)

1. Let g(t) be a continuous function on an interval between two numbers a and b. Show that

$$\int_a^b g(t) \ dt$$

lies between m(b - a) and M(b - a) if m, M are minimum and maximum values of g over the interval.

2. Use the intermediate value theorem to conclude that there is a number c in the interval such that

$$\int_a^b g(t) dt = g(c)(b - a).$$

3. Apply this to the remainder term

$$R_n = \int_a^b \frac{(b-t)^{n-1}}{(n-1)!} f^{(n)}(t) dt$$

[XIV, §3]

to conclude that there exists a number c between a and b such that

$$R_n = \frac{(b-c)^{n-1}(b-a)}{(n-1)!}f^{(n)}(c).$$

Except in some exercises, from now on we assume that the Taylor formula is taken with a = 0, so that we have

$$f(x) = f(0) + f'(0)x + \cdots + \frac{f^{(n-1)}(0)}{(n-1)!}x^{n-1} + R_n$$

with the estimate

$$|R_n| \leq M_n \frac{|x|^n}{n!}$$

if  $M_n$  is a bound for the *n*-th derivative of *f* between 0 and *x*.

# **§3.** Trigonometric functions

Let  $f(x) = \sin x$  and take a = 0 in Taylor's formula. We have already mentioned what the derivatives of  $\sin x$  and  $\cos x$  are. Thus

$$f(0) = 0$$
  $f^{(2)}(0) = 0$   
 $f'(0) = 1$   $f^{(3)}(0) = -1$ 

The Taylor formula for  $\sin x$  is therefore as follows:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots + (-1)^{n-1} \frac{x^{2n-1}}{(2n-1)!} + R_{2n+1}.$$

We see that all the even terms are 0 because  $\sin 0 = 0$ .

We can estimate  $\sin x$  and  $\cos x$  very simply, because

$$|\sin x| \leq 1$$
 and  $|\cos x| \leq 1$ 

for all x. Thus we take the bound

$$M_n = 1$$

for all n, and

$$|R_n| \leq \frac{|x|^n}{n!} \cdot$$

Thus if we look at all values of x such that  $|x| \leq 1$ , we see that  $R_n$  approaches 0 when n becomes very large.

*Example* 1. Compute  $\sin(0.1)$  to 3 decimals.

Let us estimate  $R_3$ . We let a = 0, b = a + 0.1 in Taylor's formula. We get the estimate

$$|R_3| \leq \frac{(0.1)^3}{3!} = \frac{10^{-3}}{6}$$
.

Such an error term would put us within the required range of accuracy. Hence we can just use the first term of Taylor's formula,

$$\sin(0.1) = 0.100,$$

with an error which does not exceed  $\pm \frac{1}{6} 10^{-4}$ .

We see how efficient this is for computing the sine for small values of x.

*Example 2.* Compute  $\sin\left(\frac{\pi}{6} + 0.2\right)$  to an accuracy of  $10^{-4}$ .

In this case, we take  $a = \pi/6$ . By trial and error, and guessing, we try for the remainder  $R_4$ . Thus

$$\sin (a + h) = \sin a + \cos (a) \frac{h}{1} - \sin (a) \frac{h^2}{2!} - \cos (a) \frac{h^3}{3!} + R_4$$
$$= \frac{1}{2} + \frac{\sqrt{3}}{2} (0.2) - \frac{1}{2} \frac{(0.2)^2}{2} - \frac{\sqrt{3}}{2} \frac{(0.2)^3}{6} + R_4.$$

For  $R_4$  we have the estimate

$$|R_4| \leq \frac{(0.2)^4}{4!} = \frac{16 \cdot 10^{-4}}{24} \leq 10^{-4},$$

which is within the required bounds of accuracy. You do not need to carry out the actual decimal expansion of the first four terms, which are given just as an illustration.

It is still true that the remainder term of the Taylor formula for  $\sin x$  approaches 0 when n becomes large, even when x is >1. For this we need:

THEOREM 4. Let c be any number. Then  $c^n/n!$  approaches 0 as n becomes very large.

*Proof.* Let  $n_0$  be an integer such that  $n_0 > 2c$ . Thus  $c < n_0/2$ , and  $c/n_0 < \frac{1}{2}$ . We write

$$\frac{c^n}{n!} = \frac{c \quad c \cdots c}{1 \quad 2 \cdots n_0} \frac{c \quad c \quad c}{(n_0 + 1)(n_0 + 2) \cdots n}$$
$$\leq \frac{c^{n_0}}{n_0!} \left(\frac{1}{2}\right) \cdots \left(\frac{1}{2}\right)$$
$$= \frac{c^{n_0}}{n_0!} \left(\frac{1}{2}\right)^{n-n_0}$$

As *n* becomes large,  $(1/2)^{n-n_0}$  becomes small and our fraction approaches 0. Take for instance c = 10. We write

$$\frac{10^n}{n!} = \frac{10\cdots 10}{1\cdot 2\cdots 20} \ \frac{(10)\cdots (10)}{(21)\cdots (n)} = \frac{10^{20}}{20!} \left(\frac{1}{2}\right)^{n-20}$$

and  $(1/2)^{n-20}$  approaches 0 as n becomes large.

From this we see immediately that the remainder in Taylor's formula for  $\sin x$  approaches 0.

#### Exercises

1. Write down the first 5 terms of the Taylor formula for  $\cos x$ .

2. Give an estimate for the remainder similar to that for  $\sin x$ .

3. Compute  $\cos(0.1)$  to 3 decimals.

4. Estimate the remainder  $R_3$  in the Taylor formula for  $\cos x$ , for the value x = 0.1.

5. Estimate the remainder  $R_4$  in the Taylor formula for  $\sin x$ , for the value x = 0.2.

6. Write down the terms of order  $\leq 4$  of the Taylor formula for tan x.

7. Estimate the remainder  $R_4$  in the Taylor formula for  $\tan x$ , for  $0 \le x \le 0.2$ .

8. Compute sine 31 degrees to 3 places.

Write down the terms of order  $\leq 4$  of the Taylor formula for the following functions:

9.  $\sin^2 x$  10.  $\cos^3 x$  11.  $\frac{1}{\cos x}$  12.  $\sin^3 x$ 

# §4. Exponential function

All derivatives of  $e^x$  are equal to  $e^x$  and  $e^0 = 1$ . Hence the Taylor series for  $e^x$  is

$$e^x = 1 + x + rac{x^2}{2!} + \dots + rac{x^{n-1}}{(n-1)!} + R_n.$$

Case 1. Look at values of x such that x < 0; then  $e^x < 1$  and of course  $e^x > 0$  for all x. Hence by Theorem 2, we get

$$|R_n| \leq \frac{|x|^n}{n!}$$
,

which is the same estimate we obtained for  $\sin x$ .

Case 2. Look at values of x > 0, and say  $x \leq b$  for some number b. Since  $e^x$  is strictly increasing, we know that

$$e^x \leq e^b$$

for any value of  $x \leq b$ . Hence by Theorem 5, we get

$$|R_n| \leq e^b \frac{b^n}{n!},$$

and we see again that this tends to 0 as n becomes large.

*Example 1.* Compute e to 3 decimals.

We have  $e = e^1$ . From Chapter VIII, §4 we know that e < 4. We estimate  $R_7$ :

$$|R_7| \leq e \frac{1}{7!} \leq 4 \frac{1}{5,040} \leq 10^{-3}.$$

Thus

$$e = 1 + 1 + \frac{1}{2} + \dots + \frac{1}{6!} + R_7$$
  
= 2.718....

Of course, the smaller x is, the fewer terms of the Taylor series do we need to approximate  $e^x$ .

*Example 2.* How many terms of the Taylor series do you need to compute  $e^{1/10}$  to an accuracy of  $10^{-3}$ ?

We certainly have  $e^{1/10} < 2$ . Thus

$$|R_3| \leq 2 \frac{(1/10)^3}{3!} < \frac{1}{2} 10^{-3}.$$

Hence we need just 3 terms (including the 0-th term).

#### Exercises

- 1. Write down the terms of order  $\leq 4$  of the Taylor formula  $e^{-x^2}$ .
- 2. Estimate the remainder  $R_3$  in the Taylor series for  $e^x$  for x = 1/2.
- 3. Estimate the remainder  $R_4$  for  $x = 10^{-2}$ .
- 4. Estimate the remainder  $R_3$  for  $x = 10^{-2}$ .
- 5. Write down the terms of order  $\leq 5$  of the Taylor series for  $e^{-x}$ .

6. Compute 1/e to 3 decimals, and show which remainder would give you an accuracy of  $10^{-3}$ .
LOGARITHM

7. Write down the first four terms of the Taylor series for the function  $f(x) = e^{-1/x^2}$  (when  $x \neq 0$ ), and f(0) = 0. Can you say something about the other terms in the Taylor series?

### §5. Logarithm

We leave it to you to derive the Taylor series for the logarithm. Take a = 1. We shall obtain an analogous result here by another method.

Let t be any number, and n an integer >1. Then

$$(1 - t + t^{2} - \dots + (-1)^{n-1}t^{n-1})(1 + t) = 1 + (-1)^{n-1}t^{n}.$$

This is trivially proved. We first multiply the big sum on the left by 1 and then by t, getting

$$1 - t + t^{2} - \dots + (-1)^{n-1} t^{n-1} + t - t^{2} + \dots - (-1)^{n-2} t^{n-1} + (-1)^{n-1} t^{n},$$

and add, to obtain what we want.

Suppose that  $t \neq -1$ . Then we can divide by 1 + t. Thus

$$\frac{1+(-1)^{n-1}t^n}{1+t} = \frac{1}{1+t} + \frac{(-1)^{n-1}t^n}{1+t}$$

This yields

$$\frac{1}{1+t} = 1 - t + t^2 - \dots + (-1)^{n-1} t^{n-1} + \frac{(-1)^n t^n}{1+t}$$

Consider the interval  $-1 < x \leq 1$ , and take the integral from 0 to x (in this interval). The integrals of the powers of t are well known to you. The integral

$$\int_0^x \frac{1}{1+t} dt = \log (1+x)$$

is computed by the substitution u = 1 + t, du = dt. Thus we get:

THEOREM 5. For  $-1 < x \leq 1$ , we have

$$\log (1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots + (-1)^{n-1} \frac{x^n}{n} + R_{n+1},$$

where the remainder  $R_{n+1}$  is the integral

$$(-1)^n \int_0^x \frac{t^n}{1+t} dt.$$

We shall now estimate the remainder term.

Case 1. Let a be a number with  $0 < a \leq 1$ , and consider the interval  $0 \leq x \leq a$ .

In that case,  $1 + t \ge 1$ . Thus

$$\frac{t^n}{1+t} \leq t^n,$$

and our integral is bounded by  $\int_0^x t^n dt$ . Thus in that case,

$$|R_{n+1}| \leq \int_0^x t^n dt \leq \int_0^a t^n dt \leq \frac{a^{n+1}}{n+1}.$$

(We perform the integration and use the fact that  $x \leq 1$ .) In particular, the remainder approaches 0 as n becomes large.

Case 2. Let a be a number with -1 < a < 0 and consider the interval  $a \leq x \leq 0$ .

In that case, we see that

$$0 < 1 + a \leq 1 + t$$

if t lies between x and 0.

$$-1$$
 a x 0

Thus

$$\left|\frac{t^n}{1+t}\right| = \frac{|t|^n}{1+t} \leq \frac{(-t)^n}{1+a}$$

To estimate the absolute value of the integral, we can invert the limits (we do this because  $x \leq 0$ ) and thus

$$\begin{aligned} |R_{n+1}| &\leq \int_x^0 \frac{(-t)^n}{1+a} dt \\ &\leq \frac{(-a)^{n+1}}{(n+1)(1+a)} = \frac{|a|^{n+1}}{(n+1)(1+a)} \,. \end{aligned}$$

Therefore the remainder also approaches 0 in that case.

As an exercise, compute log 1.1 to 3 decimals.

[XIV, §6]

# §6. The arctangent

We proceed as with the logarithm, except that we consider

$$\frac{1}{1+t^2} = 1 - t^2 + t^4 - \dots + (-1)^{n-1} t^{2n-2} + (-1)^n \frac{t^{2n}}{1+t^2}$$

After integration from 0 to any number x, we obtain:

THEOREM 6. The arctan has an expansion

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \dots + (-1)^{n-1} \frac{x^{2n-1}}{2n-1} + R_{2n},$$

where

$$R_{2n} = (-1)^n \int_0^x \frac{t^{2n}}{1+t^2} dt.$$

If b is a positive number such that  $|x| \leq b$  then

$$|R_{2n}| \leq \int_0^b t^{2n} dt \leq \frac{b^{2n+1}}{2n+1}$$

When  $-1 \leq x \leq 1$ , the remainder approaches 0 as n becomes large.

From our theorem, we get an expression for  $\pi/4$ :

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \cdots$$

from the Taylor formula for arctan 1. However, it takes many terms to get a good approximation to  $\pi/4$  by this expression. You will find a more clever approach in the exercises.

### Exercises

1. Prove the addition formula for the tangent:

$$\tan (x + y) = \frac{\tan x + \tan y}{1 - \tan x \tan y}$$

- 2. Prove that  $\pi/4 = \arctan \frac{1}{2} + \arctan \frac{1}{3}$ .
- 3. Verify that  $\pi = 3.14159...$ .
- 4. You will need even fewer terms if you prove that

$$\frac{\pi}{4} = 4 \arctan \frac{1}{5} - \arctan \frac{1}{239}.$$

Find the following limits as x approaches 0:

5. 
$$\frac{e^x - 1}{x}$$
 6.  $\frac{\sin(x^2)}{(\sin x)^2}$  7.  $\frac{\tan x}{\sin x}$ 

8. 
$$\frac{\arctan x}{x}$$
  
9.  $\frac{\log (1+x)}{x}$   
10.  $\frac{\log (1+2x)}{x}$   
11.  $\frac{e^x - (1+x)}{x^2}$   
12.  $\frac{\sin x - x}{x^2}$   
13.  $\frac{\cos x - 1}{x^2}$   
14.  $\frac{\log (1+x^2)}{\sin (x^2)}$   
15.  $\frac{\tan (x^2)}{(\sin x)^2}$   
16.  $\frac{\log (1+x^2)}{(\sin x)^2}$ 

## §7. The binomial expansion

Let us first consider a special case.

Let n be a positive integer, and consider the function

 $f(x) = (1+x)^n.$ 

We have no difficulty computing the derivatives:

$$f'(x) = n(1 + x)^{n-1}$$
  

$$f''(x) = n(n - 1)(1 + x)^{n-2}$$
  

$$\vdots$$
  

$$f^{(n)}(x) = n!$$
  

$$f^{(n+1)}(x) = 0.$$

Thus the Taylor formula has no remainder after the *n*-th term, and we get:

**THEOREM 7.** Let n be a positive integer. For any number x, we have

$$(1+x)^{n} = 1 + nx + \frac{n(n-1)}{2!}x^{2} + \frac{n(n-1)(n-2)}{3!}x^{3} + \dots + x^{n}$$

The coefficient of  $x^k$  is sometimes denoted by  $C_k^n$  or  $\binom{n}{k}$  and is called a *binomial coefficient*. Thus:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

If we want an expression for  $(a + b)^n$  with numbers a, b, then we let x = b/a. Then

$$\left(1+\frac{b}{a}\right)^n=\frac{1}{a^n}\left(a+b\right)^n,$$

and from this we conclude at once that

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}.$$

We shall now consider the function  $(1 + x)^s$  when s is not an integer. We define the *binomial coefficient* 

$$\binom{s}{k} = rac{s(s-1)\cdots(s-k+1)}{k!}$$

THEOREM 8. Let s be any number and let x lie in the interval -1 < x < 1. Then we have

$$(1+x)^{s} = 1 + sx + \frac{s(s-1)}{2!}x^{2} + \frac{s(s-1)(s-2)}{3!}x^{3} + \dots + {\binom{s}{k}}x^{k} + R_{k+1},$$

where the remainder term is equal to the usual integral. The remainder  $R_k$  approaches 0 as k becomes large.

*Proof.* The proof that the remainder approaches 0 is slightly more involved than our previous proofs, and will be omitted. To carry it out, one uses the form of the remainder given in Exercise 3 of §3 in case -1 < x < 0. There is of course no difficulty in verifying that when  $f(x) = (1 + x)^s$  then

$$\frac{f^{(k)}(0)}{k!} = \binom{s}{k}$$

and you should definitely do this as an exercise.

What we shall do, however, is discuss the remainder  $R_2$ . Let  $f(x) = (1 + x)^s$ . Assume for simplicity that s > 0. We have

$$f^{(2)}(x) = s(s-1)(1+x)^{s-2}$$
.

The Taylor formula gives

$$(1+x)^s = 1 + sx + R_2.$$

For small x, this means that 1 + sx should be a good approximation to the s power of 1 + x, if  $R_2$  can be proved to be small. This we can do

easily. We know that

$$|R_2| \leq \frac{M_2 |x|^2}{2!} \leq \frac{s(s-1)}{2} \frac{(1+|x|)^s}{(1-|x|)^2} |x|^2.$$

As x approaches 0, we see that  $|x|^2$  approaches 0 much more rapidly, and both 1 + |x| and 1 - |x| approach 1. Thus the expression on the right approaches 0.

This is in fact the most useful estimate in practice, when we have to compute some s power of 1 + x, with |x| small.

If additional accuracy is needed, one can carry out a similar discussion for  $R_3$ . We leave it to you.

### Exercises

- 1. Compute the cube root of 126 to 4 decimals.
- 2. Compute  $\sqrt{97}$  to 4 decimals.

3. Estimate  $R_2$  in the remainder of  $(1 + x)^{1/3}$  for x lying in the interval

$$-0.1 \leq x \leq 0.1.$$

4. Estimate the remainder  $R_2$  in the Taylor series of  $(1 + x)^{1/2}$  when x = 0.2.

5. Estimate the remainder  $R_2$  in the Taylor series for  $(1 + x)^{1/4}$  when x = 0.01.

6. Let b be a number > 0. Let f(x) be a function having n + 1 continuous derivatives in the interval -b < x < b. Assume that there are numbers  $a_0, \ldots, a_n$ , such that we can write

$$f(x) = a_0 + a_1x + \cdots + a_nx^n + g(x),$$

where g(x) is a function satisfying

$$|g(x)| \leq C|x|^{n+1}$$

for some number C > 0. Prove that

$$a_k = \frac{f^{(k)}(0)}{k!}$$

for k = 0, 1, ..., n. (In other words, the polynomial  $a_0 + \cdots + a_n x^n$  is the same as the polynomial in the Taylor formula.)

7. Prove that the expressions given in Theorems 5 and 6 for the log and arctangent satisfy the assumption stated in Exercise 6, provided 0 < b < 1.

# CHAPTER XV

# Series

Series are a natural continuation of our study of functions. In the preceding chapter we found how to approximate our elementary functions by polynomials, with a certain error term. Conversely, one can define arbitrary functions by giving a series for them. We shall see how in the sections below.

In practice, very few tests are used to determine convergence of series. Essentially, the comparison test is the most frequent. Furthermore, the most important series are those which converge absolutely. Thus we shall put greater emphasis on these.

### §1. Convergent series

Suppose that we are given a sequence of numbers

$$a_1, a_2, a_3, \ldots,$$

i.e. we are given a number  $a_n$  for each integer  $n \ge 1$ . (We picked the starting place to be 1, but we could have picked any integer.) We form the sums

$$s_n = a_1 + a_2 + \cdots + a_n.$$

It would be meaningless to form an infinite sum

$$a_1+a_2+a_3+\cdots$$

because we do not know how to add infinitely many numbers. However, if our sums  $s_n$  approach a limit, as n becomes large, then we say that the sum of our sequence *converges*, and we now define its *sum* to be that limit.

The symbols

$$\sum_{n=1}^{\infty} a_n$$

will be called a series. We shall say that the series converges if the sums  $s_n$  approach a limit as n becomes large. Otherwise, we say that it does not

211

converge, or *diverges*. If the series converges, we say that the value of the series is  $\tilde{}$ 

$$\sum_{n=1}^{\infty} a_n = \lim_{n \to \infty} s_n = \lim_{n \to \infty} (a_1 + \cdots + a_n).$$

The symbols  $\lim_{n\to\infty}$  are to be read: "The limit as n becomes large."

Example. Consider the sequence

$$1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots,$$

and let us form the sums

$$1 + \frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{2^n}$$

You probably know already that these sums approach a limit and that this limit is 2. To prove it, let  $r = \frac{1}{2}$ . Then

$$(1 + r + r^2 + \dots + r^n) = \frac{1 - r^{n+1}}{1 - r} = \frac{1}{1 - r} - \frac{r^{n+1}}{1 - r}$$

As n becomes large,  $r^{n+1}$  approaches 0, whence our sums approach

$$\frac{1}{1-\frac{1}{2}}=2.$$

Actually, the same argument works if we take for r any number such that

$$-1 < r < 1.$$

In that case,  $r^{n+1}$  approaches 0 as n becomes large, and consequently we can write

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r} \cdot$$

In view of the fact that the limit of a sum is the sum of the limits, we get:

THEOREM 1. Let  $\{a_n\}$  and  $\{b_n\}$  (n = 1, 2, ...) be two sequences and assume that the series

$$\sum_{n=1}^{\infty} a_n \quad and \quad \sum_{n=1}^{\infty} b_n$$

converge. Then  $\sum_{n=1}^{\infty} (a_n + b_n)$  also converges, and is equal to the sum of the two series.

In other words, series can be added term by term. Of course, they cannot be multiplied term by term!

We also observe that a similar theorem holds for the difference of two series.

If a series  $\sum a_n$  converges, then the numbers  $a_n$  must approach 0 as n becomes large. However, there are examples of sequences  $\{a_n\}$  for which the series does not converge. Consider for instance

$$1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}+\cdots.$$

We contend that the partial sums  $s_n$  become very large when n becomes large. To see this, we look at partial sums as follows:

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots + \frac{1}{8} + \frac{1}{9} + \dots + \frac{1}{16} + \dots$$

In each bunch of terms as indicated, we replace each term by that farthest to the right. This makes our sums smaller. Thus our expression is greater than or equal to

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8} + \dots + \frac{1}{8} + \frac{1}{16} + \dots + \frac{1}{16} + \dots$$
$$\geq 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots$$

and therefore becomes arbitrarily large when n becomes large.

### §2. Series with positive terms

Throughout this section, we shall assume that our numbers  $a_n$  are  $\geq 0$ . Then the partial sums

$$s_n = a_1 + \cdots + a_n$$

are increasing, i.e.

$$s_1 \leq s_2 \leq s_3 \leq \cdots \leq s_n \leq s_{n+1} \leq \cdots$$

If they are to approach a limit at all, they cannot become arbitrarily large. Thus in that case there is a number B such that

$$s_n \leq B$$

for all n. The collection of numbers  $\{s_n\}$  has therefore a least upper bound,

i.e. there is a smallest number S such that

 $s_n \leq S$ 

for all n. In that case, the partial sums  $s_n$  approach S as a limit. In other words, given any positive number  $\epsilon > 0$ , we have

$$S - \epsilon \leq s_n \leq S$$

for all n sufficiently large.

This simply expresses the fact that S is the least of all upper bounds for our collection of numbers  $s_n$ . We express this as a theorem.

THEOREM 2. Let  $\{a_n\}$  (n = 1, 2, ...) be a sequence of numbers  $\geq 0$  and let

$$s_n = a_1 + \cdots + a_n$$

If the sequence of numbers  $\{s_n\}$  is bounded, then it approaches a limit S, which is its least upper bound.

Theorem 1 gives us a very useful criterion to determine when a series with positive terms converges:

THEOREM 3. Let

$$\sum_{n=1}^{\infty} a_n \quad and \quad \sum_{n=1}^{\infty} b_n$$

be two series, with  $a_n \ge 0$  for all n and  $b_n \ge 0$  for all n. Assume that there is a number C > 0 such that

 $a_n \leq Cb_n$ 

for all n, and that  $\sum_{n=1}^{\infty} b_n$  converges. Then  $\sum_{n=1}^{\infty} a_n$  converges, and

$$\sum_{n=1}^{\infty} a_n \leq C \sum_{n=1}^{\infty} b_n.$$

Proof. We have

 $a_1 + \cdots + a_n \leq Cb_1 + \cdots + Cb_n = C(b_1 + \cdots + b_n) \leq C \sum_{n=1}^{\infty} b_n.$ 

This means that  $C \sum_{n=1}^{\infty} b_n$  is a bound for the partial sums  $a_1 + \cdots + a_n$ .

The least upper bound of these sums is therefore  $\leq C \sum_{n=1}^{\infty} b_n$ , thereby proving our theorem.

*Example.* Prove that the series  $\sum_{n=1}^{\infty} \frac{1}{n^2}$  converges.

Let us look at the series:

$$\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \dots + \frac{1}{8^2} + \dots + \frac{1}{16^2} + \dots + \dots$$

We look at the groups of terms as indicated. In each group of terms, if we decrease the denominator in each term, then we increase the fraction. We replace 3 by 2, then 4, 5, 6, 7 by 4, then we replace the numbers from 8 to 15 by 8, and so forth. Our partial sums are therefore less than or equal to

$$1 + \frac{1}{2^2} + \frac{1}{2^2} + \frac{1}{4^2} + \dots + \frac{1}{4^2} + \frac{1}{8^2} + \dots + \frac{1}{8^2} + \dots$$

and we note that 2 occurs twice, 4 occurs four times, 8 occurs eight times, and so forth. Hence the partial sums are less than or equal to

$$1 + \frac{2}{2^2} + \frac{4}{4^2} + \frac{8}{8^2} + \dots = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots$$

Thus our partial sums are less than or equal to those of the geometric series and are bounded. Hence our series converges.

#### EXERCISES

1. Show that the series  $\sum_{n=1}^{\infty} \frac{1}{n^3}$  converges. 2. Let  $\epsilon$  be a number >0. Show that the series  $\sum_{n=1}^{\infty} \frac{1}{n^{1+\epsilon}}$  converges.

Test the following series for convergence:

3. 
$$\sum_{n=1}^{\infty} \frac{1}{n^{1/2}}$$
4. 
$$\sum_{n=1}^{\infty} \frac{n^2}{n^4 + n}$$
5. 
$$\sum_{n=1}^{\infty} \frac{n}{n+1}$$
6. 
$$\sum_{n=1}^{\infty} \frac{n}{n+5}$$
7. 
$$\sum_{n=1}^{\infty} \frac{n^2}{n^3 + n + 2}$$
8. 
$$\sum_{n=1}^{\infty} \frac{|\sin n|}{n^2 + 1}$$
9. 
$$\sum_{n=1}^{\infty} \frac{|\cos n|}{n^2 + n}$$

SERIES

## §3. The integral test

You must already have felt that there is an analogy between the convergence of an improper integral and the convergence of a series. We shall now make this precise.

THEOREM 4. Let f be a function which is defined and positive for all  $x \ge 1$ , and decreasing. The series

$$\sum_{n=1}^{\infty} f(n)$$

converges if and only if the improper integral

$$\int_1^\infty f(x) \ dx$$

converges.

We visualize the situation in the following diagram.



Consider the partial sums

$$f(2) + \cdots + f(n)$$

and assume that our improper integral converges. The area under the curve between 1 and 2 is greater than or equal to the area of the rectangle whose height is f(2) and whose base is the interval between 1 and 2. This base has length 1. Thus

$$f(2) \leq \int_1^2 f(x) \, dx.$$

Again, since the function is decreasing, we have a similar estimate between 2 and 3:

$$f(3) \leq \int_2^3 f(x) \, dx.$$

We can continue up to n, and get

$$f(2) + f(3) + \cdots + f(n) \leq \int_{1}^{n} f(x) \, dx.$$

As n becomes large, we have assumed that the integral approaches a limit. This means that

$$f(2) + f(3) + \cdots + f(n) \leq \int_1^\infty f(x) \, dx.$$

Hence the partial sums are bounded, and hence by Theorem 2, they approach a limit. Therefore our series converges.

Conversely, assume that the partial sums

$$f(1) + \cdots + f(n)$$

approach a limit as n becomes large.

The area under the graph of f between 1 and n is less than or equal to the sum of the areas of the big rectangles. Thus

$$\int_{1}^{2} f(x) \, dx \leq f(1)(2 - 1) = f(1)$$
$$\int_{2}^{3} f(x) \, dx \leq f(2)(3 - 2) = f(2).$$

and

Proceeding stepwise, and taking the sum, we see that

$$\int_1^n f(x) \, dx \leq f(1) + \cdots + f(n-1).$$

The partial sums on the right are less than or equal to their limit. Call this limit L. Then for all positive integers n, we have

$$\int_1^n f(x) \, dx \, \leq \, L.$$

Given any number B, we can find an integer n such that  $B \leq n$ . Then

$$\int_1^B f(x) \ dx \ \leq \ \int_1^n f(x) \ dx \ \leq \ L.$$

Hence the integral from 1 to B approaches a limit as B becomes large, and this limit is less than or equal to L. This proves our theorem.

Example. Prove that the series  $\sum \frac{1}{n^2 + 1}$  converges. Let  $f(x) = \frac{1}{x^2 + 1}$ . Then f is decreasing, and  $\int_1^B f(x) dx = \arctan B - \arctan 1 = \arctan B - \frac{\pi}{4}$ .

#### SERIES

As B becomes large, arctan B approaches  $\pi/2$  and therefore has a limit. Hence the integral converges. So does the series, by the theorem.

### Exercises

1. Show that the following series diverges:  $\sum_{n=2}^{\infty} \frac{1}{n \log n}$ 2. Show that the following series converges:  $\sum_{n=1}^{\infty} \frac{n+1}{(n+2)n!}$ 

Test for convergence:

3. 
$$\sum_{n=1}^{\infty} ne^{-n^{2}}$$
4. 
$$\sum_{n=2}^{\infty} \frac{1}{n (\log n)^{3}}$$
5. 
$$\sum_{n=2}^{\infty} \frac{1}{n (\log n)^{2}}$$
6. 
$$\sum_{n=1}^{\infty} \frac{n!}{n^{n}}$$
7. 
$$\sum_{n=1}^{\infty} \frac{n}{e^{n}}$$
8. 
$$\sum_{n=1}^{\infty} \frac{n+1}{n^{3}+2}$$
9. 
$$\sum_{n=1}^{\infty} \frac{1}{n^{2}+n-1}$$
10. 
$$\sum_{n=1}^{\infty} \frac{n}{n^{3}-n+5}$$

### §4. Absolute convergence

We consider a series  $\sum_{n=1}^{\infty} a_n$  in which we do not assume that the terms  $a_n$  are  $\geq 0$ . We shall say that the series *converges absolutely* if the series

$$\sum_{n=1}^{\infty} |a_n|$$

formed with the absolute values of the terms  $a_n$  converges. This is now a series with terms  $\geq 0$ , to which we can apply the tests for convergence given in the two preceding sections. This is important, because we have:

THEOREM 5. Let  $\{a_n\}$  (n = 1, 2, ...) be a sequence, and assume that the series

$$\sum_{n=1}^{\infty} |a_n|$$

converges. Then so does the series  $\sum_{n=1}^{\infty} a_n$ .

[XV, §4]

*Proof.* Let  $a_n^+$  be equal to 0 if  $a_n < 0$  and equal to  $a_n$  itself if  $a_n \ge 0$ . Let  $a_n^-$  be equal to 0 if  $a_n > 0$  and equal to  $-a_n$  if  $a_n \le 0$ . Then both  $a_n^+$  and  $a_n^-$  are  $\ge 0$ . By assumption and comparison with  $\sum |a_n|$ , we see that each one of the series

$$\sum_{n=1}^{\infty} a_n^+ \quad \text{and} \quad \sum_{n=1}^{\infty} a_n^-$$

converges. Hence so does their difference

$$\sum_{n=1}^{\infty} a_n^+ - \sum_{n=1}^{\infty} a_n^-,$$

which is equal to

$$\sum_{n=1}^{\infty} (a_n^+ - a_n^-),$$

which is none other than  $\sum_{n=1}^{\infty} a_n$ . This proves our theorem.

We shall use one more test for convergence of a series which may have positive and negative terms.

THEOREM 6. Let 
$$\sum_{n=1}^{\infty} a_n$$
 be a series such that  $\lim_{n \to \infty} a_n = 0$ ,

such that the terms  $a_n$  are alternately positive and negative, and such that  $|a_{n+1}| \leq |a_n|$  for  $n \geq 1$ . Then the series is convergent.

*Proof.* Let us write the series in the form

$$b_1 - c_1 + b_2 - c_2 + b_3 - c_3 + \cdots,$$

with  $b_n, c_n \geq 0$ . Let

$$s_n = b_1 - c_1 + b_2 - c_2 + \dots + b_n$$
  
 $t_n = b_1 - c_1 + b_2 - c_2 + \dots + b_n - c_n$ 

Since the absolute values of the terms decrease, it follows that

 $s_1 \geq s_2 \geq s_3 \geq \cdots$ 

and

$$t_1 \leq t_2 \leq t_3 \leq \cdots$$

i.e. that the  $s_n$  are decreasing and the  $t_n$  are increasing. Let L be the

#### SERIES

greatest lower bound of the numbers  $s_n$ . In view of the fact that  $b_n$  and  $c_n$  approach 0 as n becomes large, it follows that both  $s_n$  and  $t_n$  approach L as a limit.

### Exercises

Determine whether the following series converge absolutely:

1.  $\sum \frac{\sin n}{n^3}$ 2.  $\sum \frac{1 + \cos \pi n}{n!}$ 3.  $\sum \frac{\sin \pi n + \cos 2\pi n}{n^{3/2}}$ 4.  $\sum \frac{(-1)^n}{n^2 + 1}$ 5.  $\sum \frac{(-1)^n \sin n + \cos 3n}{n^2 + n}$ 

Which of the following series converge and which do not?

6. 
$$\sum \frac{(-1)^n}{n}$$
 7.  $\sum \frac{(-1)^n}{n^2}$  8.  $\sum (-1)^n \frac{n}{n+1}$   
9.  $\sum \frac{(-1)^{n+1}}{\log (n+2)}$  10.  $\sum \frac{(-1)^n n}{n^2+1}$ 

11. For each number x, show that the series

$$\sum \frac{\sin n^2 x}{n^2}$$

converges absolutely. Let f be the function whose value at x is the above series. Show that f is continuous. Determine whether f is differentiable or not. (Remarkably enough, this does not seem to be known! Cf. J. P. Kahane, Bulletin of the American Mathematical Society, March 1964, p. 199.)

### §5. Power series

Perhaps the most important type of series are power series. Let x be any number and let  $\{a_n\}$  (n = 0, 1, ...) be a sequence of numbers. Then we can form the series

$$\sum_{n=0}^{\infty} a_n x^n.$$

The partial sums are

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

We have already met such sums when we discussed Taylor's formula.

**THEOREM 7.** Assume that there is a number r > 0 such that the series

$$\sum_{n=0}^{\infty} |a_n| r^n$$

converges. Then for all x such that |x| < r, the series

$$\sum_{n=0}^{\infty} a_n x^n$$

converges absolutely.

[XV, §5]

Proof. The absolute value of each term is

$$|a_n| |x|^n \leq |a_n|r^n.$$

Our assertion follows from the comparison Theorem 2.

The least upper bound of all numbers r for which we have the convergence stated in the theorem is called the *radius of convergence* of the series.

Theorem 5 allows us to define a function f; namely, for all numbers x such that |x| < r, we define

$$f(x) = \lim_{n \to \infty} (a_0 + a_1 x + \cdots + a_n x^n).$$

Our proofs that the remainder term in Taylor's formula approaches 0 for various functions now allow us to say that these functions are given by their Taylor series. Thus

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$
$$e^x = 1 + x + \frac{x^2}{2!} + \cdots$$

for all x. Furthermore,

$$\log\left(1+x\right) = -x + \frac{x^2}{2} - \cdots$$

is valid for -1 < x < 1.

(Here we saw that the series converges for x = 1, but it does not converge absolutely, cf. §1.)

However, we can now define functions at random by means of a power series, provided we know the power series converges absolutely, for |x| < r.

Before we give you an example, we must have some good way of finding out when a power series converges. The next theorem gives us such a way.

THEOREM 8. Let  $\{a_n\}$  (n = 1, 2, ...) be a sequence of numbers  $\geq 0$ and assume that there is a number s > 0 such that

$$a_n^{1/n} \leq s$$

for all but a finite number of integers n. Let r < 1/s, and r > 0. Then the series

$$\sum_{n=1}^{\infty} a_n r^n$$

converges.

*Proof.* For all but a finite number of integers n, we have

$$a_n^{1/n}r \leq rs < 1.$$

Hence

$$a_n r^n \leq (rs)^n$$
.

Since rs < 1, we can compare our series with the geometric series, and conclude that it converges.

On the other hand, we have:

THEOREM 9. Let  $\{a_n\}$  (n = 1, 2, ...) be a sequence of numbers  $\geq 0$ and assume that there is a number s > 0 such that  $a_n^{1/n} \geq s$  for infinitely many integers n. Let r > 1/s. Then

$$\sum_{n=1}^{\infty} a_n r^n$$

diverges.

*Proof.* In fact, we shall see that the terms do not even approach 0. We have rs > 1, whence  $ra_n^{1/n} \ge rs > 1$ . Thus  $a_n r^n > 1$  for infinitely many values of n. The series cannot converge.

Theorems 8 and 9 are particularly useful in case there is a limit

$$\lim_{n\to\infty} a_n^{1/n} \cdot$$

We see at once that if this limit is  $\neq 0$ , then the radius of convergence of the series

$$\sum_{n=1}^{\infty} a_n x^n$$

is precisely

$$\frac{1}{\lim_{n\to\infty}a_n^{1/n}}$$

Example. Prove that the series

$$\sum_{n=2}^{\infty} \frac{\log n}{n^2} x^n$$

converges absolutely for |x| < 1.

We put  $a_n = \log n/n^2$  for  $n \ge 2$ . We know that  $n^{1/n}$  approaches 1 as n becomes very large. Since  $(\log n)^{1/n} \le n^{1/n}$ , it follows that  $(\log n)^{1/n}$  also approaches 1. Furthermore,

$$(n^2)^{1/n} = (n^{1/n})^2$$

also approaches 1. Thus 1 is the radius of convergence.

222

*Remark.* If a power series  $\sum_{n=1}^{\infty} a_n x^n$  does not converge for any value of x except 0, then we agree to say that its radius of convergence is 0. If it converges for all x, then we say that the radius of convergence is *infinity*. This is the case if  $\lim_{n \to \infty} a_n^{1/n} = 0$  (trivial consequence of Theorem 8).

### Exercises

Find the radius of convergence of the following series:

1.	$\sum_{n=1}^{\infty} \frac{(2n)!}{(n!)^2} x^n$	$2. \sum_{n=1}^{\infty} \frac{n^n}{n!} x^n$
3.	$\sum_{n=1}^{\infty} \frac{\left(n!\right)^3}{(3n)!} x^n$	4. $\sum_{n=1}^{\infty} \frac{n^{5n}}{(2n)! n^{3n}} x^n$
5.	$\sum_{n=1}^{\infty} \frac{(3n)!}{(n!)^2} x^n$	$6. \sum_{n=1}^{\infty} \frac{\sin n\pi/2}{2^n} x^n$
7.	$\sum_{n=2}^{\infty} \frac{\log n}{2^n} x^n$	8. $\sum_{n=1}^{\infty} \frac{1+\cos 2\pi n}{3^n} x^n$
9.	$\sum_{n=2}^{\infty} nx^n$	$10. \sum_{n=1}^{\infty} \frac{\sin 2\pi n}{n!} x^n$
11.	$\sum_{n=2}^{\infty} n^2 x^n$	$12. \sum_{n=1}^{\infty} \frac{\cos n^2}{n^n} x^n$
13.	$\sum_{n=2}^{\infty} \frac{n}{\log n} x^n$	14. $\sum_{n=2}^{\infty} \frac{(-1)^n}{n!-1} x^n$
15.	$\sum_{n=1}^{\infty} \frac{n!}{n^n} x^n$	16. $\sum_{n=1}^{\infty} \frac{(-1)^n + 1}{n!} x^n$

(In these exercises, you may use Stirling's formula, namely

$$n! = \sqrt{2\pi n} n^n e^{-n} e^{\theta/12n}$$

with  $0 \leq \theta \leq 1$ , so that in the present applications we can replace n! by  $n^n e^{-n}$ .)

### §6. Differentiation and integration of power series

If we have a polynomial

$$a_0 + a_1 x + \cdots + a_n x^n$$

with numbers  $a_0, a_1, \ldots, a_n$  as coefficients, then we know how to find its

SERIES

derivative. It is  $a_1 + 2a_2x + \cdots + na_nx^{n-1}$ . We would like to say that the derivative of a series can be taken in the same way, and that the derivative converges whenever the series does.

THEOREM 10. Let r be a number >0 and let  $\sum a_n x^n$  be a series which converges absolutely for |x| < r. Then the series  $\sum na_n x^{n-1}$  also converges absolutely for |x| < r.

Proof. Our series is equal to

$$x\sum na_nx^n$$

and thus it will suffice to consider the series  $\sum na_n x^n$ . We may also assume that all  $a_n \ge 0$ . Let  $0 < r_1 < r$  and let c be a number, 0 < c < 1 such that  $r_1/c < r$ . Applying Theorem 9, we find that for all but a finite number of n, we have

$$a_n^{1/n} \leq c/r_1$$

(otherwise the series  $\sum a_n r^n$  would diverge), and hence for all but a finite number of n, we have

$$a_n^{1/n}r_1 \leq c.$$

On the other hand, by Corollary 3 of Theorem 10, Chapter VIII, §4 we know that  $n^{1/n}$  approaches 1, as *n* becomes large. Hence for all sufficiently large *n*,  $(na_n r_1^n)^{1/n}$  itself is less than or equal to a number <1 (possibly slightly bigger than *c*). Theorem 3 of §2 now proves that the series  $\sum na_n r_1^n$  converges. This is true for all  $r_1$  with  $0 < r_1 < r$ , and thus our theorem is proved.

A similar result holds for integration, but trivially. Indeed, if we have a series  $\sum_{n=1}^{\infty} a_n x^n$  which converges absolutely for |x| < r, then the series  $1 \sum_{n=1}^{\infty} a_n x^{n+1} \sum_{n=1}^{\infty} a_n x^n$ 

$$\frac{1}{x}\sum_{n=1}^{\infty}\frac{a_n}{n+1}x^{n+1} = \sum_{n=1}^{\infty}\frac{a_n}{n+1}x^n$$

has terms whose absolute value is smaller than in the original series.

The preceding results can be expressed by saying that an absolutely convergent power series can be integrated and differentiated term by term and still yield an absolutely convergent power series.

THEOREM 11. Let

$$f(x) = \sum_{n=1}^{\infty} a_n x^n$$

be a power series, which converges absolutely for |x| < r. Then f is differ-

entiable for |x| < r, and

$$f'(x) = \sum_{n=2}^{\infty} n a_n x^{n-1}.$$

*Proof.* Let 0 < b < r. Let  $\delta > 0$  be such that  $b + \delta < r$ . We consider values of x such that |x| < b and values of h such that  $|h| < \delta$ . We have

$$f(x+h) - f(x) = \sum_{n=1}^{\infty} a_n (x+h)^n - \sum_{n=1}^{\infty} a_n x^n = \sum_{n=1}^{\infty} a_n [(x+h)^n - x^n].$$

By the mean value theorem, there exists a number  $x_n$  between x and x + h such that the preceding expression is

$$f(x+h) - f(x) = \sum_{n=1}^{\infty} na_n x_n^{n-1} h$$

Therefore

$$\frac{f(x+h)-f(x)}{h} = \sum_{n=1}^{\infty} na_n x_n^{n-1}.$$

We have to show that the Newton quotient above approaches the value of the series obtained by taking the derivative term by term. We have

$$\frac{f(x+h)-f(x)}{h} - \sum_{n=1}^{\infty} na_n x^{n-1} = \sum_{n=1}^{\infty} na_n x_n^{n-1} - \sum_{n=1}^{\infty} na_n x^{n-1}$$
$$= \sum_{n=1}^{\infty} na_n [x_n^{n-1} - x^{n-1}].$$

Using the mean value theorem again, there exists  $y_n$  between  $x_n$  and x such that the preceding expression is

$$\frac{f(x+h)-f(x)}{h}-\sum_{n=1}^{\infty}na_nx^{n-1}=\sum_{n=2}^{\infty}(n-1)na_ny_n^{n-2}(x_n-x).$$

We have  $|y_n| \leq b + \delta < r$ , and  $|x_n - x| \leq |h|$ . Consequently,

$$\frac{\left|f(x+h)-f(x)\right|}{h} - \sum_{n=1}^{\infty} na_n x^{n-1} \le \sum_{n=2}^{\infty} (n-1)n|a_n||y_n|^{n-2}|h|$$
$$\le |h| \sum_{n=2}^{\infty} (n-1)n|a_n|(b+\delta)^{n-2}.$$

By Theorem 10 applied twice, we know that the series appearing on the right converges. It is equal to a fixed constant. As h approaches 0, it follows that the expression on the left also approaches 0. This proves that f is

SERIES

differentiable at x, and that its derivative is equal to  $\sum_{n=1}^{\infty} na_n x^{n-1}$ , for all x such that |x| < b. This is true for all b, 0 < b < r, and therefore concludes the proof of our theorem.

THEOREM 12. Let  $f(x) = \sum_{n=1}^{\infty} a_n x^n$  be a power series, which converges absolutely for |x| < r. Then the relation

$$\int f(x) \, dx = \sum_{n=1}^{\infty} \frac{a_n}{n+1} \, x^{n+1}$$

is valid in the interval |x| < r.

*Proof.* We know that the series integrated term by term converges absolutely in the interval. By the preceding theorem, its derivative term by term is the series for the derivative of the function, thereby proving our assertion.

*Example.* If we had never heard of the exponential function, we could define a function 2 3

$$f(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

Taking the derivative term by term, we see that

$$f'(x) = f(x).$$

Hence by what we know from Chapter VIII, §2, Exercise 7, we conclude that  $( \cdot ) = K x$ 

$$f(x) = Ke^{x}$$

for some constant K. Letting x = 0 shows that

$$1 = f(0) = K.$$

Thus K = 1 and  $f(x) = e^x$ .

Similarly, if we had never heard of sine and cosine, we could *define* functions

$$S(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots, \qquad C(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots.$$

Differentiating term by term shows that

$$S'(x) = C(x), \qquad C'(x) = -S(x).$$

Furthermore, S(0) = 0 and C(0) = 1. It can then be shown easily that any pair of functions S(x) and C(x) satisfying these properties must be the sine and cosine. This is actually carried out as an appendix to the second volume of this course, and there, we also show how to relate these functions to angles.

### Appendix 1

# $\epsilon$ and $\delta$

This appendix is intended to show how the notions of limits and the properties of limits can be explained and proved in terms of the notions and properties of numbers. We therefore assume the latter and carry out the proofs from there.

There remains the problem of showing how the real numbers can be defined in terms of the rational numbers, and the rational numbers in terms of integers. This takes too long to be included in this book.

Aside from the ordinary rules for addition, multiplication, subtraction, division (by non-zero numbers), ordering, positivity, and inequalities, there is one more basic property satisfied by the real numbers. This property is stated in §1. Our proofs then use only these properties.

### §1. Least upper bound

We meet again the problem of where to jump into the theory. It would be long and tedious to jump in too early. Hence we assume known the contents of Chapter I, §1 and §2. These involve the ordinary operations of addition and multiplication, and the notion of ordering, positivity, negative numbers, and inequalities. Those who are interested in seeing the logical development of these notions are referred to my book Undergraduate Analysis, Second Edition (Springer-Verlag, NY, 1997).

A collection of numbers will simply be called a *set* of numbers. This is shorter and is the usual terminology. If a set has at least one number in it, we say that it is *non-empty*.

Let S be a non-empty set of numbers. We shall say that S is bounded from above if there exists a number B such that

 $x \leq B$ 

for all x in our set S. We then call B an *upper bound* for S.

A least upper bound for S is an upper bound L such that any upper bound B for S satisfies the inequality  $B \ge L$ . If M is another least upper bound, then we have  $M \ge L$  and  $L \ge M$ , whence L = M. A least upper bound is unique. Similarly, we define the notions of bounded from below, and of greatest lower bound. (Do it yourself.)

The real numbers satisfy a property which is not satisfied by the set of rational numbers, namely:

FUNDAMENTAL PROPERTY. Every non-empty set S of numbers which is bounded from above has a least upper bound. Every non-empty set of numbers S which is bounded from below has a greatest lower bound.

PROPOSITION 1. Let a be a number such that

$$0 \leq a < \frac{1}{n}$$

for every positive integer n. Then a = 0. There is no number b such that  $b \ge n$  for every positive integer n.

*Proof.* Suppose there is a number  $a \neq 0$  such that a < 1/n for every positive integer n. Then n < 1/a for every positive integer n. Thus to prove both our assertions, it will suffice to prove the second.

Suppose there is a number b such that  $b \ge n$  for every positive integer n. Let S be the set of positive integers. Then S is bounded, and hence has a least upper bound. Let C be this least upper bound. No number strictly less than C can be an upper bound. Since 0 < 1, we have C < C + 1, whence C - 1 < C. Hence there is a positive integer n such that

$$C-1 < n.$$

This implies that C < n + 1 and n + 1 is a positive integer. We have contradicted our assumption that C is an upper bound for the set of positive integers, so no such upper bound can exist.

### §2. Limits

Let S be a set of numbers and let f be a function defined for all numbers in S. Let  $x_0$  be a number. We shall assume that S is arbitrarily close to  $x_0$ , i.e. given  $\epsilon > 0$  there exists an element x of S such that  $|x - x_0| < \epsilon$ . Let L be a number. We shall say that f(x) approaches the limit L as x approaches  $x_0$  if the following condition is satisfied:

Given a number  $\epsilon > 0$ , there exists a number  $\delta > 0$  such that for all x in S satisfying

$$|x-x_0| < \delta$$

we have

$$|f(x) - L| < \epsilon.$$

[APP. 1, §2]

LIMITS

If that is the case, then we write

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

We could also rephrase this as follows. We write

$$\lim_{h\to 0}f(x_0+h)=L$$

and say that the limit of  $f(x_0 + h)$  is L as h approaches 0 if the following condition is satisfied:

Given  $\epsilon > 0$ , there exists  $\delta > 0$  such that whenever h is a number with  $|h| < \delta$  and  $x_0 + h$  in S, then

$$|f(x_0+h)-L|<\epsilon.$$

We note that our definition of limit depends on the set S on which f is defined. Thus we should say "limit with respect to S". The next proposition shows that this is really unnecessary.

PROPOSITION 2. Let S be a set of numbers arbitrarily close to  $x_0$  and let S' be a subset of S, also arbitrarily close to  $x_0$ . Let f be a function defined on S. If

$$\lim_{h \to 0} f(x_0 + h) = L \quad \text{(with respect to S)}$$
$$\lim_{h \to 0} f(x_0 + h) = M \quad \text{(with respect to S')}$$

then L = M. In particular, the limit is unique.

*Proof.* Given  $\epsilon > 0$ , there exists  $\delta > 0$  such that whenever  $|h| < \delta$  we have

$$|f(x_0+h) - L| < \frac{\epsilon}{2}$$
$$|f(x_0+h) - M| < \frac{\epsilon}{2}$$

We have

$$|L - M| \leq |L - f(x_0 + h) + f(x_0 + h) - M| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence |L - M| is less than any  $\epsilon > 0$ , and by Proposition 1 of §1, we just have |L - M| = 0, whence L - M = 0 and L = M.

 $\epsilon$  and  $\delta$ 

*Remark.* Suppose that  $\lim_{h\to 0} f(x_0 + h) = L$ . Then there exists  $\delta > 0$  such that whenever  $|h| < \delta$  we have

$$|f(x_0 + h)| < |L| + 1.$$

Indeed, given  $\epsilon > 0$  there exists  $\delta$  such that whenever  $|h| < \delta$  we have

$$|f(x_0 + h) - L| < 1,$$

so that our assertion follows from standard properties of inequalities.

Also, note that we have trivially

$$\lim_{h\to 0} C = C$$

for any number C, viewed as a constant function on S. Indeed, given  $\epsilon > 0$ ,

$$|C-C| < \epsilon.$$

*Remark.* We mention a word about limits "when x becomes large". Let a be a number and f a function defined for all numbers  $x \ge a$ . Let L be a number. We shall say that f(x) approaches L as x becomes large, and we write

$$\lim_{x\to\infty}f(x) = L$$

if the following condition is satisfied. Given  $\epsilon > 0$  there exists a number A such that whenever x > A we have

$$|f(x) - L| < \epsilon.$$

In practice, instead of saying "when x becomes large", we sometimes say "when x approaches  $\infty$ ". We leave it to you to define the analogous notion "when x becomes large negative", or "x approaches  $-\infty$ ".

In the definition of  $\lim_{x\to\infty}$  we took a function f defined for  $x \ge a$ . If  $a_1 > a$ , and we restrict the function to all numbers  $\ge a_1$ , then the limit as x becomes very large will be the same.

Let us suppose that  $a \ge 1$ . Define a function g for values of x such that

$$0 < x \leq 1/a$$

by the rule

$$g(x) = f(1/x).$$

Then a second's thought will allow you to prove that

$$\lim_{x\to 0} g(x)$$

[APP. 1, §2]

LIMITS

exists if and only if

 $\lim_{x\to\infty}f(x)$ 

exists, and that they are equal.

Consequently all properties which we prove concerning limits as x approaches 0 (or a number) immediately give rise to similar properties concerning limits as x becomes very large. We leave their formulations to you.

THEOREM 1. Let S be a set of numbers, and f, g two functions defined for all numbers in S. Let  $x_0$  be a number. If

$$\lim_{h\to 0} f(x_0+h) = L$$

and

$$\lim_{h\to 0} g(x_0+h) = M,$$

then  $\lim_{h\to 0} (f+g)(x_0+h)$  exists and is equal to L+M.

*Proof.* Given  $\epsilon > 0$ , there exists  $\delta > 0$  such that, whenever  $|h| < \delta$  (and  $x_0 + h$  is in S), we have

$$|f(x_0+h) - L| < \frac{\epsilon}{2}$$
$$|g(x_0+h) - M| < \frac{\epsilon}{2}$$

We observe that

$$|f(x_0 + h) + g(x_0 + h) - L - M| \le$$
  
 $|f(x_0 + h) - L| + |g(x_0 + h) - M| \le \epsilon.$ 

This proves that L + M is the limit of  $(f + g)(x_0 + h)$  as h approaches 0.

In practice, we omit stating the fact that  $x_0 + h$  should be in S. It is to be understood as being so in every case.

THEOREM 2. Let S be a set of numbers, and f, g two functions defined for all numbers in S. Let  $x_0$  be a number. If

$$\lim_{h\to 0} f(x_0+h) = L$$

and

$$\lim_{h\to 0}g(x_0+h)=M,$$

then  $\lim_{h\to 0} f(x_0 + h)g(x_0 + h)$  exists and is equal to LM.

*Proof.* Given  $\epsilon > 0$  there exists  $\delta > 0$  such that, whenever  $|h| < \delta$ , we have

$$|f(x_0 + h) - L| < \frac{1}{2} \frac{\epsilon}{|M|} \quad \text{if} \quad M \neq 0$$
  
$$|f(x_0 + h) - L| < \frac{1}{2} \epsilon \quad \text{if} \quad M = 0$$
  
$$|g(x_0 + h) - M| < \frac{1}{2} \frac{\epsilon}{|L| + 1}$$
  
$$|f(x_0 + h)| < |L| + 1.$$

We have

$$\begin{aligned} |f(x_0 + h)g(x_0 + h) - LM| \\ &= |f(x_0 + h)g(x_0 + h) - f(x_0 + h)M + f(x_0 + h)M - LM| \\ &\leq |f(x_0 + h)g(x_0 + h) - f(x_0 + h)M| + |f(x_0 + h)M - LM| \\ &\leq |f(x_0 + h)| |g(x_0 + h) - M| + |f(x_0 + h) - L| |M| \\ &< (|L| + 1) \frac{1}{2} \frac{\epsilon}{|L| + 1} + \frac{1}{2} \epsilon \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &\leq \epsilon. \end{aligned}$$

COROLLARY 1. Let C be a number and let the assumptions be as in the theorem. Then C(C = 1, k) = C(C = 1, k)

$$\lim_{h\to 0} Cf(x_0+h) = CL.$$

Proof. Clear.

COROLLARY 2. Let the notation be as in Theorem 2. Then

$$\lim_{h\to 0} [f(x_0+h) - g(x_0+h)] = L - M.$$

Proof. Clear.

**THEOREM** 3. Let S be a set of numbers, and f a function defined for all numbers in S. Let  $x_0$  be a number. If

$$\lim_{h\to 0}f(x_0+h)=L$$

and  $L \neq 0$ , then the limit

$$\lim_{h\to 0}\frac{1}{f(x_0+h)}$$

exists and is equal to 1/L.

232

[APP. 1, §2]

LIMITS

*Proof.* Given  $\epsilon > 0$ , let  $\epsilon_1$  be the smaller of the numbers |L|/2 and  $\epsilon$ . There exists  $\delta > 0$  such that whenever  $|h| < \delta$  we have

$$|f(x_0+h)-L|<\epsilon_1$$

and also

$$|f(x_0+h)-L| < \frac{\epsilon|L|}{2}$$

From the first inequality, we get

$$|f(x_0 + h)| > |L| - \epsilon_1 \ge |L| - \frac{|L|}{2} = \frac{|L|}{2}$$

In particular,  $f(x_0 + h) \neq 0$  when  $|h| < \delta$ . For such h we get

$$\begin{aligned} \left| \frac{1}{f(x_0 + h)} - \frac{1}{L} \right| &= \frac{|L - f(x_0 + h)|}{|f(x_0 + h)L|} \\ &\leq \frac{2}{|L|} |L - f(x_0 + h)| \\ &< \frac{2}{|L|} \frac{\epsilon |L|}{2} = \epsilon. \end{aligned}$$

COROLLARY. Let the hypotheses be as in Theorem 2, and assume that  $L \neq 0$ . Then

$$\lim_{h \to 0} \frac{g(x_0 + h)}{f(x_0 + h)}$$

exists and is equal to M/L.

*Proof.* Use Theorem 2 and Theorem 3.

THEOREM 4. Let S be a set of numbers, and f a function on S. Let  $x_0$  be a number. Let g be a function on S such that  $g(x) \leq f(x)$  for all x in S. Assume that

 $\lim_{h\to 0} f(x_0+h) = L \quad \text{and} \quad \lim_{h\to 0} g(x_0+h) = M.$ 

Then  $M \leq L$ .

*Proof.* Let  $\varphi(x) = f(x) - g(x)$ . Then  $\varphi(x) \ge 0$  for all x in S. Also,

$$\lim_{h\to 0}\varphi(x_0+h)=L-M$$

by Corollary 2 of Theorem 2. Let K be this limit. We must show  $K \ge 0$ . Suppose K < 0. Then -K > 0 and |K| = -K. Given  $\epsilon > 0$  there exists  $\delta > 0$  such that whenever  $|h| < \delta$  we have

$$|\varphi(x_0+h)-K| < \epsilon,$$
  
 $|\varphi(x_0+h)|-K < \epsilon.$ 

whence

Since  $\varphi(x_0 + h) \ge 0$ , we get  $-K < \epsilon$  for all  $\epsilon > 0$ . In particular, for all positive integers n we get -K < 1/n. But -K > 0. This contradicts Proposition 1 of §1.

THEOREM 5. Let the notation be as in Theorem 4 and assume that M = L. Let  $\psi$  be a function on S such that

$$g(x) \leq \psi(x) \leq f(x)$$
$$\lim_{h \to 0} \psi(x_0 + h)$$

for all x in S. Then

exists and is equal to L (or M).

*Proof.* Given  $\epsilon > 0$  there exists  $\delta > 0$  such that whenever  $|h| < \delta$  we have

$$|g(x_0 + h) - L| < \frac{\epsilon}{4}$$
$$|f(x_0 + h) - L| < \frac{\epsilon}{4}$$

We also have

$$\begin{aligned} |f(x_0 + h) - \psi(x_0 + h)| &\leq |f(x_0 + h) - g(x_0 + h)| \\ &\leq |f(x_0 + h) - L + L - g(x_0 + h)| \\ &\leq |f(x_0 + h) - L| + |L - g(x_0 + h)| \\ &< \frac{\epsilon}{2} \cdot \end{aligned}$$

 $\mathbf{But}$ 

$$|L - \psi(x_0 + h)| \leq |L - f(x_0 + h)| + |f(x_0 + h) - \psi(x_0 + h)|$$
  
$$< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

This completely proves all the statements about limits we made in Chapter III.

## §3. Points of accumulation

A sequence is a function defined on a set of integers  $\geq 0$ . Usually, this set consists of all positive integers. In that case, a sequence amounts to

234

[APP. 1, §3]

giving numbers

 $a_1, a_2, a_3, \ldots$ 

for each positive integer, and we denote the sequence by  $\{a_n\}$  (n = 1, 2, ...).

If the set consists of all integers  $\geq 0$ , then we denote the sequence by  $\{a_n\}$  (n = 0, 1, 2, ...).

Let  $\{a_n\}$  (n = 1, 2, ...) be a sequence. Let C be a number. We say that C is a *point of accumulation* of the sequence if given  $\epsilon > 0$  there exist infinitely many integers n such that

$$|a_n - C| < \epsilon.$$

Let  $\{a_n\}$  (n = 1, 2, ...) be a sequence, and L a number. We shall say that L is a *limit of the sequence* if given  $\epsilon > 0$  there exists an integer N such that for all n > N we have

$$|a_n-L|<\epsilon.$$

The limit is then unique (same type of proof as we had for limits of functions).

We shall say that the sequence  $\{a_n\}$  (n = 1, 2, ...) is *increasing* if  $a_n \leq a_{n+1}$  for all positive integers n.

THEOREM 6. Let  $\{a_n\}$  (n = 1, 2, ...) be an increasing sequence, and assume that it is bounded from above. Then the least upper bound L is a limit of the sequence.

*Proof.* Given  $\epsilon > 0$  the number  $L - (\epsilon/2)$  is not an upper bound for the sequence. Hence there exists some number  $a_N$  such that  $L - (\epsilon/2) \leq a_N$ . This inequality is also satisfied for all n > N, since the sequence is increasing. But

$$a_n \leq L$$

because L is an upper bound. Thus

$$|L-a_n|=L-a_n\leq rac{\epsilon}{2}<\epsilon$$

for all n > N, thereby proving our assertion.

COROLLARY. Let  $\{a_n\}$  (n = 1, 2, ...) be a sequence, and let A, B be two numbers such that  $A \leq a_n \leq B$  for all positive integers n. Then there exists a point of accumulation C of the sequence with C between A and B.

*Proof.* For each integer n we let  $b_n$  be the greatest lower bound of the set of numbers  $\{a_n, a_{n+1}, a_{n+2}, \ldots\}$ . Then  $b_n \leq b_{n+1} \leq \cdots$ , i.e.  $\{b_n\}$   $(n = 1, 2, \ldots)$  is an increasing sequence. Let L be its limit, as in Theorem

6. We leave it to you as an exercise to prove that this limit is a point of accumulation.

One can reduce the notion of limit of a sequence to that of limits defined previously.

Let S be the set of numbers

1, 
$$\frac{1}{2}$$
,  $\frac{1}{3}$ , ...,  $\frac{1}{n}$ , ...,

i.e. the set of numbers which can be written as 1/n, where n is a positive integer.

If  $\{a_n\}$  (n = 1, 2, ...) is a sequence, we let f be a function defined on S by the rule

$$f(1/n) = a_n$$

Then you will verify immediately that

$$\lim_{n \to \infty} a_n$$
$$\lim_{h \to 0} f(h)$$

exists if and only if

 $h \rightarrow 0$ 

exists, and in that case the two limits are equal. We say that a sequence  $\{a_n\}$  approaches a number L when n becomes large if  $L = \lim a_n$ .

Thus properties concerning limits in the sense of §2 immediately give rise to properties concerning limits of sequences (for instance limits of sums, products, quotients). We leave their translations to you.

### §4. Continuous functions

Let f be a function defined on a set of numbers S. Let  $x_0$  be a number in S. Then S is arbitrarily close to  $x_0$ . We say that f is continuous at  $x_0$  if

$$\lim_{h\to 0}f(x_0+h)=f(x_0).$$

Note that there may be two numbers a, b with  $a < x_0 < b$  such that  $x_0$ is the only point which is in the interval and lies also in S. (In this case, one could say that  $x_0$  is an *isolated* point of S.)

It follows at once from our definition that if  $\{a_n\}$  (n = 1, 2, ...) is a sequence of numbers in S such that

$$\lim_{n\to\infty}a_n=x_0$$

[APP. 1, §4]

then

 $\lim_{n\to\infty}f(a_n)=f(x_0).$ 

It is immediate that the sum, product, quotient of continuous functions are again continuous. (In the quotient, we have to assume that  $f(x_0) \neq 0$ , of course.) Every constant function is continuous. The function f(x) = xis continuous for all x. This is trivially verified. From the quotient theorem, we see that the function f(x) = 1/x (defined for  $x \neq 0$ ) is continuous.

If g is a continuous function, and f is a function which is defined at all values of g, then the composite function  $f \circ g$  is continuous. This is an easy exercise, which we leave to you.

THEOREM 7. Let f be a continuous function on a closed interval  $a \leq x \leq b$ . Then there exists a point c in the interval such that f(c) is a maximum, and there exists a point d in the interval such that f(d) is a minimum.

*Proof.* We shall first prove that f is bounded, i.e. that there exists a number M such that  $f(x) \leq M$  for all x in the interval.

If f is not bounded, then for every positive integer n we can find a number  $x_n$  in the interval such that  $f(x_n) > n$ . The sequence of such  $x_n$  has a point of accumulation C in the interval. We have

$$|f(x_n) - f(C)| \ge |f(x_n)| - |f(C)|$$
$$\ge n - f(C).$$

Given  $\epsilon > 0$ , there exists a  $\delta > 0$  such that, whenever

$$|x_n - C| < \delta$$

we have  $|f(x_n) - f(C)| < \epsilon$ . This has to happen for infinitely many n, since C is an accumulation point. Our statements are contradictory, and we therefore conclude that the function is bounded (from above).

Let  $\beta$  be the least upper bound of the set of values f(x) for all x in the interval. Then given a positive integer n, we can find a number  $z_n$  in the interval such that

$$|f(z_n) - \beta| < \frac{1}{n}.$$

Let c be a point of accumulation of the sequence of numbers  $\{z_n\}$ (n = 1, 2, ...). Then  $f(c) \leq \beta$ . We contend that  $f(c) = \beta$  (this will prove our theorem).

Given  $\epsilon > 0$ , there exists  $\delta > 0$  such that whenever  $|z_n - c| < \delta$  we have

$$|f(z_n) - f(c)| < \epsilon.$$

This happens for infinitely many n, since c is a point of accumulation of the sequence  $\{z_n\}$ . But

$$\begin{aligned} |f(c) - \beta| &\leq |f(c) - f(z_n)| + |f(z_n) - \beta| \\ &< \epsilon + \frac{1}{n} \cdot \end{aligned}$$

This is true for every  $\epsilon$  and infinitely many positive integers *n*. Hence  $|f(c) - \beta| = 0$  and  $f(c) = \beta$ .

The proof for the minimum is similar and will be left as an exercise.

THEOREM 8. Let f be a continuous function on a closed interval  $a \leq x \leq b$ . Let  $\alpha = f(a)$  and  $\beta = f(b)$ . Let  $\gamma$  be a number such that  $\alpha < \gamma < \beta$ . Then there exists a number c between a and b such that  $f(c) = \gamma$ .

*Proof.* Let S be the set of numbers x in the interval such that  $f(z) \leq \gamma$  for all numbers z in the interval, with  $z \leq x$ . Then S is not empty because a is in it, and b is an upper bound for S. Let c be its least upper bound. Then c is in our interval. We contend that  $f(c) = \gamma$ .

Given a positive integer n, there exists a number  $a_n$  such that

$$|a_n - c| < \frac{1}{n}$$

 $a \leq a_n \leq c$ 

and  $f(a_n) \leq \gamma$ , because c is a least upper bound of our set S. Since  $\lim a_n = c$  we get  $\lim f(a_n) = f(c)$ . Hence  $f(c) \leq \gamma$ .

Suppose  $f(c) < \gamma$ . Then  $c \neq b$ . We have by continuity

$$\lim_{h\to 0}f(c+h)=f(c).$$

Let  $\epsilon > 0$  be such that  $f(c) + \epsilon < \gamma$ . There exists  $\delta > 0$  such that whenever  $|h| < \delta$  we have

$$|f(c+h) - f(c)| < \epsilon,$$

whence in particular

$$f(c+h) < f(c) + \epsilon < \gamma$$
.

This means that c cannot be an upper bound for our set S (which would contain for instance c + h for h > 0 and  $h < \delta$ ).

# Appendix 2

# **Physics and Mathematics**

Mathematics consists in discovering and describing certain objects and structures. It is essentially impossible to give an all-encompassing description of these. Hence, instead of such a definition, we simply state that the objects of study of mathematics as we know it are those which you will find described in the mathematical journals of the past two centuries, and leave it at that. There are many reasons for studying these objects, among which are aesthetic reasons (some people like them), and practical reasons (some mathematics can be applied).

Physics, on the other hand, consists in describing the empirical world by means of mathematical structures. The empirical world is the world with which we come into contact through our senses, through experiments, measurements, etc. What makes a good physicist is the ability to choose, among many mathematical structures and objects, the ones which can be used to describe the empirical world. I should of course immediately qualify the above assertion in two ways: First, the description of physical situations by mathematical structures can only be done within the degree of accuracy provided by the experimental apparatus. Second, the description should satisfy certain aesthetic criteria (simplicity, elegance). After all, a complete listing of all results of all experiments performed is a description of the physical world, but is quite a distinct thing from giving at one single stroke a general principle which will account simultaneously for the results of all these experiments.

For psychological reasons, it is impossible (for most people) to learn certain mathematical theories without seeing first a geometric or physical interpretation. Hence in this book, before introducing a mathematical notion, we frequently introduce one of its geometric or physical interpretations. These two, however, should not be confused. Thus we might make two columns, as shown on the following page.

As far as the logical development of our course is concerned, we could omit the second column entirely. The second column is used, however, for many purposes: To motivate the first column (because our brain is made up in such a way that to understand something in the first column, it needs the second). To provide applications for the first column, other than pure aesthetic satisfaction (granting that you like the subject).

Mathematics	Physics and geometry
numbers	points on a line
derivative	slope of a line rate of change
$\frac{df}{dx} = Kf(x)$	exponential decay
integral	årea work moments

Nevertheless, it is important to keep in mind that the derivative (as the limit of

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

and the integral (as a unique number between upper and lower sums), are not to be confused with a slope or an area respectively. It is simply our mind which interprets the mathematical notion in physical or geometric terms. Besides, we frequently assign several such interpretations to the same mathematical notion (viz. the integral being interpreted as an area, or as the work done by a force).

And by the way, the above remarks which are about physics and mathematics belong neither to physics nor to mathematics. They belong to philosophy.
### **ANSWERS TO EXERCISES**

I am much indebted to D. Levine for the answers to the exercises.

Chapter I, §2 **1.** -3 < x < 3 **2.**  $-1 \le x \le 0$  **3.**  $-\sqrt{3} \le x \le -1$  or  $1 \le x \le \sqrt{3}$ 4. x > 2 5. -1 < x < 2 6. x < -1 or x > 1 7. -5 < x < 58.  $-1 \leq x \leq 0$  9.  $x \geq 1$  or x = 0 10.  $x \leq -10$  or x = 511.  $x \leq -10$  or x = 5 12.  $x \geq 1$  or  $x = -\frac{1}{2}$  13. x < -4Chapter I, §3 1.  $\frac{4}{3}, -\frac{3}{2}$  2.  $\frac{1}{(2x+1)}$  3. 0, 2, 108 4.  $2z - z^2, 2w - w^2$ 5.  $x \neq \sqrt{2}$  or  $-\sqrt{2}$ .  $f(5) = \frac{1}{23}$  6. All x. f(27) = 37. (a) 1 (b) 1 (c) -1 (d) -1 8. (a) 1 (b) 4 (c) 0 (d) 0 9. (a) -2 (b) -6 (c)  $x^2 + 4x - 2$  10.  $x \ge 0, 2$ Chapter I, §4 1. 8 and 9 2.  $\frac{1}{5}$  and -1 3.  $\frac{1}{16}$  and 2 4.  $\frac{1}{9}$  and  $2^{1/3}$  5.  $\frac{1}{16}$  and  $\frac{1}{2}$ 6. 9 and 8 7.  $-\frac{1}{3}$  and -1 8.  $\frac{1}{4}$  and  $\frac{1}{4}$  9. 1 and  $-\frac{1}{4}$  10.  $-\frac{1}{512}$  and  $\frac{1}{3}$ Chapter II, §1 **3.** x negative, y positive **4.** x negative, y negative Chapter II, §3 5.  $y = -\frac{8}{3}x - \frac{5}{3}$  6.  $y = -\frac{3}{2}x + 5$  7.  $x = \sqrt{2}$ 8.  $y = \frac{9}{\sqrt{3}+3}x + 4 - \frac{9\sqrt{3}}{\sqrt{3}+3}$  9. y = 4x - 3 10. y = -2x + 211.  $y = -\frac{1}{2}x + 3 + \frac{\sqrt{2}}{2}$  12.  $y = \sqrt{3}x + 5 + \sqrt{3}$ Chapter II, §4 1.  $\sqrt{97}$  2.  $\sqrt{2}$  3.  $\sqrt{52}$  4.  $\sqrt{13}$  5.  $\sqrt{5/4}$  6. (4, -3) 7. 5 and 5 Chapter II, §7 5.  $(x-2)^2 + (y+1)^2 = 25$  6.  $x^2 + (y-1)^2 = 9$ 7.  $(x + 1)^2 + y^2 = 3$ 10.  $y + 4 = (x - 1)^2$ 11.  $(x + 1)^2 + (y - 2)^2 = 2$ 13.  $(x + 1)^2 + (y - 2)^2 = 2$ 12.  $(x-2)^2 + (y-1)^2 = 2$  13.  $x + \frac{25}{8} = 2(y+\frac{1}{4})^2$ 14.  $x - 1 = (y + 2)^2$ Chapter III, §1 **1.** 4 **2.** -2 **3.** 2 **4.**  $\frac{3}{4}$  **5.**  $-\frac{1}{4}$  **6.** 0

Chapter III, §2

	Tangent line at $x = 2$	
1. 2x	y = 4x - 3	
<b>2.</b> $3x^2$	y = 12x - 16	
3. $6x^2$	y = 24x - 32	
4. 6x	y = 12x - 12	
5. $2x$	y = 4x - 9	
<b>6.</b> $4x + 1$	y = 9x - 8	
7. $4x - 3$	y = 5x - 8	
8. $\frac{3x^2}{2} + 2$	y = 8x - 8	
9. $-\frac{1}{(x+1)^2}$	$y = -\frac{1}{9}x + \frac{5}{9}$	
10. $-\frac{2}{(x+1)^2}$	$y = -\frac{2}{9}x + \frac{10}{9}$	

11. The slopes are 4, 12, 24, 12, 4, 9, 5, 8,  $-\frac{1}{9}$ ,  $-\frac{2}{9}$ . The tangent lines at the point x = 2 are indicated near the corresponding problem.

12. -1. Left derivative -1. No right derivative.

13. No. f'(x) exists for all other values of x.

14. Left derivative 0. No right derivative. f'(x) = 0 if x < 0 and f'(x) = 1 if x > 1.

15. 0, 0, 0

Chapter III, §3

**1.** 
$$4x + 3$$
 2.  $-\frac{2}{(2x+1)^2}$  **3.**  $\frac{1}{(x+1)^2}$  **4.**  $2x + 1$  **5.**  $-\frac{1}{(2x-1)^2}$   
**6.**  $9x^2$  **7.**  $4x^3$  **8.**  $5x^4$  **9.**  $6x^2$  **10.**  $\frac{3x^2}{2} + 1$ 

Chapter III, §4

1. 
$$x^4 + 4x^3h + 6x^2h^2 + 4xh^3 + h^4$$
 2.  $4x^3$   
3. (a)  $\frac{2}{3}x^{-1/3}$  (b)  $-\frac{3}{2}x^{-5/2}$  (c)  $\frac{7}{6}x^{1/6}$  4.  $y = 9x - 8$   
5.  $y = \frac{1}{3}x + \frac{4}{3}$ , slope  $\frac{1}{3}$  6.  $y = \frac{-3}{29}x + \frac{7}{32}$ , slope  $-3/2^9$   
7.  $y = \frac{1}{2\sqrt{3}}x + \frac{\sqrt{3}}{2}$ , slope  $\frac{1}{2\sqrt{3}}$   
8. (a)  $\frac{1}{4}5^{-3/4}$  (b)  $-\frac{1}{4}7^{-5/4}$  (c)  $\sqrt{2}(10\sqrt{2}-1)$  (d)  $\pi7^{\pi-1}$ 

Chapter III, §5

1.  $\frac{2}{3}x^{-2/3}$  2.  $55x^{10}$  3.  $-\frac{3}{8}x^{-7/4}$  4.  $21x^2 + 8x$  5.  $-25x^{-2} + 6x^{-1/2}$ 6.  $\frac{6}{5}x - 16x^7$  7.  $(x^3 + x) + (3x^2 + 1)(x - 1)$ 8.  $(2x^2 - 1)4x^3 + 4x(x^4 + 1)$  9.  $(x + 1)(2x + \frac{15}{2}x^{1/2}) + (x^2 + 5x^{3/2})$ 10.  $(2x - 5)(12x^3 + 5) + 2(3x^4 + 5x + 2)$ 11.  $(x^{-2/3} + x^2)\left(3x^2 - \frac{1}{x^2}\right) + (-\frac{2}{3}x^{-5/3} + 2x)\left(x^3 + \frac{1}{x}\right)$ 

12.  $(2x+3)(-2x^{-3}-x^{-2})+2(x^{-2}+x^{-1})$ 13.  $\frac{9}{(x+5)^2}$  14.  $\frac{(-2x^2+2)}{(x^2+3x+1)^2}$ 15.  $\frac{(t+1)(t-1)(2t+2) - (t^2+2t-1)2t}{(t^2-1)^2}$ 16.  $\frac{(t^2+t-1)(-5/4)t^{-9/4}-t^{-5/4}(2t+1)}{(t^2+t-1)^2}$ 17.  $\frac{5}{49}$ ,  $y = \frac{5}{49}t + \frac{4}{49}$  18.  $\frac{1}{2}$ ,  $y = \frac{1}{2}t$ Chapter III, §6 **1.**  $8(x+1)^7$  **2.**  $\frac{1}{2}(2x-5)^{-1/2} \cdot 2$  **3.**  $3(\sin x)^2 \cos x$  **4.**  $5(\log x)^4 \frac{1}{2}$ 5.  $(\cos 2x)^2$  6.  $\frac{1}{x^2+1}^2 2x$  7.  $e^{\cos x}(-\sin x)$  8.  $\frac{1}{e^x+\sin x}(e^x+\cos x)$ 9.  $\cos\left(\log x + \frac{1}{r}\right)\left(\frac{1}{r} - \frac{1}{r^2}\right)$  10.  $\frac{\sin 2x - (x+1)(\cos 2x)2}{(\sin 2x)^2}$ **11.**  $3(2x^2+3)^2(4x)$  **12.**  $-\sin(\sin 5x)(\cos 5x)5$  **13.**  $\frac{1}{\cos 2x}(-\sin 2x)2$ 14.  $\cos((2x+5)^2)(2x+5)(2x+5)(x+1))(-\sin(x+1))$ **16.**  $(\cos e^{x})e^{x}$  **17.**  $-\frac{1}{(3x-1)^{8}}[4(3x-1)^{3}] \cdot 3$  **18.**  $-\frac{1}{(4x)^{6}} \cdot 3(4x)^{2} \cdot 4$ 19.  $-\frac{1}{(\sin 2x)^4} 2(\sin 2x)(\cos 2x) \cdot 2$  20.  $-\frac{1}{(\cos 2x)^4} 2(\cos 2x)(-\sin 2x)2$ Chapter III, §7 **1.** 18x **2.**  $5(x^2+1)^4 \cdot 2 + 20(x^2+1)^3 4x^2$  **3.** 0 **4.** 5040 **5.** 0 **6.** 6 7.  $-\frac{1}{(\sin 3x)^2}\cos 3x \cdot 3$  8.  $-\sin^2 x + \cos^2 x$  9.  $(x^2 + 1)e^x + 2xe^x$ 10.  $(x^3 + 2x) \cos 3x \cdot 3 + (3x^2 + 2) \sin 3x$ 11.  $-\frac{1}{(\sin x + \cos x)^2}(\cos x - \sin x)$  12.  $\frac{2e^x \cos 2x - (\sin 2x)e^x}{e^{2x}}$ 13.  $\frac{(x^2+3)/x - (\log x)(2x)}{(x^2+3)^2}$  14.  $\frac{\cos 2x - (x+1)(-\sin 2x) \cdot 2}{\cos^2 2x}$ 15.  $(2x-3)(e^x+1)+2(e^x+x)$  16.  $(x^3-1)(e^{3x}\cdot 3+5)+3x^2(e^{3x}+5x)$ 17.  $\frac{(x-1)3x^2-(x^3+1)}{(x-1)^2}$  18.  $\frac{(2x+3)2x-(x^2-1)2}{(2x+3)^2}$ **19.**  $2(x^{4/3} - e^x) + (\frac{4}{3}x^{1/3} - e^x)(2x+1)$ **20.**  $(\sin 3x)\frac{1}{4}x^{-3/4} + 3(\cos 3x)(x^{1/4} - 1)$  **21.**  $[\cos (x^2 + 5x)](2x+5)$ **22.**  $e^{3x^2+8}(6x)$  **23.**  $\frac{-1}{[\log (x^4+1)]^2} \cdot \frac{1}{x^4+1} \cdot 4x^3$ 24.  $\frac{-1}{[\log (x^{1/2} + 2x)]^2} \frac{1}{(x^{1/2} + 2x)} (\frac{1}{2}x^{-1/2} + 2)$  25.  $\frac{2e^x - 2xe^x}{e^{2x}}$ 27. 0 **28.** 240 in<sup>3</sup>/sec **29.**  $36\pi$  in<sup>3</sup>/sec **30.**  $2\pi r$ ,  $\frac{\pi d}{2}$ ,  $\frac{c}{2\pi}$  **31.** -3/16 units/sec

Chapter IV, §1

**1.** 
$$1/\sqrt{2}$$
 **2.**  $\sqrt{3}/2$  **3.**  $\frac{1}{2\sqrt{2}}(\sqrt{3}+1)$  **4.**  $\frac{1}{2}$  **5.**  $-\sqrt{3}/2$  **6.**  $-\frac{1}{2}$   
**7.**  $\sqrt{3}/2$  **8.**  $-\frac{1}{\sqrt{2}}$  **9.** 1 **10.**  $\sqrt{3}$  **11.**  $-1$  **12.**  $-1$ 

18.  $\pi - a + 2\pi n$ , where n is an integer.

Chapter IV, §4

**1.**  $1 + \tan^2 x$  **2.**  $\cos (3x) \cdot 3$  **3.**  $-\sin (5x) \cdot 5$  **4.**  $\cos (4x^2 + x)(8x + 1)$ **5.**  $\sec^2 (x^3 - 5)(3x^2)$  **6.**  $\sec^2 (x^4 - x^3)(4x^3 - 3x^2)$  **7.**  $\sec^2 (\sin x) \cos x$ **8.**  $\cos (\tan x)(1 + \tan^2 x)$  **9.**  $-\sin (\tan x)(1 + \tan^2 x)$  **10.** -1

**11.** -1 **12.**  $\sqrt{3}/2$  **13.**  $-\frac{2}{\sqrt{2}}$  **14.** 2 **15.**  $-2\sqrt{3}$ 

18.  $\pi - a + 2\pi n$  or  $a + 2\pi n$ , where n is an integer.

Chapter IV, §5

**1.** 2 **2.** 3 **3.**  $\frac{1}{3}$  **4.** 1 **5.** 1 **6.** 0 **7.** 0 **8.** 1 **9.** 2 **10.**  $\frac{1}{2}$  **11.**  $\frac{2}{3}$ 

Chapter V, §1

1. 1 2.  $\frac{3}{4}$  3.  $\frac{1}{6}$  4. 1 5.  $\frac{3}{4}$  6. 0 7.  $\pm 1$ 8.  $\frac{\pi}{4} + 2n\pi$  and  $\frac{5\pi}{4} + 2n\pi$ , n = integer. 9.  $n\pi$ , n = integer 10.  $\frac{\pi}{2} + n\pi$ , n = integer 11. Base =  $\sqrt{C/3}$ , height =  $\sqrt{C/12}$ 12. Radius =  $\sqrt{C/3\pi}$ , height =  $\sqrt{C/3\pi}$ 13. Base =  $\sqrt{C/6}$ , height =  $\sqrt{C/6}$ ; Radius =  $\sqrt{C/6\pi}$ , height =  $2\sqrt{C/6\pi}$ 

Chapter V, §3

**1.**  $(\frac{13}{3})^{1/2}$  **2.** 0 **3.**  $\sqrt{7/3}$  **4.** 1

Chapter V, §4

- 1. Increasing for all x.
- 2. Increasing for  $x \leq 2 1/\sqrt{3}$ , and  $x \geq 2 + 1/\sqrt{3}$ . Decreasing for  $2 1/\sqrt{3} \leq x \leq 2 + 1/\sqrt{3}$ .
- **3.** Decreasing for  $x \leq \frac{1}{2}$ , increasing for  $x \geq \frac{1}{2}$ .
- 4. Decreasing for  $\pi/4 \leq x \leq 5\pi/4$ , increasing for  $0 \leq x \leq \pi/4$  and  $5\pi/4 \leq x \leq 2\pi$ . In general, add  $2n\pi$  to these intervals.
- 5. Decreasing for  $\pi/4 \leq x \leq 3\pi/4$  and  $5\pi/4 \leq x \leq 7\pi/4$ , increasing for  $0 \leq x \leq \pi/4$  and  $3\pi/4 \leq x \leq 5\pi/4$  and  $7\pi/4 \leq x \leq 2\pi$ .
- 6. Decreasing  $x \leq -\sqrt{3/2}$  and  $0 \leq x \leq \sqrt{3/2}$ . Increasing  $-\sqrt{3/2} \leq x \leq 0$ and  $\sqrt{3/2} \leq x$ .
- 7. Increasing all x.
- 8. Decreasing  $x \leq -\sqrt{2/3}$  and  $x \geq \sqrt{2/3}$ . Increasing  $-\sqrt{2/3} \leq x \leq \sqrt{2/3}$ .
- 9. Increasing all x.
- 10. Decreasing for  $x \leq 0$ . Increasing for  $x \geq 0$ . 11. All increasing.

# Chapter VI, §1

1. 0, 0 2. 0, 0 3. 0, 0 4.  $\frac{1}{\pi}$ ,  $\frac{1}{\pi}$  5. 0, 0 6.  $\infty$ ,  $-\infty$  7.  $-\infty$ ,  $\infty$ 8.  $-\frac{1}{2}$ ,  $-\frac{1}{2}$  9.  $-\infty$ ,  $+\infty$  10. 0, 0 11.  $\infty$ ,  $-\infty$  12.  $-\infty$ ,  $\infty$ 13.  $\infty$ ,  $\infty$  14.  $-\infty$ ,  $-\infty$  15.  $\infty$ ,  $-\infty$  16.  $-\infty$ ,  $\infty$  17.  $\infty$ ,  $\infty$ 18.  $-\infty$ ,  $-\infty$ 

19.	$x \to \infty$	$a_n > 0$	$a_n < 0$
	n odd	ø	<u> </u>
	n even	80	∞

$x \rightarrow -\infty$	$a_n > 0$	$a_n < 0$
n odd	<b>—∞</b>	×
n even	œ	<b>— %</b>

Chapter VI,  $\S2$ 

	c.p.	Increasing	Decreasing	
1.	$3 \pm \sqrt{11}$	$x \leq 3 - \sqrt{11}$ and	$3-\sqrt{11} \leq x < 3$ and	
2.	$3 \pm \sqrt{10}$	$x \ge 3 + \sqrt{11}$ $3 - \sqrt{10} \le x \le 3 + \sqrt{10}$	$3 < x \le 3 + \sqrt{11}$ $3 + \sqrt{10} \le x \text{ and}$	
3.	$-1 \pm \sqrt{2}$	$-1 - \sqrt{2} \le x \le -1 + \sqrt{2}$	$\begin{vmatrix} x \leq 3 - \sqrt{10} \\ x \leq -1 - \sqrt{2} \text{ and} \\ x \geq -1 + \sqrt{2} \end{vmatrix}$	
4.	$n\pi/2$	$0 \leq x \leq \pi/2$ and add $n\pi$	$\pi/2 \leq x \leq \pi$ and add $n\pi$	
5.	$n\pi/2$	$\pi/2 \leq x \leq \pi$ and add $n\pi$	$0 \leq x \leq \pi/2$ and add $n\pi$	
6.	None	x < 0 and $x > 0$	Never	
7.	$n\pi$	$0 \leq x < \pi/2$ and add $n\pi$	$-\pi/2 < x \leq 0$ and add $n\pi$	
8.	$\frac{1}{2}$	$x \ge \frac{1}{2}$	$x \leq \frac{1}{2}$	
9.	$0, \frac{3}{2}$	$x \ge \frac{3}{2}$	$x \leq \frac{3}{2}$	
10.	0	$x \leq 0$	$\sqrt{2} < x \text{ and } 0 \leq x < \sqrt{2}$	
11.	None	$x < -\frac{1}{3}$	$x > -\frac{1}{3}$	
12.	1	$x \ge -1$	$x \leq -1$	
13.	None	All x	Never	
14.	-1	$x \ge -1$	$x \leq -1$	
15.	None	All x	Never	
16.	$-(\frac{1}{8})^{1/7}$	$x \ge -(\frac{1}{8})^{1/7}$	$x \leq -(\frac{1}{8})^{1/7}$	
17. (a), (e) 18. (c)				

Chapter VI, §3

**3.** (a) 
$$(\sqrt{2}, \pi/4)$$
 (b)  $(\sqrt{2}, 5\pi/4)$  (c)  $(6, \pi/3)$  (d)  $(1, \pi)$ 

Chapter VII, §2 (Alternative answers depend on the choice of intervals.)

1. 
$$\frac{1}{3}$$
 2.  $\frac{1}{8}$  3.  $\frac{1}{3}$  or  $-\frac{1}{3}$  4.  $-1$  or 1 5. 1 or  $-1$   
6.  $-\frac{1}{2}$  or  $-\frac{1}{10}$ , or  $\pm \frac{1}{2\sqrt{2}}$  7.  $\frac{1}{4}$  8.  $-1$  or  $\frac{1}{2} \pm \frac{3}{10}\sqrt{5}$  9.  $\frac{1}{24}$   
10.  $\frac{1}{10\sqrt{2}}$  or  $\frac{-1}{10\sqrt{2}}$ 

Chapter VII, §3 **2.**  $-1/\sqrt{1-x^2}$  **3.**  $2/\sqrt{3}, \sqrt{2}, \pi/6, \pi/4$  **4.**  $-2/\sqrt{3}, -\sqrt{2}, \pi/3, \pi/4$ 5. Let  $y = \sec x$  on interval  $0 < x < \pi/2$ . Then  $x = \operatorname{arcsec} y$  is defined on 1 < y, and  $dx/dy = \frac{1}{y\sqrt{y^2 - 1}}$ 6.  $-\pi/2$  7. 0 8.  $\pi/2$  9.  $\pi/2$  10.  $-\pi/4$ Chapter VII, §4 1.  $\pi/4$ ,  $\pi/6$ ,  $-\pi/4$ ,  $\pi/3$  2.  $\frac{1}{2}$ ,  $\frac{3}{4}$ ,  $\frac{1}{2}$ ,  $\frac{1}{4}$  3.  $1/(1+y^2)$ 4. (a)  $-\pi/4$  (b) 0 (c)  $-\pi/6$  (d)  $\pi/6$ Chapter VIII, §1 1. (a)  $y = \frac{1}{2}x + \log 2 - 1$  (b)  $y = \frac{1}{5}x + \log 5 - 1$ (c)  $y = 2x - \log 2 - 1$ 2. (a)  $y = -x + \log 2 - 1$  (b)  $y = \frac{4}{5}x + \log 5 - \frac{8}{5}$ (c)  $y = -\frac{3}{5}x + \log 10 - \frac{9}{5}$ 3. (a)  $\frac{\cos x}{\sin x}$  (b)  $\cos (\log (2x+3)) \frac{1}{2x+3} \cdot 2$  (c)  $\frac{1}{x^2+5} \cdot 2x$ (d)  $\frac{(\sin x)1/x - (\log 2x)\cos x}{\sin^2 x}$ 4.  $y = \frac{1}{4}x + \log 4 - \frac{3}{4}$  5.  $y = \frac{2x}{2} + \log 3 - \frac{8}{3}$ Chapter VIII, §2 1. (a)  $y = 2e^2x - e^2$  (b)  $y = 2e^{-4}x + 5e^{-4}$  (c) y = 2x + 1**2.** (a)  $y = \frac{1}{2}e^{-2}x + 3e^{-2}$  (b)  $y = \frac{1}{2}e^{1/2}x + \frac{1}{2}e^{1/2}$  (c)  $y = \frac{1}{2}x + 1$ **3.**  $y = 3e^2x - 4e^2$ 4. (a)  $e^{\sin 3x}(\cos 3x)3$  (b)  $\frac{1}{e^x + \sin x}(e^x + \cos x)$  (c)  $\cos (e^{x+2})e^{x+2}$ (d)  $4 \cos{(e^{4x-5})e^{4x-5}}$ Chapter VIII, §3 **1.**  $10^x \log 10$ ,  $7^x \log 7$  **2.**  $3^x \log 3$ ,  $\pi^x \log \pi$  **3.**  $x^x(1 + \log x)$ 4.  $x^{(x^x)}[x^{x-1} + (\log x)x^x(1 + \log x)]$  5. e 6. Yes 8. y = x9.  $y = (\log 10)x + 1, y = (\log 7)x + 1$ 10.  $y = (9 \log 3)x - 18 \log 3 + 9$  and  $y = (\pi^2 \log \pi)x - 2\pi^2 \log \pi + \pi^2$ Chapter VIII, §4 13, 14. All derivatives are 0 at 0. Chapter VIII. §5 1.  $-\log 25$  2.  $5e^{-4}$  3.  $e^{-(\log 10)10^{-6}t}$  4. 20/e 5.  $-(\log 2)/K$ 6.  $(\log 3)/4$  7.  $12 \log 10/\log 2$  8.  $\frac{-3 \log 2}{\log 9 - \log 10}$ Chapter IX, §1 1.  $-(\cos 2x)/2$  2.  $\frac{\sin 3x}{2}$  3.  $\log (x+1), x > -1$ 4.  $\log (x+2), x > -2$ 

Chapter IX, §3 1.  $\frac{624}{4}$  2. 2 3. 2 4.  $\log 2$  5.  $\log 3$  6.  $\frac{2}{5}$  7. e-1Chapter X,  $\S1$ 1.  $\frac{63}{2}$  2. 0 3. 0 4. 0 Chapter X,  $\S2$ 1.  $x^4$  2.  $3x^5/5 - x^6/6$  3.  $-2\cos x + 3\sin x$  4.  $\frac{9}{5}x^{5/3} + 5\sin x$ 5.  $5e^x + \log x$  6. 0 7. 0 8.  $e^2 - e^{-1}$  9.  $4 \cdot 28/3$  11.  $\frac{1}{2}$  12.  $\frac{1}{12}$ 13. 70  $-\frac{139}{2}$  14.  $\frac{8}{3} + \frac{5}{12}$  15.  $\sqrt{2} - 1$ Chapter X, §4 1. Yes 2. No 3. Yes 4. No 5. No 6. Yes 7.  $-\frac{1}{2}e^{-2B} + \frac{1}{2}e^{-4}$ . Yes,  $\frac{1}{2}e^{-4}$ . Chapter XI, §1 1.  $e^{x^2/2}$  2.  $-\frac{1}{4}e^{-x^4}$  3.  $\frac{1}{6}(1+x^3)^2$  4.  $(\log x)^2/2$ 5.  $\frac{(\log x)^{-n+1}}{1-n}$  if  $n \neq 1$ , and  $\log(\log x)$  if n = 1. 6.  $\log(x^2 + x + 1)$ 7.  $x - \log(x+1)$  8.  $\frac{\sin^2 x}{2}$  9.  $\frac{\sin^3 x}{3}$  10. 0 11.  $\frac{2}{5}$  12.  $-\arctan(\cos x)$ 13.  $\frac{1}{2}(\arctan x)^2$  14.  $\frac{2}{15}$  15.  $-\frac{1}{4}\cos(\pi^2/2) + \frac{1}{4}$ 16.  $-\frac{1}{2}e^{-B^2} + \frac{1}{2}$ . Yes,  $\frac{1}{2}$  17.  $-\frac{1}{3}e^{-B^3} + \frac{1}{3}$ . Yes.  $\frac{1}{3}$ 18.  $2\sqrt{1+e^x} - \log(\sqrt{1+e^x}+1) + \log(\sqrt{1+e^x}-1)$ **19.**  $x - \log (1 + e^x)$  **20.**  $\arctan (e^x)$ **21.**  $-\log(\sqrt{1+e^x}+1) + \log(\sqrt{1+e^x}-1)$ Chapter XI, §2 1.  $x \arcsin x + \sqrt{1 - x^2}$  2.  $x \arctan x - \frac{1}{2} \log (x^2 + 1)$ 3.  $\frac{e^{2x}}{12}$  (2 sin 3x - 3 cos 3x) 4.  $\frac{1}{10}e^{-4x}$  sin 2x -  $\frac{1}{5}e^{-4x}$  cos 2x 5.  $x(\log x)^2 - 2x \log x + 2x$  6.  $(\log x)^3 x - 3 \int (\log x)^2 dx$ 7.  $x^2e^x - 2xe^x + 2e^x$  8.  $-x^2e^{-x} - 2xe^{-x} - 2e^{-x}$  9.  $-x\cos x + \sin x$ 10.  $x \sin x + \cos x$  11.  $-x^2 \cos x + 2 \int x \cos x \, dx$ 12.  $x^2 \sin x - 2 \int x \sin x \, dx$  13.  $\frac{1}{2} [x^2 \sin x^2 + \cos x^2]$ 14.  $-\frac{1}{3}x^4(1-x^2)^{3/2} - \frac{4}{3+5}x^2(1-x^2)^{5/2} - \frac{8}{3+5+7}(1-x^2)^{7/2}$ 15.  $\frac{1}{3}x^3 \log x = \frac{1}{3}x^3$  16.  $(\log x) \frac{x^4}{4} - \frac{x^4}{16}$  17.  $(\log x)^2 \frac{x^3}{3} - \frac{2}{3} \int x^2 \log x \, dx$ 18.  $-\frac{1}{2}x^2e^{-x^2} - \frac{1}{2}e^{-x^2}$  19.  $\frac{1}{4}x^4\left(\frac{1}{1-x^4}\right) + \frac{1}{4}\log(1-x^4)$  20.  $-4\pi$ **21.**  $-Be^{-B} - e^{-B} + 1$ . Yes, 1 **22.** Yes **23.** Yes **24.**  $-\frac{1}{\log B} + \frac{1}{\log 2}$  Yes,  $1/\log 2$ . **25.** Yes,  $1/3(\log 3)^3$ Chapter XI, §3 1.  $-\frac{1}{4}\sin^3 x \cos x - \frac{3}{8}\sin x \cos x + \frac{3}{8}x$  2.  $\frac{1}{3}\cos^2 x \sin x + \frac{2}{3}\sin x$ 3.  $\frac{\sin^3 x}{2} - \frac{\sin^5 x}{5}$  4.  $3\pi$  5.  $8\pi$  6.  $\pi ab$  (If a, b > 0). 7.  $\pi r^2$ 

13. 
$$-\log \cos x$$
 14.  $\arcsin \frac{x}{3}$  15.  $\arcsin \frac{x}{\sqrt{3}}$  16.  $\frac{1}{2} \arcsin (\sqrt{2} x)$   
17.  $\frac{1}{b} \arcsin \frac{bx}{a}$   
Chapter XI, §4  
1.  $C_1 = -\frac{3}{100}, C_2 = -\frac{11}{100}, C_3 = -\frac{130}{100}, C_4 = -\frac{110}{100}, C_5 = \frac{11}{100}$   
2.  $\frac{x}{2(x^2 + 1)} + \frac{1}{2} \arctan x$   
3. (a)  $\frac{1}{2}[\log (x - 3) - \log (x + 2)]$  (b)  $\log (x + 1) - \log (x + 2)$   
4.  $-\frac{1}{2}\log (x + 1) + 2\log (x + 2) - \frac{3}{2}\log (x + 3)$   
5.  $-\log (x^2 + x) + 3\log x$  6.  $\log (x + 1) + \frac{1}{x + 1}$   
7.  $\frac{1}{2} \frac{-1}{x^2 + 9} + \frac{1}{18} \frac{x}{x^2 + 9} + \frac{1}{54} \arctan \frac{x}{3}$  8.  $\frac{1}{8} \frac{x}{x^2 + 16} + \frac{1}{32} \arctan \frac{x}{4}$   
9.  $-\log (x + 1) + \log (x + 2) - \frac{2}{x + 2}$   
10.  $\frac{1}{4} \frac{x}{(x^2 + 1)^2} + \frac{3}{8} \frac{x}{(x^2 + 1)} + \frac{3}{8} \arctan x$   
11.  $-\frac{1}{8}\log (x - 1) + \frac{1}{8}\log (x + 7)$   
Chapter XII, §1  
2. (a)  $4/e^2$  (b)  $2^{2}5^{5}e^{-4}/3^{3}$   
Chapter XIII, §1  
2.  $2\pi r$  3.  $\sqrt{2}(e - 1)$  4. (a)  $\frac{3}{4}$  (b) 3  
5.  $\sqrt{1 + e^2} - \sqrt{2} + \frac{1}{2}\log \frac{\sqrt{1 + e^2} - 1}{\sqrt{1 + e^2} + 1} + \frac{1}{2}\log \frac{\sqrt{2} + 1}{\sqrt{2} - 1}$   
6.  $\frac{1}{27}(3^{13/2} - 13^{3/2})$  7.  $e - \frac{1}{e}$  8. 4 9.  $\pi$  10.  $2\sqrt{3}$  11.  $2\pi$   
13.  $\sqrt{2}(e^2 - e)$  14.  $\sqrt{2}(e^{\theta_2} - e^{\theta_1})$   
Chapter XIII, §3  
1.  $\frac{4}{3}\pi r^3$  2.  $\pi$  3.  $\frac{\pi^2}{8} - \frac{\pi}{4}$  4.  $\frac{\pi^2}{8} + \frac{\pi}{4}$  5.  $\frac{2 \cdot 5^4\pi}{3}$  6.  $\pi(e - 2)$  7.  $\pi e^2$   
8.  $\pi[2(\log 2)^2 - 4\log 2 + 2]$  9.  $12\pi$  10.  $\frac{\pi}{2} \left[\frac{1}{e^2} - \frac{1}{e^{2B}}\right]$ , Yes,  $\frac{\pi}{2e^2}$   
11.  $\pi r^2h/3$   
Chapter XIV, §3  
1.  $1 - \frac{x^2}{2!} + \frac{x^4}{4!}$  2.  $|R_n| \le \frac{|x|^n}{n!}$  3.  $1 - \frac{0.01}{2} = .995$  4.  $|R_3| \le \frac{1}{8}10^{-3}$ 

5. 
$$|R_4| \leq \frac{2}{3}10^{-4}$$
 6.  $\tan x = x + \frac{x^3}{3}$   
7.  $|R_4| \leq 10^{-4}$  by crude estimates  
8.  $\sin\left(\frac{\pi}{6} + \frac{\pi}{180}\right) = \frac{1}{2} + \frac{\sqrt{3}}{2} \cdot \frac{\pi}{180} = 0.515$  9.  $x^2 - \frac{1}{8}x^4$   
10.  $1 - \frac{3x^2}{2} + \frac{7}{8}x^4$  11.  $1 + \frac{x^2}{2} + \frac{5}{24}x^4$  12.  $x^3$ 

Chapter XIV, §4

1.  $1 - x^2 + \frac{x^4}{2}$  2.  $|R_3| \le 1/24$  3.  $|R_4| \le 10^{-9}$  4.  $|R_3| \le \frac{1}{3}10^{-6}$ 5.  $1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \frac{x^4}{4!} - \frac{x^5}{5!}$  6. Use  $|R_7| \le \frac{1}{5040}$  7. All equal to 0.

### Chapter XIV, §6

5. 1 6. 1 7. 1 8. 1 9. 1 10. 2 11.  $\frac{1}{2}$  12. 0 13.  $-\frac{1}{2}$  14. 1 15. 1 16. 1

# Chapter XIV, §7

**1.**  $5 + \frac{1}{75}$  **2.**  $10 - \frac{3}{20} - \frac{9}{8000}$  **3.**  $|R_2| \le 2 \times 10^{-2}$  **4.**  $|R_2| \le 0.09$ **5.**  $|R_2| \le 2 \times 10^{-5}$ 

Chapter XV, §2

3. No 4. Yes 5. No 6. No 7. No 8. Yes 9. Yes

### Chapter XV, §3

3. Yes 4. Yes 5. Yes 6. Yes 7. Yes 8. Yes 9. Yes 10. Yes Chapter XV, §4

1. Yes 2. Yes 3. Yes 4. Yes 5. Yes 6. Yes 7. Yes 8. No 9. Yes 10. Yes

# Chapter XV, §5

1.  $\frac{1}{4}$  2. 1/e 3. 27 4.  $4/e^2$  5. 0 6. 2 7. 2 8. 3 9. 1 10.  $\infty$ 11. 1 12.  $\infty$  13. 1 14.  $\infty$  15. e 16.  $\infty$ 

## SUPPLEMENTARY EXERCISES

# CHAPTER I

1. Prove that the sum of a rational number and an irrational number is always irrational. What about the product?

2. Find all solutions of the equation

$$x + |x - 2| = 1 + |x|.$$

3. A function f is called *even* if f(x) = f(-x) for all x, and is called *odd* if f(-x) = -f(x) for all x.

(a) Which of the following functions are even? Odd? Neither?

$$f(x) = x^{2}, \quad f(x) = x^{17}, \quad f(x) = x^{2} + 2x + 1, \quad f(x) = \frac{1}{1 + x^{2}},$$
  
$$f(x) = \sin x, \quad f(x) = \cos x, \quad f(x) = \tan x.$$

(b) Let f be a function defined for all x. Show that the functions g and h defined by

$$g(x) = f(x) + f(-x)$$
 and  $h(x) = f(x) - f(-x)$ 

are even and odd respectively. Show that any function f can be expressed as the sum of an even function and an odd function.

(c) Does there exist a function which is both even and odd? Is it unique?

4. How many functions are there which are defined for the numbers  $\{1, 2, 3\}$  and whose values are in the set of integers n with  $1 \le n \le 10$ ?

5. Let D be a set of m numbers and E a set of n numbers. How many functions are there defined for all numbers in D and taking their values in E?

6. Let x, y be numbers > 0. Assume that x < y. Prove by induction that for every positive integer n we have  $x^n < y^n$ . Prove that  $x^{1/n} < y^{1/n}$  (where  $x^{1/n}$  is the unique number whose nth power is x).

# CHAPTER III

1. Let f be the function defined for all numbers as follows. If x is not a rational number then f(x) = 0. If x is a rational number, which can be written as a fraction p/q, with integers q, p, and if this fraction is in lowest form, q > 0, then  $f(x) = 1/q^3$ . Show that f is not differentiable at any rational number x.

2. Let  $\alpha$  be an irrational number having the following property. There exists a number c > 0 such that for any rational number p/q (in lowest form) with q > 0, we have

 $\left| lpha - rac{p}{q} 
ight| > rac{c}{q^2}$   $\left| q lpha - p 
ight| > rac{c}{q}$ 

or equivalently,

Show that the function f in Exercise 1 is differentiable at 
$$\alpha$$

3. Let  $\alpha = \sqrt{2}$  (or more generally  $\sqrt{a}$ , where *a* is a positive integer such that its square root is irrational). Prove that  $\alpha$  has the property stated in Exercise 2,

namely there exists a number c > 0 such that for all integers q, p with q > 0we have  $|q\alpha - p| > c/q$ .

(More generally, let  $\alpha$  be an algebraic number, i.e. a number which is a root of a polynomial of degree  $\geq 1$  with rational coefficients, and suppose that  $\alpha$  is not the root of a quadratic polynomial with rational coefficients. Does  $\alpha$  still have the above property? As far as I know, the answer is not known, and is a good research problem. A similar question can be asked of all the common numbers vou know, like  $e, \pi$ , etc.)

4. Let m, n be integers  $\neq 0$ . Show that the number  $\alpha = m + n\sqrt{2}$  has the property stated in Exercise 2.

5. Let q be the function defined for all numbers as follows:

If x is not a rational number, then g(x) = 0. If x is a rational number which can be written as a fraction p/q, with integers q, p, and if this fraction is in lowest form, q > 0, then g(x) = 1/q.

(a) Show that g is not continuous at any rational number.

(b) Investigate the differentiability of g at numbers  $\alpha$  having the property stated in Exercise 2.

(c) Show that q is continuous at all irrational numbers.

6. Let  $\alpha$  be an irrational number. Given a positive integer n, show that there exist integers p, q such that

$$|q\alpha - p| < \frac{1}{n} \quad 0 < q \leq n,$$
$$|\alpha - \frac{p}{n}| < \frac{1}{n}.$$

and hence that

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2} \cdot$$

[*Hint*: Cut up the interval between 0 and 1 into n equal segments of length 1/n, and consider the n + 1 numbers

$$0, \alpha, 2\alpha, \ldots, n\alpha.$$

Show that there exist integers r, s with  $0 \leq r \leq n$  and  $0 \leq s \leq n, r \neq s$ , such that  $r\alpha - p_1$  and  $s\alpha - p_2$  lie in the same segment for suitable integers  $p_1, p_2$ .]

7. Let  $\alpha$  be an irrational number. Let w be any number, and let  $\epsilon > 0$ . Show that there exist integers q, p such that  $|q\alpha - p - w| < \epsilon$ . (In other words, the numbers of type  $q\alpha - p$  come arbitrarily close to w.)

### CHAPTER V

1. (a) Let a, b be numbers  $\geq 0$ . Show that

$$(ab)^{1/2} \leq \frac{a+b}{2}$$

(b) Let  $a_1, \ldots, a_n$  be numbers  $\geq 0$ . Show that

$$(a_1\cdots a_n)^{1/n} \leq \frac{a_1+\cdots+a_n}{n}$$

[*Hint*: Take the *n*th power. Fix  $a_1, \ldots, a_{n-1}$  and replace  $a_n$  by a variable x. Use induction.]

2. (a) Let  $a_0, \ldots, a_n$  be numbers such that for all x we have

$$a_0+a_1x+\cdots+a_nx^n=0.$$

Show that  $a_i = 0$  for all *i*. [*Hint*: Divide by  $x^n$  and let x become very large.] (b) Let f, g be polynomials,

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n$$
  
$$g(x) = b_0 + b_1 x + \cdots + b_n x^n.$$

If f(x) = g(x) for all x, prove that  $a_i = b_i$  for i = 1, ..., n. [Hint: Subtract g from f.]

3. Let  $f(x) = a_n x^n + \cdots + a_0$  be a polynomial, with  $a_n \neq 0$ . Let  $c_1 < c_2 < \cdots < c_r$  be numbers such that  $f(c_i) = 0$  for  $i = 1, \ldots, r$ . Show that  $r \leq n$ . [Hint: Show that f' has at least r - 1 roots, and continue to take the derivative.]

4. Let f be a function which is infinitely differentiable. Let  $c_1 < c_2 < \cdots < c_r$  be numbers such that  $f(c_i) = 0$  for all i. Show that f' has at least r - 1 zeros [i.e. numbers b such that f'(b) = 0].

5. Let a, b be numbers, a < b. Let f be a continuous function defined over the interval [a, b]. Assume that f' and f'' exist on the interval a < x < b, and that f''(x) > 0 for all x in this interval. Prove that the graph of f(x) in the open interval lies below the line segment joining the two points of the graph whose coordinates are (a, f(a)) and (b, f(b)). [Hint: Let  $\varphi(x)$  be the difference between the straight line joining the two points, and f(x). Show that  $\varphi'(x) = f'(c) - f'(x)$  for some number c, a < c < b. Applying the mean value theorem to f' show that  $\varphi$  is increasing to the left of c and decreasing to the right of c. Use the value of  $\varphi$  at the end points of the interval.]

6. Let the hypotheses be as in Exercise 5. Let d be a number, a < d < b. Show that the tangent line to the curve y = f(x) at x = d lies below the graph of f except at x = d where it touches the graph.

7. Let the hypotheses be as in Exercise 5, and let  $x_1, x_2$  be two numbers in the interval [a, b] such that  $x_1 < x_2$ . Show that

$$f\left(\frac{x_1+x_2}{2}\right) \leq \frac{f(x_1)+f(x_2)}{2}$$

Generalize to n numbers.

8. In Exercise 5, assume that  $f''(x) \leq 0$  instead of  $\geq 0$ , and let the other assumptions and notations be the same. What can you conclude about the relative positions of the graph of f and the straight line between the two points?

9. (a) Find the equation for the tangent to the hyperbola xy = 1 at the point  $(\frac{1}{2}, 2)$ .

(b) Prove that the tangent line is below the hyperbola for all x > 0 (except  $x = \frac{1}{2}$  where it touches).

10. Apply Exercises 5 and 8 to the functions  $\sin x$  and  $\cos x$  over suitable intervals.

11. Prove that a polynomial of odd degree has a root. [In other words, if f is a polynomial of odd degree, then there exists a number a such that f(a) = 0.]

252

12. Let  $a_1, \ldots, a_n$  be numbers. Determine x so that

$$\sum_{i=1}^n \left(a_i - x\right)^2$$

is a minimum.

# CHAPTER VI

1. Sketch the curves  $y = x + \frac{1}{x}$  and  $y = \sqrt{x + \frac{1}{x}}$ . 2. Sketch the curves  $y = \frac{x}{x^2 + 1}$ , and  $y = \sqrt{\frac{x}{x^2 + 1}}$ . 3. Sketch the curves  $y = \frac{x}{x^2 - 1}$ , and  $y = \sqrt{\frac{x}{x^2 - 1}}$ .

4. Show that among all triangles with given area, the equilateral triangle has the least perimeter.

5. Show that among all triangles with given perimeter, the equilateral triangle has the maximum area.

- 6. Prove that  $\tan x \ge x$  if  $0 \le x \le \frac{\pi}{2}$ .
- 7. Sketch the curves:
- (a)  $y = \sqrt{(x-1)(x-2)}$  (b)  $y = \sqrt{(x-1)(x-2)(x-3)}$ (c)  $y^2 = (x-1)(x-2)(x-3)$ 8. Let a, b, c be numbers with a < b < c. Sketch the curves: (a)  $y^2 = (x-a)(x-b)(x-c)$  (b)  $y = \sqrt{(x-a)(x-b)(x-c)}$ (b)  $y = \sqrt{(x-1)(x-2)(x-3)}$
- 9. Sketch the following curves:

(a) 
$$y = \sqrt{\frac{x-1}{x+1}}$$
 (b)  $y = \sqrt{\frac{x^2+1}{x+1}}$   
(c)  $y = \sqrt{\frac{x-1}{x^2+1}}$  (d)  $y = \sqrt{\frac{x^2+1}{x^2-1}}$ 

#### CHAPTER VIII

1. Let f be the function such that f(x) = 0 if  $x \leq 0$ , and  $f(x) = e^{-1/x}$  if x > 0. Show that f is infinitely differentiable at 0, and that all its derivatives are equal to 0 at 0.

2. Let a, b be numbers, a < b. Let f be the function such that f(x) = 0 if  $x \leq a \text{ or } x \geq b$ , and

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

if a < x < b. Sketch the graph of f. Show that f is infinitely differentiable at both a and b.

3. Let a, b be numbers, a < b. Let y = f(x) be the equation of the straight line between the points  $(a, e^a)$  and  $(b, e^b)$  lying on the curve  $y = e^x$ . Show that  $f(x) > e^x$  for all numbers x such that a < x < b.

4. Let a, b be numbers, 0 < a < b. Let y = f(x) be the equation of the straight line between the points  $(a, \log a)$  and  $(b, \log b)$  lying on the curve  $y = \log x$ . Show that  $\log x > f(x)$  for all numbers x such that a < x < b.

5. Let n be an integer  $\geq 1$ . Let  $f_0, \ldots, f_n$  be polynomials such that

$$f_n(x)e^{nx} + f_{n-1}(x)e^{(n-1)x} + \cdots + f_0(x) = 0$$

for all numbers x. Show that  $f_0, \ldots, f_n$  are identically 0. [Hint: Cf. Problem 2 of Chapter V.]

6. Let n be an integer  $\geq 1$ . Let  $f_0, \ldots, f_n$  be polynomials such that

$$f_n(x)(\log x)^n + f_{n-1}(x)(\log x)^{n-1} + \cdots + f_0(x) = 0$$

for all numbers x > 0. Show that  $f_0, \ldots, f_n$  are identically 0.

7. Show that e is not a rational number. To do this, use the series for e given later in the text, namely

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} \cdot$$

8. If a is a number >1 and x > 0, show that

$$x^a - 1 \ge a(x - 1).$$

9. Let p, q be numbers  $\geq 1$  such that  $\frac{1}{p} + \frac{1}{q} = 1$ . If  $x \geq 1$ , show that

$$x^{1/p} \leq \frac{x}{p} + \frac{1}{q} \cdot$$

10. Let  $\alpha, \beta$  be positive numbers such that  $\alpha/\beta \ge 1$ , and let p, q be as in Exercise 9. Show that

$$\alpha^{1/p}\beta^{1/q} \leq \frac{\alpha}{p} + \frac{\beta}{q}$$

11. Let  $\alpha$ ,  $\beta$  be two positive numbers, and p, q two nonzero numbers with p < q. Show that for any number t with 0 < t < 1, we have

$$[t\alpha^{p} + (1-t)\beta^{p}]^{1/p} \leq [t\alpha^{q} + (1-t)\beta^{q}]^{1/q}.$$

12. Show that the equality sign in the preceding inequality holds if and only if  $\alpha = \beta$ .

13. Let  $\alpha$ ,  $\beta$  be two numbers >0, and 0 < t < 1. Show that

$$\alpha^t\beta^{1-t} \leq t\alpha + (1-t)\beta,$$

and that equality holds if and only if  $\alpha = \beta$ .

14. Let a be a number >0. Find the minimum and maximum of the function  $f(x) = x^2/a^x$ .

15. Let  $a_1, \ldots, a_n$  be numbers  $\geq 0$  and let  $0 < r \leq s$ . Show that

$$\left[\frac{a_1^r + \dots + a_n^r}{n}\right]^{1/r} \leq \left[\frac{a_1^s + \dots + a_n^s}{n}\right]^{1/s}.$$

16. Hyperbolic functions

(a) Define functions

$$\cosh t = \frac{e^{t} + e^{-t}}{2}$$
 and  $\sinh t = \frac{e^{t} - e^{-t}}{2}$ .

Show that their derivatives are given by

$$\cosh' = \sinh$$
 and  $\sinh' = \cosh$ 

(b) Show that for all t we have

$$\cosh^2 t - \sinh^2 t = 1.$$

(c) Sketch the graph of the curve  $x^2 - y^2 = 1$ .

(d) If you let  $x = \cosh t$  and  $y = \sinh t$ , which portion of the curve in (c) is parametrized by these functions?

(e) For a suitable interval of values of t, determine inverse functions for  $\cosh t$  and  $\sinh t$ , and determine their derivatives.

## CHAPTER IX

1. Let

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} ((x^2 - 1)^n).$$

Show that

$$\int_{-1}^{1} P_n(x) P_m(x) \ dx = 0 \quad \text{if} \quad m \neq n.$$

2. Show that

$$\int_{-1}^{1} P_n^2(x) \, dx = \frac{2}{2n+1} \, \cdot \,$$

3. Show that

$$\int_{-1} x^m P_n(x) \, dx = 0 \quad \text{if} \quad m < n$$

4. Evaluate

$$\int_{-1}^{1} x^n P_n(x) \, dx.$$

5. Let a, b be numbers with a < b. If f, g are continuous functions on the interval [a, b] let

$$\langle f,g\rangle = \int_a^b f(x) g(x) dx.$$

Show that the symbol  $\langle f, g \rangle$  satisfies the following properties:

(a) If  $f_1, f_2, g$  are continuous on [a, b] then

c1

$$\langle f_1 + f_2, g \rangle = \langle f_1, g \rangle + \langle f_2, g \rangle.$$

If c is a number, then

$$\langle cf, g \rangle = c \langle f, g \rangle.$$

- (b) We have  $\langle f, g \rangle = \langle g, f \rangle$ .
- (c) We have  $\langle f, f \rangle \ge 0$ , and equality holds if and only if f = 0.
- 6. Let the notation be as in Exercise 5. For any number  $p \ge 1$  define

$$||f||_{p} = \left[\int_{a}^{b} |f(x)|^{p} dx\right]^{1/p}.$$

Let q be such that  $\frac{1}{p} + \frac{1}{q} = 1$ . Prove that

 $|\langle f, g \rangle| \leq ||f||_p ||g||_q.$ 

[*Hint*: If  $||f||_p$  and  $||g||_q \neq 0$  let  $\alpha = |f|^p / ||f||_p^p$  and  $\beta = |g|^q / ||g||_q^q$ .] 7. Notation being as in the preceding exercise, prove that

$$||f + g||_p \leq ||f||_p + ||g||_p$$

[Hint: Let I denote the integral. Show that

$$||f + g||_{p}^{p} \leq I(|f + g|^{p-1}|f|) + I(|f + g|^{p-1}|g|)$$

and apply Exercise 6.]

8. Let  $\alpha$  be a number >0. Let

$$a_n = \frac{\alpha(\alpha+1)\cdots(\alpha+n)}{n!n^{\alpha}}$$

Show that  $\{a_n\}$  is monotonically decreasing for sufficiently large values of n, and hence approaches a limit.

Let f be a continuous function, defined for all numbers. We say that the integral  $c^{\infty}$ 

$$\int_{-\infty}^{\infty} f(t) dt$$

converges absolutely if the limits

$$\lim_{B\to\infty}\int_0^B |f(t)| \ dt \qquad \text{and} \qquad \lim_{B\to\infty}\int_0^{-B} |f(t)| \ dt$$

exist. If the integral converges absolutely, then it can be shown that the similar limits exist without inserting the absolute value sign, and hence that

$$\lim_{\substack{B\to\infty\\C\to\infty}}\int_{-C}^{B}f(t)\ dt$$

exists (no matter in which order we take the two limits). We may assume this in the following exercises.

9. Let P be a polynomial, and let a be a number >0. Show that the improper integral  $c^{\infty}$ 

$$\int_{-\infty}^{\infty} P(t) e^{-at^2} dt$$

converges absolutely.

10. Let the notation be as in the preceding exercise. Show that the improper integral  $\int_{-\infty}^{\infty}$ 

$$\int_{-\infty}^{\infty} P(t) e^{-a|t|} dt$$

converges absolutely.

In the following exercise, you may assume that

$$\int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}.$$

11. (a) Let k be an integer  $\geq 0$ . Let P(t) be a polynomial, and let c be the coefficient of its term of highest degree. Integrating by parts, show that the integral

$$\int_{-\infty}^{\infty} \left(\frac{d^k}{dt^k} e^{-t^2}\right) P(t) dt$$

is equal to 0 if deg P < k, and is equal to  $(-1)^k k! c \sqrt{\pi}$  if deg P = k.

(b) Show that

$$\frac{d^k}{dt^k} (e^{-t^2}) = P_k(t)e^{-t^2}$$

where  $P_k$  is a polynomial of degree k, and such that the coefficient of  $t^k$  in  $P_k$  is equal to

$$a_k = (-1)^k 2^k.$$

(c) Let m be an integer  $\geq 0$ . Let  $H_m$  be the function defined by

$$H_m(t) = e^{t^2/2} \frac{d^m}{dt^m} (e^{-t^2}).$$

Show that

$$\int_{-\infty}^{\infty} H_m(t)^2 dt = (-1)^m m! a_m \sqrt{\pi}$$

and that if  $m \neq m'$  then

$$\int_{-\infty}^{\infty} H_m(t) H_{m'}(t) dt = 0.$$

# CHAPTER XV

1. Prove: If the series

$$\sum_{n=1}^{\infty} a_n$$

with decreasing positive terms converges, then  $\lim_{n\to\infty} na_n = 0$ .

2. Prove that if  $\sum a_n$  converges, and if  $b_1, b_2, b_3, \ldots$ , is a bounded monotone sequence of numbers, then  $\sum a_n b_n$  converges.

3. Prove that if the partial sums of the series  $\sum a_n$  oscillate between finite bounds, and if  $b_n$  is a monotone sequence of numbers tending to 0, then  $\sum a_n b_n$  converges.

4. Formulate Exercises 2 and 3 for integrals instead of series.

5. Show that the integral

$$\int_1^\infty \sin t \, \frac{\log t}{t} \, dt$$

converges.

6. Show that the integral

$$\int_0^\infty (\cos t) e^{-t} dt$$

converges absolutely.

7. Let  $a_n$  be a sequence of numbers  $\geq 0$ , and assume that the series

$$\sum \frac{a_n}{n^x}$$

converges when  $x = x_0$ . Show that it converges for any  $x > x_0$ .

8. Determine which of the following series converge:

(a) 
$$\sum \frac{(\log n)^5}{n^{1+\epsilon}}$$
 (with  $\epsilon$  fixed > 0) (b)  $\sum \frac{\log \log n}{n(\log n)^{1+\epsilon}}$   
(c)  $\sum \frac{1}{n(\log n)(\log \log n)}$  (d)  $\sum \frac{1}{n(\log n)(\log \log n)^{1+\epsilon}}$ 

(All above sums are taken from 2 to  $\infty$ .)

9. Let  $\sum a_n$  be a series with numbers  $a_n \ge 0$ . If there exist infinitely many n such that  $a_n > 1/n$ , show that the series does not converge.

## LIMITS

1. Find the limits (for a fixed x > 0):

(a) 
$$\lim_{n\to\infty} x^n$$
 (b)  $\lim_{h\to0} \frac{\sqrt{x+h}-\sqrt{x}}{h}$  (c)  $\lim_{n\to\infty} x^{1/n}$ 

(If any limit does not exist, say so.)

2. For  $x \neq -1$ , show that the following limit exists:

$$f(x) = \lim_{n \to \infty} \left( \frac{x^n - 1}{x^n + 1} \right)^2.$$

- (a) What is f(1), f(1/2), f(2)?
- (b) What is  $\lim_{x\to 1} f(x)$ ?
- (c) What is  $\lim_{x \to 0^+} f(x)$ ?
- $x \rightarrow -1$

(d) For which values of  $x \neq -1$  is f continuous? Is it possible to define f(-1) in such a way that f is continuous at -1?

3. Let

$$f(x) = \lim_{n\to\infty} \frac{x^n}{1+x^n}.$$

- (a) What is the domain of f, i.e. for which numbers x does the limit exist?
- (b) Give explicitly the values f(x) of f for the various x in the domain of f.(c) For which x in the domain is f continuous at x?

4. Let a be a number. Let f be a function defined for all numbers x < a. Assume that when x < y < a we have f(x) < f(y), and also that f is bounded from above. Prove that  $\lim_{x\to a} f(x)$  exists.

5. Using only the properties of numbers related to addition, subtraction, multiplication, division, ordering, and the least upper bound (and greatest lower

bound) axiom, prove that if a is a number >0, then there exists a number b > 0 such that  $b^2 = a$ .

6. Let  $a_1, a_2, \ldots, a_n, \ldots$ , be a sequence of numbers. We shall say that it is a *Cauchy* sequence if given  $\epsilon > 0$  there exists a positive integer N such that, for all integers m, n > N we have  $|a_n - a_m| < \epsilon$ . Prove that every Cauchy sequence has a limit.

7. Let S be a set of numbers. We shall say that S is closed if whenever x is a number such that S is arbitrarily close to x, then x lies in S.

(a) If S, S' are closed, show that  $S \cap S'$  is closed.

(b) If  $\{S_i\}$  is a family of closed sets, show that the intersection of all the sets  $S_i$  in the family is closed.

(c) Prove that a set S is closed if and only if every Cauchy sequence in S has a limit in S.

8. Let S be a set and let f be a function defined on S. We shall say that f is bounded if there exists a number C > 0 such that  $|f(x)| \leq C$  for all  $x \in S$ .

(a) If f, g are bounded functions on S, show that f + g, f - g, fg are bounded. In particular, if a is a number, then af is bounded.

(b) Let f be a bounded function on S. Define ||f|| to be the least upper bound of the set of numbers |f(x)|, with x in S. If f, g are bounded functions on S, show that  $||f + g|| \le ||f|| + ||g||$ ,  $||fg|| \le ||f|| ||g||$ , and if a is a number, then ||af|| = |a| ||f||.

9. Let S be a set of numbers such that every sequence  $\{a_n\}$  of numbers in S has a point of accumulation in S. Let f be a function defined on S. Show that f is bounded, and has a maximum in S (i.e. that there exists c in S such that  $f(c) \ge f(x)$  for all x in S).

10. (Tate) Let f be a function defined for all numbers. Assume that there exists a number C > 0 such that for all numbers x, y we have

$$|f(x+y) - f(x) - f(y)| \leq C.$$

Show that there exists a unique function g such that

$$g(x+y) = g(x) + g(y)$$

for all x, y and such that f - g is bounded. [Hint: Let

$$g(x) = \lim_{n \to \infty} \frac{f(2^n x)}{2^n} \cdot$$

Of course, you must prove that this limit exists.]

11. Let S be a set of numbers such that any Cauchy sequence in S has a limit in S. Let  $f: S \to S$  be a map of S into itself, and assume that there exists a number c, 0 < c < 1 such that for all x, y in S we have

$$|f(x) - f(y)| \leq c|x - y|.$$

Prove the following.

(a) Given any  $x \in S$ , the sequence  $\{f^n(x)\}$  is Cauchy. (Here  $f^n$  means the iterate of f taken n times.)

(b) Let  $z = \lim_{n \to \infty} f^n(x)$ . Then f(z) = z.

(c) If w is an element of S such that f(w) = w, then w = z [the element in (b)].

### INDEX

Absolutely convergent, 1 Absolute value, 6 Addition formula, 72 Arccosine, 113 Arcsine, 111 Arctangent, 114 Area in polar coordinates, 188 Axes, 14 Binomial coefficient, 209 Circle, 28 Closed interval, 80 Composite function, 58 Constant function, 20 Continuous, 137 Converge, 160, 211 Coordinate axes, 14 Coordinates, 16 Cosine, 65 Critical point, 79 Curve, 28 Decreasing, 90 Definite integral, 145 Degree, 67 Derivative, 40 Differentiable, 40 Distance, 27 Divergent, 212 Ellipse, 31 Even, 3 Exponential function, 123Fraction, 1 Function, 8

Graph, 17

Greater than, 5 Greatest lower bound, 147

Half closed, 80 Half open, 80 Hyperbola, 33

Improper integral, 159 Increasing, 89 Indefinite integral, 134 Inner function, 58 Integration by parts, 166 Intermediate value theorem, 92 Interval, 80 Inverse function, 107

Least upper bound, 147 Left derivative, 42 Length, 185 Limits, 228 Local maximum, 83 Local minimum, 83 Logarithm, 118 Lower bound, 147 Lower integral, 147 Lower sum, 144

Maximum, 80 Minimum, 82 Moment, 192

Negative integer, 1 Negative number, 4 Newton quotient, 40 Number, 2

Odd, 3 Open interval, 80 Outer function, 58 260 Parabola, 32 Parametrization, 104 Partition, 142 Polar coordinates, 102 Positive integer, 1 Positive number, 4 Power series, 220

Radians, 66 Radius of convergence, 221 Rational number, 1 Rational point, 105 Real number, 2 Remainder, 196 Riemann sum, 142 Right derivative, 42

Sequence, 211 Series, 211 Sine, 65 Slope, 23 Stirling's formula, 179 Straight line, 23 Strictly decreasing, 90 Strictly increasing, 90 Substitution, 163

Taylor formula, 196 Trivial partition, 142

Upper bound, 146 Upper integral, 148 Upper sum, 143

Value, 9 Volume of revolution, 189

Wallis product, 181 Work, 190

### **Undergraduate Texts in Mathematics**

Abbott: Understanding Analysis. Anglin: Mathematics: A Concise History and Philosophy. Readings in Mathematics. Anglin/Lambek: The Heritage of Thales. Readings in Mathematics. **Apostol:** Introduction to Analytic Number Theory. Second edition. Armstrong: Basic Topology. Armstrong: Groups and Symmetry. Axler: Linear Algebra Done Right. Second edition. Beardon: Limits: A New Approach to Real Analysis. Bak/Newman: Complex Analysis. Second edition. **Banchoff/Wermer:** Linear Algebra Through Geometry. Second edition. Berberian: A First Course in Real Analysis. Bix: Conics and Cubics: A Concrete Introduction to Algebraic Curves. Brémaud: An Introduction to Probabilistic Modeling. **Bressoud:** Factorization and Primality Testing. Bressoud: Second Year Calculus. Readings in Mathematics. Brickman: Mathematical Introduction to Linear Programming and Game Theory. Browder: Mathematical Analysis: An Introduction. Buchmann: Introduction to Cryptography. Buskes/van Rooij: Topological Spaces: From Distance to Neighborhood. Callahan: The Geometry of Spacetime: An Introduction to Special and General Relavitity. Carter/van Brunt: The Lebesgue-Stieltjes Integral: A Practical Introduction. Cederberg: A Course in Modern Geometries. Second edition.

Childs: A Concrete Introduction to Higher Algebra. Second edition. **Chung:** Elementary Probability Theory with Stochastic Processes. Third edition. Cox/Little/O'Shea: Ideals, Varieties, and Algorithms. Second edition. **Croom:** Basic Concepts of Algebraic Topology. Curtis: Linear Algebra: An Introductory Approach. Fourth edition. Devlin: The Joy of Sets: Fundamentals of Contemporary Set Theory. Second edition. Dixmier: General Topology. Driver: Why Math? Ebbinghaus/Flum/Thomas: Mathematical Logic. Second edition. Edgar: Measure, Topology, and Fractal Geometry. Elaydi: An Introduction to Difference Equations. Second edition. Exner: An Accompaniment to Higher Mathematics. Exner: Inside Calculus. Fine/Rosenberger: The Fundamental Theory of Algebra. Fischer: Intermediate Real Analysis. Flanigan/Kazdan: Calculus Two: Linear and Nonlinear Functions. Second edition. Fleming: Functions of Several Variables. Second edition. Foulds: Combinatorial Optimization for Undergraduates. Foulds: Optimization Techniques: An Introduction. Franklin: Methods of Mathematical Economics. Frazier: An Introduction to Wavelets Through Linear Algebra. Gamelin: Complex Analysis. Gordon: Discrete Probability. Hairer/Wanner: Analysis by Its History. Readings in Mathematics.

### **Undergraduate Texts in Mathematics**

Halmos: Finite-Dimensional Vector Spaces. Second edition. Halmos: Naive Set Theory. Hämmerlin/Hoffmann: Numerical Mathematics. Readings in Mathematics. Harris/Hirst/Mossinghoff: Combinatorics and Graph Theory. Hartshorne: Geometry: Euclid and Beyond. Hijab: Introduction to Calculus and Classical Analysis. Hilton/Holton/Pedersen: Mathematical Reflections: In a Room with Many Mirrors. Hilton/Holton/Pedersen: Mathematical Vistas: From a Room with Many Windows. **Iooss/Joseph:** Elementary Stability and Bifurcation Theory. Second edition. Isaac: The Pleasures of Probability. Readings in Mathematics. James: Topological and Uniform Spaces. Jänich: Linear Algebra. Jänich: Topology. Jänich: Vector Analysis. Kemeny/Snell: Finite Markov Chains. Kinsey: Topology of Surfaces. Klambauer: Aspects of Calculus. Lang: A First Course in Calculus. Fifth edition. Lang: Calculus of Several Variables. Third edition. Lang: Introduction to Linear Algebra. Second edition. Lang: Linear Algebra. Third edition. Lang: Short Calculus: The Original Edition of "A First Course in Calculus." Lang: Undergraduate Algebra. Second edition. Lang: Undergraduate Analysis. Lax/Burstein/Lax: Calculus with Applications and Computing. Volume 1.

LeCuyer: College Mathematics with APL. Lidl/Pilz: Applied Abstract Algebra. Second edition. Logan: Applied Partial Differential Equations. Macki-Strauss: Introduction to Optimal Control Theory. Malitz: Introduction to Mathematical Logic. Marsden/Weinstein: Calculus I. II. III. Second edition. Martin: Counting: The Art of Enumerative Combinatorics. Martin: The Foundations of Geometry and the Non-Euclidean Plane. Martin: Geometric Constructions. Martin: Transformation Geometry: An Introduction to Symmetry. Millman/Parker: Geometry: A Metric Approach with Models. Second edition. Moschovakis: Notes on Set Theory. **Owen:** A First Course in the Mathematical Foundations of Thermodynamics. Palka: An Introduction to Complex Function Theory. Pedrick: A First Course in Analysis. Peressini/Sullivan/Uhl: The Mathematics of Nonlinear Programming. Prenowitz/Jantosciak: Join Geometries. Priestley: Calculus: A Liberal Art. Second edition. Protter/Morrey: A First Course in Real Analysis. Second edition. Protter/Morrey: Intermediate Calculus. Second edition. Roman: An Introduction to Coding and Information Theory. Ross: Elementary Analysis: The Theory of Calculus. Samuel: Projective Geometry. Readings in Mathematics. Saxe: Beginning Functional Analysis

# **Undergraduate Texts in Mathematics**

Stanton/White: Constructive Scharlau/Opolka: From Fermat to Minkowski. Combinatorics. Schiff: The Laplace Transform: Theory Stillwell: Elements of Algebra: Geometry, and Applications. Numbers, Equations. Stillwell: Mathematics and Its History. Sethuraman: Rings, Fields, and Vector Spaces: An Approach to Geometric Second edition. Constructability. Stillwell: Numbers and Geometry. Readings in Mathematics. Sigler: Algebra. Silverman/Tate: Rational Points on Strayer: Linear Programming and Its Elliptic Curves. Applications. Toth: Glimpses of Algebra and Geometry. Simmonds: A Brief on Tensor Analysis. Second Edition. Second edition. Readings in Mathematics. Singer: Geometry: Plane and Fancy. Singer/Thorpe: Lecture Notes on Troutman: Variational Calculus and Elementary Topology and Optimal Control. Second edition. Valenza: Linear Algebra: An Introduction Geometry. Smith: Linear Algebra. Third edition. to Abstract Mathematics. Whyburn/Duda: Dynamic Topology. Smith: Primer of Modern Analysis. Wilson: Much Ado About Calculus. Second edition.