A PRIMER OF REAL FUNCTIONS

RALPH P. BOAS, JR.

FOURTH EDITION

REVISED AND UPDATED
BY
HAROLD P. BOAS

THE MATHEMATICAL ASSOCIATION OF AMERICA

The Carus Mathematical Monographs
Number 13

A PRIMER OF REAL FUNCTIONS

Fourth edition

Library of Congress Number 96-77785

Print ISBN 978-0-88385-029-9
Electronic ISBN 978-1-61444-013-0

Printed in the United States of America

Current Printing (last digit):
10 9 8 7 6 5 4 3 2 1

The Carus Mathematical Monographs

NUMBER THIRTEEN

A PRIMER OF REAL FUNCTIONS

RALPH P. BOAS

revised and updated by

HAROLD P. BOAS

THE
CARUS MATHEMATICAL MONOGRAPHS

Published by
THE MATHEMATICAL ASSOCIATION OF AMERICA

The following monographs have been published:

1. *Calculus of Variations*, by G. A. Bliss (out of print)
2. *Analytic Functions of a Complex Variable*, by D. R. Curtiss (out of print)
3. *Mathematical Statistics*, by H. L. Rietz (out of print)
4. *Projective Geometry*, by J. W. Young (out of print)
5. *A History of Mathematics in America before 1900*, by D. E. Smith and Jekuthiel Ginsburg (out of print)
6. *Fourier Series and Orthogonal Polynomials*, by Dunham Jackson (out of print)
7. *Vectors and Matrices*, by C. C. MacDuffee (out of print)
8. *Rings and Ideals*, by N. H. McCoy (out of print)
9. *The Theory of Algebraic Numbers*, second edition, by Harry Pollard and Harold G. Diamond
10. *The Arithmetic Theory of Quadratic Forms*, by B. W. Jones (out of print)
11. *Irrational Numbers*, by Ivan Niven
12. *Statistical Independence in Probability, Analysis and Number Theory*, by Mark Kac
13. *A Primer of Real Functions*, fourth edition, by Ralph P. Boas, Jr., revised and updated by Harold P. Boas
14. *Combinatorial Mathematics*, by Herbert J. Ryser
15. *Noncommutative Rings*, by I. N. Herstein (out of print)
16. *Dedekind Sums*, by Hans Rademacher and Emil Grosswald
17. *The Schwarz Function and its Applications*, by Philip J. Davis
18. *Celestial Mechanics*, by Harry Pollard
19. *Field Theory and its Classical Problems*, by Charles Robert Hadlock
20. *The Generalized Riemann Integral*, by Robert M. McLeod
21. *From Error-Correcting Codes through Sphere Packings to Simple Groups*, by Thomas M. Thompson
22. *Random Walks and Electric Networks*, by Peter G. Doyle and J. Laurie Snell
23. *Complex Analysis: The Geometric Viewpoint*, by Steven G. Krantz
24. *Knot Theory*, by Charles Livingston
25. *Algebra and Tiling: Homomorphisms in the Service of Geometry*, by Sherman Stein and Sándor Szabó

MAA Service Center
P.O. Box 91112
Washington, DC 20090-1112
1-800-331-1MAA   FAX: 1-301-206-9789

Preface to the Fourth Edition

This book is an heirloom and a memorial. The first edition, dedicated "To My Epsilons," appeared in the year that I (the youngest of my father's "little ones") turned six years old. Now, many years later, my siblings being of a nonmathematical bent, the familial interest in the book resides in me. The most fitting tribute I can pay to my father's memory is to pass on a new edition of this monograph to the next generation of epsilons.

The principal change in this edition is the addition of a chapter on integration and some of its applications. This topic was deliberately omitted from previous editions, "reluctantly, because of the many technical details that are needed before one gets to the interesting results." My father eventually decided that it would be acceptable and enjoyable to present some of the interesting results without all the technical details, on the principle that one need not understand the inner workings of the motor to appreciate a drive in the country. Chapter 3 is my reworking of a draft that my father left at his death. Besides revising this material, I have added most of the notes, exercises, and solutions for Chapter 3.

Resetting the book in LaTeX afforded the opportunity to renumber the exercises and to relocate the notes from the end of the volume to the ends of individual sections. I have also made minor revisions throughout Chapters 1 and 2.

I thank Robert B. Burckel for reading the manuscript with great care and for suggesting many improvements.

Harold P. Boas

Texas A&M University
July 1995

Preface to the Third Edition

I. To the beginner. In this little book I have presented some of the concepts and methods of "real variables" and used them to obtain some interesting results. I have not sought great generality or great completeness. My idea is to go reasonably far in a few directions with a minimum amount of special terminology. I hope that in this way I have been able to preserve some of the sense of wonder that was associated with the subject in its early days but has now largely been lost. I hope also that someone who has read this book will be able to go on to one of the many more forbidding systematic treatises, of which there is no lack.

No previous knowledge of the subject is assumed, but the reader should have had at least a course in calculus. In general, each topic is developed slowly but rises to a moderately high peak; a reader who finds the slope too steep may skip to the beginning of the next section.

Since this is not a handbook, but more in the nature of a course of informal lectures, I have not been at all consistent either about the proportion of detailed proof to general discussion or about strict logical arrangement of material.

All phrases like "it is clear," "plainly," "it is trivial" are intended as abbreviations for a statement something like "It should seem reasonable, you should be able to supply the proof, and you are invited to do so." On the other hand, "It can be shown ... " is usually to suggest that the proof is too complicated to give here, or depends on notions that are not discussed here, and that you are not expected to try to supply the proof yourself.

In stating definitions, I have frequently used "if" where I should really have used "if and only if." For example, "If a set is both bounded above and bounded below, it is called *bounded*." This definition is to be understood to carry an additional clause, "and if it is not both bounded above and bounded below, it is not called bounded."

There are a number of exercises, some of which merely supply illustrative material, and some of which are essential parts of the book. An exercise that merely states a proposition is to be interpreted as a demand for a proof of the proposition. Answers to all exercises are given at the end of the book.

Paragraphs in small type deal either with peripheral material or with more difficult questions.

I apologize in advance for whatever mistakes the alert reader may be able to detect. None were intentionally included; nevertheless, the detection and rectification of mistakes is a good exercise, and fosters a healthy skepticism about the printed word.

II. To the expert. Experts are not supposed to read this book at all; since this statement will doubtless be taken as an invitation for them to do so, I must explain what I have tried (and not tried) to do. I have set out to tell readers with no previous experience of the subject some of the results that I find particularly interesting. I have therefore tried to present the material that seemed essential for the results I had in mind, together with as much related material as seemed interesting and not too complicated. Since this is not a systematic treatise, I have deliberately tried not to introduce any concepts or notations, however significant or convenient, that I did not really need to use. I have omitted integration, reluctantly, because of the many technical details that are needed before one gets to the interesting results.

Since this is not a treatise it has not been written like one. The style is deliberately wordy. The axiom of choice is frequently used but never mentioned; this book is not the place to discuss philosophical questions, and, in any case, after Gödel's results, the assumption of the axiom of choice can do no mathematical harm that has not already been done. On the other hand, according to the more recent work of P. J. Cohen, by assuming the axiom of choice rather than its negative we are selecting one kind of mathematics rather than another, say Zermelonian rather than non-Zermelonian. With this selection, there is no point in avoiding the axiom of choice whenever it seems natural to use it, even in cases where it is known to be avoidable.

III. Acknowledgments. I am indebted to my teachers, J. L. Walsh and D. V. Widder, for introducing me to this kind of mathematics; to M. L. Boas and to E. F. and R. C. Buck for criticizing early drafts of the book; and to H. M. Clark and H. M. Gehman for help with the proofreading. I am grateful to several people who have pointed out oversights or suggested improvements, and especially to Richard L. Baker, E. M. Beesley, H. P. Boas, A. M. Bruckner, G. T. Cargo, S. Haber, A. P. Morse, C. C. Oehring, J. M. H. Olmsted, J. C. Oxtoby, A. C. Segal, A. Shuchat, and H. A. Thurston.

******

In preparing this revision, I have tried to resist the temptation to insert additional material; but a considerable number of references have been added to the notes.

Ralph P. Boas, Jr.

Northwestern University
March 1960
June 1965
March 1972
October 1979

Contents

Chapter 1

Sets

1. Sets. In order to read anything about our subject, you will have to learn the language that is used in it. I have tried to keep the number of technical terms as small as possible, but there is a certain minimum vocabulary that is essential. Much of it consists of ordinary words used in special senses; this practice has both advantages and disadvantages, but has in any case to be endured since it is now too late to change the language completely. Much of the standard language is taken from the theory of sets, a subject with which we are not concerned for its own sake. The theory of sets is, indeed, an independent branch of mathematics. It has its own basic undefined concepts, subject to various axioms; one of these undefined concepts is the notion of "set" itself.

From an intuitive point of view, however, we may think of a set as being a collection of objects of some kind, called its *elements*, or *members*, or *points*. We say that a set contains its elements, or that the elements belong to the set or simply are in the set. The normal usage of set, as in "a set of dishes" or "a set of the works of Bourbaki," is fairly close to what we should have in mind, although the second phrase suggests some sort of arrangement of the elements which is irrelevant to the mathematical concept. Sets may, for example, consist of ordinary geometrical points, or of functions, or indeed of other sets. We shall use the words *class*, *aggregate*, and *collection* interchangeably with set, especially to make complicated situations clearer: thus we may speak of a collection of aggregates of sets rather than a set of sets of sets.

If E is a set,* a set H is called a *subset* of E if every element of H is also an element of E. For example, if E is the set whose elements are the numbers 1, 2, and 3, then there are eight subsets of E. Three of them contain one element each; three contain two elements each; one is the set E itself (a subset does not have to be, in any sense, "smaller" than the original set); the eighth subset of E is, by convention, the *empty set*, which is the set that has no elements at all. If H is a subset of E, we write $H \subset E$ or $E \supset H$; sometimes we say that E contains H or that E covers H. If H is a subset of E but is not all of E, we call H a *proper* subset of E.

We write $x \in E$ to mean that x is an element of E. We often say that x is in E, or that x belongs to E, or that E contains x, meaning the same thing. Since the elements of sets are usually things of a different kind from the sets themselves, we should distinguish between the element x and the set whose only element is x. It is often convenient to denote the latter set by $\{x\}$. The notations $x \in E$ and $\{x\} \subset E$ mean the same thing.

A *space* is a set that is being thought of as a universe

* It is traditional to call sets "E," presumably because the French for "set" is "ensemble."

from which sets can be extracted. If Ω is a space and $E \subset \Omega$, then the *complement* of E (with respect to Ω) is the set consisting of all the elements of Ω that are not elements of E. The complement of E is denoted by $C(E)$. For example, if Ω consists of the letters of the alphabet and E of the consonants (including y as a consonant), then $C(E)$ consists of the vowels. If, however,† E consists of the single letter a, then $C(E)$ consists of the letters $b, c, \ldots, z$. If E consists of the entire alphabet, then $C(E)$ is empty. If E is empty, then $C(E) = \Omega$.

Exercise 1.1. Show that $C(C(E)) = E$.

If E and F are two sets, there are two other sets that can be formed by using them, and that occur so frequently that they have special names. One of these sets is the *union* of the two sets, written $E \cup F$ (sometimes called their sum, and written $E + F$); it consists of all elements that are in E or in F (or in both; an element that is in both is counted only once). The other is the *intersection* of the two sets, written $E \cap F$ (sometimes called their product, and written $E \cdot F$ or EF); it consists of all elements that are in both E and F. If $E \cap F$ is empty, then E and F are called *disjoint*; that is, E and F are disjoint if they have no element in common.

Exercise 1.2. Let Ω consist of the 26 letters of the alphabet. Let E consist of all the consonants (including y), and F of all the letters that occur in the words *real functions* (the n is counted only once). Show that (a) $E \cup F = \Omega$; (b) $F \supset C(E)$; (c) $C(F) \subset E$; (d) $F \cap E$ and $C(E)$ are disjoint.

† For simplicity of notation we frequently use, as here, a letter that has just been used as the name of a set, that we are now through with, to denote a different set.

There are various logical difficulties inherent in the uncritical use of the terminology of the theory of sets, and they have given rise to a great deal of discussion. Fortunately, however, they arise only at a higher level of abstraction than we shall attain in the rest of this book, and in contexts that we should consider rather artificial, so that we may safely ignore them hereafter. Some forms of words which appear to define sets may not actually do so, somewhat as some combinations of letters which might well represent English words (for instance, "frong") do not actually do so. For example, although we can safely speak of sets whose elements are sets, we cannot safely talk about the set of all sets whatsoever. Supposing that we could, the set of all sets would necessarily have itself as one of its elements. This is a peculiar property, although there are other ostensible sets that have it, for example, the set of all objects definable in fewer than thirteen words (since this "set" is itself defined in fewer than thirteen words). We might well decide to exclude from consideration those sets that are elements of themselves. The remaining sets do not have themselves as elements; form the aggregate A of all such acceptable sets. Now is A one of the sets that we accept, or one of the sets that we exclude? If we accept A, it does not have itself as an element and so must be included in the aggregate of all sets with this property; that is, A belongs to A, and therefore we do not accept A. On the other hand, if we do not accept A, then A is an element of itself; then since all elements of A are sets that are not elements of themselves, and so are acceptable, we must accept A. Thus if A is a set at all, we are involved in a logical contradiction. The only way out seems to be to declare that the words that seem to define A do not actually define a set.

Another paradoxical property of "the set of all sets" will turn up in §3.

Exercise 1.3. A librarian proposes to compile a bibliography listing those, and only those, bibliographies that do not list themselves. Comment on this proposal.

2. Sets of real numbers. Since we have to start somewhere, I assume that you are familiar with the system of real numbers. I will take completely for granted its algebraic properties—those connected with addition, subtraction, multiplication, and division, and with inequalities. However, there is one property of the real numbers that is less familiar to most people, even though it underlies concepts, such as limit and convergence, that are fundamental in calculus. This property can be stated in many equivalent forms, and the particular one that we select is a matter of taste. I shall take as fundamental the so-called *least upper bound property.*

Before we can state what this property is, we need some more terminology. Let E be a nonempty set of real numbers. We say that E is *bounded above* if there is a number M such that every x in E satisfies the inequality $x \leq M$. For example, the set of all real numbers less than 2 is bounded above, and we can take $M = 2$, or $M = \pi$, or $M = 100$. On the other hand, the set of all positive integers is not bounded above. If E is bounded above, then its *least upper bound,* or *supremum,* is B if B is the smallest M that can be used in the preceding definition. In our example, where E is the set of all real numbers less than 2, the supremum of E is 2. Another way of stating the definition of the supremum of E is to say that it is a number B such that every x in E satisfies $x \leq B$, while if $A < B$ there is at least one x in E satisfying $x > A$. The supremum of E may or may not belong to E. In the example just given, it does not. However, if we change the example so that E consists of all numbers not greater than 2, the supremum of E is still 2, and now it belongs to E.

So far, although we have talked about the supremum of a set, we have not known (except in our illustrative examples) whether there is any such thing. The least upper bound property, which we take as one of the axioms about real numbers, is just that *every nonempty set E that is bounded above does in fact have a supremum.* In other words, if we form the collection of all upper bounds of E, this collection has a smallest element (hence the name). We denote the supremum of E by $\sup E$ or $\sup_{x\in E} x$. When $\sup E$ belongs to E, we sometimes write $\max E$ instead. Thus $\max E$ is the largest element of E if E has a largest element. The greatest lower bound, or *infimum*, of E, denoted by $\inf E$, and the minimum $\min E$, are defined similarly. (See Exercise 2.2.)

An *interval* is a set consisting of all the real numbers between two given numbers, or of all the real numbers on one side or the other of a given number. More precisely, an interval consists of all real numbers x that satisfy an inequality of one of the forms $a < x < b$, $a \leq x < b$, $a < x \leq b$, $a \leq x \leq b$ (where $a < b$), $x > a$, $x \geq a$, $x < a$, or $x \leq a$. Using a square bracket to suggest $\leq$ or $\geq$ and a parenthesis to suggest $<$ or $>$, we shall often use the following notations for the corresponding intervals: (a,b), $[a,b)$, $(a,b]$, $[a,b]$, (a,∞), $[a,\infty)$, $(-\infty,a)$, $(-\infty,a]$. Thus $(0,1]$ means the set of all real numbers x such that $0 < x \leq 1$. (The use of the symbol ∞ in the notation for intervals is simply a matter of convenience and is not to be taken as suggesting that there is a number ∞.)

Exercise 2.1. For each of the sets E indicated below, describe the set of all upper bounds, the set of all lower bounds, $\sup E$, and $\inf E$.

(a) E is the interval $(0,1)$.
(b) E is the interval $[0,1)$.
(c) E is the interval $[0,1]$.
(d) E is the interval $(0,1]$.

(e) E consists of the numbers $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$.

(f) E is the set containing the single point 0.

(g) $E = \{n + (-1)^n : n = 1, 2, \ldots\}$.

(h) $E = \{n + (-1)^n/n : n = 1, 2, \ldots\}$.

(i) E is the set of real numbers x between 0 and π such that $\sin x > \frac{1}{2}$.

Exercise 2.2. Give a detailed definition of $\inf E$, formulate a greatest lower bound property, and prove that it is equivalent to the least upper bound property.

If E is not bounded above, we write $\sup E = +\infty$; if E is not bounded below, we write $\inf E = -\infty$. These are convenient abbreviations, but are not to be interpreted as implying that there are real numbers $+\infty$ and $-\infty$; there are not. We can, if we like, create such infinite numbers and adjoin them to the real number system, but for most purposes it is undesirable to do so. No matter how we introduce infinite numbers, we are bound to make arithmetic worse than it already is: there is one impossible operation to begin with (division by zero), but if we make this operation possible we render even more operations impossible.

Exercise 2.3. Explore the consequences of introducing numbers $+\infty$ and $-\infty$ such that $a/0 = +\infty$ if $a > 0$, and $a/0 = -\infty$ if $a < 0$. Can a reasonable meaning be given to $+\infty + (-\infty)$? To $0 \cdot (+\infty)$?

If a set is both bounded above and bounded below, it is called *bounded.* A bounded nonempty set E is characterized by having $\inf E$ and $\sup E$ both finite, or equivalently by being contained in some finite interval (a, b).

We have supposed in our discussion of upper and lower bounds that we have been considering nonempty sets. So that we shall not have to make reservations about whether or not our sets have elements, we make the (rather peculiar) convention that if E is empty, then $\sup E = -\infty$ and $\inf E = +\infty$. This convention allows us to say, for example, that $\sup(E \cup F)$ is the larger of $\sup E$ and $\sup F$, without having to consider whether E or F may be empty.

Exercise 2.4. If E is not empty, then $\inf E \leq \sup E$; there is strict inequality if E contains at least two points.[‡]

3. Countable and uncountable sets. If we have a set E with (say) five elements, for instance, the fingers on a hand, we can count these elements (or, for short, count E). This means exactly what we might expect: we can point to the elements of E, one by one, naming the integers 1, 2, 3, 4, 5, successively, as we do so. In slightly more formal language, we label all the elements of E by using the integers 1, 2, 3, 4, 5 once each. In still more formal language, we put the elements of E into *one-to-one correspondence* with the integers 1, 2, 3, 4, 5.

It turns out to be useful to extend the concept of counting to sets that have infinitely many elements. Suppose, for example, that E is now the set of all even positive integers. We can no longer point to all the elements of E, one after another, naming successive integers, because E has too many elements. However, we can still imagine labeling all the elements of E with all the positive integers, so that each element of E has a different integer attached to it: we just have to label each even integer with the integer whose double it is, that is, we label the integer $2n$ with the label n. Since the even integers can be put into

[‡] An exercise that simply makes a statement calls for a proof of that statement.

one-to-one correspondence with all the integers in this way, we may reasonably say that there are the same number of each, without, however, committing ourselves as to what the number of integers may be.

It is interesting that there actually are sets that cannot be counted (as we shall soon prove), so that we can classify sets according to whether they can be counted or not. The ones that cannot be counted can be thought of as "bigger" than those that can be counted. (As we have just seen, a set is not necessarily bigger, in this sense, than one of its proper subsets.) Before establishing the existence of uncountable sets, we shall introduce some more terminology and give additional examples of sets that can be counted.

To count a set means to put its elements into one-to-one correspondence with some set of consecutive positive integers starting from 1; this set is not necessarily to be a proper subset of the set of all positive integers. If a set can be counted we call it *countable.* (The empty set is also called countable: here the relevant subset of the positive integers can be thought of, with a little effort, as the empty set.) We can write the elements of a nonempty countable set in some such form as $x_1, x_2, x_3, \ldots$, where a typical element of the set would be denoted by x and the subscripts are the consecutive integers that are used as labels in counting the set. If we start out counting a countable set, either eventually we find a last element, or else the counting process continues indefinitely. In the first case, the set is called *finite*; in the second, *countably infinite.*

The set of all the positive and negative integers together is countable, since we can use the odd positive integers to label all the positive integers and the even positive integers to label all the negative integers, thus:

elements	...	-3	-2	-1	1	2	3	...
labels	...	6	4	2	1	3	5	...

Exercise 3.1. Show similarly that *the union of any two countably infinite sets is countable.*

A less obviously countable set is the set of *lattice points* in the plane: these are the points with both coordinates integral, for example $(1,2)$ or $(-5,18)$. It is easy to see how to count them from the diagram (which, for convenience, shows only the lattice points in the first quadrant; we can count all the lattice points by using Exercise 3.1 repeatedly). Without drawing the picture, we can think of first grouping all the lattice points (m,n) for which $m+n$ is 0, 1, 2, 3, ... , and then counting the groups, one after another. We can alternatively view the diagram as representing the lattice points in the first quadrant as the union of a countable collection of countable sets: the sets are the successive horizontal rows.

$$
\begin{array}{ccccccc}
\vdots & & \vdots & & \vdots & & \\
& \nwarrow & & \searrow & & \nwarrow & \\
(0,2) & & (1,2) & & (2,2) & & \cdots \\
\uparrow & \searrow & & \nwarrow & & \searrow & \\
(0,1) & & (1,1) & & (2,1) & & \cdots \\
& \nwarrow & & \searrow & & \nwarrow & \\
(0,0) & \rightarrow & (1,0) & & (2,0) & \rightarrow & \cdots
\end{array}
$$

Some rather more complicated sets can be discussed by using the fact that *every subset of a countable set is countable.* Expressed in another way, this theorem says that a set whose elements can be labeled with some of the positive integers, using each only once, can equally well be labeled with consecutive positive integers (possibly all of them). To see this, observe that each element of the given subset of a countable set is labeled with some positive integer. Take the element with the smallest label and relabel it 1; then take the element with the smallest label from the rest of the subset (if there is any more of the subset) and relabel it 2; and so on.

It is now easy to see that *the positive rational numbers form a countable set.* A positive rational number can be represented as a fraction p/q in lowest terms, where p and q are positive integers. If we associate the fraction $3/11$ with the lattice point $(3, 11)$, and generally p/q with (p, q), we have the rational numbers in one-to-one correspondence with a subset of the lattice points, that is, with a subset of a countable set. Hence the positive rational numbers form a countable set.

Exercise 3.2. Show that *the union of any countable collection of countable sets is countable.*

Exercise 3.3. The set of all points inside a circle is called a *disk*. Let S be a set of nonoverlapping disks in the plane—that is, no disk in S has a nonempty intersection with any other disk in S. Show that S must be countable.

A still less obvious example is furnished by the algebraic numbers: these are the numbers (real or complex) that can be roots of polynomials with integral coefficients (for example, all rational numbers, $\sqrt{2}$, i, $2\sqrt[3]{3} + \sqrt[5]{7}$). To see that *the set of algebraic numbers is countable*, we notice first that there are only countably many linear polynomials with integral coefficients, countably many quadratic polynomials with integral coefficients, and so on.

Exercise 3.4. Why is this?

The polynomials of a given degree n with integral coefficients have at most n roots each, thus countably many altogether. The aggregate of roots of polynomials of every degree with integral coefficients is accordingly a countable collection of countable sets, and so countable.[1]

A more abstract example is the class of all finite subsets of a given countable set. For, the class of subsets with one element each is countable, the class of subsets with two elements each is countable, and so on. We again have a countable collection of countable sets. (As we shall see later (page 16), the set of *all* subsets of a countably infinite set is not countable.)

The notion of countability can sometimes be used to show the existence of things of a particular kind. For a simple illustration, we prove that *not all real numbers are algebraic.* (A number that is not algebraic is called transcendental.) The real algebraic numbers are, as we know, countable; our first step is to suppose that they have been counted, and represented as decimals. It will simplify the notation somewhat, and do no harm, if we consider only real numbers between 0 and 1.

Every real number between 0 and 1 has an ordinary decimal expansion, for example,

$$\frac{1}{7} = 0.142857142857142857 14 \ldots ,$$
$$\pi - 3 = 0.14159265358979323846 \ldots .$$

Conversely, every such expansion defines a real number between 0 and 1; if we write down, for example,

$$x = 0.12345678910111213141 5 \ldots ,$$

then x is certainly a real number between 0 and 1, although we cannot connect it in any simple way with more familiar numbers.

Now suppose that the real algebraic numbers between 0 and 1 have been counted, so that there are a first, a second, a third, and so on; call them a_1, a_2, a_3, and so on. We then have a column of decimals, which might start like this:

$$\begin{aligned} a_1 &= 0.215367\ldots \\ a_2 &= 0.652489\ldots \\ a_3 &= 0.061259\ldots \\ a_4 &= 0.300921\ldots \\ &\vdots \end{aligned}$$

This list is supposed to contain *all* the real algebraic numbers between 0 and 1; that is, any such number will appear in the list if we go far enough. It is now easy to construct a decimal that does not appear anywhere in this list and so cannot be an algebraic number. For example, if we write down 0.5655, then we have the start of a decimal that differs from a_1 in the first decimal place, from a_2 in the second, from a_3 in the third, and from a_4 in the fourth; obviously it is not going to be any of these four numbers. We can go on in the same way, putting 5 in the nth decimal place if a_n has anything except 5 there, and putting 6 in the nth place if a_n does have 5 there. The resulting decimal differs from each a_n in the nth decimal place and so cannot appear in our hypothetical list of all algebraic numbers, so it is not algebraic.

We can describe the construction more concisely if we let the digits of the nth algebraic number be $a_{n,1}$, $a_{n,2}$, $a_{n,3}$, ... , and construct a new number $b = 0.b_1b_2b_3\ldots$ by taking $b_n = 5$ if $a_{n,n} \neq 5$, and $b_n = 6$ if $a_{n,n} = 5$. (There is no particular significance to the numbers 5 and 6.)

It is sometimes alleged that a proof of this kind is only a "pure existence proof" and furnishes no explicit example of a transcendental number. This is not the case. At least in principle, it is possible to count the algebraic numbers explicitly, find their decimal expansions, and so write down as many digits as we like of at least one transcendental number. The reason that the number π, say, seems more concrete is that π occurs in more contexts than the number we have just been talking about, so that more is known about it; in particular, people have been interested enough to compute billions[2] of decimal places of π already.

We have shown how to find a transcendental number; to show that some given number is transcendental is much harder: it is a problem in number theory and requires deeper methods than the simple argument used here. The transcendence of e is fairly difficult, that of π considerably harder, e^π and $2^{\sqrt{2}}$ much harder again;[3] and it is not known whether π^e is transcendental or not, or even whether it is irrational. Another transcendental number is $0.10100100000010\ldots$, where there are $n!$ zeros after the nth 1. This transcendental number is in a sense simpler than π or e, since we could say without much trouble what any particular digit, say the quadrillionth, is, whereas we cannot, at least at present, do this for π or e.

If we look over the proof of the existence of transcendental numbers, we see that no use was made of the fact that $a_1, a_2, \ldots,$ were algebraic numbers beyond the fact that the algebraic numbers form a countable set. The same argument applies, word for word, to show that if E is any given countable set of real numbers between 0 and 1, then there is a real number between 0 and 1 that is not in E.

Hence no countable set can exhaust the set of real numbers between 0 and 1, or in other words *the set of real numbers between* 0 *and* 1 *cannot be countable.*

Exercise 3.5. The particular interval $(0, 1)$ is not significant; modify the preceding argument, or use the result, to show that the set of real numbers in any interval, however short, is not countable.

Exercise 3.6. Show that the set of real numbers in $(0, 1)$ whose decimal expansions do not contain any 3's is not countable. (The number 3 has no particular significance here.)

Exercise 3.7. Criticize the following "proof" that the set of real numbers between 0 and 1 is countable: First count the decimals that have only one nonzero digit, then those with at most two nonzero digits, and so on; we have then broken the set into a countable collection of countable sets.

We defined a set to be finite if it is countable but not countably infinite. We naturally call a set *infinite*, whether it is countable or not, as long as it is not finite. *Every infinite set contains a countably infinite subset.* To see this, choose a first element x_1, quite arbitrarily. The set with x_1 removed is still infinite (why?); choose an element x_2 from this reduced set; and so on. This process cannot terminate (again, why?), so our set contains the countably infinite subset $x_1, x_2, \ldots$.

Exercise 3.8. Supply answers to the two questions in the preceding paragraph.

Exercise 3.9. Show that if E is any infinite set and F is E with one point deleted, then E and F can be put into one-to-one correspondence with each other. Thus *every infinite set can be put into one-to-one correspondence with a proper subset of itself.*

Exercise 3.10. Establish a one-to-one correspondence between a finite interval and the set of all real numbers.

As another application of the kind of reasoning that we have been using, we prove that *the aggregate A of subsets of any given nonempty set E is "larger" than E, in the sense that A cannot be put into one-to-one correspondence with E, or indeed with any subset H of E.* We shall not make any use of this fact, but it helps to justify the remark on page 4 about the paradoxical character of "the set of all sets."

Exercise 3.11. Use the theorem just stated to show that the notion of the set of all sets *is* paradoxical.

For finite sets we could show without much trouble that there are two subsets of a set with one element (the set itself and the empty set); four subsets of a set with two elements (the whole set, two sets containing one element each, and the empty set); eight subsets of a set with three elements; and generally 2^n subsets of a set with n elements. So our statement is true for finite sets; the general proof will, in fact, cover all sets with at least one element.

Suppose, then, that E is a set with at least one element, and suppose that the collection of all the subsets of E can be put into one-to-one correspondence with a subset H of E. In other words, suppose that the subsets of E can be labeled, say as F_x, where x runs through the elements of H, so that every subset is labeled and no element of H is used twice. We are now going to deduce a contradiction, and this contradiction will show that the alleged one-to-one correspondence cannot exist. We form a subset G of E in the following way. For each x in H, we look at F_x and see whether F_x contains x. If F_x does not contain x, put x in G. (In particular, the x for which F_x is empty is put in G; the x for which F_x is E is not put in G.) Then G is a proper subset of E, and so by assumption corresponds to some z in H, that is, G is F_z. However, by construction, if z is in F_z, then we did not put the element z into G, so G is not

F_z; if, on the other hand, z is not in F_z, then G contains z and F_z does not, so again G is not F_z. We have thus deduced from our initial assumption the contradictory statements that G is F_z and that G is not F_z, so that the initial assumption is untenable.

The preceding theorem shows, for example, that *there are more sets of real numbers than there are real numbers.*

In a rather similar way we could show that there are more real-valued functions, with domain the real numbers, than there are real numbers.

Exercise 3.12. Prove the preceding statement.

We now establish the fact, which seems surprising at first sight, that there are just as many points in a line segment as in a square area: that is, *the real numbers between* 0 *and* 1 *can be put into one-to-one correspondence with the points in a square.* (The points in a square are ordered pairs of real numbers, their coordinates; see page 22.) The general idea of the correspondence is easy to grasp: if we have two real numbers, represented as decimals, we can interlace their digits to get a single real number; conversely, given a real number, we can dissect its decimal expansion to get a pair of real numbers. The details are not quite obvious, however. Suppose, to make decimal representations unique, that we select the nonterminating decimal when there is a choice: thus we take $0.243999\ldots$ instead of $0.244000\ldots$. The obvious procedure of making (p, q) with $p = 0.p_1p_2p_3\ldots$ and $q = 0.q_1q_2q_3\ldots$ correspond to $0.p_1q_1p_2q_2\ldots$ does not work because, for example, the decimal $0.13201020\ldots$ would correspond to (p, q) with $p = 0.1212\ldots$ and $q = 0.300\ldots$, and the latter is a decimal of the kind we are not allowing. Once we recognize this difficulty, we can easily avoid it, however. All that is necessary is to attach to each nonzero digit any string of consecutive zeros that immediately precedes it, and treat these groups of digits as units. Thus $0.13201020\ldots$ now corresponds to (p, q) with $p = 0.1202\ldots$ and $q = 0.301\ldots$; and to (p, q) with $p = 0.003100054\ldots$ and $q = 0.100359\ldots$ corresponds the real number $0.003110030005549\ldots$.

It is often quite hard to exhibit a one-to-one correspondence between two sets explicitly; it is sometimes easier to show that each set can be put into one-to-one correspondence with a subset of the other. The following proposition, known as the Schroeder-Bernstein theorem, is useful in such situations. *If A and B are sets, if A can be put into one-to-one correspondence with a subset of B, and if B can be put into one-to-one correspondence with a subset of A, then A and B can be put into one-to-one correspondence with each other.*[4]

We may suppose that initially the subsets of B and A with which we are concerned are not B and A themselves, since if they are, there is nothing to prove. We are supposed to have two one-to-one correspondences, one (call it S) between A and a subset of B, the other (call it T) between B and a subset of A. Take any element a_1 of A, find its image b_1 in B under S, find the image a_2 of b_1 under T, and so on. This process may lead us back to a_1 after a finite number of steps, or it may go on indefinitely. What it cannot do is to give a chain of elements that crosses itself, for example, so that $a_5 = a_2$; for, if this happened, T would carry b_1 into a_2 and also carry b_4 into a_2. This would contradict the assumption that T is one-to-one unless $b_1 = b_4$, in which case $a_1 = a_4$. In addition, it may happen that a_1 occurs as the image of some element of B under T, and in this case we can prolong the chain backwards from a_1, possibly indefinitely. If any elements of A remain, pick one and start a new chain.

In this way, the elements of A fall into disjoint classes: A_1 consists of elements that belong to chains with one pair of elements (symbolically, $a_1 \xrightarrow{S} b_1 \xrightarrow{T} a_1$); A_2 consists of elements that belong to chains with two pairs of elements ($a_1 \xrightarrow{S} b_1 \xrightarrow{T} a_2 \xrightarrow{S} b_2 \xrightarrow{T} a_1$); and so on. There may be, in addition, elements of A that belong to infinite chains. There are three kinds of infinite chains: those with a first element in A that has no antecedent in B to produce it under T; those with a first element in B; and those with no first element at all. We call these classes A_0, A_{-1}, A_∞ respectively. The classes A_{-1}, A_0, A_1, A_2, ... , A_∞ are all disjoint from each other, and every element of A is in one of them. Let B_k be the class of elements of B that belong to chains containing elements of A_k. Then the B_k are also disjoint and every element of B is in one of them (since we can start a chain from an element of B as well as from an element of A).

Now A_1 and B_1 are obviously in one-to-one correspondence already. Since A_2 consists of pairs of elements of A connected with pairs of elements of B_2, we can put A_2 and B_2 into one-to-one correspondence by pairing off the elements in the obvious way (the first element of a pair in A with the first element of the corresponding pair in B, and so on). We proceed similarly with A_k and B_k for $k = 3, 4, \ldots$. We put A_0 into one-to-one correspondence with B_0 by operating on each chain separately: pair the first element a_1 with its image b_1 under S; pair a_2 with b_2; and so on. In other words, use S on A_0 to carry A_0 into B_0. Here we use the fact that chains of elements in A_0 do not terminate. Similarly, for A_{-1} we use T to establish the one-to-one correspondence. Finally, A_∞ and B_∞ consist of chains that are infinite in both directions; here we can use either S or T. Thus we have established one-to-one correspondences between each A_k and the corresponding B_k, and hence a one-to-one correspondence between all of A and all of B.

As an application of the Schroeder-Bernstein theorem we show that *there are just as many sets of positive integers as there are real numbers* (in contrast to the fact, noted on page 12, that there are only countably many *finite* sets of positive integers). By Exercise 3.10, we only need to consider the real numbers between 0 and 1. If we have a real number r between 0 and 1, we can represent it as a nonterminating decimal, for instance, as $0.20015907\ldots$. To this real number we assign the set of integers 20, 10000, 500000, 9000000, ... In the general case, if the digit $a \neq 0$ occurs in the nth decimal place of r, we incorporate in our set the integer whose decimal representation is a, followed by n zeros. In this way the set consists of different integers, and two different real numbers r generate different sets of integers.

It may seem at first sight that we get only a relatively small proportion of all the possible sets of integers in this way. However, let us now consider an arbitrary set S of positive integers. To S we assign a unique real number as follows. First write the decimal $u = 0.123456789101112\ldots$ (formed by writing down all the positive integers in their natural order). If the integer n occurs in S, replace it in u by a string of zeros. For example, if $S = \{1, 8, 12, 13, 17\}$, the corresponding decimal will be

$$0.02345670910110000141516001819\ldots.$$

If S consists of all even positive integers, its decimal representative is $0.10305070900110013\ldots.$ Thus we have a one-to-one correspondence between all sets of positive integers and a set of real numbers, and another one-to-one correspondence between the set of all real numbers and a class of sets of positive integers. Then, by the Schroeder-Bernstein theorem, there is a one-to-one correspondence between the class of all sets of positive integers and the class of all real numbers.

Exercise 3.13. Show that there are just as many sequences of real numbers as there are real numbers.

NOTES

[1]For a direct proof of the countability of the polynomials with integral coefficients, see Alan Frank, The countability of the rational polynomials, *American Mathematical Monthly* 87 (1980), 810–811.

[2]The third edition of this book cited a paper of D. Shanks and J. W. Wrench, Jr., Calculation of π to 100,000 decimals, *Mathematics of Computation* 16 (1962), 76–99, and noted that one million decimal digits of π were known, but not widely available. Nowadays, finding one million digits of π is a feasible computation on a typical office workstation at a typical university. For an entertaining

account of how the Chudnovsky brothers calculated over two *billion* digits of π on a home-made supercomputer, see Richard Preston, The Mountains of Pi, *The New Yorker*, March 2, 1992, pp. 36–67. The Chudnovskys' own account of their work is D. V. and G. V. Chudnovsky, The computation of classical constants, *Proceedings of the National Academy of Sciences* 86 (1989), 8178–8182. For information about modern algorithms for computing π, see J. M. Borwein and P. B. Borwein, *Pi and the AGM: a Study in Analytic Number Theory and Computational Complexity*, Wiley, New York, 1987.

[3] See Carus Mathematical Monograph number 11, *Irrational Numbers*, by Ivan Niven.

[4] First proved by Felix Bernstein as a 19-year-old student, this result is also known as the Bernstein equivalence theorem. For alternative proofs see M. Reichbach, Une simple démonstration du théorème de Cantor-Bernstein, *Colloquium Mathematicum* 3 (1955), 163; M. S. Hellmann, A short proof of an equivalent form of the Schroeder-Bernstein theorem, *American Mathematical Monthly* 68 (1961), 770; R. H. Cox, A proof of the Schroeder-Bernstein theorem, *American Mathematical Monthly* 75 (1968), 508; and Keith Devlin, *The Joy of Sets*, second edition, Springer, New York, 1993, p. 78.

4. Metric spaces. A *space* is just another name for a set, with emphasis on the possibility of considering its subsets. However, when we call a set a space we usually intend to imply that some sort of additional conditions are to be imposed on the points of the set, which, of course, need not be points in the ordinary sense. A *metric space* is a (nonempty) set in which we can speak of the distance between two points. It is a generalization of the ordinary lines, planes, and three-dimensional spaces of geometry, but only some of the geometrical properties have been preserved.

We require the distance between two points to satisfy the following conditions (which the ordinary distance in Euclidean geometry certainly does satisfy): the distance is a nonnegative real number, zero only if the points coincide; it is the same in either direction; and the sum of two sides of a triangle is at least as much as the third side. In symbols, if $d(x, y)$ denotes the distance between the points x and y, we are to have

(1) $d(x, y) \geq 0$; $d(x, x) = 0$; $d(x, y) > 0$ if $x \neq y$ (positivity);

(2) $d(x, y) = d(y, x)$ (symmetry);

(3) $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality).

We often refer to the distance function for the space as the *metric* of the space.

It turns out that a good deal of geometry depends only on these three properties of distance. Consequently, many facts about ordinary space can be carried over to other spaces that at first sight are very different because their points are not points in the ordinary sense, but may, for example, be functions. The possibility of using geometrical language in metric spaces makes many of their properties more intuitive, although, of course, it may on occasion be misleading as well.

Here are some examples of metric spaces.

$\mathbf{R}_1$, Euclidean one-dimensional space, is the set of all real numbers, with $d(x, y) = |x - y|$.

$\mathbf{R}_2$, Euclidean two-dimensional space, is the ordinary plane of analytic geometry, and the distance is the ordinary distance. The points are ordered pairs of real numbers ("ordered" means that (x_1, x_2) is not the same as (x_2, x_1)). While in elementary mathematics it is usual to denote points in the plane as (x, y), where x is the abscissa and y is the ordinate, it is more in keeping with our notation for other metric spaces to use the letter x to denote a point of the plane with coordinates (x_1, x_2), and similarly y to denote a point of the plane with coordinates (y_1, y_2). The distance from (x_1, x_2) to (y_1, y_2) is

$$\left\{(x_1 - y_1)^2 + (x_2 - y_2)^2\right\}^{1/2}.$$

$\mathbf{R}_n$, Euclidean n-dimensional space, is defined similarly.

Exercise 4.1. Are the following objects metric spaces, or not?

(a) The Euclidean plane with points $x = (x_1, x_2)$, but with distance defined by $d(x, y) = |x_1 - y_1|$.

(b) The set of cities in the United States that have airports, with $d(A, B) =$ "scheduled air travel time from A to B."

(c) The positive real numbers with $d(x, y) = x/y$.

(d) As in (c) but with $d(x, y) = |\log(x/y)|$.

(e) The set of all positive numbers represented by terminating decimals, with $d(x, y) = |x - y|$ rounded to 10 decimal places.

Exercise 4.2. If we use the same points, but change the definition of distance, we get a new space. For instance, change the distance in $\mathbf{R}_2$ by saying that the distance from (x_1, x_2) to (y_1, y_2) is the sum $|x_1 - y_1| + |x_2 - y_2|$. Show that the result is a metric space; show that there is now a nondegenerate triangle such that the sum of two sides is equal to the third side; draw the locus of points that are at unit distance from $(0, 0)$. Do the same things if the distance from (x_1, x_2) to (y_1, y_2) is taken to be the larger of $|x_1 - y_1|$ and $|x_2 - y_2|$.

In our next three examples, the elements of the space will be infinite sequences of numbers. Since the elements of $\mathbf{R}_n$ are sequences of n numbers, these "sequence spaces" may be thought of as infinite-dimensional generalizations of $\mathbf{R}_n$.

The space c_0 is the space of sequences that converge to zero. Its points are sequences of numbers: $\{x_1, x_2, x_3, \dots\}$, where $\lim x_n = 0$. If we denote such a sequence by the single letter x, the distance $d(x, y)$ is defined to be $\sup_{n \geq 1} |x_n - y_n|$. For example, if $x = \{1, \frac{1}{2}, -\frac{1}{3}, -\frac{1}{4}, \frac{1}{5}, \dots\}$ and $y = \{1, -1, 0, 0, \dots\}$, then $d(x, y) = \frac{3}{2}$.

The space m is the space of bounded sequences. Its elements are again sequences of numbers, but now required only to be bounded. The distance is the same as for c_0. Thus we could have $x = \{1, 0, 1, 0, \dots\}$ and $y = \{1, 4, 1, 5, 9, 2, 6, 5, \dots\}$, where no y_n is to be negative or exceed 9. Here the distance $d(x, y)$ cannot be computed until we know more about the law of formation of y. Given that $y_9 = 3$, $y_{10} = 5$, $y_{11} = 8$, and $y_{12} = 9$, however, we see that $d(x, y) = 9$, the largest possible value.

The space l^2 is the space of sequences for which the sum of the squares of the components converges. Its elements are sequences $x = \{x_1, x_2, \dots\}$ with $x_1^2 + x_2^2 + x_3^2 + \cdots < \infty$. We take $d(x, y) = \{\sum_{n=1}^{\infty} (x_n - y_n)^2\}^{1/2}$. To verify the triangle inequality in this case, we need Minkowski's inequality (see page 184):

$$\begin{aligned} d(x, z) &= \left\{\sum (x_n - z_n)^2\right\}^{1/2} \\ &= \left\{\sum \left[(x_n - y_n) + (y_n - z_n)\right]^2\right\}^{1/2} \\ &\le \left\{\sum (x_n - y_n)^2\right\}^{1/2} + \left\{\sum (y_n - z_n)^2\right\}^{1/2} \\ &= d(x, y) + d(y, z). \end{aligned}$$

Next we have some examples of metric spaces whose points are functions.

The space C is the space of continuous functions defined on the closed interval $[0, 1]$. Its elements are continuous functions $x = x(t)$, where $0 \le t \le 1$, with distance

$$d(x, y) = \max_{0 \le t \le 1} |x(t) - y(t)|.$$

For example, we could have $x(t) = \cos \pi t$ and $y(t) = 2t - 1$, and then $d(x, y) = 2$.

The space B is the space of bounded functions defined on $(0,1)$, with $d(x,y) = \sup_{0 \leq t \leq 1} |x(t) - y(t)|$. (We have to write sup instead of max here since $|x(t) - y(t)|$ may fail to attain its maximum.) For example, $x(t)$ could be defined in the interval $[1/(n+1), 1/n)$ as $1 - n^{-1}$ for $n = 1, 2, \ldots$; if 0 stands for the function that takes the single value 0, then $d(x, 0) = 1$.

The set of continuous functions on $[0,1]$ can be made into a metric space different from the metric space C by providing it with the metric

$$d(x,y) = \left\{ \int_0^1 |x(t) - y(t)|^2 \, dt \right\}^{1/2}$$

Exercise 4.3. Any (nonempty) subset of a metric space is again a metric space with the metric that it inherits from the original space.

Exercise 4.4. Any nonempty set whatsoever can be made into a metric space if we introduce a metric by $d(x,y) = 1$ if $x \neq y$, and $d(x,x) = 0$.

5. Open and closed sets. There are several special kinds of sets in metric spaces that occur so frequently that they need to have names. Sections 5, 6, and 7 are mainly devoted to introducing them and making them seem familiar.[1]

A *neighborhood* of a point x is a generalization of a circular disk (the inside of a circle) with center at x: it is[2] the set of all points y with distance less than some positive r from the point x; in symbols, $d(x,y) < r$; we can call the number r the *radius* of the neighborhood. In fact, in $\mathbf{R}_2$ the neighborhoods of x are just circular disks centered at x. In $\mathbf{R}_1$, they are intervals centered at x; in $\mathbf{R}_3$, they are (open) solid spheres. If $r < 1$, the neighborhoods in the space of Exercise 4.4 are single points.

A set is called *bounded* if it is contained in some neighborhood. Thus the interval $(0, 1)$ in $\mathbf{R}_1$ is bounded, but the interval $(1, \infty)$ is not. In $\mathbf{R}_1$, this definition agrees with the one that we used earlier, that a set is bounded if it is both bounded above and bounded below. In $\mathbf{R}_1$, a bounded set has a least upper bound and a greatest lower bound, but there is nothing corresponding to this property in general metric spaces.

Exercise 5.1. Describe a neighborhood in the space C of continuous functions (page 24).

Exercise 5.2. Describe the neighborhoods in the space consisting of the points of $\mathbf{R}_2$ that have two integral coordinates, with the $\mathbf{R}_2$ metric.

Exercise 5.3. If N is a neighborhood of x, and y is another point of N, then N contains a neighborhood of y.

If E is a set in a metric space and x is a point of E, we say that x is an *interior* point of E if some neighborhood of x (possibly a small one) consists exclusively of points of E. The idea of the definition is to give a set an interior that corresponds fairly closely to the intuitive notion of "interior," and in a space like $\mathbf{R}_2$ it succeeds fairly well. For example, the set in $\mathbf{R}_2$ of points (x, y) such that $0 \leq x \leq 1$ and $0 \leq y \leq 1$ is a square, and its interior consists of all the points in the square that are not on its perimeter. On the other hand, the set of rational points of $\mathbf{R}_1$ has no interior points at all. In Exercise 4.4 we considered the space consisting of arbitrary points with $d(x, y) = 1$ or 0

according as $x \neq y$ or $x = y$. In this space every point is an interior point of every set that contains it. If any set in a metric space is regarded as a new metric space in itself, with the original metric, all its points become interior points of the new space. Thus the notion of interior point depends not only on the set we are considering, but also on the space in which the set lies.

Again let E be a set in a metric space; if x is not necessarily a point of E, but every neighborhood of x (with emphasis on its possible smallness) contains both at least one point of E (possibly only x itself) and at least one point of the complement $C(E)$ (again, possibly only x itself), then x is called a *boundary point* of E. The *boundary* of E means the set of all boundary points of E. For a square in $\mathbf{R}_2$, the boundary is just what we might expect: the perimeter. In $\mathbf{R}_1$, the boundary of the interval $[a, b]$ or of the interval (a, b) consists of the two points a and b; so does the boundary of the set consisting of the two points a and b.

The term *frontier point* is sometimes used instead of boundary point, and would perhaps be preferable: the idea of a boundary point has nothing to do with the idea of boundedness. An unbounded set can (and often does) have a nonempty boundary. For example, the interval $(0, \infty)$ in $\mathbf{R}_1$ has the point 0 as its boundary; while $\mathbf{R}_1$, considered as a subset of $\mathbf{R}_2$, has itself as boundary. On the other hand, a nonempty bounded set may have an empty boundary (although this cannot happen in $\mathbf{R}_1$ or $\mathbf{R}_2$).

Exercise 5.4. Describe the interior points and the boundary points of each set:

(i) In $\mathbf{R}_2$, the circumference $x^2 + y^2 = 1$.

(ii) In $\mathbf{R}_1$, the union of the open intervals $(1/(n+1), 1/n)$ as n runs over the positive integers.

(iii) In $\mathbf{R}_2$, the union of the open rectangles of height 1 standing on the intervals of (ii).

(iv) The set in $\mathbf{R}_2$ indicated in the picture on page 33.

Exercise 5.5. For the space of Exercise 5.2, show that the boundary of every set is empty.

Exercise 5.6. Show that E and $C(E)$ have the same boundary.

Exercise 5.7. If E is a set, and B is the boundary of E, then the boundary of B is a subset of B; it may be a proper subset.

Exercise 5.8. Let N be a neighborhood of x of radius r. What can be said about its boundary (a) if the underlying space is $\mathbf{R}_2$? (b) if the underlying space is an arbitrary metric space?

A set all of whose points are interior points is called *open*; a set that contains all its boundary points is called *closed.* As we shall see, a set may be neither open nor closed, and a set may be simultaneously open and closed. These notions depend on the space in which the set lies, as well as on the set itself.

Exercise 5.9. In $\mathbf{R}_1$, the interval (a, b) is open (it is called an open interval for this reason), and the interval $[a, b]$ is closed (and called a closed interval).

Exercise 5.10. Are the intervals $[a, b]$ and (a, b) open, closed, or neither, if considered as subsets of $\mathbf{R}_2$?

Exercise 5.11. Show that the interval $[0, 1)$ is neither open nor closed in $\mathbf{R}_1$.

Exercise 5.12. Neighborhoods are open sets.

Exercise 5.13. Show that the empty set and the whole space are always both open and closed.

Exercise 5.14. Consider the metric space consisting of the intervals $(n, n+\frac{1}{2})$ in $\mathbf{R}_1$, $n = 0, \pm 1, \pm 2, \ldots$, with the $\mathbf{R}_1$ metric. Show that this space has many sets that are simultaneously open and closed.

Exercise 5.15. Is the set of all rational points in $\mathbf{R}_1$ open, closed, or neither?

Exercise 5.16. If the set of the preceding exercise is considered as a space by itself, with the $\mathbf{R}_1$ metric, then it has many subsets that are simultaneously open and closed.

Exercise 5.17. Show that all sets in the space of Exercise 5.2 are simultaneously open and closed.

Exercise 5.18. Show that E is open if (and only if) every point of E is contained in an open subset of E.

The next three exercises give alternative definitions of open set and closed set.

Exercise 5.19. *A set is open if and only if it contains none of its boundary points.*

Exercise 5.20. *A set is open if and only if its complement is closed.*

Exercise 5.21. *A set is closed if and only if its complement is open.*

Exercise 5.22. Define a *limit point* of E as a point x (whether in E or not) such that every neighborhood of x (again, with emphasis on the possible smallness of the neighborhood) contains at least one point of E other than x. The more descriptive term *cluster point* is also used. *A set is closed if and only if it contains all its limit points.*

Exercise 5.23. Every neighborhood of a limit point of E contains infinitely many points of E.

Exercise 5.24. The set of limit points of E is closed; so is the set of boundary points of E.

Exercise 5.25. Find the limit points of the following sets in $\mathbf{R}_1$: (a) the interval $(0,1)$; (b) the set consisting of 1, $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, ... ; (c) the set of rational points in $(0,1)$.

Exercise 5.26. Find the limit points of the following sets in $\mathbf{R}_2$:

(i) The set indicated in the picture on page 33.

(ii) The set of all points with coordinates $(1/m, 1/n)$, where m and n run through the positive integers.

(iii) The set of all points with polar coordinates $(r, 1/n)$, where $0 \leq r \leq 1$ and $n = 1, 2, \ldots$.

If f is a continuous real-valued function on a real interval (see §§12 and 13 for a formal definition), and c is a given real number, then the set of points x for which $f(x) < c$ is open, and the sets where $f(x) = c$ or where $f(x) \geq c$ are closed (see page 88).

Open sets in $\mathbf{R}_1$ have an especially simple structure: if they are not empty, then they *consist of countably many disjoint open intervals.* The word countable is redundant here; any collection of disjoint intervals in $\mathbf{R}_1$ is countable, since each interval contains a rational number that is in no other interval, and so our collection of intervals is in one-to-one correspondence with a subset of the rational numbers.

To show that a given nonempty open set G in $\mathbf{R}_1$ is a union of disjoint intervals is rather tedious in detail, but the idea of the proof is simple. Since G is open and not empty, it contains a point and then a neighborhood of that point. We let this neighborhood, which is an open interval, expand until it is as large as possible. If the enlarged interval does not exhaust G, pick a new point and a neighborhood of it in G, and repeat the process; and so on. No point of G can escape, since if one did we could proceed to enclose it in an interval as before.

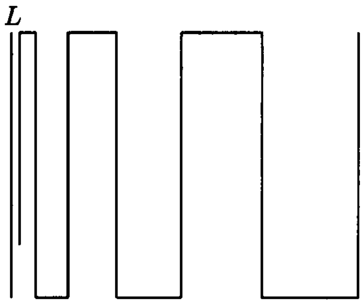
To do this carefully, it is convenient to suppose first that G is bounded. If we show that there is a largest open interval contained in G and containing a given point x, then there is a largest open interval contained in G (for there are only a finite number of intervals of length greater than 1, a finite number of intervals of length greater than $\frac{1}{2}$, and so on). If there is more than one largest interval, we can arrange them according to the magnitudes of their left-hand endpoints. Then we do the same for the next largest intervals, and so on. In this way we can see that the representation of G as a union of disjoint open intervals is unique.

It remains to show that there is a largest open interval that is a subset of G and contains a given point x. There is certainly some interval (a, b) (a neighborhood of x) that is in G. Let B be the least upper bound of numbers b such that the interval (x, b) is in G. Similarly, let A be the greatest lower bound of numbers a such that the interval (a, x) is in G. Since G is bounded, A and B will be finite. Then B is not a point of G, since if it were, G would contain a neighborhood of B, and B would not be an upper bound of the set used in defining it. A similar argument applies to A. Hence the interval (A, B) is in G and cannot be enlarged without including points outside G.

Finally, suppose that G is unbounded. If $G \neq \mathbf{R}_1$ (when $G = \mathbf{R}_1$ there is nothing to prove), then we may as well assume that G is bounded from one side; for if a is a point of the complement of G, we may treat the parts of G to the right of a and to the left of a separately. Suppose, then, that G is bounded below, and let a denote its greatest lower bound. If the complement of G has a finite supremum, say b, then G consists of the interval (b, ∞) together with the countably many disjoint open intervals that make up the bounded part of G between a and b. If $C(G)$ is not bounded above, then we can choose an increasing sequence of points of $C(G)$ tending to ∞, and the part of G between each pair of these points is a countable union of disjoint open intervals by the argument for the bounded case.

If we look for subsets (other than the empty set and the whole space) of $\mathbf{R}_1$ or $\mathbf{R}_2$ that are both open and closed, a little experiment will convince us that there are none. That this conclusion is in fact correct will be proved shortly. The property of $\mathbf{R}_1$ and $\mathbf{R}_2$ (and generally of $\mathbf{R}_n$) that allows only trivial sets to be both open and closed is called *connectedness.* We define this property first for open sets. An open set (in particular, the whole space) is *connected* if it cannot be represented as the union of two disjoint open sets, neither of which is empty. Thus, for example, in $\mathbf{R}_1$, the union of the two open intervals $(0, \frac{1}{2})$ and $(\frac{1}{2}, 1)$ is not connected, since the two intervals are each open and not empty, and are disjoint.

More generally, a set E (not necessarily open) is connected if it cannot be covered by two open sets whose intersections with E are disjoint and not empty. This notion of connectedness is perhaps not completely in accord with intuition. We shall show presently that $\mathbf{R}_1$ and $\mathbf{R}_2$ are connected. The set of rational points in $\mathbf{R}_1$ is not connected, since, for example, it is covered by the open sets defined by the inequalities $x > \sqrt{2}$ and $x < \sqrt{2}$. On the other hand, a set such as the one indicated in the figure, consisting of an oscillating curve condensing toward a line segment, together with the line segment, is connected. (The graph of $y = \sin(1/x)$, together with the segment $-1 \leq y \leq 1$,



has a similar character.) We might think that the line segment L at the left could be separated from the rest of the set, just as two abutting open intervals can be separated from each other. However, an open set that covers any point of L would have to contain a neighborhood of a point of L, and the oscillating part of the graph enters any such neighborhood. Indeed, for the same reason, the set is still connected if we include only the rational points on L, or only the irrational points on L. By combining two sets of this kind it is possible to construct two connected sets inside a square, one of which joins two opposite corners of the square, and the other of which joins the other two corners, although the two sets have no point in common.

It is also possible to have a set that is connected but has the property that after one particular point is removed, the remainder has no connected subsets containing more than one point. (A construction is outlined on page 42.)

A set may be open or not according to the space in which it is considered. However, we must not suppose, merely because connectedness was defined by using open sets, that a set can change from being connected to not being connected if it is considered as a subset of a different space. In fact, the property of being connected, unlike the property of being open, is an intrinsic property of the set. That is, if a set is connected when considered in one space, it is still connected when considered in another space, as long as the metric remains the same on points of the set.

We shall show (equivalently) that the property of not being connected depends only on the set. Suppose that E is a set in a metric space S, that E is not connected, and that S_1 is a subspace of S and S is a subspace of S_2, with $E \subset S_1$. We have to show that, whether we have added points to S (to get S_2) or taken points away from S (to get S_1), the set E is still disconnected.

We start with the assumption that $E \subset A \cup B$, where A and B are open (in S), and the intersections $A \cap E$ and $B \cap E$ are disjoint and nonempty.

To show that E is disconnected in the space S_1, we replace the sets A and B by subsets A_1 and B_1 of S_1, where A_1 contains all the points of A that are in S_1, and B_1 contains all the points of B that are in S_1. These sets still cover E (since $E \subset S_1$), and their intersections with E are still nonempty and disjoint. We need to show that A_1 and B_1 are open subsets of S_1. If p is a point of A_1, then p is certainly a point of A, and since A is open in S, all points of S at distance less than some r from p belong to A. Consequently, all points of S_1 at distance less than r from p belong to A_1. Thus every point of A_1 is an interior point of A_1 (with respect to S_1). Similarly B_1 is open with respect to S_1. Hence E is still disconnected in S_1.

To show that E is disconnected in the space S_2, we inflate the covering sets A and B into subsets A_2 and B_2 of S_2 as follows. If p is a point of A, then since A is open in S, all points of S at distance less than some r from p belong to A. We adjoin to A all points of S_2 at distance less than r from p. Do this for each point p in A, and denote the enlarged set by A_2. Notice that A_2 is an open set relative to S_2 by Exercise 5.18, because each point of A_2 belongs to a neighborhood in S_2 that is contained in A_2. We obtain the open set B_2 by similarly enlarging B. The sets A_2 and B_2 still cover E and have nonempty intersections with E. It may happen that A_2 and B_2 have some points in common, but their intersections with E are still disjoint, because the points we added to A and B are points that lie outside S, and in particular lie outside E. Hence E is still disconnected in S_2.

It is easy to show that $\mathbf{R}_1$ *is connected.* If it were not, it would be the union of two disjoint nonempty open sets. These sets would also be closed (Exercise 5.20), since each is the complement of the other. Hence it is enough to prove that $\mathbf{R}_1$ has no nonempty subset that is both open and closed and is not all of $\mathbf{R}_1$. Suppose there is such a subset; call it G. Then G is made up of countably many open intervals whose endpoints are not in G, as we saw on page 30. On the other hand, these endpoints are boundary points (also limit points) of G, and since G is also closed, they belong to G. This contradiction shows that G cannot exist.

To show that $\mathbf{R}_2$ is connected, it is sufficient to show that it contains no set, other than itself and the empty set, that is both open and closed. Let E be such a set; let $P \in E$ and $Q \in C(E)$. Consider the infinite straight line through P and Q, regarded as a space L, with the distance between two points in L equal to their distance in $\mathbf{R}_2$. Then L is a copy of $\mathbf{R}_1$. The sets $E \cap L$ and $C(E) \cap L$ are both open and closed in L, and neither of them is empty (since one contains P and the other contains Q). This contradicts the connectedness of $\mathbf{R}_1$.

Exercise 5.27. Show that every nonempty set in $\mathbf{R}_2$, except for $\mathbf{R}_2$ itself, has a nonempty boundary.

Exercise 5.28. The *closure* of a set E is the union of E and the set of all limit points of E. Show that it is also the union of E and the set of all boundary points of E, and that it is closed.

Exercise 5.29. If E is not finite, must some point of E be an interior point of the closure of E?

Exercise 5.30. What are the closures of the sets in Exercise 5.25?

Exercise 5.31. A neighborhood of x consists of the points y such that $d(x,y) < r$. Show that in $\mathbf{R}_1$ or $\mathbf{R}_2$ the closure of this neighborhood is the set of points y such that $d(x,y) \leq r$. Is this true in every metric space?

An important fact about closed sets is that *the union of two closed sets is still closed*, and *the intersection of two closed sets is still closed.* Since a closed set is characterized by containing all its limit points, to prove the first statement we consider two closed sets E_1 and E_2, and any limit point p of $E_1 \cup E_2$; we are to show that $p \in E_1 \cup E_2$. I claim that p is a limit point of either E_1 or E_2. For if p is not a limit point of E_1, then there is a neighborhood of p containing no point of E_1 (other than p), and if p is not a limit point of E_2, then there is a neighborhood of p containing no point of E_2 (other than p); the smaller of these neighborhoods contains no point of either E_1 or E_2 (other than p), contradicting that p is a limit point of $E_1 \cup E_2$. Thus p is a limit point of at least one of E_1 and E_2. Since E_1 and E_2 are both closed, $p \in E_1$ if p is a limit point of E_1, and $p \in E_2$ if p is a limit point of E_2. Hence p belongs to at least one of E_1 and E_2, that is, p belongs to $E_1 \cup E_2$. Therefore $E_1 \cup E_2$ contains all its limit points and so is closed.

To see that the intersection of two closed sets is closed, consider a limit point q of $E_1 \cap E_2$. Every neighborhood of q contains points, other than q, belonging both to E_1

and to E_2. This property makes q a limit point of E_1 and a limit point of E_2. Since E_1 and E_2 are both closed, q belongs to both and so to $E_1 \cap E_2$. Therefore $E_1 \cap E_2$ contains all its limit points and so is closed.

Exercise 5.32. Prove by induction that the union of any finite number of closed sets is closed.

Exercise 5.33. Show that *the intersection of any collection (possibly uncountable) of closed sets is closed.*

Exercise 5.34. Find an example to show that the union of a countably infinite collection of closed sets is not necessarily closed.

Exercise 5.35. By arguing in a similar way, or by considering the complements of the sets concerned, show that *the union of any collection of open sets is open*; that the intersection of finitely many open sets is open; and that the intersection of a countably infinite collection of open sets need not be open.

Exercise 5.36. Let x be a given point of a metric space. If N_1 is the neighborhood of x consisting of all y such that $d(x, y) < r$, and N_2 is the neighborhood of the same x consisting of all y such that $d(x, y) < r/2$, show that the closure of N_2 is a subset of N_1. Must it be a proper subset?

A set E is called *perfect* if either it is empty, or it is closed and every point of E is a limit point of E. A closed interval in $\mathbf{R}_1$ is perfect; so is the union of a finite number of closed intervals. In the next section, we shall meet examples of more general perfect sets.

NOTES

[1] Two references for the concepts in §§5–7 are K. Kuratowski, *Topology*, English translation, Academic Press, New York, 1966; and A. Wilansky, *Topology for Analysis*, Krieger Publishing, Melbourne, FL, 1983.

[2] Many people use the word "neighborhood" to mean any set that *contains* one of our neighborhoods.

6. Dense and nowhere dense sets. A set E is *everywhere dense* (or, for short, just *dense*) if its closure is the whole space. In particular (in $\mathbf{R}_n$, equivalently), E is dense when every point of the space is a limit point of E. A set is *nowhere dense* if its closure contains no neighborhoods. In other words, E is nowhere dense either if E is empty, or if every neighborhood in the space contains a subneighborhood that is disjoint from E. The rational points in $\mathbf{R}_1$ form a dense set. A set consisting of a finite number of points of $\mathbf{R}_1$ is nowhere dense.

Note that "nowhere dense" is not the negative of "everywhere dense." A set that fails to be everywhere dense must have the property that its closure fails to fill some neighborhood (possibly small). If a set fails to be nowhere dense, its closure must fill some neighborhood, but not necessarily the whole space.

Occasionally we need to say that a set E is dense in an interval, or in some other set, or is nowhere dense in an interval. Such phrases are self-explanatory.

Exercise 6.1. Consider the space Ω whose elements are the integral points $1, 2, 3, \ldots$ of $\mathbf{R}_1$ with the $\mathbf{R}_1$ metric. Describe the neighborhoods in Ω. Is the set containing the single point 1 a nowhere dense set in Ω?

Exercise 6.2. If a closed set contains no neighborhoods, it is nowhere dense.

Since every point of a perfect set is a limit point of the set, it would appear that a nonempty perfect set must have a great many points. It is therefore somewhat surprising that a nonempty set can be both nowhere dense and perfect.

Exercise 6.3. $\mathbf{R}_1$, considered as a subset of $\mathbf{R}_2$, is nowhere dense and perfect.

In $\mathbf{R}_1$, there are no nowhere dense perfect sets of such a simple structure. An example of a nowhere dense perfect set in $\mathbf{R}_1$ is the *Cantor set*, which may be used as the basis for constructing many examples of sets and functions with remarkable properties. This set is constructed as follows.

Consider the closed interval $[0, 1]$ in $\mathbf{R}_1$. Remove the open middle third, that is, the interval $(\frac{1}{3}, \frac{2}{3})$. Next remove the open middle thirds of the two remaining intervals, that is, remove $(\frac{1}{9}, \frac{2}{9})$ and $(\frac{7}{9}, \frac{8}{9})$. Then remove the open middle thirds of the four remaining intervals; and so on indefinitely.

What remains? In the first place, what has been removed is a union of open sets (indeed, of open intervals), and so is open; what remains is its complement (with respect to $[0, 1]$), and so is a closed set. The endpoints of the various middle thirds were not removed, so they remain; and since the remaining set is closed, every limit point of endpoints remains. For example, if we start from $\frac{1}{3}$ and take the closest endpoint in the second step $(\frac{1}{3} - \frac{1}{9} = \frac{2}{9})$, then the closest endpoint in the third step $(\frac{1}{3} - \frac{1}{9} + \frac{1}{27})$, and so on, the (only) limit point of this set of points is $\frac{1}{3} - \frac{1}{9} + \frac{1}{27} - \cdots = \frac{1}{4}$. Thus there are, in fact, limit points of endpoints which are not endpoints.[1] The Cantor set is the set that remains after we have removed all the middle thirds: it consists of all the endpoints and of their limit points.

Exercise 6.4. Does the Cantor set contain the point $\sqrt{\pi} - 1 = 0.77245\ldots$?

We have just observed that the Cantor set is closed. It also contains no interval, since the total length of the intervals removed is $\frac{1}{3}+\frac{2}{9}+\frac{4}{27}+\cdots=1$. Therefore the Cantor set is nowhere dense (Exercise 6.2). To show that it is perfect, we have to show that each of its points is a limit point. Since the limit points of endpoints are naturally limit points of the set, it is merely a question of showing that the endpoints are limit points. Consider, for example, the point $\frac{1}{3}$. To the left of it there is an interval of length $\frac{1}{3}$ from which we remove the middle third, leaving an interval of length $\frac{1}{9}$ adjacent to the point $\frac{1}{3}$; then we remove the interval $(\frac{7}{27}, \frac{8}{27})$, leaving an interval of length $\frac{1}{27}$ adjacent to $\frac{1}{3}$; and so on. In any neighborhood of $\frac{1}{3}$ there will always be a short interval that is not removed at some step, and this interval will contain an endpoint belonging to a subsequent step. Hence $\frac{1}{3}$ is a limit point of endpoints. A similar argument applies to any other endpoint.

It will be useful later on to have a purely arithmetical construction for the Cantor set. We use the expansion of real numbers in "decimals" in bases 2 and 3 (binary and ternary expansions), instead of in base 10. For example,

$$0.10010110\ldots \qquad \text{(base 2)}$$

means

$$\frac{1}{2}+\frac{0}{2^2}+\frac{0}{2^3}+\frac{1}{2^4}+\frac{0}{2^5}+\frac{1}{2^6}+\frac{1}{2^7}+\frac{0}{2^8}+\cdots,$$

while $0.10010110\ldots$ (base 3) means

$$\frac{1}{3}+\frac{0}{3^2}+\frac{0}{3^3}+\frac{1}{3^4}+\frac{0}{3^5}+\frac{1}{3^6}+\frac{1}{3^7}+\frac{0}{3^8}+\cdots.$$

In base 2 we have only the digits 0 and 1, whereas in base 3 we have the digits 0, 1, and 2. Thus $0.020202\ldots$ (base 3) equals $\frac{1}{4}$. (The reasoning is the same that is used in summing a repeating decimal in base 10: if $x = .0.0202\ldots$ (base 3), then multiplying by 3^2 moves the ternary point two places to the right, so $9x = 2.0202\ldots = 2 + x$, whence $8x = 2$.) Note that this is not the expansion of $\frac{1}{4}$ that we used above in showing that $\frac{1}{4}$ belongs to the Cantor set; but it is equivalent since

$$\begin{aligned}\frac{2}{3^2} + \frac{2}{3^4} + \cdots &= \frac{3}{3^2} - \frac{1}{3^2} + \frac{3}{3^4} - \frac{1}{3^4} + \cdots \\ &= \frac{1}{3} - \frac{1}{3^2} + \frac{1}{3^3} - \frac{1}{3^4} + \cdots .\end{aligned}$$

Now let us express all the numbers between 0 and 1 in base 3. The numbers whose first digit is 1 are between $\frac{1}{3}$ and $\frac{2}{3}$ (inclusive), so they fill the first interval whose interior was discarded in forming the Cantor set. The numbers whose first digit is 0 fill the interval $[0, \frac{1}{3}]$, and the subset of these whose second digit is 1 are the numbers between $\frac{1}{9}$ and $\frac{2}{9}$, that is, they fill one of the intervals whose interior was excluded in the second step of the construction. Every number excluded so far has a 1 in the first or second place in its ternary expansion; so do the endpoints, but they also have expansions that have no 1's. For example, $\frac{1}{3} = 0.10000\ldots = 0.02222\ldots$ (just as $0.1000\ldots = 0.09999\ldots$ in base 10), and $\frac{2}{3} = 0.12222\ldots = 0.2000\ldots$. By continuing in this way, we see that the Cantor set can be described as consisting precisely of those numbers that have expansions in base 3 containing no 1's. The endpoints are the numbers of this kind whose expansions in base 3 end in all 0's or all 2's.

Exercise 6.5. Show that the Cantor set is uncountable.

Actually we can say more: the Cantor set can be put into one-to-one correspondence with the set of all real numbers between 0 and 1. Recall that the points of the Cantor set are just the numbers that can be written in base 3 using only 0's and 2's. With each such number x associate the number obtained by halving each digit in the ternary expansion of x and interpreting the result in base 2. In this way we get every number between 0 and 1 from some point of the Cantor set, and the endpoints of excluded intervals give rise to two different representations of the same number, for example, $\frac{1}{3} = 0.022\ldots$ in base 3 and $\frac{2}{3} = 0.2000\ldots$ in base 3 both yield $\frac{1}{2} = 0.0111\ldots = 0.1000\ldots$ in base 2. The correspondence is one-to-one between the Cantor set, less the countable set of endpoints of excluded intervals, and the set of all real numbers, less the countable set of numbers that have double representations in base 2. By making these countable sets correspond to each other, we obtain a one-to-one correspondence between the Cantor set and the set of real numbers between 0 and 1.

Exercise 6.6. Does the Cantor set contain any irrational points? If so, find one explicitly.

We can use the Cantor set to construct the set mentioned on page 33, which becomes totally disconnected after the removal of a single point. Let P be the point in the plane with coordinates $(\frac{1}{2}, 1)$, and join P to the points of the Cantor set by straight line segments. Now delete the points with irrational ordinates on the lines going to endpoints of complementary intervals of the Cantor set, and delete the points with rational ordinates on the lines going to the other points of the Cantor set. The resulting set is connected, but when P is removed it contains no connected subsets except for single points. It is known as the Cantor teepee.[2]

A metric space is called *separable* if it contains a countable set that is everywhere dense. For example, $\mathbf{R}_1$ is separable because the rational numbers form a countable dense set.

Exercise 6.7. Show that $\mathbf{R}_2$ is separable.

The space c_0 (sequences tending to zero; page 23) is separable. As a countable dense set we may choose the set consisting of all sequences of rational numbers in which only a finite number of elements are different from 0 (for instance, $(1, 0, 0, \dots)$ or $(\frac{2}{3}, -\frac{5}{2}, \frac{3}{4}, 0, 0, \dots)$). This set is countable for the same reason that the set of all finite subsets of the rational numbers is countable (page 12). To show that it is dense in c_0, we recall that the distance between two points $x = (x_1, x_2, \dots)$ and $r = (r_1, r_2, \dots)$ in c_0 is $\sup |x_k - r_k|$. Let x be any point in c_0. We want to select a point r, where all r_k are rational and only finitely many of them are not zero, so that $\sup |x_k - r_k|$ is small. Taking an arbitrary positive ϵ to measure the desired degree of smallness, choose N so large that $|x_n| < \epsilon$ for $n > N$ (compare the more detailed discussion of convergence in §8). Let r_1, r_2, ... , r_N be rational numbers such that $|x_n - r_n| < \epsilon$ for $n = 1, 2, \dots, N$. Then $r = (r_1, r_2, \dots, r_N, 0, 0, \dots)$ is the required point.

In the space C of continuous functions, the set of all polynomials is everywhere dense, as we shall show in §19.

Exercise 6.8. Show that there are uncountably many polynomials.

Exercise 6.9. Show that there are countably many polynomials all of whose coefficients are rational.

Exercise 6.10. Show that if the set of all polynomials is dense in C, so is the set of all polynomials all of whose coefficients are rational. Deduce that C is separable.

The space m of bounded sequences of real numbers (page 24) is an example of a space that is not separable. We can see this as follows.

Exercise 6.11. Show that there is an uncountable set S in m, all the points of S being represented by sequences containing only 0's and 1's.

Exercise 6.12. What is the distance in m between any two of the points of the set S of Exercise 6.11?

Take any everywhere dense set E in m. We shall put the set S (described in Exercise 6.11) into one-to-one correspondence with a subset of E. This will show that E contains an uncountable subset and so is uncountable. Therefore m can contain no countable dense set. To put S into one-to-one correspondence with a subset of E, we proceed as follows. Take any point p of S. There is a point q of E at distance less than $\frac{1}{2}$ from p, since E is everywhere dense. The distance from q to any other point s of S is more than $\frac{1}{2}$, since $1 = d(p, s) \leq d(p, q) + d(q, s) < \frac{1}{2} + d(q, s)$. In this way we have associated a different point of E with each point of S, so that E has at least as many points as S, and hence uncountably many.

Exercise 6.13. Show similarly that the space B of bounded functions (page 25) is not separable.

NOTES

[1] When I was a sophomore, an older student showed me the Cantor set, but admitted that he had never been able to find a limit point that was not an endpoint. He later left mathematics for biology.

[2] See, for example, Lynn Arthur Steen and J. Arthur Seebach, Jr., *Counterexamples in Topology*, second edition, Springer-Verlag, New York, 1978, pp. 145–147; J. Cobb and W. Voxman, Dispersion points and fixed points, *American Mathematical Monthly* 87 (1980), 278-281.

7. Compactness. It is frequently desirable to be able to assert that a set possesses a limit point, even though we may not be able to say just where that limit point is. Suppose, for example, we are trying to prove that a bounded, continuous, real-valued function f, defined on a given set E in $\mathbf{R}_1$, has a maximum. That is, we want to show that there is a point x of E such that $f(x)$ is actually equal to the supremum of all numbers $f(y)$ for y in E. (You are supposed to have at least a rough idea already of what a continuous function is; the formal definition will be discussed in §13.) We are then trying to show that there is an x in E such that $f(x) \geq f(y)$ for all y in E.

We supposed that f is bounded, that is, the values $f(y)$ for y in E form a bounded set of real numbers. Such a set has a finite least upper bound M, so $f(y) \leq M$ for all y in E. There must then be points x_n in E such that $f(x_n) > M - 1/n$, since otherwise some smaller number than M would also be an upper bound. Moreover, we can suppose for the same reason that the x_n are all different, if E is not a finite set (and if E is finite there is nothing to prove). If the x_n have a limit point x in E, the continuity of f makes $f(x) \geq M$ because $f(x_n) \geq M - 1/n$. Since $f(x) \leq M$ in any case, by the definition of M, it follows that $f(x) = M$.

Now under what circumstances can we assert that a set $\{x_n\}$ of infinitely many different points of E has a limit point in E? This cannot always be true: for example, in $\mathbf{R}_1$ the set $E_1 = \{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots\}$ has the limit point 0, but this limit point is not in E_1. The set $E_2 = \{1, 2, 3, 4, \dots\}$ has no limit point at all in $\mathbf{R}_1$. The *Bolzano-Weierstrass theorem* exists for the purpose of furnishing simple conditions that ensure that a set contains a limit point. It says that *an infinite bounded set in* $\mathbf{R}_1$ *has a limit point in* $\mathbf{R}_1$; if the set is also closed, it then contains a limit point. Assuming the truth of this theorem, for the moment, we see that *a bounded, continuous, real-valued function attains a maximum on any closed bounded set in* $\mathbf{R}_1$. The examples (i) $f(x) = x$, with E the interval $(0, 1)$; and (ii) $f(x) = 1 - x^{-1}$, with E the interval $[1, \infty)$; show that both "closed" and "bounded" are essential conditions on the set E.

Exercise 7.1. Deduce from the Bolzano-Weierstrass theorem that the boundedness of the *function* is redundant: it follows from the other hypotheses.

We may prove the Bolzano-Weierstrass theorem by a process that has been suggested as a method for catching a lion in the Sahara Desert.[1] We surround the desert by a fence, and then bisect the desert by a fence running (say) north and south. The lion is in one half or the other; bisect this half by a fence running east and west. The lion is now in one of two quarters; bisect this by a fence; and so on: the lion ultimately becomes enclosed by an arbitrarily small enclosure. The idea is actually employed in the Heligoland bird trap.[2]

The essential point in applying this idea to our problem is that if a set E has infinitely many points and lies in some finite interval I, then at least one half of I must contain infinitely many points of E. Let I_2 be one of the halves, containing infinitely many points of E, and bisect I_2. Again, one of the halves of I_2 contains infinitely many points of E; call such a half I_3. Continue this process. We obtain a nested sequence of intervals I_2, I_3, ..., each containing infinitely many points of E. The left-hand endpoints of the intervals I_n form a set which is bounded above (since it is in I) and so has a least upper bound x. Every neighborhood of x contains some I_n, since the length of I_n tends to 0, and so contains infinitely many points of E. That is, x is a limit point of E.

Exercise 7.2. Prove the Bolzano-Weierstrass theorem for $\mathbf{R}_2$.

As an application of the Bolzano-Weierstrass theorem together with the uncountability of the set of real numbers, we prove a theorem about the approximation of a function by the partial sums of its power series. *Let $f(x) = \sum_{k=0}^{\infty} a_k x^k$, with the series converging in $|x| < 1$. Suppose that, for each x in $[0,1)$, $f(x)$ coincides with some partial sum of its power series, that is, for each x there is an n such that $\sum_{k=n+1}^{\infty} a_k x^k = 0$. Then f is a polynomial.*[3]

Let E_n be the set of points x in $[0, \frac{1}{2}]$ for which the partial sum $\sum_{k=0}^{n} a_k x^k = f(x)$. Since there are countably many integers and uncountably many points in $[0, \frac{1}{2}]$, some E_n must be uncountable, and so infinite. Then E_n has a limit point in $[0, \frac{1}{2}]$, and f coincides on E_n with a polynomial, the same at every point of E_n. But an analytic function cannot coincide with a polynomial on a set having a limit point inside the interval of convergence without being itself a polynomial. (See §24 for more about analytic functions.)

The statement that every infinite bounded set has a limit point makes sense in any metric space, although it may fail to be true. For example, it fails for the space consisting of the rational points of $\mathbf{R}_1$ with the $\mathbf{R}_1$ metric. We can see this by considering the set consisting of the rational approximations 1, 1.4, 1.41, 1.414, 1.4142, ... to $\sqrt{2}$. This set is bounded; it is closed (since $\sqrt{2}$ is not in the space); but it contains no limit point. The Bolzano-Weierstrass property fails here because the space has, so to speak, too few points. It can also fail for a space that has too many points. For example, take the space B whose points are bounded functions on $[0, 1]$. We have seen (Exercise 6.13) how to construct infinitely many such functions, each at distance 1 from the others; such a set of points of B cannot have a limit point.

A set E with the property that every infinite subset of E has a limit point in E used to be called compact; we have just seen that bounded closed sets in $\mathbf{R}_1$ or $\mathbf{R}_2$ have this property. However, the term *compact* is now usually applied to sets with a less intuitive property (formerly called bicompactness). A set is called compact if, whenever it is covered by a collection of open sets, it is also covered by a finite subcollection of these sets. (To say that E is covered by a collection of sets $\{G\}$ means that each point of E is in at least one G.)

To see how the property of compactness can be applied, we shall use it to show that *a continuous real-valued function f defined on a compact set E in a metric space has a maximum on E.* First we show that the function is bounded. To each x in E we assign a neighborhood N with center at x such that $f(y) < f(x) + 1$ for all y in N. We can do this because f, being continuous, does not change much if we change x only a little. These neighborhoods are open sets, and every x is in at least one of them, so since E is compact, a finite number of them cover E. Let these be $N_1, N_2, \dots, N_n$. If x_k is the center of N_k, then $f(x)$ cannot exceed the largest of the *finite* class of numbers $f(x_k) + 1$; so f is bounded above. Similarly f is bounded below.

Now we suppose that f does not attain a maximum on E and deduce a contradiction. The values $f(x)$ for x in E form a bounded set, as we have just seen, so they have a supremum M, which we are supposing is not attained. To each x we can then assign a neighborhood N such that $f(y) < f(x) + \frac{1}{2}(M - f(x))$ for all y in N (by the continuity of the function f). Again finitely many of these neighborhoods, $N_1, N_2, \ldots, N_n$ (not the same neighborhoods as before), cover E. Let M' ($< M$) be the largest of the numbers $f(x_k)$, where x_k is the center of N_k. Then for every y in E we find, by taking an x_k that is in the same N_k as y, that

$$\begin{aligned} f(y) &< f(x_k) + \tfrac{1}{2}(M - f(x_k)) = \tfrac{1}{2}f(x_k) + \tfrac{1}{2}M \\ &\leq \tfrac{1}{2}(M' + M). \end{aligned}$$

Thus the values of $f(y)$, for y in E, have the upper bound $\frac{1}{2}(M' + M)$, which is less than M, contradicting the definition of M as the supremum of f.

Exercise 7.3. If E is a set in $\mathbf{R}_1$, and E is covered by a finite number of open intervals, then we can reduce the number of intervals to arrange that no point of E is in more than two of them, and the reduced set still covers E.

Exercise 7.4. Show that a closed subset of a compact set is compact.

The preceding proof indicates that it is desirable to be able to recognize a compact set when we meet one. In $\mathbf{R}_1$ this is easy: the *Heine-Borel theorem* states that *a set in $\mathbf{R}_1$ is compact if it is closed and bounded.* The proof is almost the same as the proof of the Bolzano-Weierstrass theorem. Suppose that the conclusion of the Heine-Borel theorem is false. Then we have a set E which is closed and bounded, and a collection $\{G\}$ of open sets that covers E, whereas no finite subcollection of sets G covers E. The set E lies in some finite interval I; bisect I. The part of E in one of the halves of I must fail to be covered by a finite subcollection of the sets G; for, if both parts of E could be so covered, so could the whole set E. Let this half of I

be I_2. Now bisect I_2 and continue the process as before to define a limiting point x. As with the Bolzano-Weierstrass theorem, we see that every neighborhood of x contains an interval I_n that in turn contains a part of E that cannot be covered by a finite subcollection of the sets G (and hence is infinite). On the other hand, x is in E since E is closed, and so x can be covered by one of the sets G. Since G is open it contains a neighborhood of x, and this neighborhood contains an interval I_n if n is large enough. The part of E in this I_n is covered by a finite number (namely, one) of the sets G. We have thus arrived at a contradiction by supposing the Heine-Borel theorem false.

Exercise 7.5. Prove the Heine-Borel theorem in $\mathbf{R}_2$.

Exercise 7.6. Let S be a compact set in $\mathbf{R}_2$ containing at least three noncollinear points. (a) Show that there is a triangle of largest area with vertices in S. (b) Does the diameter of S have to equal the length of one of the sides of this triangle?

You should have noticed that the hypotheses on the set in the Heine-Borel and Bolzano-Weierstrass theorems are the same. The similarity of both hypotheses and proofs should suggest a close relationship between the theorems. As a matter of fact, the theorems stand or fall together: for a given metric space, if either of them is true, so is the other. However, we shall omit the proof.[4]

Exercise 7.7. Show directly the noncompactness of the two examples given on page 45 of infinite sets that do not contain limit points.

Exercise 7.8. In $\mathbf{R}_1$, let E be the interval $(0,1]$. With each x associate the open interval $(\frac{1}{2}x, 2x)$. These intervals cover E. Show that no finite number can cover E, and explain why this fact does not contradict the Heine-Borel theorem.

Exercise 7.9. The set $[0, \infty)$ in $\mathbf{R}_1$ is covered by the open intervals $(n-1, n+1)$, $n = 0, 1, 2, \ldots$. No finite number of these intervals can cover the set. Explain why this fact does not contradict the Heine-Borel theorem.

Exercise 7.10. The set E in $\mathbf{R}_1$ consisting of the rational numbers between 0 and 1 is not closed. It is covered by open sets in the following way: let x be covered by an open interval of length $\frac{1}{10}$ centered at x. A finite number of these open intervals do cover E. Explain why this fact does not contradict the Heine-Borel theorem.

Exercise 7.11. Let the set E in $\mathbf{R}_1$ be the closed interval $[0, 1]$. Let each $x \neq 0$ in E be covered by the interval $[\frac{1}{2}x, 2x]$, and let 0 be covered by $[0, \frac{1}{10}]$. The covering intervals are not open, yet a finite number of them do cover E. Explain why this fact does not contradict the Heine-Borel theorem.

It is also worth observing that not only is a set in $\mathbf{R}_1$ (or in any $\mathbf{R}_k$) compact if it is closed and bounded, but *if it is compact it is necessarily closed and bounded.* To see this, suppose that E is a nonempty compact set in $\mathbf{R}_1$. It cannot be all of $\mathbf{R}_1$, so its complement contains a point x. Consider all open finite intervals G such that the closure of G does not contain x. Among these intervals G there are certainly neighborhoods of each point of E, since if $y \in E$, a neighborhood of y that reaches only halfway to x will not contain x in its closure. Therefore E is covered by the open sets G. Since E is compact, a finite collection of sets G covers E; let these sets be $G_1, \ldots, G_n$. Hence E is, in the first place, bounded, since it is contained in the union of a finite number of finite intervals. Since the closures of $G_1, \ldots, G_n$ do not contain x, the complements of these closures all do contain x, and therefore so does their intersection. The closures of the G_k are closed, their complements are open, so the intersection of the complements is open. Accordingly, x is in an open subset of $C(E)$; since x can be any point of $C(E)$, it follows that every point of $C(E)$ is in an open subset of $C(E)$. Therefore $C(E)$ is open (Exercise 5.18). Hence E is closed (Exercise 5.21).

NOTES

[1]H. Pétard, A contribution to the mathematical theory of big game hunting, *American Mathematical Monthly* 45 (1938), 446–447.

[2]See John Sack, *Report from Practically Nowhere*, Harper, New York, 1959, p. 23.

[3]Proof suggested by W. C. Fox and R. R. Goldberg.

[4]See, for example, M. E. Munroe, *Introduction to Measure and Integration*, Addison-Wesley, Cambridge, MA, 1953, p. 31; E. T. Copson, *Metric Spaces*, Cambridge University Press, 1968, p. 79.

8. Convergence and completeness.

8. Convergence and completeness. A great deal of analysis is concerned with infinite series or sequences of functions. The idea of an infinite series of numbers is more intuitive: we write, for instance, $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots$, meaning that we add up the terms one by one, thus forming the successive "partial sums"

$$\frac{1}{2}, \ \frac{1}{2} + \frac{1}{4}, \ \frac{1}{2} + \frac{1}{4} + \frac{1}{8}, \ \ldots,$$

and call the limit of these (if there is one) the sum of the infinite series. (It is assumed that you already have some idea of the meaning of "limit"; see, however, page 54.) In this case the partial sums are $\frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \ldots$, and it is clear that the sum of the series must be 1. It is even clearer if we use the formula for the sum of a geometric progression:

$$\frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^n} = \frac{\dfrac{1}{2} - \dfrac{1}{2^{n+1}}}{1 - \frac{1}{2}} = 1 - \frac{1}{2^n}.$$

In general, if we write the infinite series

$$a_1 + a_2 + a_3 + \cdots ,$$

we expect to calculate first a_1, then $a_1 + a_2$, then $a_1 + a_2 + a_3$, and so on, and call the limit (if any) of these partial sums the sum of the series. Note that, for example, $1-1+1-1+\cdots$ and $(1-1)+(1-1)+(1-1)+\cdots$ must be different infinite series, since the first has successive partial sums 1, 0, 1, 0, ... , whereas all the partial sums of the second are 0.

However, we have not actually defined "an infinite series" or even a particular infinite series: we have merely suggested a way of attaching a numerical value to a formula that has no meaning a priori. In order to see how we could actually give a definition, we notice that what was really used in the suggested calculation is the sequence of partial sums. Technically speaking, a sequence is a function from the positive integers to some space: see §12. However, in less formal language, we may think of a sequence of numbers as being a collection of numbers that have been labeled with the positive integers, preserving their order, and are not necessarily all different: thus 1, 0, 1, 0, ... , is a sequence (the labeling being implicit); generally a sequence can be written in some such way as a_1, a_2, a_3, ... , or $\{a_n\}_{n=1}^{\infty}$, or simply $\{a_n\}$ if it is clear from the context that we are talking about a sequence and not a set. We must always distinguish a sequence from the set of numbers that appear in it. Any countably infinite set can be arranged as a sequence (in many ways), but a sequence need have only a finite number of different elements in it. (Note that a "finite sequence" such as $\{5, 12, 13\}$ is not a sequence according to our definition.) We can now define the infinite series $a_1 + a_2 + a_3 + \cdots$ as meaning no more and no less than the sequence whose elements are the partial sums $s_1 = a_1$, $s_2 = a_1 + a_2$, $s_3 = a_1 + a_2 + a_3$, and so on.

Conversely, any sequence of numbers defines a corresponding infinite series of which it is the sequence of partial sums. For example, the sequence 1, 0, 1, 0, ... is the sequence of partial sums of the series $1 - 1 + 1 - 1 + \cdots$. Generally, the sequence $s_1, s_2, s_3, \ldots$ is the sequence of partial sums of the series $s_1 + (s_2 - s_1) + (s_3 - s_2) + \cdots$.

The notion of sequence is more general than the notion of series since we can have a sequence whose elements are sets, or, indeed, points of any space we like; but there is no associated series unless there is an operation of addition for points of the space.

We shall naturally say that *an infinite series converges if its associated sequence of partial sums converges; otherwise the series is said to diverge.* To make this definition precise, we must therefore define what we are to mean by the convergence of a sequence. If $\{s_n\}$ is a sequence of real numbers, we say that it converges to the limit L if $|s_n - L|$ eventually becomes and remains as small as we please. We then write $s_n \to L$. In more formal language, $s_n \to L$ if, given any positive number ϵ (with emphasis on its possible smallness), there is an integer N (usually rather large) such that $|s_n - L| < \epsilon$ provided that $n > N$. This definition extends immediately to sequences whose elements are points of any metric space: we have only to replace $|s_n - L|$ by $d(s_n, L)$. Thus the sequence of points $\{(\cos(1/n), \sin(1/n))\}$ of $\mathbf{R}_2$ converges to $(1, 0)$; if elements x_n of the space C of continuous functions are defined by $x_n(t) = t^n(1-t)^n$, $0 \le t \le 1$, then the sequence $\{x_n\}$ converges to the element 0 of C (since $t(1-t) \le \frac{1}{4}$).

Although defining the infinite series $a_1 + a_2 + \cdots$ to mean the sequence $\{s_1, s_2, s_3, \ldots\}$ of its partial sums seems natural enough, we are under no compulsion to use this definition, and indeed it is not always the most reasonable definition to use. For some purposes it is better to define $a_1 + a_2 + \cdots$ to mean some other sequence, for example,

$$\frac{s_1}{1}, \quad \frac{s_1 + s_2}{2}, \quad \frac{s_1 + s_2 + s_3}{3}, \quad \ldots$$

or

$$\frac{s_1 + s_2}{2}, \quad \frac{s_2 + s_3}{2}, \quad \frac{s_3 + s_4}{2}, \quad \ldots .$$

It can be shown that either of these definitions preserves the sum of any convergent series.[1] In addition, either one makes some divergent series converge; for example, the divergent series $1-1+1-1+\cdots$ has $s_1 = 1$, $s_2 = 0$, $s_3 = 1$, and so on, so that either of the suggested definitions would give it the sum $\frac{1}{2}$.

We are now going to discuss some properties of sequences of points in a metric space; the idea of infinite series has served to motivate the introduction of sequences, but we have no further use for infinite series just now. (Some theorems are more conveniently formulated in terms of series than in terms of sequences; we shall have an example on page 171.)

If a sequence converges to a limit L, its elements eventually become and remain close to each other. Indeed, let N be so large that, for $n > N$, we have $d(s_n, L) < \epsilon/2$; let $m > N$; then $d(s_m, L) < \epsilon/2$ also. By the triangle inequality, $d(s_m, s_n) \le d(s_n, L) + d(s_m, L) < \epsilon$. In other words, $d(s_m, s_n)$ can be made as small as we please by taking m and n simultaneously sufficiently large.

A sequence $\{s_n\}$ with the property that its elements eventually become and remain close to each other, in the sense just described, is called a *Cauchy sequence*. It may or may not converge to a limit in the space. For example, in the metric space of rational numbers with the $\mathbf{R}_1$ distance, the sequence

$$\{1, 1.4, 1.41, 1.414, 1.4142, \dots\}$$

of decimal approximations to $\sqrt{2}$ is a Cauchy sequence. In fact, if $m > N$ and $n > N$, then s_m and s_n agree at least through the Nth decimal place, and so $|s_m - s_n| < 10^{-N}$. However, the sequence does not converge to a point of the space, because $\sqrt{2}$ is an irrational number.

Exercise 8.1. A sequence $\{s_n\}$ of real numbers whose consecutive terms eventually become close together (that is to say, $|s_n - s_{n+1}| \to 0$) need not be a Cauchy sequence. Give an example.

A metric space for which every Cauchy sequence converges to a point of the space is called *complete.* The metric space of rational numbers is not complete. However, $\mathbf{R}_1$ is complete, as we shall see shortly. It is, in fact, always possible to make a metric space complete by adding new points to it to make a larger space, somewhat as the real numbers can be constructed from the rational numbers. We shall not discuss this construction here, however.[2]

We shall now show that the completeness of $\mathbf{R}_1$ follows from the least upper bound property that we took as fundamental in §2. Let $\{s_n\}$ be a Cauchy sequence. Then if ϵ is a given positive number, there is an N such that $|s_m - s_n| < \epsilon$ if $n > N$ and $m > N$. In the first place, the different numbers s_n must form a bounded set. To see this, take $\epsilon = 1$, find the corresponding N, and take a convenient $m > N$. Then $s_n = s_m + (s_n - s_m)$, whence $|s_n| \le |s_m| + 1$ if $n > N$. Since the finite set

$$\{s_1, s_2, s_3, \dots, s_N\}$$

is bounded, so is the whole set of numbers s_n.

Now let L_k be the supremum of the set consisting of all the different s_n for $n > k$. Since taking a larger k means that we are considering the supremum of a smaller set of numbers, we have $L_k \geq L_{k+1} \geq \cdots$. Let L be the infimum of the set of all L_k. We shall show that $s_n \to L$. Let ϵ be any positive number, and take the N corresponding to this ϵ, so that if $n > N$ and $m > N$, we have $|s_m - s_n| < \epsilon$. By the definition of infimum, there must be an L_k, with $k > N$, between L and $L + \epsilon$. Since L_k is the supremum of the set of s_n with $n > k$, there is an s_m (with $m > k$) between $L_k - \epsilon$ and L_k; hence $L - \epsilon \leq s_m \leq L + \epsilon$. Since s_n differs from s_m by at most ϵ, we have $L - 2\epsilon \leq s_n \leq L + 2\epsilon$. Since 2ϵ is just as arbitrary as ϵ, we now know that s_n can be made arbitrarily close to L by taking n large enough; that is, $s_n \to L$.

Exercise 8.2. Show, conversely, that if the completeness of $\mathbf{R}_1$ is assumed, then the least upper bound property follows.

Exercise 8.3. If $\{s_n\}$ is a sequence of points of $\mathbf{R}_1$ such that $s_n \leq s_{n+1}$, and $s_n \leq M$ for all n, then $\{s_n\}$ converges. In other words, *an increasing bounded sequence has a limit.* (This is often taken as the fundamental property of $\mathbf{R}_1$, in place of the least upper bound property.)

Exercise 8.4. Prove that $\mathbf{R}_2$ is complete.

We shall see in §17 that the space C is complete; and, more generally, the space of continuous functions on any given compact set is complete.

We noticed in Exercise 4.3 that a subset of a metric space is again a metric space (if the same metric is used).

Exercise 8.5. A subset of a complete metric space is not necessarily a complete metric space. Give an example of this.

However, we can show that *a nonempty closed subset E of a complete metric space S is also a complete metric space* (again, using the original distance). Let $\{x_k\}$ be a Cauchy sequence of points of E. It is also a Cauchy sequence of points of S, since the distances in E and S are the same. Since S is complete, $x_k \to x_0$, where $x_0 \in S$. We have only to show that $x_0 \in E$. There are two cases to consider. In the first case, all but a finite number of the x_k coincide. Clearly they must then coincide with x_0, which accordingly belongs to E. In the second case, there are infinitely many different x_k's. The condition $x_k \to x_0$ then shows that x_0 is a limit point (in S) of the set consisting of these x_k, and hence x_0 is a limit point of the set E (considered as a subset of S). Since E is closed, it contains its limit points (Exercise 5.22); so $x_0 \in E$.

We need to distinguish carefully between the limit of a sequence and a limit point of the set consisting of the (different) elements of the sequence. For example, the sequence $\{0, 0, 0, \dots\}$ in $\mathbf{R}_1$ has the limit 0, but the set of elements of the sequence has only one point and so has no limit point. On the other hand, the set of elements of the sequence $\{0, 1, \frac{1}{2}, \frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \frac{7}{8}, \dots\}$ in $\mathbf{R}_1$ has two limit points (0 and 1), but the sequence has no limit. The necessity of this distinction was what forced us to consider two cases in the preceding proof.

However, there is a close connection between the limit of a sequence and the limit points of the set consisting of the different elements of the sequence.

Exercise 8.6. Show that if a sequence converges to L and has infinitely many different elements, then the set consisting of all the different elements of the sequence has L as a limit point, indeed as its only limit point.

Exercise 8.7. Hence show that if a sequence converges to a limit and its elements belong to a closed set, the limit of the sequence belongs to the same set.

Exercise 8.8. Hence show that if E is a nonempty compact set in $\mathbf{R}_1$, then E has a largest element.

We may conclude from Exercise 8.6 that if a sequence converges, the set of its elements does its best to have the limit of the sequence as a limit point. In the other direction, we can easily see that *if a set E has a limit point L, there is a sequence of points of E having L as its limit.* In fact, there is a point x_1 in E at distance less than 1 from L; then there is a point x_2 in E at distance less than $\frac{1}{2}$ from L; and so on. In applications, the set E often consists of the elements of a sequence; if these elements have a limit point, then there is a subsequence converging to the limit point. We shall refer to this as the *subsequence principle*; it has many uses.

For example, one of the ways to prove the fundamental theorem of algebra is to show (A) the absolute value of a nonconstant polynomial $P(z)$ has infimum zero in the complex plane, so that there is a bounded sequence $\{z_n\}$ on which $P(z_n) \to 0$; (B) by the subsequence principle, a subsequence of $\{z_n\}$ has a limit z_0, and hence $P(z_0) = 0$. Some nineteenth-century proofs stop after step (A), apparently regarding (B) as self-evident.

Exercise 8.9. If E is any bounded sequence in $\mathbf{R}_1$ ("bounded" means that the elements of E form a bounded set), then E contains at least one convergent subsequence. (This is an analogue for sequences of the Bolzano-Weierstrass theorem for sets.)

As an illustration of the use of the subsequence principle, we discuss the diameter of a set E. The *diameter* is defined to be the supremum of the distances between points of E; in symbols, $\operatorname{diam} E = \sup d(x, y)$ for x and y in E. For example, in $\mathbf{R}_2$ the diameter of a circle of radius 1 is 2; so is the diameter of the (open) area inside the circle, and so is the diameter of the closed area. The diameter of the set that consists of the three points $(0, 0)$, $(0, 1)$, and $(1, 0)$ is $\sqrt{2}$. It may happen that there are no points x and y in E for which $d(x, y) = \operatorname{diam} E$, even when E is bounded and so has a finite diameter. For example, when E is a neighborhood in $\mathbf{R}_2$, there are no two points of E that are $\operatorname{diam} E$ apart.

However, there must be two points of E that are $\operatorname{diam} E$ apart if E is a *compact* nonempty set in $\mathbf{R}_1$ or $\mathbf{R}_2$. Since E is bounded, $\operatorname{diam} E$ is finite. We must be able, by the definition of diameter, to find pairs of points x_1, y_1; x_2, y_2; ... such that $d(x_n, y_n) > \operatorname{diam} E - 1/n$. Infinitely many of the x_n or of the y_n will be different (otherwise we have already found x and y such that $d(x, y) = \operatorname{diam} E$). If there are infinitely many different x_n, then they have a limit point (by the Bolzano-Weierstrass theorem), and we can select a subsequence having this limit point as a limit. If there are only a finite number of different x_n, one of them, say x_1, will occur infinitely often, and then a sequence all of whose elements are equal to x_1 will have x_1 as a limit. We can proceed similarly with the y_n that correspond to the x_n already selected. The result is that we have sequences, which for simplicity of notation we may call $\{x_n\}$ and $\{y_n\}$ again, such that $x_n \to x_0$, and $y_n \to y_0$, and $d(x_n, y_n) \to \operatorname{diam} E$. Then we must have $d(x_0, y_0) = \operatorname{diam} E$. For, on the one hand, $d(x_0, y_0)$ cannot exceed $\operatorname{diam} E$, since E is closed and so x_0 and y_0 are in E. On the other hand, the triangle inequality shows that

$$d(x_n, y_n) \le d(x_n, x_0) + d(y_n, y_0) + d(x_0, y_0),$$

so that $\operatorname{diam} E \leq d(x_0, y_0)$.

Exercise 8.10. Define the distance between two sets F and G to be $\inf d(x, y)$ for x in F and y in G. Show that if F and G are closed and nonempty sets in $\mathbf{R}_2$, and F is bounded, then there are points x in F and y in G such that $d(x, y)$ is the distance between F and G.

Exercise 8.11. If N is a neighborhood of y, consisting of all x such that $d(x, y) < r$, is $\operatorname{diam} N = 2r$? (Consider (a) $\mathbf{R}_1$ or $\mathbf{R}_2$; (b) general metric spaces.)

Exercise 8.12. Show that E and its closure have the same diameter.

NOTES

[1] Any method, other than convergence, for attaching a sum to an infinite series is called a method of summability. See O. Szász, *Introduction to the Theory of Divergent Series*, Hafner, New York, 1948; G. H. Hardy, *Divergent Series*, Clarendon Press, Oxford, 1949; K. Zeller and W. Beekmann, *Theorie der Limitierungsverfahren*, second edition, Springer, New York, 1970.

[2] See, for example, E. T. Copson, *Metric Spaces*, Cambridge University Press, 1968, §35; Irving Kaplansky, *Set Theory and Metric Spaces*, second edition, Chelsea, New York, 1977, pp. 90–93.

9. Nested sets and Baire's theorem.

Suppose that we have two sets E_1 and E_2, that $E_1 \supset E_2$, and that E_2 is not empty. Then there is at least one point that is simultaneously in both sets, since $E_1 \cap E_2 = E_2$. Similarly, if we have a finite number of sets that are *nested*: $E_1 \supset E_2 \supset E_3 \supset \cdots \supset E_n$, and if the last set E_n is not empty, then there is at least one point that is simultaneously in all the sets. There is nothing corresponding to this when we have infinitely many nested sets, none of which is empty; the intersection of all the sets may quite well be empty nevertheless. Consider the following three examples: (i) E_n is the open interval $(0, 1/n)$ in $\mathbf{R}_1$; (ii) E_n is the closed interval $[n, \infty)$ in $\mathbf{R}_1$; (iii) E_n is the set of those points x in the metric space of rational points in $\mathbf{R}_1$ that are subject to the inequality $|x - \sqrt{2}| < 1/n$. In each of these cases, the intersection of all the sets E_n is empty.

We now state conditions that prevent a nested collection of sets from having an empty intersection. *Cantor's nested set theorem*: *If* $E_1 \supset E_2 \supset E_3 \supset \cdots$; *if each* E_n *is closed and not empty*; *if the underlying space is complete*; *and if* $\operatorname{diam} E_n \to 0$; *then there is exactly one point in the intersection of all the* E_n.

Cantor's theorem involves three conditions besides the condition that the sets are nested and not empty: closedness and small diameter of the sets, and completeness of the space. In each of our three examples of nested sets with empty intersection, a different one of these conditions fails.

To prove Cantor's theorem, let x_n be any point belonging to E_n. The sequence $\{x_n\}$ is a Cauchy sequence since if $m > n$, then $x_m \in E_n$, and so $d(x_n, x_m) \le \operatorname{diam} E_n$, which approaches zero. Since the space is complete, $\{x_n\}$ has a limit in the space. If we select any E_n, then the x_k belong to E_n when $k \ge n$, so the limit belongs to E_n because E_n is closed. That is, the limit is in every E_n. Finally, there cannot be two points that are in every E_n, since $\operatorname{diam} E_n$ is at least as large as the distance between any two of the points of E_n.

It is sometimes useful to have the following weaker theorem: *if we keep all the hypotheses of Cantor's theorem except that we no longer require that* $\operatorname{diam} E_n \to 0$, *but require instead that the* E_n *are compact, we still can say that the intersection of the* E_n *is not empty* (although it may now contain more than one point). Since we have kept the hypothesis that E_n is closed, in any $\mathbf{R}_k$ our new hypothesis amounts to supposing that E_n is also bounded. In $\mathbf{R}_k$, the generalized theorem is an easy application of the subsequence principle: a sequence $\{x_n\}$ consisting of one point from each set has a subsequence that has a limit, and this limit is a point of the required kind.

In the general case we have to proceed differently. Let us cover E_1 by neighborhoods of all its points, each of diameter at most 1. Because E_1 is compact, a finite number of these, say $N_1, \ldots, N_p$, cover E_1. One of the N_k must contain points of all the E_n (from $n = 1$ onward). For if N_1 is disjoint from E_{m_1}, and N_2 is disjoint from E_{m_2}, and so on, then, because the E_n are nested, no N_k contains points of E_n for $n = \max(m_1, \ldots, m_p)$, contradicting that the N's cover E_1. Thus it must be true, as asserted, that some N contains points of all the E_n for $n = 1$, 2, and so on. We now replace each E_n by the closure of $N \cap E_n$ to get closed nested sets of diameter at most 1. Repeat this reasoning with neighborhoods of diameter at most $\frac{1}{2}$ covering E_2; then with neighborhoods of diameter at most $\frac{1}{3}$ covering E_3; and so on. We obtain nested subsets of the original E_n, with diameters approaching zero, and we can then apply Cantor's theorem in its original form.

We can sometimes use Cantor's theorem to show that a set in which we are interested cannot be empty. If we want to know that there are things with a certain property, we can be sure that there are some if we can exhibit them as the intersection of a nested collection of sets satisfying the hypotheses of Cantor's theorem. However, it is often more efficient not to use the nested sets directly, but to use instead another theorem that is a consequence of Cantor's theorem. To state this new theorem we need to introduce a new class of sets: namely, sets that can be represented as unions of countably many sets, each nowhere dense. ("Countably many" includes none, or one, or a finite number, as well as a countable infinity.) A set that can be represented as the union of countably many nowhere dense sets is called a set of *first category.* (Since the name is not at all descriptive, the name *meager set* has been suggested as an alternative; the reason for using this name will appear shortly.)

In $\mathbf{R}_1$, any set consisting of a finite number of points is of first category. So is any countable set, for example, the set of all rational numbers, since although this set is everywhere dense, it is the union of countably many sets, each consisting of a single point. The Cantor set, being nowhere dense, is of first category but uncountable. If we form the union of the Cantor set with the set of all rational points, we obtain a set of first category which is both everywhere dense and uncountable.

Sets that are not of first category are said to be of *second category.* Since the empty set is nowhere dense, it is of first category; hence *a set of second category cannot be empty.* This fact is the basis for the principal use of the notion of category: if we can show that a set is of second category, then it must contain points. We can sometimes exhibit the aggregate of things of a particular kind as a set of second category; there must then be things of this kind. Some examples will be given in §10. The technique for applying this idea depends on *Baire's theorem,* which states that *a complete metric space is of second category.*

Before proving Baire's theorem we make a few remarks. First, the completeness of the metric space is an essential part of the theorem. The metric space whose points are the rational points of $\mathbf{R}_1$, with the $\mathbf{R}_1$ distance, is not complete; each point of the space, regarded as a set, is nowhere dense; the whole space is therefore the union of countably many nowhere dense sets.

We cannot state without reservation that a countable set is of first category, although the preceding example may make this seem plausible. As we noted in Exercise 6.1, a single point need not form a nowhere dense set. In particular, single points are neighborhoods in any space that contains only a finite number of points. The next exercise considers the other extreme.

Exercise 9.1. Show that if all the points of a space are limit points, any set containing only a single point is nowhere dense.

Exercise 9.2. Hence show that Baire's theorem implies that both $\mathbf{R}_1$ and the Cantor set are uncountable.

We now prove Baire's theorem. Let $\{E_n\}$ be a sequence of nowhere dense sets in a complete metric space. We are to prove that there is at least one point of the space that is in none of the E_n. The idea of the proof is that since E_1 is nowhere dense, its complement contains a neighborhood N_1; then N_1, in turn, contains a subneighborhood N_2 that is in the complement of E_2 as well as in the complement of E_1; and so on. In this way we obtain a nested sequence of neighborhoods that are disjoint from more and more of the E_k, and a common point of all the neighborhoods cannot be in any E_k.

To show that there actually is a common point, we must take a certain amount of care so that we can apply Cantor's theorem. First select a neighborhood N_1 in $C(E_1)$ whose radius is less than 1. Let M_1 be the closure of the concentric subneighborhood of N_1 whose radius is half the radius of N_1. Then M_1 is a subset of N_1 (by Exercise 5.36), and so M_1 is disjoint from E_1. Next, since E_2 is nowhere dense, M_1 contains a neighborhood N_2 that is in $C(E_2)$ (as well as in $C(E_1)$). Let M_2 be the closure of the concentric subneighborhood of N_2 whose radius is half the radius of N_2. Continuing in this way, we obtain nested closed sets M_k, whose diameters tend to 0, which are not empty, and which have the property that M_k is disjoint from E_1, $E_2, \ldots, E_k$. The common point of all the M_k is a point of the required kind, since it cannot be in any E_k.

10. Some applications of Baire's Theorem.

In this section, I assume that you are familiar with elementary notions of derivatives and integrals (antiderivatives) as discussed in a first calculus course.

(i) A PROPERTY OF REPEATED INTEGRALS. Let f be a continuous real-valued function on a real interval, say $[0, 1]$. Let f_1 be any integral of f (that is, the constant of integration may be chosen arbitrarily), f_2 any integral of f_1, and so on. If some f_k vanishes identically, so does f: we have only to differentiate f_k repeatedly. The following proposition generalizes this simple fact: *if for each x there is an integer k, possibly differing from one x to another, such that $f_k(x) = 0$, then f vanishes identically.*

To prove this theorem, let E_k be the set of points x for which $f_k(x) = 0$; then our hypothesis says that every x in $[0, 1]$ is in some E_k. By Baire's theorem, not every E_k is nowhere dense. Hence there is some k for which the closure of E_k fills an interval I_k. For this particular k, since f_k is continuous and vanishes on E_k, we must have $f_k(x) = 0$ for every x in I_k; and as we observed above, this implies that f vanishes identically on I_k. If I_k is not all of $[0, 1]$, we repeat this argument with any remaining part of $[0, 1]$, and so on. In this way we have $f(x) = 0$ for all points x of an everywhere dense set; and since f is continuous, it then follows that $f(x) = 0$ for every x in $[0, 1]$.

Thus if $f(x) \not\equiv 0$, then no matter how the integrals f_k are selected, there must be some x (indeed, a somewhere dense set of x) such that $f_k(x) \neq 0$ for every k.

(ii) A CHARACTERIZATION OF POLYNOMIALS. Again consider a continuous real-valued function f on $[0,1]$. If f has an nth derivative that is identically zero, it is easily proved, for example by Taylor's theorem with remainder (see page 187), that f coincides on $[0,1]$ with a polynomial (of degree at most $n-1$). The following theorem generalizes this in the spirit of example (i). *Let f have derivatives of all orders on $[0,1]$, and suppose that at each point some derivative of f is zero. That is, for each x there is an integer $n(x)$ such that $f^{(n(x))}(x) = 0$. Then f coincides on $[0,1]$ with some polynomial.*[1]

We can start the proof just as in (i). Let E_n be the set of points x for which $f^{(n)}(x) = 0$. By hypothesis, every x is in at least one E_n. By Baire's theorem, there is an interval I in which some E_n is everywhere dense. Since $f^{(n)}$ is a continuous function, $f^{(n)}(x) \equiv 0$ in I, and f coincides in I with a polynomial. We may expand I as much as possible, so that there is no larger interval containing I in which $f^{(n)}$ is identically zero. If I is not all of $[0,1]$, repeat the reasoning in any remaining part of $[0,1]$, and so on. In this way, we see that there is a set of intervals whose union is everywhere dense and in each of which f coincides with a polynomial. We still have to show that f coincides with the same polynomial in all the intervals.

To do this, we are going to apply Baire's theorem again to the nowhere dense set H that is left when we remove the interiors of our dense set of intervals from $[0,1]$. (Here "interior" means interior with respect to the metric space $[0,1]$. Thus, the points 0 and 1 could possibly get removed.) We first need to show that H is perfect. In the first place H is closed, since it is obtained by removing a collection of open sets from a metric space. If H is not perfect, then H must have a point y that is not a limit point. This point is the common endpoint of two maximal intervals in each of which f coincides with some polynomial. Then if n equals the larger of the degrees of the two polynomials, $f^{(n+1)}(x) = 0$ for x in both intervals, and at the endpoint by the continuity of $f^{(n+1)}$. Therefore the intervals were not maximal, and the point y does not belong to H after all.

The preceding discussion shows that H is perfect, and we may suppose it to be not empty (otherwise there was only one interval to begin with, and there is nothing to prove). Consider H as a complete metric space. By Baire's theorem for H, some E_n is everywhere dense in some neighborhood in H, that is, in the part of H that is in some open interval J. In other words, there is an open interval J that contains points of H, and $f^{(n)}(x) = 0$ for every x in $J \cap H$ (with the same n). Since every point of $J \cap H$ is a limit of other points of $J \cap H$, it follows from the definition of the derivative as a limit of difference quotients that all derivatives of f of order greater than n are zero at all points of $J \cap H$. Now J also contains intervals complementary to H, and in each such interval K we have $f^{(m)}(x) = 0$ for some m (depending on K). If $m \leq n$, we have $f^{(n)}(x) = 0$ in K, by differentiating. If $m > n$, we have $f^{(n)}(x) = f^{(n+1)}(x) = \cdots = f^{(m)}(x) = 0$ at the endpoints of K, since these are points of $J \cap H$; so by integrating $f^{(m)}$ repeatedly, we get $f^{(n)}(x) = 0$ throughout K. This reasoning applies to every interval K that is complementary to H and is in J; so $f^{(n)}(x) = 0$ throughout J. Thus J contains no points of H after all. But we arrived at J as an interval containing points of H on the assumption that H was not empty. This contradiction means that H must be empty, there was only one interval $I = [0, 1]$ to begin with, and f coincides with a single polynomial throughout this interval.

Exercise 10.1. Prove the proposition on page 47 without appealing to properties of analytic functions.

(iii) CONTINUOUS EVERYWHERE OSCILLATING FUNCTIONS. In our next application of Baire's theorem, the underlying metric space will be the space C of continuous functions on a real interval. It will be shown later (page 112) that this space is complete. Let us seek first to construct a continuous function that is not monotonic in any interval. This can actually be done more directly,[2] but it is a good illustration of the use of Baire's theorem in a fairly uncomplicated situation. The intervals with two rational endpoints form a countable set. Let them be I_1, I_2, I_3, ... , in some order, and let E_n be the set of elements of the space C that are monotonic on I_n. We are going to show that each set E_n is nowhere dense in C; it then will follow from Baire's theorem that there is an element of C that is not in any E_n. In other words, there is a continuous function that is not monotonic on any I_n, and hence not monotonic on any interval (since every interval in $\mathbf{R}_1$ contains an interval with rational endpoints).

The technique for showing that E_n is nowhere dense is one that is useful in many applications: we show that the complement $C(E_n)$ is open and everywhere dense.

Exercise 10.2. Show that a closed set with an everywhere dense complement is nowhere dense.

We first show that $C(E_n)$ is open. If f is in $C(E_n)$, then f is not monotonic in I_n. This means that there are three points x, y, and z in I_n, with $x < y < z$, and $f(x) < f(y)$ and $f(z) < f(y)$ (or else $f(x) > f(y)$ and $f(z) > f(y)$). Recalling that the distance between elements f and g of C is $\max|f(x) - g(x)|$, we see that if g is closer to f than half the smaller of the numbers $f(y) - f(x)$ and $f(y) - f(z)$, we also have $g(x) < g(y)$ and $g(z) < g(y)$, so that g is not monotonic in I_n. That is, all elements g sufficiently close to f are not in E_n, which is to say that $C(E_n)$ is open.

To say that the complement of E_n is everywhere dense is to say that in every neighborhood in C there exists a function f that is not monotonic in I_n. It is rather intuitive that there is a very wiggly function close to any given continuous function. To justify this intuition, we can (for example) let g be the center of the given neighborhood in C; we can approximate g in the metric of C, as closely as we like, by a polynomial p (§19). Now p has a bounded derivative, so if we add to p a small saw-tooth function with very steep teeth we obtain a function f that is as close as we like to g and is not monotonic in I_n.

(iv) EXISTENCE OF CONTINUOUS NOWHERE DIFFERENTIABLE FUNCTIONS. The fact that a continuous function may fail to have a derivative at any point came as a shocking surprise to the mathematicians of the nineteenth century. However, it turns out that "most" continuous functions have this property, and we should rather be surprised that any continuous functions are differentiable at all.[3] Still more surprising is the fact that a function may be everywhere oscillating and still have a finite derivative at every point. Unfortunately all known examples of this last phenomenon are too complicated to give here.[4]

What we are going to show is that[5] *the elements of the space C that have, even at one point, a finite derivative, even on one side* (see page 140) *form a set of first category in C.* This theorem shows that all the functions that are ordinarily encountered in calculus form only a set of first category in C.[6] We do not exclude the possibility that "most" elements of C might have infinite one-sided derivatives at most points (geometrically, their graphs might have cusps). We shall see later (page 152) that a continuous function actually cannot have a vertical tangent at *all* the points of an interval. Although there really are nowhere differentiable functions that do not even have one-sided infinite derivatives, they are much harder to construct;[7] the greater difficulty may be connected with the fact (which we cannot prove here) that these functions form only a set of first category.[8] The "typical" continuous function has an everywhere dense set of cusps (like the graph of $y = |x|^{1/2}$ at the origin). A nowhere differentiable function must be everywhere oscillating, since a monotonic function has a derivative at most points (page 165).

We now prove that the nowhere differentiable functions form a set of second category in C. Consider the set E_n formed of all those elements f of C such that, for some point x in the interval $[0, 1 - 1/n]$,

$$\left| \frac{f(x+h) - f(x)}{h} \right| \leq n$$

if $0 < h < 1/n$. Clearly a function f that has a finite right-hand derivative at some x belongs to E_n for some n. Hence the union of all the E_n contains all the elements of C that have a finite right-hand derivative at some point. We shall show that each E_n is nowhere dense; then the union of the E_n is of first category (and so cannot be all of C). As in the preceding example, we do this by showing that E_n is closed and has an everywhere dense complement.

That E_n is closed follows as in example (iii), or from the fact that the inequality defining E_n is preserved under convergence in C. That the complement of E_n is everywhere dense follows just as in (iii): if f is an arbitrary continuous function, we find a function close to f with a bounded derivative, and then add to this function a small continuous function the slope of whose graph has large absolute value.

(v) Decomposition of a Closed Interval.

Exercise 10.3. Show that a closed interval cannot be written as the union of a countably infinite number of disjoint, closed, nonempty sets.[9]

NOTES

[1] Ernest Corominas and Ferran Sunyer i Balaguer, Condiciones para que una función infinitamente derivable sea un polinomio, *Revista Matemática Hispano-Americana* (4) 14 (1954), 26–43. Several extensions of the theorem are given in this paper. For a generalization to functions of several variables, see A. B. Boghossian and P. D. Johnson, Jr., A pointwise condition for an infinitely differentiable function of several variables to be a polynomial, *Journal of Mathematical Analysis and Applications* 151 (1990), no. 1, 17–19.

[2] See Bernard R. Gelbaum and John M. H. Olmsted, *Counterexamples in Analysis*, Holden-Day, San Francisco, 1964, p. 29.

[3] Nowhere differentiable functions arise naturally in the mathematical theory of "Brownian motion." With probability 1, a Brownian path is nowhere differentiable.

[4] C. E. Weil (On nowhere monotone functions, *Proceedings of the American Mathematical Society* 56 (1976), 388–389) gave a short proof of the existence of such functions via Baire's theorem, and references to other constructions. See also A. Denjoy, Sur les fonctions dérivées sommables, *Bulletin de la Société Mathématique de France* 43 (1915), 161–248, especially pp. 228 ff.

[5] S. Banach, Über die Baire'sche Kategorie gewisser Funktionenmengen, *Studia Mathematica* 3 (1931), 174–180.

[6] Our existence proof does not produce any specific example of a continuous nowhere differentiable function. For explicit constructions, see John McCarthy, An everywhere continuous nowhere differentiable function, *American Mathematical Monthly* 60 (1953), 709; T. H. Hildebrandt, A simple continuous function with a finite derivative at no point, *American Mathematical Monthly* 40 (1933), 547–548; Bernard R. Gelbaum and John M. H. Olmsted, *Counterexamples in Analysis*, Holden-Day, San Francisco, 1964, pp. 38–39.

[7]The simplest example was constructed by A. P. Morse, A continuous function with no unilateral derivatives, *Transactions of the American Mathematical Society* 44 (1938), 496–507.

[8]S. Saks, On the functions of Besicovitch in the space of continuous functions, *Fundamenta Mathematicae* 19 (1932), 211–219. On the other hand, we might consider the continuous, nowhere differentiable functions that lack one-sided derivatives (even infinite ones) at all points except for a set of measure zero: these functions form a set of second category. See S. A. Shkarin, On a class of continuous nowhere differentiable functions, *Russian Academy of Sciences. Izvestiya. Mathematics* 45 (1995), no. 2, 423–432; translated from *Rossiiskaya Akademiya Nauk. Izvestiya. Seriya Matematicheskaya* 58 (1994), no. 5, 195–205.

[9]This is a special case of Sierpiński's theorem that a compact connected set containing more than one point is never a countable union of disjoint closed sets; for some references see the solution to problem E 2613, *American Mathematical Monthly* 84 (1977), 827–828.

11. Sets of measure zero. A set of first category may be thought of as having relatively few points, principally because it cannot exhaust a complete metric space. It may, on the other hand, seem rather large if looked at from other points of view. It may be everywhere dense, like the set of rational points in $\mathbf{R}_1$. It may be uncountable, as the Cantor set is. It may be both everywhere dense and uncountable, as we observed on page 64.

There is another—quite different—kind of "sparse" set that has many uses. Suppose that a subset E of $\mathbf{R}_1$ has the property that it can be covered by a countable collection of open intervals whose total length is arbitrarily small. Then E is called a *set of measure zero.* There is a similar definition in any $\mathbf{R}_n$. Sets of measure zero are just the sets that are negligible in the theory of Lebesgue integration (see Chapter 3), and the name comes from that theory. If something happens except on a set of measure zero, it is said to happen *almost everywhere*, or for almost all points.

The union of two, or of finitely many, or even of a countable infinity of sets of measure zero is again of measure zero.

Obviously a countable subset of $\mathbf{R}_1$ is a set of measure zero, so a set of measure zero can be everywhere dense. However, a set of measure zero need not be countable, or even of first category; and, on the other hand, a set of first category may fail to be of measure zero. Let us justify these statements by examples. First, the Cantor set E is uncountable but of measure zero.[1] For, if ϵ is a small positive number, take enough complementary intervals of E so that their total length exceeds $1 - \epsilon/2$. The rest of the unit interval containing E can be covered by a finite set of intervals of length not exceeding ϵ, and so E is certainly of measure zero.

To construct a set that is of first category, but not of measure zero, modify the construction of the Cantor set as follows. Let $\{a_n\}$ be a sequence of positive numbers such that $\sum a_n = \epsilon < 1$. Remove from the center of the unit interval an open interval of length a_1; then from the center of each remaining interval remove an open interval of length $\frac{1}{2}a_2$; and so on. After the nth step, the remaining set E_n consists of 2^n closed intervals each of length less than 2^{-n}. The set $E = \bigcap_n E_n$ is therefore a closed set containing no interval. Thus E is nowhere dense, hence of first category. On the other hand, E cannot be of measure zero, since if it were covered by a countable set of intervals of total length less than $1-\epsilon$, we should have a unit interval covered by a set of intervals of total length less than 1.

To construct a set that is both of second category and of measure zero, construct a generalized Cantor set E of the kind just described with $\epsilon = \frac{1}{2}$. Then construct similar sets with $\epsilon = \frac{1}{4}$, $\epsilon = \frac{1}{8}$, and so on. The complement of the union of all the generalized Cantor sets has measure zero; but since each Cantor set is nowhere dense, this complement is of second category.[2]

Since an interval in $\mathbf{R}_1$ is not of measure zero, another way to show that a set of points is not empty is to exhibit it as the complement of a set of measure zero. For instance, every countable subset of $\mathbf{R}_1$ is of measure zero, and so $\mathbf{R}_1$ cannot be countable. A set of measure zero, like a set of first category, is "small" in the sense that its complement cannot be empty.[3]

Exercise 11.1. If E is a subset of $\mathbf{R}_1$ that covers at most a fixed fraction of every interval, then E covers almost none of every interval. In other words, if there is a positive q less than 1 such that, for every interval (a, b), the set $E \cap (a, b)$ can be covered by countably many intervals whose total length is at most $q(b - a)$, then E is a set of measure zero.[4]

NOTES

[1] Another example of an uncountable set of measure zero is the set of zeroes of a typical (in the sense of Baire category) continuous function that has zeroes: see Tomás Domínguez Benavides, How many zeroes does a continuous function have?, *American Mathematical Monthly* 93 (1986), 464–466.

[2] Another interesting example of a measure-zero, second-category set is the set of so-called *Liouville numbers* (roughly speaking, transcendental numbers that can be rapidly approximated by rational numbers). See W. M. Priestley, Sets thick and thin, *American Mathematical Monthly* 83 (1976), 648–650.

[3] For some other notions of size of sets and examples of sets that are small in one sense and large in another sense, see Harvey Diamond and Gregory Gellès, Relations among some classes of subsets of $\mathbf{R}$, *American Mathematical Monthly* 91 (1984), 19–22. There is a sort of "duality" between the notions of first category and measure zero that is discussed in John C. Oxtoby, *Measure and Category*, second edition, Springer-Verlag, New York, 1980.

[4] However, it is possible for a set and its complement both to cover a positive fraction of every interval if this fraction is not fixed. See Walter Rudin, Well-distributed measurable sets, *American Mathematical Monthly* 90 (1983), 41–42; F. S. Cater, A partition of the unit interval, *American Mathematical Monthly* 91 (1984), 564–566.

Chapter 2

Functions

12. Functions. In elementary mathematics, it is customary to say that y is a function of x if, when x is given, y is determined (uniquely; we are not concerned with "multiple-valued functions"). This is a good working definition and one that suffices for most practical purposes. However, we should realize that it does not define "function," although it does give a definite meaning to some phrases containing this word. (In a somewhat similar way, we are accustomed to attaching a definite meaning to the phrase "$y \to \infty$" even though ∞ by itself has no meaning.) However, it is interesting, and sometimes helpful, actually to define a function as a genuine mathematical entity. Consider two sets E and F of real numbers, neither of which is empty, and form a class of ordered pairs (x, y) with $x \in E$ and $y \in F$, where each x occurs exactly once and each y occurs at least once. Such a class of ordered pairs is called a *function* with *domain* E and *range* F, or a function from E to F; or, on occasions when it is unnecessary to say precisely what F is, a function with domain E and values in $\mathbf{R}_1$, or a function from E into $\mathbf{R}_1$, or a real-valued function with domain E, etc. For example, let E be all of $\mathbf{R}_1$, let F be the closed interval $[-1,1]$, and let the ordered pairs be $(x, \sin x)$ for each x in $\mathbf{R}_1$. This is what is usually spoken of as "the function $\sin x$." Notice that if we take E to be, for instance, the interval $[0, 2\pi]$ instead of all of $\mathbf{R}_1$, the set of ordered pairs $(x, \sin x)$ for $x \in E$ is a different function, the *restriction* of the original function to $[0, 2\pi]$.

For another example, let E consist of the positive integers $1, 2, 3, \ldots$, and let F be the set $\{1, 4, 9, 16, \ldots\}$. The function whose ordered pairs are $(1,1)$, $(2,4)$, $(3,9)$, $\ldots$ is an example of the special kind of function that we usually call a *sequence*, specifically a sequence of real numbers.

Exercise 12.1. An equation in x and y determines a set of ordered pairs (x, y) whose components satisfy the equation. Accordingly, it may (or may not) determine a function. Which of the following equations do determine functions?

(a) $x^2 + y^2 = 25$,	(b) $x^2 + y^2 = 0$,
(c) $y = \lvert x\rvert$,	(d) $\lvert x\rvert + \lvert y\rvert = -2$,
(e) $y = \cos x$,	(f) $x = \cos y$.

If F consists of a single point, say the point 3, the function whose ordered pairs are $(x, 3)$ is a *constant* function; ideally, it should be distinguished notationally from the number 3.

The definition of function is easily extended to a more general setting. Let E be a nonempty set of points in some space S (which might be $\mathbf{R}_n$, or C, or anything else; it need not be a metric space). Let F be another nonempty set of points belonging to some space T, in general quite different from S. The class of all ordered pairs (x, y) with x in S

and y in T is called the *Cartesian product* of S and T. A function from E to F is a subset of this Cartesian product consisting of pairs (x, y), where each x in E occurs exactly once and each y in F occurs at least once. It is also called a function from E into T. We speak of the points of F as the *values* of the function, and of F as the *image* of E. It is often convenient to call the function a *mapping* of E onto F.

The Cartesian product of two $\mathbf{R}_1$'s is the space $\mathbf{R}_2$ (if we introduce the right metric in the product). A function whose domain and range are in $\mathbf{R}_1$ is just the set of points of $\mathbf{R}_2$ that are in what we ordinarily call the *graph* of the function. "A real function of a real variable" ordinarily means a function with domain and range in $\mathbf{R}_1$. For the purposes of this book it is unnecessary to attempt to define a "variable."

The advantage of the abstract definition of a function as a class of ordered pairs is that it gives us a mathematical object defined in terms of concepts we have already had. Every proof about functions can be phrased in terms of it. Its disadvantage is that it loses most of the intuitive content of the notion of function. For many purposes, and especially for a student's first introduction to the concept, it is better to think of a function as a mapping or transformation or operator (emphasizing the process by which values are obtained from points of the domain); as a rule (emphasizing the correspondence between the values and the points of the domain); or in physicists' language as a field (emphasizing the domain, and the association of a value with each point of the domain—usually in physics the domain is in $\mathbf{R}_3$ and the range is in $\mathbf{R}_1$ (scalar field) or in $\mathbf{R}_3$ (vector field)). We should, perhaps, think of a class of ordered pairs as being a *model* of a function rather than the function itself. On the other hand, if "function" is taken as a primitive notion, then "ordered pair" can be defined, or modeled, in terms of it (as a function on the set $\{1, 2\}$).[1]

A *sequence* is a function whose domain consists of all the positive integers. If we are talking about sequences, the domain is understood to be fixed, and so we can specify the sequence by listing the points of the range in the order imposed by the points of the domain. We then often speak of the points of the range (repeated as often as necessary) as the elements of the sequence, rather than as the values of the function. This brings us back to the informal definition of a sequence that we have already used (page 53). Thus "the sequence $\{2, 4, 8, 16, \dots\}$" or "the sequence $\{2^n\}$" means the set of ordered pairs $(1, 2)$, $(2, 4)$, $(3, 8)$, $\dots$. "The sequence $\{1\}$" means the set of ordered pairs $(1, 1)$, $(2, 1)$, $(3, 1)$, $\dots$. A *subsequence* of a given sequence is a restriction of the sequence to a subset of the integers; it can be identified with the sequence obtained by renumbering its elements. For instance, $\{2^n\}$ is a subsequence of $\{n\}$. (As noted on page 53, a "finite sequence" is not a sequence in our sense of the word.)

In careful work, it is usually helpful to distinguish notationally between (say) f, the name of a function, and $f(x)$, "the value of the function f at the point x." In other words, f is to denote a set of ordered pairs, while $f(x)$ stands for the point of the range that is paired with the point x of the domain. For example, the logarithmic function consists of the pairs $(x, \log x)$; one of these pairs is $(e, 1)$, and $\log(e) = 1$. The distinction between a function and its values is not made consistently in most books.[§] It looks a little odd, although it is both correct and unambiguous, to speak of "the function log" (of course, we must specify the domain of the function; here it would presumably be $(0, \infty)$). This would usually be called "the function $\log x$." We often want to talk about more com-

[§] This one is no exception!

plicated functions, such as the function whose value at x is $\log(\sin x)$. Here we are forced into a certain amount of awkwardness if we are to avoid the possibly misleading "the function $\log(\sin x)$." The same problem arises when we want to talk about functions that are so simple that they have no generally accepted names. The identity function, whose ordered pairs are (x, x) (for x in a specified domain) should not be called "the function x," but may correctly be described as the function I such that $I(x) = x$. The gain in clarity is well worth the loss of brevity.

If the domain of a function f is in $\mathbf{R}_1$, its values are usually written $f(x)$. If the domain is in $\mathbf{R}_2$, the values are usually written $f(x, y)$, although $f((x, y))$ would be more in accordance with our general principles. The elements of a sequence are conventionally written as s_n, rather than as $s(n)$, and we ought to refer to the sequence as s. However, as we have already remarked, it is usually more convenient to call the sequence $\{s_n\}$, that is, to specify what its elements are (since its domain is understood to be the set of positive integers). Thus the sequence whose ordered pairs are $(1, 1)$, $(2, 4)$, $(3, 9)$, $\ldots$ is usually written $\{n^2\}$, and restrictions of it are clearly indicated by $\{n^2\}_{n=3}^{\infty}$, $\{n^2\}_{n \text{ odd}}$, $\{n^2\}_{n=3}^{8}$, and the like. Similarly we could describe the function f such that $f(x) = \sin 2x$ (with domain $\mathbf{R}_1$) as $\{\sin 2x\}$, reserving the notation $\sin 2x$ for the particular point of $\mathbf{R}_1$ (as range) that is paired with the point x of the domain. The restriction of this function to $(0, 2\pi)$ would be $\{\sin 2x\}_{0<x<2\pi}$. If the domain is understood, we might refer to this function as $x \mapsto \sin 2x$. Various other notations are in use;[2] it will not be worthwhile to adopt any of them systematically for the purposes of this book.

Originally, a function was thought of as being defined by a formula, but for many years there has been little worry about this aspect of functions. Many functions that are defined in an apparently arbitrary way can be represented by formulas, although only by formulas of a rather elaborate kind. (We must, however, recognize that the simple notation $f(x) = \sin x$ conceals a nontrivial limiting process, which we forget because the function is so familiar.) For example, let a function f be defined by putting $f(x) = 1$ for rational x in $\mathbf{R}_1$ and $f(x) = 0$ for irrational x. Then

$$f(x) = \lim_{m\to\infty} \lim_{n\to\infty} |\cos(m!\,\pi x)|^n.$$

Exercise 12.2. Verify this.

A more complicated example of the same kind is[3]

$$f(x) = \lim_{r\to\infty} \lim_{m\to\infty} \lim_{n\to\infty} \sum_{\nu=0}^{m} \left[1 - (\cos\{(\nu!)^r \pi/x\})^{2n}\right],$$

which for positive integers x yields the largest prime factor of x.

The functions that can be represented by only one limiting process starting from continuous functions are rather special (see §18). Moreover, some functions having domain and range in $\mathbf{R}_1$ cannot be represented by formulas, at least not by formulas starting with continuous functions and involving only countably many limiting processes.[4] We shall not attempt to construct an example of this phenomenon here.

There are nontrivial, and even interesting, properties of completely general functions with a given domain and range in a given space (see pages 85, 87, and 155), but the most interesting and generally useful properties are possessed only by functions that belong to more or less special classes. In other words, the interesting thing is to impose some special property and see what other properties follow as consequences. From the point of view of the general theory, it is only a fortunate accident that the functions that arise naturally in the applications of mathematics are frequently continuous, or differentiable, and so on. However, since such special functions are, in fact, frequently encountered, it is both desirable and interesting to know some of their important properties.

NOTES

[1]See also G. J. Minty, On the notion of "function," *American Mathematical Monthly* 78 (1971), 188–189; Israel Kleiner, Evolution of the function concept: a brief survey, *College Mathematics Journal* 20 (1989), 282–300.

[2]See especially J. Barkley Rosser, *Logic for Mathematicians*, second edition, Chelsea, New York, 1978, chap. X, §5; Karl Menger, *Calculus, a Modern Approach*, Ginn, Boston, 1955, chap. IV.

[3]G. H. Hardy, A formula for the prime factors of any number, *Messenger of Mathematics* 35 (1906), 145–146.

[4]W. Sierpiński, Sur un exemple effectif d'une fonction non représentable analytiquement, *Fundamenta Mathematicae* 5 (1924), 87–91.

13. Continuous functions. We are now going to define what we shall mean by continuous functions, and we shall then discuss some of their properties. This is not how the concept of continuous function originally got into mathematics. The term *continuous* came first, and then people sought a definition that would fit their intuitive feelings about it. For a real-valued function whose domain is an interval in $\mathbf{R}_1$ (the most familiar case), it was, for example, felt at one time that a continuous function should be defined as one that takes on every value between any two values that it assumes; in other words, as a function such that the image of every interval in its domain is an interval or a point. We call this the *intermediate value property.* It does not, as it turns out, force a function to have all the properties that a continuous function might reasonably be expected to have.[1] For example, the function defined by $f(x) = \sin(1/x)$ for all real x except 0, and $f(0) = 0$, has the intermediate value property, but it does not impress most people as being continuous at 0. By a rather more elaborate construction, we can exhibit a function that has the intermediate value property in every interval, however small, but does not look continuous since it manages to have the intermediate value property simply by taking every value between 0 and 1 in every interval.

We construct this function, with domain $(0, 1)$, as follows.[2] Let x be expanded as an ordinary decimal, $x = 0.a_1a_2\ldots$ (take the nonterminating decimal if there is a choice), and consider the number $z = 0.a_1a_3a_5\ldots.$ If z is not a periodic decimal, put $f(x) = 0$. If, however, z is periodic, with its first period beginning with a_{2n-1}, put

$$f(x) = 0.a_{2n}a_{2n+2}a_{2n+4}\cdots.$$

This defines the required function f. For, if an interval I in $(0, 1)$ is given, we can find n so large that I contains a terminating decimal $0.a_1a_2\ldots a_{2n-1}$ and all the numbers

$$0.a_1a_2a_3\ldots a_{2n-1}a_{2n}\cdots$$

starting with the same first $2n - 1$ digits. Now let $y = 0.b_1b_2\ldots$ be any number in $(0, 1)$. We can choose a_{2n+1}, $a_{2n+3}, \ldots$ so that $0.a_1a_3\ldots a_{2n-1}a_{2n+1}a_{2n+3}\cdots$ is periodic with its first period starting at a_{2n-1}, and then according to our construction the number

$$x = 0.a_1a_2\ldots a_{2n-1}b_1a_{2n+1}b_2a_{2n+3}\cdots$$

has $f(x) = y$.

It is interesting in this connection that every function from an interval in $\mathbf{R}_1$ into $\mathbf{R}_1$ can be written as the sum of two functions, each of which takes on every real value in every subinterval;[3] and every function can have its values modified on a first-category, measure-zero set to produce a nowhere continuous function that has the intermediate value property.[4]

For functions from an interval in $\mathbf{R}_1$ into $\mathbf{R}_1$, you are probably familiar with the definition of continuity that has come to be accepted. We say that f is *continuous at* x_0 if, given any positive number ϵ, we can find a positive number δ such that if $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \epsilon$; and we say that f is *continuous on an interval* if it is continuous at every point of the interval. The intuitive idea behind this definition is that a small change in the position of the point that we are looking at in the domain should produce a small change in the position of the image point in the range. It must be conceded that this definition still does not correspond as closely as we might wish to the intuitive idea of a continuous function. For example, we cannot necessarily draw a satisfactory graph of a given continuous function on paper with a pencil: think, for example, of the everywhere oscillating functions of §10. Indeed, the intuitive notion of a continuous function is more nearly that of a function whose graph has no jumps and is made up of a finite number of increasing or decreasing pieces.

Exercise 13.1. The "ruler function" f is defined on the interval $[0, 1]$ by $f(x) = 0$ if x is an irrational number, and $f(x) = 1/q$ if x is a fraction that has denominator q when written in lowest terms. Show that the ruler function is continuous at each irrational number and discontinuous at each rational number.

The definition of continuity can be carried over to the case where the domain and the range are in any two metric spaces. In the simplest case, the domain of f contains a neighborhood of the point x_0; then we say that f is continuous at x_0 if, given any positive number ϵ, we can find a positive number δ such that if $d(x, x_0) < \delta$, then $d(f(x), f(x_0)) < \epsilon$. (Of course, the two d's refer, in general, to different metrics.)

If we want to consider more general situations, we must face up to the fact that a function may be continuous or not according to the space in which its domain is supposed to lie. For a simple example, consider a constant function with domain $\mathbf{R}_1$. This function is certainly continuous at each point of its domain. However, if $\mathbf{R}_1$ is regarded as a subset of $\mathbf{R}_2$, we cannot say that the function is continuous at any point, since it is not even defined in any neighborhood in $\mathbf{R}_2$. Indeed, this function is the restriction to $\mathbf{R}_1$ of many functions with domains in $\mathbf{R}_2$; some of these functions are continuous and some are not.

We can always take the domain of a given function as a metric space in itself, and ask whether the function is then continuous; we may add, "with respect to its domain," for emphasis. A new idea enters when we consider the restriction of a function to a subset of its domain. It may happen that the restriction is continuous (with respect to its own domain) whereas the original function is not continuous. For example, the function f (page 82) that has the value 1 for rational points of $\mathbf{R}_1$ and the value 0 for irrational points is clearly discontinuous at each point of $\mathbf{R}_1$. The restriction of the same function to the set P of rational points of $\mathbf{R}_1$ is a function g that is continuous (on P as space). Some authors say that the original function is discontinuous at every point of P, but continuous on P with respect to P. The confusion created by statements of this kind can best be avoided by recognizing that to define a function we must say what its domain is as well as how its values are to be calculated. Care in specifying the domain is especially important when the function is defined by a formula.

To say that a function f is continuous at a point x_0 of a set E, with respect to E, then, is to say that the restriction of f to E is continuous at x_0 with E as space. An equivalent definition is obtained by using the (ϵ, δ) definition (page 85) with the additional requirement that $x \in E$. In particular, let f be a function whose domain is an interval in $\mathbf{R}_1$ containing x_0 in its interior, and let g be the restriction of f to an interval $[x_0, b)$ to the right of x_0. If g is continuous at x_0, it is often said that f is continuous on the right at x_0. This is the same as saying that f satisfies the definition of continuity at x_0 except that only right-hand neighborhoods of x_0 are considered. To have an illustration, let us define functions f_1, f_2, f_3, respectively, all to have the value -1 for $x < 0$ and the value $+1$ for $x > 0$, whereas $f_1(0) = 1$, $f_2(0) = -1$, and $f_3(0) = 0$. Then f_1 is continuous on the right at 0, f_2 is continuous on the left, f_3 is continuous on neither side, and all three functions are discontinuous at 0.

Exercise 13.2. If f is simultaneously continuous on the right at x_0 and continuous on the left at x_0, then f is continuous at x_0.

In this connection, it is interesting that for any real-valued function whatsoever (whose domain is an interval) there is a dense (but countable) set E such that f, restricted to E, is continuous on E.[5] On the other hand, there is a function whose restriction to every set E that can be put into one-to-one correspondence with $\mathbf{R}_1$ is discontinuous.[6]

Exercise 13.3. Show that if y is a point of a metric space, then the function f defined by $f(x) = d(x, y)$ is continuous on the space.

Exercise 13.4. Let E be a closed set in a metric space; let D be the function such that $D(x)$ is, for each point x in the space, the distance (see Exercise 8.10) from x to E. Show that D is continuous.

If we want to consider continuous functions on spaces that are not metric, a definition of continuity in terms of distance naturally will not do. Although we shall consider only metric spaces in this book, we shall rephrase the definition of continuity in a form that can be extended to more general spaces, if only because it is often a convenient form to use even in metric spaces. This more sophisticated definition reads: *f is continuous on its domain if and only if the inverse image of each open set in the range space is an open set in the domain.* (Here the domain of f is to be regarded as the space with respect to which open sets are defined.) The *inverse image* of a set E means, of course, the set of points of the domain whose image points are in E. For example, if $f(x) = \sin x$ with domain $\mathbf{R}_1$, then the inverse image of the interval $(0, 2)$ consists of the union of the intervals $(0, \pi)$, $(2\pi, 3\pi)$, $(-2\pi, -\pi)$, $\ldots$, which is an open set. If, however, $f(x) = 1$ for $x > 0$, $f(0) = 0$, and $f(x) = -1$ for $x < 0$, then the inverse image of the open interval $(-\frac{1}{2}, \frac{1}{2})$ contains the single point 0, and so is not open.

Exercise 13.5. Give an example to show that the image of an open set under a continuous function is not necessarily open.

To verify the equivalence of the two definitions of continuity in a metric space, suppose first that f is continuous under the original definition. Let E be an open set in the range space, and let x_0 be a point in the inverse image of E. Then $f(x_0) \in E$, and if ϵ is small enough, every y such that $d(f(x_0), y) < \epsilon$ belongs to E (since E is open). Since f is continuous, there is a positive δ such that $d(x, x_0) < \delta$ implies $d(f(x), f(x_0)) < \epsilon$. Hence all points x sufficiently near x_0 have their images in E; that is to say, the inverse image of E contains a neighborhood of each of its points, so it is open. Conversely, suppose that the inverse image of every open set is open. In particular, the inverse image of any neighborhood in the range space, defined by an inequality $d(y, f(x_0)) < \epsilon$, is open and so contains a neighborhood defined by an inequality $d(x, x_0) < \delta$; this is an appropriate neighborhood to associate with ϵ in the original definition.

We have proved a little more than we set out to do, namely that f *is continuous at* x_0 *if and only if the inverse image of every neighborhood of* $f(x_0)$ *contains a neighborhood of* x_0.

Exercise 13.6. Show that if the range of f is in $\mathbf{R}_1$, and f is continuous at x_0, and $f(x_0) \neq 0$, then there is a neighborhood of x_0 in which $|f(x)|$ has a positive lower bound, that is, $|f(x)| \geq m > 0$.

Exercise 13.7. Show that if f has an interval in $\mathbf{R}_1$ as its domain, has its range in $\mathbf{R}_1$, and is continuous at x_0, then f is bounded in some neighborhood of x_0.

Exercise 13.8. If a function f with range in $\mathbf{R}_1$ is discontinuous at x_0, then there are a positive ϵ and a sequence $\{x_n\}$ with limit x_0 such that $|f(x_n) - f(x_0)| > \epsilon$.

NOTES

[1] The distinction between continuity and the intermediate value property was first clarified by G. Darboux, Memoire sur les fonctions

discontinues, *Annali della Scuola Normale Superiore di Pisa, Classe di Scienze* 4 (1875), 57–112. Therefore functions possessing the intermediate value property are sometimes called Darboux functions.

[2] H. Lebesgue, *Leçons sur l'intégration et la recherche des fonctions primitives*, 2nd ed., Gauthier-Villars, Paris, 1928, p. 97.

[3] H. Fast, Une remarque sur la propriété de Weierstrass, *Colloquium Mathematicum* 7 (1959), 75–77.

[4] V. Pambuccian, Another example of an exotic function, *American Mathematical Monthly* 96 (1989), 913–914. For more information about functions with the intermediate value property, see Chapter 1 of *Differentiation of Real Functions* by Andrew Bruckner, Providence, American Mathematical Society, 1994, and its references.

[5] H. Blumberg, New properties of all real functions, *Transactions of the American Mathematical Society* 24 (1922), 113–128.

[6] W. Sierpiński and A. Zygmund, Sur une fonction qui est discontinue sur tout ensemble de puissance du continu, *Fundamenta Mathematicae* 4 (1923), 316–318.

14. Properties of continuous functions. It is a commonplace that sums, products, and quotients of continuous functions are continuous (provided that the divisor in the quotient is not zero). More precisely, we should suppose that the two functions f and g have the same domain, and that their values are in $\mathbf{R}_1$. Then we can define $f+g$, fg, and f/g in the usual (and natural) way, provided in the last case that g is not zero anywhere on the common domain of the functions. Then if f and g are both continuous at x_0, so are $f+g$, fg, and f/g if it is defined. However, it is usually not necessary to be so meticulous. When f and g have different domains, we are likely to write $f+g$ for the sum of their restrictions to the intersection of the domains of f and g, and similarly f/g for the quotient of the restrictions of f and g to the part of the intersection of their domains where g does not take the value 0. Note in this connection that the function f_1 defined by $f_1(x) = x/x$ and the function f_2 defined by $f_2(x) = 1$ are different functions since their domains are different; we cannot say that f_1 is continuous at 0. In discussing the continuity of a product, it is convenient to use Exercise 13.7.

Exercise 14.1. Show that if f is continuous at x_0 and g is not, then $f + g$ is not. Can $f + g$ be continuous at x_0 if neither f nor g is continuous at x_0?

Exercise 14.2. Carry out the detailed proofs for the continuity of $f + g$, fg, and f/g when f and g are continuous.

A function f is said to be *univalent*, or *one-to-one*, if in the set of ordered pairs that constitute the function, not only does no x in the domain occur twice, but also no y in the range occurs twice. In this case the ordered pairs (y, x) with y in the range and x in the domain also constitute a function, the *inverse* of f, often denoted by f^{-1} (not to be confused with $1/f$, which is the reciprocal of f); its domain is the range of f and its range is the domain of f. It is often useful to know that *under certain circumstances the inverse of a continuous univalent function is continuous.* This statement is true *if the domain of the function is a compact set in some* $\mathbf{R}_n$, or, more generally, whenever the domain has the property that every infinite subset of it has a limit point (the conclusion of the Bolzano-Weierstrass theorem).

To verify the preceding assertion, we have to show that the images of open sets are open, since these will be the inverse images of open sets under f^{-1}. It is an equivalent statement that the images of closed sets are closed, a statement which is somewhat more convenient for us to use.

Exercise 14.3. Establish the equivalence asserted in the preceding sentence.

Let, then, E be a closed set in the domain of f, let F be the image of E, and let y_0 be a limit point of F; we have to show that $y_0 \in F$. Let $\{y_n\}$ be a sequence of distinct points of F with $y_n \to y_0$, and let $y_n = f(x_n)$. There is, by univalence, just one x_n for each y_n, and the x_n are all different since the y_n are all different. Then the *set* whose points are the x_n has a limit point x_0, and a subsequence $\{x_{n_k}\}$ has x_0 as its limit. (See page 59.) We have $x_0 \in E$ since E is closed. Since f is continuous, $y_{n_k} = f(x_{n_k}) \to f(x_0)$; but $y_{n_k} \to y_0$, so $y_0 = f(x_0) \in F$.

Although the intermediate value property that was discussed in §13 does not characterize continuous functions, it is a property that (under some conditions) continuous functions do have. It is possessed by a continuous function with values in $\mathbf{R}_1$ if its domain is a connected set in a metric space S. To see this, let $f(a) = A$ and $f(b) = B$, and suppose $A < C < B$. Consider the sets E_1 and E_2 in S consisting, respectively, of the points x of S for which $f(x) < C$ and of the points x of S for which $f(x) > C$. These sets are disjoint, since $f(x)$ cannot (for the same x) be simultaneously less than C and greater than C. They are not empty, since $a \in E_1$ and $b \in E_2$. They are open, since they are inverse images of open sets. Since the domain of f was assumed connected, it cannot be the union of two disjoint, nonempty, open sets. Hence the domain of f must contain at least one point c that is in neither E_1 nor E_2. The only possible value for $f(c)$ is C.

A continuous function whose domain is compact and whose range is in $\mathbf{R}_1$ has a largest and a smallest value (page 48).

Exercise 14.4. A more elegant proof of the preceding statement can be given by supposing that the least upper bound M of the values of f is not a value of f, and then considering the function $1/(M - f)$.

Exercise 14.5. Deduce that if the domain is both compact and connected, then the range of a continuous function with values in $\mathbf{R}_1$ is a closed bounded interval or a point.

Compactness is of course not necessary for the existence of a maximum.

Exercise 14.6. Let f be a nonnegative continuous function with domain $[a, \infty)$ and range in $\mathbf{R}_1$, and suppose that $f(x) \to 0$ as $x \to \infty$. Then f has a maximum on $[a, \infty)$.

Exercise 14.7. If f is as in the preceding exercise, but strictly positive, then there is a sequence $\{x_n\}$ such that $x_n \to \infty$, and $f(x) < f(x_n)$ for all $x > x_n$. That is, as $x \to \infty$ the function f never again takes as large a value as it had at x_n.

We now give some interesting applications of the intermediate value property.

Consider a function f, from an interval in $\mathbf{R}_1$ into $\mathbf{R}_1$, that has the intermediate value property on every interval in its domain and has a discontinuity at the point c. Then (Exercise 13.8) there is a sequence $\{x_n\}$ with limit c such that for some positive ϵ, either $f(x_n) > f(c) + \epsilon$ for all n or $f(x_n) < f(c) - \epsilon$ for all n; say the former. Since f has the intermediate value property on every interval, it takes every value between $f(c)$ and $f(c) + \epsilon$. Furthermore it does this infinitely often, since we can always consider a smaller neighborhood of c. Hence *if a function from an interval into* $\mathbf{R}_1$ *has the intermediate value property on every subinterval and is discontinuous, it must take some values infinitely often.* (Thus the discontinuous function with the intermediate value property, constructed on page 84, illustrates the typical situation better than one might have expected.) As a corollary, we have that *if a function has the intermediate value property on every interval and takes no value more than once, then it is continuous.* In particular, a function must be continuous if it takes each value between $f(a)$ and $f(b)$ exactly once, in every interval $[a, b]$ in its domain.[1]

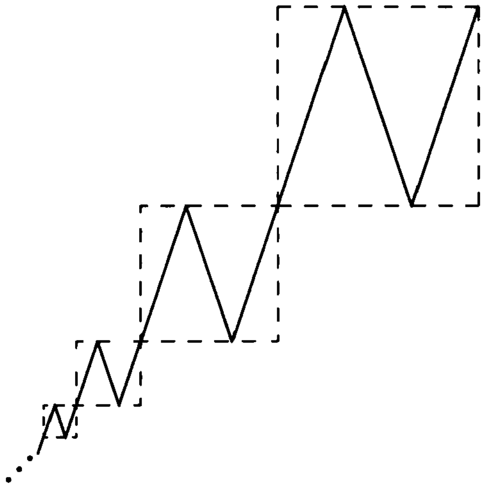
A strictly monotonic function (page 158) is an example of a function that takes no value more than once, but not every function with this property is strictly monotonic. (Consider $f(x) = x + 1$ for $-1 < x \le 0$, and $f(x) = x - 1$ for $0 < x < 1$.) However, a continuous function that takes no value more than once is indeed strictly monotonic. For, if it were not strictly monotonic, then there would have to be points $x_1 < x_2 < x_3$ with $f(x_1) \le f(x_2)$ and $f(x_3) \le f(x_2)$ (or the corresponding inequalities reversed); and so f would have (in the first case) a maximum at some point c between x_1 and x_3 (because f is continuous), and the maximum would be proper (since f takes no value more than once). Then $f(x) < f(c)$ for values of x on both sides of c, and so by the intermediate value property f would take some value near $f(c)$ twice, contradicting the hypothesis. (The same reasoning shows that a continuous function is monotonic between its consecutive local maxima and minima: that is, if f has a maximum at x_1 and a minimum at $x_2 > x_1$, and neither a maximum nor a minimum in (x_1, x_2), then f is monotonic in (x_1, x_2).)

We have just seen that the continuous functions (from an interval in $\mathbf{R}_1$ into $\mathbf{R}_1$) that take each value exactly once are strictly monotonic. What sort of continuous functions take each value exactly twice? The answer is that there are no such functions.[2]

In fact, it is just as easy to show that no continuous function on a closed bounded interval can take each of its values exactly n times, where $n > 1$. For suppose that there is such a function f. Then f has an absolute maximum and an absolute minimum, and each must be attained at interior points except in the case $n = 2$, when the minimum (say) might be attained at both endpoints. Hence we may suppose that the maximum is attained at an interior point c_1. Let $c_2, c_3, \ldots, c_n$ be the other points where f takes its maximum value. Then $f(x) < f(c_1)$ in a deleted neighborhood of each c_k (a one-sided neighborhood if c_k is an endpoint). For a sufficiently small positive ϵ, the line $y = f(c_1) - \epsilon$ intersects the graph of f twice near c_1 and at least once near each other c_k, by the intermediate value theorem, so there is a value that f takes at least $n + 1$ times, a contradiction.

Since f cannot be continuous and exactly two-to-one, let us suppose now that f takes each value *at most* twice. Then we may conclude that the graph of f falls into at most three monotonic pieces[3] (like the graph on page 99, with a small piece removed at one end). It adds interest to this theorem to observe that nothing similar holds for continuous functions that take each value at most three times: the graph of such a function does not have to be composed of a finite number of monotonic pieces,[4] as is indicated by the sketch, where the nth square from the right has side a_n, and $\sum a_n < \infty$.

Let us suppose, then, that f is continuous on $[a, b]$ and at most two-to-one. By the principle mentioned above, f is monotonic between any two consecutive local maxima and minima. Thus if maxima and minima of f are attained only at the endpoints, then f is itself monotonic. If it is not monotonic, then it has at least one maximum or minimum inside the interval, say a maximum at c. If f has no other interior maxima or minima, it must be monotonic on (a, c) and on (c, b). The next case is the one where f has



precisely one interior maximum at c_1 and one interior minimum at c_2, say with $a < c_1 < c_2 < b$. Then f is monotonic on each of the three intervals (a, c_1), (c_1, c_2), (c_2, b). Finally suppose that f has more than a total of two interior maxima and minima. Let a maximum, a minimum, and a maximum (not necessarily consecutive) occur at c_1, c_2, c_3, where $c_1 < c_2 < c_3$, and $f(c_1) > f(c_2)$, $f(c_3) > f(c_2)$; suppose for definiteness that $f(c_1) \geq f(c_3)$. Then f takes some value slightly less than $f(c_3)$ at least twice near c_3; taking this value to be greater than $f(c_2)$, it will be less than $f(c_1)$, and by the intermediate value property it will be taken again between c_1 and c_2, for a total of three times, contradicting the hypothesis.

We turn now to a different application of the intermediate value property. Let f be a continuous function from an interval in $\mathbf{R}_1$ into $\mathbf{R}_1$. We say that f has a *horizon-*

tal chord of length a (where $a > 0$) if there is a point x such that x and $x + a$ are both in the domain of f and $f(x) = f(x + a)$. This means that there is a horizontal line segment of length a having both ends on the graph of the function; we do not care whether or not the segment has other points in common with the graph. For example, if $f(x) = 1$ for every x, then f has horizontal chords of all lengths; the segment from $(-1, 1)$ to $(1, 1)$ is a horizontal chord of length 2 of the function f defined by $f(x) = x^3 - x + 1$; the function f defined by $f(x) = x^3$ has no horizontal chords.

A function f whose domain is all of $\mathbf{R}_1$ is said to be *periodic* with period p if $f(x + p) = f(x)$ for all x. (Of course $p \neq 0$; normally p is understood to be positive.)

Exercise 14.8. A continuous periodic function is bounded.

Exercise 14.9. A continuous periodic function has a maximum.

Exercise 14.10. If f is continuous and has period p, then the integral $\int_a^{a+p} f(x)\,dx$ has a value independent of a.

We first notice that *a continuous periodic function f has horizontal chords of all lengths.* That is, if f has period p, and a is any real number, then there is an x such that $f(x + a) - f(x) = 0$.

To see this, consider $\int_0^p [f(x + a) - f(x)]\,dx$, which is zero (by Exercise 14.10). Hence the integrand cannot have a fixed sign on the interval $(0, p)$: it must be equal to zero at some point. The periodic integrand must return to the same value at p as it has at 0, so it must actually be equal to zero at least twice on the interval $[0, p)$. Accordingly, there are at least two points x in $[0, p)$ for which $f(x+a) = f(x)$, and we see that *a continuous function of period p has two* *horizontal chords of any given length, with their left-hand endpoints at different points of* $[0, p)$.[5]

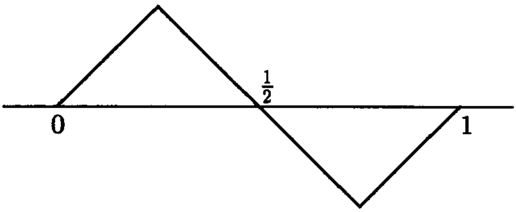
Exercise 14.11. Show that a periodic continuous function always has a chord (not necessarily horizontal), of prescribed span, with its midpoint on the graph: that is, for each positive a there is an x such that $f(x+a) - f(x) = f(x) - f(x-a)$.

For functions that are not periodic, the situation is quite different. A given continuous function f, say with domain $[0, 1]$, may have no horizontal chords at all. (For instance, f could be strictly increasing.) However, let us suppose to begin with that f does have one horizontal chord. More specifically, *suppose that* $f(0) = f(1)$, *so that* f *has a horizontal chord of length* 1. The universal chord theorem[6] states that *there are then horizontal chords of lengths* $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$, *but not necessarily a horizontal chord of any given length that is not the reciprocal of an integer.*

To prove the positive half of this theorem, let k be a positive integer and consider the continuous function g, defined by $g(x) = f(x + 1/k) - f(x)$, whose domain is $[0, 1 - 1/k]$. I assert that 0 is in the range of g. If not, then g would be either positive for all x in its domain or else negative for all x in its domain (by the intermediate value property), and so $g(0) + g(1/k) + g(2/k) + \cdots + g(1 - 1/k)$ would be either positive or negative; on the other hand, this sum "telescopes" and is equal to $f(1) - f(0) = 0$.

Alternatively, if $g(x) > 0$ for all x, that is, if $f(x) < f(x + 1/k)$ for all x, then $f(0) < f(1/k) < f(2/k) < \cdots < f((k-1)/k) < f(k/k) = f(1) = f(0)$, a contradiction.

The figure indicates a function that has a horizontal chord of length 1 but no horizontal chord of length a for $\frac{1}{2} < a < 1$. Naturally a function of this kind has horizontal chords of *some* lengths that are not reciprocals of integers;



the negative part of the universal chord theorem asserts that, for each number b that is not the reciprocal of an integer, there is a continuous function that has a horizontal chord of length 1 but does not have a horizontal chord of the particular length b.

Exercise 14.12. Show how to construct an example for the negative part of the universal chord theorem for every $b \neq 1/k$.

An interesting complement to the universal chord theorem is that[7] *a continuous function f with a horizontal chord of length 1 always has either a horizontal chord of length a or two different ones of length $1-a$* (if $0 < a < 1$). To see this, suppose that $f(0) = f(1)$, and form a new function g by repeating f with period 1. As a periodic continuous function, g necessarily has two horizontal chords of length a with left-hand endpoints in $[0,1)$ (see page 97). A horizontal chord of length a for g, starting at x, is a horizontal chord for f unless $x + a > 1$. If $x + a > 1$, then $0 < x+a-1 < 1$, and so a horizontal chord of length $1-a$ for g (also for f) starts at $x + a - 1$ and ends at x.

Exercise 14.13. Prove the universal chord theorem by induction, starting from the fact that if $f(0) = f(1)$, then f has either a horizontal chord of length a or one of length $1 - a$.

In the same way, we can show that if f has a derivative f' in $[0,1]$, and $f'(0) = f'(1)$, then for any positive integer n there are points x and $x + n^{-1}$ such that f has the same slope at both points. This depends on the fact (page 143) that a derivative always has the intermediate value property; this property (for g, not merely for f) was all that was used in the proof of the universal chord theorem.

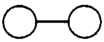
Another application of the same idea yields a simple fixed-point theorem. This says that *a continuous mapping of a compact interval into (part or all of) itself has at least one fixed point*; that is, there is at least one point that coincides with its image. Another way of saying this is that if f is a continuous function on $[0,1]$ such that $0 \le f(x) \le 1$, then the graph of the function must cross the line $y = x$.

Exercise 14.14. Prove the preceding statement; that is, prove that if f is continuous in $[a,b]$ and all the values of f are in $[a,b]$, then there is an x in $[a,b]$ for which $f(x) = x$.

Exercise 14.15. Suppose that a clock runs irregularly, but at the end of 24 hours it has neither gained nor lost overall. Is there some hour for which this clock shows an elapsed time of exactly one hour? Is there some continuous 576 minutes during which it shows an elapsed time of 576 minutes? (Assume that the indicated time is a continuous increasing function.)[8]

Another closely related theorem states that if a continuous periodic function, of period $2p$, has the property that $f(x) = -f(x+p)$ for every x (like the graph of $y = \sin x$, with $p = \pi$), then $f(x) = 0$ for at least one x. This is obvious from the intermediate value theorem. If formulated in a different way, it becomes a simple case of Borsuk's antipodal-point theorem, which is much harder to prove in higher-dimensional situations.[9] In this formulation we consider a continuous function f whose domain is the circumference of a circle in $\mathbf{R}_2$ and whose range is in $\mathbf{R}_1$. Suppose that the image of every pair of antipodal points (points at opposite ends of a diameter) is a pair of points that are symmetric with respect to the origin. Then some point of the circumference maps into the origin.

Still another application of the intermediate value theorem shows that a (two-dimensional) pancake of arbitrary shape can be bisected by a knife cut in any specified direction. At least if the boundary of the pancake is sufficiently simple, it is easy to see that the part of the pancake lying on one side of a line in the given direction has an area that varies continuously as the line is moved parallel to itself. Since this area can be either 0 or the total area of the pancake, it must at some time be exactly half the total area.

We now indicate how two pancakes in a plane can be bisected simultaneously by some line. Let the two pancakes be A and B. As we have just seen, there is a line in any given direction bisecting B. For the direction determined by any point P on a given circle in $\mathbf{R}_2$, with center O, find a line bisecting B, and let $f(P)$ be the signed difference between the area of the part of A on the left of this line and the area of the part of A on the right. We have a function whose domain is a circumference; whose range is in $\mathbf{R}_1$; and which maps antipodal points into points symmetric about the origin, since replacing P by its antipode interchanges left and right. If we can show that f is continuous, then the special case of Borsuk's theorem, noted above, shows that $f(P) = 0$ for some P, which is to say that a line in the direction of OP bisects both A and B. The continuity of f is plausible, but not quite obvious. It is enough to show that a small change in the position of P produces not only a small change in the direction of the line bisecting B, but also a small change in its position, for example, a small change in its intercept on one of the coordinate axes. For, if this statement is true, $f(P)$ will change by only a small amount. Now the preceding statement about lines bisecting B is true without reservation only if we impose some restriction on the admissible shape of a pancake; for example, a "pancake" shaped like  can be bisected by many vertical lines. If we restrict ourselves to convex pancakes there is no difficulty, as is clear from a figure.[10]

Exercise 14.16. Given a convex closed curve in the plane, there is a line that simultaneously bisects the curve and the area that it surrounds.[11]

There are similar theorems in more dimensions, but they are harder to prove. The three-dimensional fixed-point theorem can be stated in picturesque, if misleading, language: if a cup of coffee is stirred in any continuous fashion, there is at least one molecule that ends up in its original position. (This is correct only if the coffee occupies every point inside the cup and "molecule" is interpreted as "point.") The three-dimensional antipodal-point theorem states that a continuous mapping of the surface of a sphere into the space $\mathbf{R}_1$ that carries every pair of antipodal points into a pair of points symmetric with respect to the origin must carry some point of the surface of the sphere into the origin. This can be used to show that any three volumes in space can be bisected simultaneously by some plane (the "ham sandwich" theorem).[12]

NOTES

[1] This also follows from the definition of continuity on page 88, since the property in question requires the inverse image of every

open interval to be an open interval (and hence the inverse image of every open set to be an open set). For a more thorough treatment see J. B. Diaz, Discussion and extension of a theorem of Tricomi concerning functions which assume all intermediate values, *Journal of Mathematics and Mechanics* 18 (1968/69), 617–628; also the review of this paper in *Mathematical Reviews* 39 #370. For a localized version of the property just discussed see E. W. Chittenden, Note on functions which approach a limit at every point of an interval, *American Mathematical Monthly* 25 (1918), 249–250.

[2]More generally there is no continuous transformation of an interval such that each image point has exactly two inverses. See O. G. Harrold, The non-existence of a certain type of continuous transformation, *Duke Mathematical Journal* 5 (1939), 789–793; and for extensions, J. H. Roberts, Two-to-one transformations, *Duke Mathematical Journal* 6 (1940), 256–262; P. Civin, Two-to-one mappings of manifolds, *Duke Mathematical Journal* 10 (1943), 49–57.

There is a substantial amount of literature on related topics. See, for example, *Mathematical Reviews* 94b:54094, 94b:54093, 90e:54021, 48#472, 36#3320, 32#434, 30#2476, 28#2547, 26#4310, 26#3021, 24#A2954, 24#A1709, 2,324d; and Problems E 1094 and E 1715, *American Mathematical Monthly* (1954, 425; 1965, 784).

[3]D. C. Gillespie, A property of continuity, *Bulletin of the American Mathematical Society* 28 (1922), 245–250.

[4]In the paper cited in the preceding note, Gillespie gives a formula for such a function: $f(x) = \pi x + x^2 \sin(\pi/x)$, $0 < x \leq 1$.

[5]J. B. Diaz and F. T. Metcalf, A continuous periodic function has every chord twice, *American Mathematical Monthly* 74 (1967), 833–835, give a different proof. For extensions to almost periodic and more general functions, see J. C. Oxtoby, Horizontal chord theorems, *American Mathematical Monthly* 79 (1972), 468–475.

[6]Although T. M. Flett (*Bulletin of the Institute of Mathematics and its Applications* 11 (1975), 34) has discovered that the positive part of the universal chord theorem was proved by A. M. Ampère in 1806 (see *Mathematical Reviews* 56 #5805), the modern history of the theorem begins with P. Lévy, Sur une généralisation du théorème de Rolle, *Comptes Rendus de l'Académie des Sciences Paris* 198 (1934), 424–425; it has been repeatedly rediscovered. See H. Hopf, Über die Sehnen ebener Kontinuen und die Schliefen geschlossener Wege, *Commentarii Mathematici Helvetici* 9 (1937), 303–319, for extensions. For a discussion of the possible lengths of horizontal chords for a given function, see Hopf's paper; also Oxtoby's paper cited in the preceding note, and R. J. Levit, The finite difference

extension of Rolle's theorem, *American Mathematical Monthly* 70 (1963), 26–30. For further information and some applications see J. T. Rosenbaum, Some consequences of the universal chord theorem, *American Mathematical Monthly* 78 (1971), 509–513. Rosenbaum also gave a picturesque interpretation of the theorem: How to build a picnic table for use on a mountain range of known period, *Notices of the American Mathematical Society* 16 (1969), 94.

Another application was shown to me by J. C. Oxtoby: if f is continuous and $f(x+y) = g(f(x), y)$ for all x and y, then f is either strictly monotonic or constant. (This is problem E2246, *American Mathematical Monthly* 78 (1971), 676–677.) For if f is not strictly monotonic, then $f(x_0 + h) = f(x_0)$ for some x_0 and some $h > 0$. Then for all x,

$$\begin{aligned} f(x+h) &= f((x_0+h) + (x - x_0)) = g(f(x_0+h), x - x_0) \\ &= g(f(x_0), x - x_0) = f(x_0 + (x - x_0)) = f(x). \end{aligned}$$

Since f has arbitrarily short horizontal chords (by the universal chord theorem), this formula shows that f has arbitrarily short periods, and so is constant. Oxtoby remarks that conversely, if f is continuous and either strictly monotonic or constant, then there is a function g such that $f(x+y) = g(f(x), y)$.

[7] Hopf, preceding note.

[8] Suggested by J. D. Memory, Kinematics problem for joggers, *American Journal of Physics* 41 (1973), 1205–1206.

[9] For an elementary and detailed exposition of this and related theorems, see A. W. Tucker, Some topological properties of disk and sphere, *Proceedings of the First Canadian Mathematical Congress 1945*, University of Toronto Press, 1946, pp. 285–309.

[10] See W. G. Chinn and N. E. Steenrod, *First Concepts of Topology* (New Mathematical Library, no. 18), Random House, New York, 1966, p. 65.

[11] J. G. Brennan, A property of a plane convex region, *Mathematical Gazette* 42 (1958), 301–302; A. C. Zitronenbaum, Bisecting an area and its boundary, *Mathematical Gazette* 43 (1959), 130–131.

[12] For a proof, references, and extensions, see A. H. Stone and J. W. Tukey, Generalized "sandwich" theorems, *Duke Mathematical Journal* 9 (1942), 356–359; Chinn and Steenrod, book cited above, p. 120.

15. Upper and lower limits. We shall need to use a generalization of the notion of limit of a sequence of real numbers. If $\{s_n\}$ is such a sequence, and the numbers s_n form a bounded set, then (Exercise 15.1) there is always a number L with the following property: Given any positive ϵ, we have $s_n \leq L + \epsilon$ whenever n is sufficiently large, and in addition an infinite number of s_n satisfy $s_n \geq L - \epsilon$. This number L is called the *upper limit* of $\{s_n\}$, and is written $\limsup s_n$ or $\overline{\lim}\, s_n$. If $\{s_n\}$ is not bounded above, then no such number L exists, and we write $\limsup s_n = +\infty$. If $\{s_n\}$ is bounded above but unbounded below, L may or may not exist; in the case that it fails to exist, we write $\limsup s_n = -\infty$.

Consider some examples. (i) Let $s_n = (-1)^n$, so that our sequence is $-1, +1, -1, +1, \ldots$. Then $\limsup s_n = 1$. (ii) Let $s_n = n$; the sequence is 1, 2, 3, ... Then we have $\limsup s_n = +\infty$. (iii) Let $\{s_n\} = \{-1, -2, \ldots\}$. Then $\limsup s_n = -\infty$. (iv) Let $s_n = 1/n$, $n = 1, 2, 3, \ldots$. Then $\limsup s_n = \lim s_n = 0$. (v) Let

$$\{s_n\} = \left\{ \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{4}{5}, \ldots \right\};$$

then $\limsup s_n = 1$. (vi) Let

$$\{s_n\} = \left\{ -1, \frac{1}{2}, -3, \frac{1}{4}, -5, \frac{1}{6}, \ldots \right\};$$

then $\limsup s_n = 0$.

Exercise 15.1. By considering the numbers L_n defined by $L_n = \sup s_k$ for $k \geq n$, show that $\limsup s_n$ exists (finite) if $\{s_n\}$ is bounded. Thus $\limsup s_n = \lim_{n\to\infty} (\sup_{k \geq n} s_k)$.

Exercise 15.2. Define the lower limit, lim inf, or $\underline{\lim}$, similarly, and determine it for the six examples given above for lim sup.

The definition of upper limit resembles that of least upper bound, from which it differs in that finite subsets of the elements $\{s_n\}$ are disregarded. For example, the value of $\limsup s_n$ is unchanged if the first thousand s_n are replaced by other numbers. We could also define $\limsup s_n$ as the largest limit obtainable by picking convergent subsequences out of $\{s_n\}$.

We have $\limsup(s_n + t_n) \leq \limsup s_n + \limsup t_n$ when both the quantities on the right are finite; but strict inequality may occur here. However, $\limsup(s_n + t_n) = \limsup s_n + \lim t_n$ if $\lim t_n$ exists.

Exercise 15.3. Prove the statements in the preceding paragraph.

The inequality extends to finite sums, but not to infinite sums. For example, if we let $s^{(k)}$ denote the sequence $\{0, 0, \ldots, 0, 1, 0, \ldots\}$ with elements $s_n^{(k)}$ equal to 0, except that the kth element $s_k^{(k)}$ of the kth sequence is 1, we have (as $n \to \infty$) that $\limsup(s_n^{(1)} + s_n^{(2)} + \ldots) = \limsup 1 = 1$, whereas $\limsup s_n^{(k)} = 0$ for each k.

Upper and lower limits can also be defined for functions whose ranges are in $\mathbf{R}_1$ but whose domains are more general. If the domain of f is a set S in a metric space, and x_0 is a limit point of S, by $\limsup_{x \to x_0} f(x) = L$ we mean that given any positive ϵ we have $f(x) \leq L + \epsilon$ whenever x is in a sufficiently small neighborhood of x_0, and in addition there is a sequence of points $\{x_n\}$ with limit x_0 such that $f(x_n) \geq L - \epsilon$. Slight modifications have to be made if $L = \pm\infty$. In a similar way we can define $\limsup_{x \to +\infty} f(x)$ if the domain of f is a subset of $\mathbf{R}_1$ that is unbounded above. Examples:

(i) If $f(x) = \sin x$ for $x \in \mathbf{R}_1$, then

$$\limsup_{x\to+\infty} f(x) = +1 \quad \text{and} \quad \liminf_{x\to+\infty} f(x) = -1.$$

(ii) If $f(x) = e^x \sin x$ for $x > 0$, then

$$\limsup_{x\to+\infty} f(x) = +\infty \quad \text{and} \quad \liminf_{x\to+\infty} f(x) = -\infty.$$

(iii) If $f(x) = e^{1/x}$ for $x \neq 0$, then $\limsup_{x\to 0^+} f(x) = +\infty$ and $\liminf_{x\to 0^-} f(x) = 0$.

Exercise 15.4. Show that if $\limsup s_n = \liminf s_n = L$, and L is finite, then $\lim s_n = L$ according to the definition given on page 54.

Exercise 15.5. If $\limsup s_n \leq L$ and $\liminf s_n \geq L$, then $\lim s_n$ exists (and equals L).

Up to this point, although we have considered limits of sequences, we have not had to consider limits of more general functions. For functions with values in $\mathbf{R}_1$, it is simplest to define $\lim_{x\to x_0} f(x)$ to be the common value (if there is one) of $\limsup_{x\to x_0} f(x)$ and $\liminf_{x\to x_0} f(x)$. If the domain of f includes a right-hand neighborhood of x_0, we define as $\limsup_{x\to x_0^+} f(x)$ the upper limit of the restriction of f to a right-hand neighborhood of x_0; similarly for lim inf. The common value (if any) of these right-hand upper and lower limits is denoted by $\lim_{x\to x_0^+} f(x)$, or more compactly by $f(x_0^+)$. There is a similar definition for $f(x_0^-)$.[1]

NOTES

[1] For an interesting analysis of the concept of continuity in terms of limits, see K. P. Williams, Note on continuous functions, *American Mathematical Monthly* 25 (1918), 246–249.

16. Sequences of functions. Before we go on to consider the special properties of further classes of functions, it will be convenient to discuss several kinds of convergence of sequences of functions. This subject was a source of confusion to mathematicians in the first half of the nineteenth century. Even the famous Cauchy, before he hit on the notion we call "Cauchy sequence," mistakenly thought that the limit of a convergent series of continuous functions must be continuous.[1] Nowadays, we can use the terminology of metric spaces to help clarify the different notions of convergence.

Suppose that we have a sequence whose elements are functions s_n with a common domain and with values $s_n(x)$ belonging to $\mathbf{R}_1$. We may fix our attention either on the sequence $\{s_n\}$, whose elements are the functions themselves, or on the sequences $\{s_n(x)\}$ whose elements are the values of the functions at the individual points x of the domain. We say that $\{s_n\}$ *converges pointwise* on a set S if the sequences $\{s_n(x)\}$ of real numbers converge for each x in S. For example, let $s_n(x) = x^n$, with $0 \leq x \leq 1$. For each x in $[0,1]$, the sequence $\{s_n(x)\}$ converges; the limit defines the discontinuous function L such that $L(x) = 0$ for $0 \leq x < 1$, but $L(1) = 1$. On the other hand, we may consider the same functions s_n as points of the metric space C. Then the sequence $\{s_n\}$ does not converge; indeed, $d(s_n, s_{2n}) = \max_{0\leq x\leq 1}(x^n - x^{2n})$, and if we take $x = 2^{-1/n}$ we see that we certainly have $d(s_n, s_{2n}) \geq x^n - x^{2n} = \frac{1}{4}$. Hence $\{s_n\}$ cannot be a Cauchy sequence in the space C.

Exercise 16.1. Examine $\{s_n\}$ for convergence in C, (a) if $s_n(x) = x^n(1 - x)$, (b) if $s_n(x) = nx^n(1 - x)$; in each case $0 \leq x \leq 1$.

We see that a sequence of continuous functions can converge pointwise even though the same sequence of functions, considered as elements of the space C, does not converge. It is clear that convergence of a sequence of elements of C does imply the pointwise convergence of the corresponding sequence of functions. There are, however, other metric spaces whose elements are functions for which convergence of a sequence of elements of the space does not guarantee the pointwise convergence of the sequence of functions. An example is the space, mentioned on page 25, whose elements are continuous functions on $[0, 1]$, with metric given by

$$d(x, y) = \left\{ \int_0^1 |x(t) - y(t)|^2 \, dt \right\}^{1/2}$$

Consider a sequence of elements of this space defined as follows: if $2^n \leq k < 2^{n+1}$, then $x_k(t) = 0$ except in the interval $(\frac{k}{2^n} - 1, \frac{k+1}{2^n} - 1)$; in this interval, the graph of x_k is a triangle with base of length 2^{-n} and altitude 1; this triangle moves back and forth as k increases, thus preventing pointwise convergence; but $d(x_k, 0) < 2^{-n/2} \to 0$.

Whether or not a sequence of functions converges pointwise, we can always define the pointwise upper and lower limits $\limsup s_n(x)$ and $\liminf s_n(x)$ if the functions have a common domain, and range in $\mathbf{R}_1$. If these upper and lower limits are finite, they define two functions, which we can call $\limsup s_n$ and $\liminf s_n$.

It is often convenient to generalize the space C of continuous functions by considering functions whose domains are sets more general than intervals in $\mathbf{R}_1$. In defining C, we used the existence of the maximum of a continuous function whose domain is a compact interval. Since every continuous function whose domain is a compact set has a maximum, we can use just the same definition: let E be a compact set in a metric space; then the space C_E consists of continuous functions f with domain E and range in $\mathbf{R}_1$, and metric $d(f,g)$ defined by $\max_E |f(x) - g(x)|$. Similarly, we can define the space B_E of bounded functions with domain E by taking the metric $d(f,g)$ to be $\sup_E |f(x) - g(x)|$; here E does not have to be compact.

If a sequence of continuous functions, regarded as a sequence of elements of C_E, converges, it is customary to say that it *converges uniformly* on E. More generally, suppose that we have a sequence of functions (bounded or not, continuous or not); if the difference of each two of the functions is bounded, we can form the distance between them in the metric of B_E; and if the sequence of functions is a Cauchy sequence in this metric, we say that it converges uniformly on E. For example, if $f(x) = 1/x$, then the sequence $\{f, f, f, \dots\}$ converges uniformly on $0 < x \leq 1$. The sequence $\{s_n\}$ with $s_n(x) = x^n$ converges uniformly on any specified interval $[0, a]$ with $0 < a < 1$, but not on $[0, 1]$. Notice that it does not converge uniformly on the half-closed interval $[0, 1)$ either. It is important to realize that "uniform convergence on each closed subinterval of an open interval" is not the same as "uniform convergence on the open interval."

Another, more conventional, way of saying that a sequence $\{s_n\}$ is uniformly convergent on E is to say that, given a positive ϵ, there is an N such that, for n and m exceeding N, we have $|s_n(x) - s_m(x)| < \epsilon$ simultaneously for every x in E; we emphasize that N is to be independent of which x in E is being considered. If N is allowed to depend on x, we have the definition of pointwise convergence again.

Suppose we can find numbers M_n with the property that $|s_n(x) - s_{n+1}(x)| \leq M_n$ for all x in E; that is, we have

$\sup_{x\in E} |s_n(x) - s_{n+1}(x)| \le M_n$. Then if $\sum M_n$ converges, the sequence $\{s_n\}$ converges uniformly on E. For, if $m > n$ we have $|s_n(x) - s_m(x)| \le M_n + M_{n+1} + \cdots + M_{m-1}$, and the sum on the right is small when m and n are large. This is known as the *Weierstrass M-test*,[2] and is usually stated in the equivalent form that it assumes when the s_n are the partial sums of a series. This reads as follows: *If c_n are functions with domain E, and $|c_n(x)| \le M_n$ for all x in E* (where we emphasize that M_n is independent of x), *then $\sum c_n$ converges uniformly if $\sum M_n$ converges.*

It is sometimes useful to consider another kind of convergence: pointwise convergence together with uniform boundedness, that is, boundedness in the metric of B_E. We call this *bounded convergence.* For example, the sequence $\{s_n\}$ with $s_n(x) = x^n$ is boundedly convergent on $[0, 1]$, although it is uniformly convergent only on each $[0, a]$, $0 < a < 1$. If $s_n(x) = nx^n$, we still have $\{s_n\}$ uniformly convergent on each $[0, a]$, $0 < a < 1$, but $\{s_n\}$ is not boundedly convergent on $[0, 1)$.

Exercise 16.2. Show that the limit of a boundedly convergent sequence of functions is a bounded function.

NOTES

[1] For a counterexample, see page 121. The error appears in various places in Cauchy's earlier writings. See, for example, his 1821 *Cours d'analyse* reprinted in *Oeuvres Complètes*, II série, Tome III, p. 120. (These collected works of Cauchy were published in 27 volumes by Gauthier-Villars, Paris, starting in 1882.) Also see the reprint of the first part of Cauchy's *Cours d'analyse* with extensive notes by Umberto Bottazzini, Cooperativa Libraria Universitaria Editrice Bologna, Bologna, 1992.

[2] "M" for "majorant."

17. Uniform convergence. One of the useful properties of uniform convergence is that *the limit of a uniformly convergent sequence of continuous functions is continuous.* The same fact can be stated more concisely by saying that *the space C_E is complete.* This is easily proved. In fact, we prove a slightly more general statement that is sometimes useful: *if each s_n is continuous at x_1, and if $\{s_n\}$ converges uniformly in a neighborhood of x_1, then the function L defined by $L(x) = \lim s_n(x)$ is continuous at x_1.* In the first place, uniform convergence implies pointwise convergence, so there is a function L. Let D be the distance function in the space B_E of bounded functions. Then

$$\begin{aligned}|L(x_1) - L(x_2)| &\le |L(x_1) - s_n(x_1)| + |L(x_2) - s_n(x_2)| \\ &\quad + |s_n(x_1) - s_n(x_2)| \\ &\le 2D(L, s_n) + |s_n(x_1) - s_n(x_2)|.\end{aligned}$$

If a positive ϵ is given, then $D(L, s_n)$ can be made less than $\epsilon/3$ by choosing n large enough, since $\{s_n\}$ converges uniformly to L. After choosing n this large, fix n. Then, since s_n is continuous at x_1, there is a positive δ such that $|s_n(x_1) - s_n(x_2)| < \epsilon/3$ when $d(x_1, x_2) < \delta$. Consequently, $d(x_1, x_2) < \delta$ implies that $|L(x_1) - L(x_2)| < \epsilon$, and thus L is continuous at x_1.

Of course, the limit of a nonuniformly convergent sequence of continuous functions is not necessarily discontinuous. For instance, the sequence $\{s_n\}$, where $s_n(x) = nx^n(1 - x)$, is not uniformly convergent on $[0, 1]$ (Exercise 16.1), but its limit is the continuous function 0. However, under additional restrictions it is possible to conclude that if the limit function is continuous, then the convergence is uniform. For example, *if a sequence of continuous functions converges monotonically to a continuous function on a compact interval in $\mathbf{R}_1$, the convergence is necessarily uniform.*[1] The hypothesis of monotonic convergence means that either $s_n(x) \ge s_{n+1}(x)$ for every n and every x in the interval, or else $s_n(x) \le s_{n+1}(x)$ for every n and every x in the interval.

Let us take the hypothesis in the form $s_n(x) \ge s_{n+1}(x)$. Call the limit function L; then $s_n(x) - L(x) \ge 0$. If the convergence is not uniform, then $\max_x[s_n(x) - L(x)]$ does not approach zero, so there is a sequence of values of n for which $\max_x[s_n(x) - L(x)] > b > 0$. Since $s_n - L$ is a continuous function, it attains its maximum at a point x_n; from the set $\{x_n\}$ we can, by Exercise 8.9, select a sequence $\{y_k\}$ with a limit z. Then we have $s_k(y_k) - L(y_k) > b$, and consequently $s_n(y_k) - L(y_k) > b$ for each $n \le k$ (it is only here that we make essential use of the inequality $s_n(x) \ge s_{n+1}(x)$). Letting $k \to \infty$, with fixed n, we infer (by the continuity of $s_n - L$) that $s_n(z) - L(z) \ge b$ for every n. On the other hand, $s_n(z) - L(z) \to 0$, since we have convergence at the point z. Thus the assumption of nonuniform convergence leads to a contradiction.

Another condition that produces the same result is that the functions s_n are monotonic (even if they are not necessarily continuous). More precisely, *let $s_n \to L$ pointwise on a compact interval $[a, b]$, let L be continuous, and let all the s_n be nondecreasing functions. Then $s_n \to L$ uniformly.*

The proof of this theorem requires a fact from §19, but since the theorem fits in well here, I prove it now. Fix a positive number ϵ, and choose a finite set of points x_k in $[a, b]$ such that $a = x_1 < x_2 < \cdots < x_m = b$, and $0 \le L(x_k) - L(x_{k-1}) < \epsilon$ for $k = 2, 3, \ldots, m$. Since L is uniformly continuous (see page 127) and nondecreasing, these inequalities will certainly hold if the distances between consecutive x_k's are small enough. Moreover, since L is nondecreasing, we have $0 \le L(x_k) - L(x) < \epsilon$ for $x_{k-1} \le x \le x_k$. Now since the sequence of functions $\{s_n\}$ converges pointwise, and there are only a finite number of points x_k, we can choose n so large that $|s_n(x_k) - L(x_k)| < \epsilon$ for all k. Since each x is in some $[x_{k-1}, x_k]$, and L is nondecreasing, we have

$$s_n(x) \le s_n(x_k) < L(x_k) + \epsilon < L(x) + 2\epsilon,$$

where we use successively the hypothesis that s_n is nondecreasing, the inequality $s_n(x_k) < L(x_k) + \epsilon$, and the inequality $L(x_k) < L(x) + \epsilon$. Similarly,

$$L(x) - 2\epsilon < L(x_{k-1}) - \epsilon < s_n(x_{k-1}) \leq s_n(x).$$

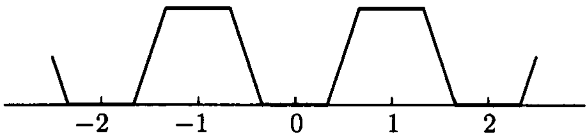
These two sets of inequalities together imply that, if n is large enough, $|s_n(x) - L(x)| < 2\epsilon$ for every x in $[a, b]$; this is the statement of the uniform convergence of $\{s_n\}$.

Exercise 17.1. The limit of a sequence of discontinuous functions may be either continuous or discontinuous, whether or not the convergence is uniform.

One reason for the importance of the idea of uniform convergence is that a good way to construct a function with some special property is often to exhibit it as the uniform limit of functions that do not quite have the property.

As an illustration of this principle (sometimes called the condensation of singularities), we shall exhibit *a continuous curve that passes through every point of a plane area.* (Such space-filling curves are known as Peano curves.) We must of course decide in advance what the phrase "continuous curve" is to mean; the lesson of our construction is that a plausible definition of continuous curve may lead to an object that does not fit the intuitive idea of what a continuous curve should look like.[2]

One natural way of defining a continuous curve in $\mathbf{R}_2$ is to say that it is a continuous image of a line segment, that is, the set of values of a continuous function from a convenient closed interval (say $[0, 1]$) into $\mathbf{R}_2$. Of course, different functions may lead to the same image, but this is irrelevant here; we are going to show that there is at least one function for which the image of the interval covers every point of a whole square, indeed, covers some points more than once. We may represent the points p in the image by their coordinates, letting the image point $(x(t), y(t))$



correspond to the point t of the domain. This amounts to saying that we are thinking of a continuous curve as defined by a pair of parametric equations, $x = x(t)$, $y = y(t)$, with continuous functions x and y.

We are now going to construct a continuous curve, in this sense, that passes through every point of the square where $0 \leq x \leq 1$ and $0 \leq y \leq 1$. As a matter of fact, the curve that we shall construct passes through some points of the square five times. It is possible to refine the construction so that the curve has nothing worse than triple points, but farther than this we cannot go, as is shown in topology.[3] We shall verify presently that the curve must have double points at least, that is, there is no continuous one-to-one mapping of a line segment onto a square.

We base our construction[4] on the properties of a continuous function f that is even, is periodic with period 2, has the value 0 on $[0, \frac{1}{3}]$, has the value 1 on $[\frac{2}{3}, 1]$, and is linear in $(\frac{1}{3}, \frac{2}{3})$. Define two functions x and y by

$$\begin{aligned} x(t) &= \frac{1}{2}f(t) + \frac{1}{2^2}f(3^2t) + \frac{1}{2^3}f(3^4t) + \cdots, \\ y(t) &= \frac{1}{2}f(3t) + \frac{1}{2^2}f(3^3t) + \frac{1}{2^3}f(3^5t) + \cdots. \end{aligned}$$

Both series are uniformly convergent (by the M-test), and hence x and y are continuous functions.[5]

Let $0 \leq x_0 \leq 1$ and $0 \leq y_0 \leq 1$, and represent x_0 and y_0 by "decimals" in base 2 (compare page 40):

$$x_0 = 0.a_0 a_2 a_4 \ldots \quad \text{(base 2)},$$
$$y_0 = 0.a_1 a_3 a_5 \ldots \quad \text{(base 2)}.$$

Now define a number t_0 by taking its expansion in base 3 to be $t_0 = 0.(2a_0)(2a_1)(2a_2)\ldots$ (base 3). That is, t_0 is constructed by doubling the binary digits of x_0 and y_0, interlacing them, and interpreting the result in base 3. We shall now show that $x(t_0) = x_0$ and $y(t_0) = y_0$, so that the curve whose parametric equations are $x = x(t)$, $y = y(t)$ passes through (x_0, y_0).

To do this, we show that $f(3^k t_0) = a_k$ for $k = 0$, $1, 2, \ldots$; it will then be obvious from the definition of $x(t_0)$ and $y(t_0)$ that $x(t_0) = x_0$ and $y(t_0) = y_0$. Now a_k is either 0 or 1. If $a_k = 0$, the number represented by $0.(2a_k)(2a_{k+1})\ldots$ (base 3) is between 0 and $\frac{1}{3}$, and so

$$f(3^k t_0) = f(0.(2a_k)(2a_{k+1})\ldots) = 0;$$

if $a_k = 1$, the number represented by $0.(2a_k)(2a_{k+1})\ldots$ (base 3) is between $\frac{2}{3}$ and 1, and so $f(3^k t_0) = 1$.

We now show that *there cannot be a continuous curve which passes through every point of a square precisely once.* If there were such a curve, it would be the image of a continuous univalent function with domain an interval in $\mathbf{R}_1$, say $[0, 1]$, and range a square in $\mathbf{R}_2$. By the theorem on page 91, this function has a continuous inverse. Since both the function and its inverse have the property that the inverse images of open sets are open, it is equally true that the images of open sets are open. Consider the sets $[0, \frac{1}{2})$ and $(\frac{1}{2}, 1]$, which are open relative to the domain $[0, 1]$. Their images are two sets E_1 and E_2 which are open relative to the range square and which fill the square except for one point. If we take a neighborhood in E_1 and a neighborhood in E_2, we can obviously draw a line segment connecting some point of one neighborhood with some point of the other neighborhood, remaining in the square and not going through the missing point. Thus we obtain two disjoint sets, both open, and neither empty, covering a line segment; this contradicts the connectedness of $\mathbf{R}_1$. Hence our hypothetical continuous curve cannot exist.

On the other hand, we showed on page 17 that there is a one-to-one correspondence between an interval and a square; in the light of what we have just proved, it cannot be continuous.

A useful theorem that uses the idea of uniform convergence can be stated concisely in the following form: *A uniformly convergent sequence can be integrated term by term.* More precisely, if f_n are functions from a finite real interval I into $\mathbf{R}_1$, if $f_n \to f$ uniformly on I, and if each f_n is integrable in the ordinary (Riemann) sense over I, then

$$\int_I f_n(x)\,dx \to \int_I f(x)\,dx.$$

The proof is immediate if we know that f is integrable: if $I = [a, b]$, then

$$\begin{aligned}\left|\int_I f_n(x)\,dx - \int_I f(x)\,dx\right| &= \left|\int_I [f_n(x) - f(x)]\ dx\right| \\ &\leq \int_I |f_n(x) - f(x)|\,dx \\ &\leq (b-a)\sup_x |f_n(x) - f(x)| \\ &\to 0.\end{aligned}$$

The same proof shows that the sequence of indefinite integrals $\int_a^y f_n(x)\,dx$ converges uniformly itself. If each f_n is, for example, continuous, then f is continuous (page 112) and so integrable.

We will not complete the proof by showing that the uniform limit of (Riemann) integrable functions is (Riemann) integrable, for later on (page 212) we shall see a much more powerful convergence theorem for the more general Lebesgue integral. However, even the simplest case, that in which we have a uniformly convergent sequence of continuous functions, has many interesting applications. Here is one of them.

Let f be a function with derivatives of all orders, all of which are then necessarily continuous functions (Exercise 21.1). Suppose that the sequence $\{f^{(n)}\}_{n=1}^{\infty}$ of derivatives converges uniformly to a function L; we might call L a derivative of f of infinite order. As $n \to \infty$, we have on the one hand that

$$\int_a^x f^{(n)}(t)\,dt = f^{(n-1)}(x) - f^{(n-1)}(a) \to L(x) - L(a),$$

and on the other hand, by our theorem on integrating uniformly convergent sequences,

$$\int_a^x f^{(n)}(t)\,dt \to \int_a^x L(t)\,dt.$$

Consequently $\int_a^x L(t)\,dt = L(x) - L(a)$. This implies that $L(x) = L'(x)$, and therefore $L(x) = ce^x$. Thus *any derivative of infinite order is necessarily a simple exponential*[6] (including the case where it is identically zero), no matter what function gave rise to it.

The same theorem on uniform convergence can be used to establish a theorem on termwise differentiation of a sequence of functions. *If the functions f_n have continuous derivatives on a bounded interval I, if $\{f_n(a)\}$ converges* *for some a in I, and if $\{f_n'\}$ converges uniformly, then f_n converges uniformly to a limit f which is differentiable, and* $\lim f_n' = f'$.

For suppose $f_n' \to g$ uniformly. Then

$$f_n(x) - f_n(a) = \int_a^x f_n'(t)\,dt \to \int_a^x g(t)\,dt.$$

Since $f_n(a) \to f(a)$, it follows that $f_n(x)$ converges for each x. Moreover, $|f_n(x) - f_m(x) + f_m(a) - f_n(a)|$ equals

$$\left|\int_a^x (f_n'(t) - g(t))\,dt - \int_a^x (f_m'(t) - g(t))\,dt\right|$$
$$\leq (\text{length}\, I)(\sup_{t\in I} |f_n'(t) - g(t)| + \sup_{t\in I} |f_m'(t) - g(t)|),$$

and since the right-hand side is independent of x and tends to zero, $\{f_n\}$ converges uniformly, to some limit f. Now

$$f_n(x) - f_n(a) \to f(x) - f(a),$$

whence

$$f(x) - f(a) = \int_a^x g(t)\,dt.$$

Hence f is differentiable, and $f'(x) = g(x)$; we take for granted the fact that the indefinite integral of a continuous function has the integrand as derivative (compare page 214).

Later (page 146) we shall prove a more general theorem in which no continuity is assumed for f_n'. There is some point in doing so, since a function can have a derivative at every point with the derivative not integrable, either in the Riemann or the Lebesgue sense (page 215).

Exercise 17.2. If f is a function from $\mathbf{R}_1$ into $\mathbf{R}_1$, having derivatives of all orders, and the series

$$\cdots + \int_0^x dt \int_0^t f(u)\,du + \int_0^x f(t)\,dt + f(x) + f'(x) + \cdots$$

converges uniformly, what is its sum?[7]

Our applications of uniform convergence have been precise formulations under various circumstances of the principle that we can take limits term by term in a uniformly convergent sequence. Here is another one.

Exercise 17.3. Tannery's theorem:[8] if $f_n(k) \to L_n$ for each n as $k \to \infty$; and $|f_n(k)| \le M_n$ for all k, where $\sum M_n$ converges; and if $p = p(k) \to \infty$ when $k \to \infty$; then

$$\lim_{k\to\infty} \{f_1(k) + f_2(k) + \cdots + f_p(k)\} = \sum_{n=1}^{\infty} L_n .$$

Exercise 17.4. As an application of Tannery's theorem, prove that $\lim_{k\to\infty} (1 + x/k)^k = \sum_{n=0}^{\infty} x^n/n!$.

There are fairly simple situations in which the theorem on term-by-term integration of a uniformly convergent sequence is inadequate. For example, $1 - x + x^2 - \cdots + (-x)^n = [1-(-x)^{n+1}]/(1+x)$, and hence $1-x+x^2-\cdots = 1/(1+x)$ if $|x| < 1$. Formal integration suggests that

$$\begin{aligned}\log 2 &= \int_0^1 \frac{dx}{1+x} = \int_0^1 dx - \int_0^1 x\,dx + \int_0^1 x^2\,dx - \cdots \\ &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots .\end{aligned}$$

We cannot justify this calculation by the theorem on uniform convergence, since the sequence $\{f_n\}$ with $f_n(x) =$

$[1-(-x)^n]/(1+x)$ does not converge uniformly on $[0,1]$ (since it does not converge at 1), and does not converge uniformly even on $[0,1)$. Indeed, if it did, the supremum on $[0,1)$ of $|x^n|/(1+x)$ would approach zero as $n \to \infty$; but since $|x|^n/(1+x) \geq \frac{1}{2}|x|^n$, the supremum in question is at least $\frac{1}{2}$ and so cannot approach zero. In this case it is easy to verify the result of the formal calculation:

$$1 - \frac{1}{2} + \frac{1}{3} - \cdots + \frac{(-1)^{n-1}}{n} = \int_0^1 \frac{1-(-x)^n}{1+x}\,dx$$

$$= \int_0^1 \frac{dx}{1+x} - \int_0^1 \frac{(-x)^n}{1+x}\,dx,$$

$$\left|1 - \frac{1}{2} + \frac{1}{3} - \cdots + \frac{(-1)^{n-1}}{n} - \log 2\right| = \int_0^1 \frac{x^n}{1+x}\,dx$$

$$< \int_0^1 x^n\,dx = \frac{1}{n+1} \to 0.$$

This calculation shows simultaneously that the series $1 - \frac{1}{2} + \frac{1}{3} - \cdots$ converges and that its sum is $\log 2$.

Similarly, it is not very difficult to show (although we shall not do it) that

$$\frac{x}{2} = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}\sin nx}{n}, \qquad -\pi < x < \pi.$$

Niels Abel gave this example[9] to refute Cauchy's mistaken early belief that a convergent series of continuous functions must have a continuous sum: the series converges at $x = \pi$, but to 0, not to $\pi/2$. Formal integration yields

$$\frac{\pi^2}{4} = \int_0^{\pi} \frac{x}{2}\,dx = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \int_0^{\pi} \sin nx\,dx$$

$$= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}(1 - \cos n\pi)}{n^2} = \sum_{n=1}^{\infty} \frac{2}{(2n-1)^2},$$

so that we have a simple way of summing the numerical series

$$\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \cdots .$$

Since the original series is not uniformly convergent on the interval $[0, \pi]$ (because the sum is discontinuous at π), the justification of our evaluation of the numerical series eludes the elementary theory. We will consider some more sophisticated convergence theorems for integrals in §28.

NOTES

[1] This is known as Dini's theorem. For it and the following theorem, see G. Pólya and G. Szegő, *Problems and Theorems in Analysis*, vol. I, Springer, New York, 1972, pp. 81 and 269–270, problems 126 and 127 of part II.

[2] For analyses of the notion of continuous curve from different points of view, see G. T. Whyburn, What is a curve?, *American Mathematical Monthly* 49 (1942), 493–497; J. W. T. Youngs, Curves and surfaces, *American Mathematical Monthly* 51 (1944), 1–11.

[3] W. Hurewicz, Über dimensionserhöhender stetige Abbildungen, *Journal für die Reine und Angewandte Mathematik* 169 (1933), 71–78.

[4] I. J. Schoenberg, On the Peano curve of Lebesgue, *Bulletin of the American Mathematical Society* 44 (1938), 519. Another simple construction was given by Liu Wen, A space filling curve, *American Mathematical Monthly* 90 (1983), 283.

[5] It can be shown that there is no point where x and y are simultaneously differentiable, that is, the curve has a tangent at no point. James Alsina, The Peano curve of Schoenberg is nowhere differentiable, *Journal of Approximation Theory* 33 (1981), 28–42.

[6] For extensions, see Lee Lorch, Derivatives of infinite order, *Pacific Journal of Mathematics* 3 (1953), 773–778.

[7] Pólya and Szegő, book cited above, pp. 37 and 214, problem 165 of part I.

[8]See T. J. I'A. Bromwich, *An Introduction to the Theory of Infinite Series*, second edition, MacMillan, London, 1926, reprinted 1965, pp. 136–137.

[9]N. Abel, Untersuchungen über die Reihe: $1+\frac{m}{1}x+\frac{m(m-1)}{1\cdot 2}x^2+\frac{m(m-1)(m-2)}{1\cdot 2\cdot 3}x^3+\cdots$, *Journal für die Reine und Angewandte Mathematik* 1 (1826), 311–339; a French version appears in Abel's *Oeuvres Complètes*, Grøndahl, Christiania, 1881, pp. 219–250.

18. Pointwise limits of continuous functions. We consider functions from an interval in $\mathbf{R}_1$ into $\mathbf{R}_1$. Although a *pointwise limit of continuous functions* is not necessarily continuous, it cannot be extremely discontinuous, as we shall now show: *its points of continuity must at least form an everywhere dense set.*[1] (Hence, for example, the everywhere discontinuous function mentioned on page 82, which was obtained from continuous functions by two successive limiting processes, cannot be obtained by one limiting process involving only continuous functions.)

We begin by observing that if a function f is discontinuous at a point x, then the images of arbitrarily small neighborhoods of x do not have arbitrarily small diameters. That is, there is an integer n such that the diameter of the image of each neighborhood of x is at least $1/n$. (The images of larger neighborhoods have, if anything, larger diameters, so we can say "every neighborhood" instead of "small neighborhoods.") Now suppose that f is discontinuous at every point of an interval, and let E_n be the set of points x in this interval for which the diameter of the image of every neighborhood of x is at least $1/n$. As we have just seen, every x belongs to some E_n. Moreover, E_n is a closed set. For, if y is a limit point of E_n, every neighborhood of y contains some point x of E_n, and hence contains a neighborhood of x, and so the diameter of the image of every neighborhood of y is also at least $1/n$.

We now use Baire's theorem, which tells us that some E_M is dense in some subinterval J. Since E_M is closed, E_M contains J. That is, we have found an interval J with the property that the image of every subinterval of J has diameter at least $1/M$. The existence of such an interval J is, then, a consequence of the property of being discontinuous at every point of some interval. We shall now show that a pointwise limit of continuous functions cannot have such intervals J, and hence cannot be discontinuous at every point of an interval.

The range of f, being some subset of $\mathbf{R}_1$, can be covered by countably many intervals $I_n = (a_n, b_n)$, each of length less than $1/M$. Let us look at the inverse images H_n of these I_n. The sets H_n collectively cover the interval J, but none of them can contain a subinterval of J, since the images of subintervals of J all have diameters greater than $1/M$. On the other hand, by Baire's theorem one of the H_n is dense in a subinterval of J. If we knew that the H_n were closed, we should have a contradiction, since a closed set that is dense in an interval contains that interval.

Even when f is a pointwise limit of continuous functions, there is no reason to suppose that the H_n are closed. However, we can show that each H_n is a countable union of closed sets, and this property will do just as well. For, if the H_n have this property, then, by still another application of Baire's theorem, one of the closed sets is dense in a subinterval of J, and so contains that subinterval. Since the closed sets are subsets of H_n, it follows that H_n will also contain a subinterval of J.

We have thus reduced the proof of our theorem to showing that if f is a pointwise limit of continuous functions f_k, then the sets H_n are countable unions of closed sets. We recall that H_n is the inverse image under f of the interval (a_n, b_n); that is, H_n is the set of points x such that $a_n < f(x) < b_n$. Consider an x in H_n. Then if j is large enough, we have $a_n + 2/j \leq f(x) \leq b_n - 2/j$, since $a_n < f(x) < b_n$. Since $f_k(x) \to f(x)$, we then have $a_n + 1/j \leq f_k(x) \leq b_n - 1/j$ for all sufficiently large k. Let $E_{k,j}$ be the set on which this last inequality holds, and let $F_{m,j}$ be the intersection of $E_{m,j}$, $E_{m+1,j}$, $E_{m+2,j}$, and so on. Because f_k is continuous, the sets $E_{k,j}$ are closed (they are inverse images of closed intervals), and the sets $F_{m,j}$ are closed (they are intersections of closed sets). We have just seen that if x is in H_n, then x is in some $F_{m,j}$. That is, H_n is a subset of the union of all the $F_{m,j}$. On the other hand, if x is in some $F_{m,j}$, we have $a_n + 1/j \leq f_k(x) \leq b_n - 1/j$ for some j and all sufficiently large k; since $f_k(x) \to f(x)$, this implies that $a_n + 1/j \leq f(x) \leq b_n - 1/j$, and so $x \in H_n$. Thus H_n is precisely the union of all the $F_{m,j}$ for all positive integers m and j, so we have succeeded in representing H_n as a countable union of closed sets. This is what was required to complete the proof.

Only slight modifications of the proof show that a limit of continuous functions has points of continuity everywhere dense in every nonempty perfect set. In this form the condition can be shown to be reversible: a function whose points of continuity are everywhere dense in every nonempty perfect set can be represented as a limit of continuous functions.

Baire described discontinuous limits of continuous functions as of class 1; limits of functions of class 1, not themselves of class 1, as of class 2; and so on.[2] There are functions that do not belong to any Baire class.

An interesting example of a function of Baire class 1 is any discontinuous function on $\mathbf{R}_2$ that is continuous on each line parallel to a coordinate axis.[3]

Exercise 18.1. Construct an example of such a function.

It is interesting that although we asserted only that a pointwise limit of continuous functions has an everywhere dense set of points of continuity, more than this is true: its set of points of discontinuity must form only a set of the first category. This property is really independent of the fact that the function in question is a pointwise limit of continuous functions. In fact, we shall show that a real-valued function f whose domain is an interval in $\mathbf{R}_1$, if continuous at the points of an everywhere dense set, is continuous except on a set of first category.

Consider the sets E_n of points x such that there exists a sequence $\{y_k\}$ whose limit is x and which has the property that $|f(y_k)-f(x)| > 1/n$. Every point of discontinuity is in some E_n;

that is, the set of points of discontinuity is a subset of the union of the E_n. Since there are countably many E_n, our theorem is proved if we show that each E_n is nowhere dense. If some E_n failed to be nowhere dense, some point x where f is continuous would be a limit point of that E_n. If we choose a positive δ such that $|y - x| < \delta$ implies that $|f(y) - f(x)| < 1/(2n)$, and then choose a point w of E_n so that $|w - x| < \frac{1}{2}\delta$, and let y_k be the points occurring in the definition of E_n, we have

$$|f(y_k) - f(w)| \le |f(y_k) - f(x)| + |f(x) - f(w)| < 1/n$$

for sufficiently large k, contradicting the definition of E_n.

NOTES

[1] W. F. Osgood. For more on the subject of this section, see E. R. Lorch, Continuity and Baire functions, *American Mathematical Monthly* 78 (1971), 748–762; A. C. M. van Rooij and W. H. Schikhof, *A Second Course on Real Functions*, Cambridge University Press, 1982, §§10–11 and 16–17; C. de La Vallée Poussin, *Intégrales de Lebesgue, Fonctions d'ensemble, Classes de Baire*, second edition, Gauthier-Villars, Paris, 1934, chap. VII–VIII.

[2] René Baire, *Leçons sur les Fonctions Discontinues*, Gauthier-Villars, Paris, 1905.

[3] H. Lebesgue. For a simple proof, see F. W. Carroll, Separately continuous functions are Baire functions, *American Mathematical Monthly* 78 (1971), 175.

19. Approximations to continuous functions. We have seen (page 69) that a continuous function from $\mathbf{R}_1$ into $\mathbf{R}_1$ may have a rather irregular graph, at least to the extent of having oscillations in every interval. On the other hand, there is always a quite smooth function whose graph is very close to the graph of a given continuous function. More precisely, *if the domain of a continuous function is a compact interval, then we can find, as close as we please to the function, each of the following: a step function, a continuous polygonal function, and a polynomial.* A step function has a graph made up of a finite number of horizontal line segments; a polygonal function has a graph made up of a finite number of line segments of any orientation (not vertical). Here "as close as we please" is to be interpreted in the metric of the space B. In other words, if f is the continuous function, and ϵ is a given positive number, there are a step function f_1, a continuous polygonal function f_2, and a polynomial f_3 such that $|f(x) - f_k(x)| < \epsilon$ (for $k = 1$, 2, and 3) for all x in the interval in question.

The property that makes such approximations possible is called *uniform continuity.* A function f is continuous at x if, given a positive ϵ, there is a positive δ such that $|f(x) - f(y)| < \epsilon$ if $|x - y| < \delta$. Here we shall in general have to take smaller and smaller δ's as we consider different x's. If it is always possible (for a given f) to find a δ that will work simultaneously for all x in a given set, the function f is said to be *uniformly continuous* on that set. We now show that *a continuous function is uniformly continuous on any compact set in its domain.* Indeed, we shall establish this theorem in a more general setting: a continuous function whose domain and range are in metric spaces is uniformly continuous on any compact subset S of its domain.

The idea of the proof is that if a certain δ works for a certain x, then a slightly smaller δ works for all nearby x's. Using compactness, we can find a finite set of δ's one of which works for any given x; the minimum of this finite set of δ's works simultaneously for all x.

To make this idea precise, fix a positive number ϵ. Attach to x first a neighborhood N such that $d(f(x), f(y)) < \epsilon/2$ for all y in N, and then the neighborhood M of center x and half the radius of N. The neighborhoods M cover S, and since S is compact, some finite number of them still cover S. Let these covering neighborhoods be M_1, M_2,

$\dots, M_n$. Let δ be the smallest radius of any M_k. Now let x and y be any two points of S such that $d(x, y) < \delta$. Since x is in some M_k, there is a point z that is the center of an M_k in which x lies. By the triangle inequality, $d(y, z) \leq d(x, y) + d(x, z) < d(x, z) + \delta$.

Since δ is the smallest radius of an M_k, this inequality shows that y is in the N_k whose center is z. Hence both $d(f(x), f(z)) < \epsilon/2$ and $d(f(y), f(z)) < \epsilon/2$, so that by the triangle inequality $d(f(x), f(y)) < \epsilon$.

Exercise 19.1. Can a function f be simultaneously uniformly continuous and unbounded on $0 < x \leq 1$?

Exercise 19.2. If f is continuous for $x \geq 0$ and $f(x) \to L$ (finite) as $x \to +\infty$, must f be uniformly continuous for $x \geq 0$?

Exercise 19.3. If f is uniformly continuous for $x \geq 0$, must $f(x)$ approach a limit (finite or infinite) as $x \to +\infty$?

We now construct the three approximating functions whose existence was asserted above.

Let S be a compact interval $[a, b]$ in $\mathbf{R}_1$. We have just seen that if a positive tolerance ϵ is prescribed, we can find a positive δ such that the given function f varies by less than ϵ on each subinterval of $[a, b]$ of length less than δ. To construct a step function f_1 that approximates f within ϵ, divide $[a, b]$ into a finite number of subintervals, each of length less than δ, and over each subinterval draw a horizontal line segment whose height is (for example) equal to the value of f at the midpoint of the subinterval.

To construct a continuous polygonal approximation f_2, shorten each step of f_1 by a small amount at each end, and then join the right-hand endpoint of each reduced step to the left-hand endpoint of the next reduced step by a line segment.

To construct a polynomial approximation f_3 is somewhat harder.[1] One procedure is as follows. We may suppose, merely in order to simplify the formulas, that the domain of our given continuous function is an interval $[h, 1-h]$, where $0 < h < \frac{1}{2}$. We can extend our function in an obvious way so that the extended function f is continuous on $\mathbf{R}_1$ and is zero outside $(\frac{1}{2}h, 1 - \frac{1}{2}h)$. Now consider the function I defined by

$$I(x) = c_n \int_0^1 f(t)[1 - (x-t)^2]^n \, dt,$$

where

$$1/c_n = \int_{-1}^1 (1 - t^2)^n \, dt.$$

Evidently I is a polynomial of degree (at most) $2n$. The factor in brackets in the integrand has a peak at $t = x$ and (when n is large) is small when t is not near x, so it may seem plausible that $I(x)$ should be close to $f(x)$. We now show that this is really the case.

We can write

$$I(x) = c_n \int_{x-1}^{x} f(x-s)(1-s^2)^n \, ds,$$

and since $f(t) = 0$ when $t < 0$ or $t > 1$, this is the same as

$$I(x) = c_n \int_{-1}^{1} f(x-s)(1-s^2)^n \, ds.$$

Because of the way c_n was defined, we can write

$$I(x) - f(x) = c_n \int_{-1}^{1} [f(x-s) - f(x)](1-s^2)^n \, ds.$$

We now break the integral up into three parts,

$$I_1 = c_n \int_{-1}^{-\delta}, \quad I_2 = c_n \int_{-\delta}^{\delta}, \quad I_3 = c_n \int_{\delta}^{1},$$

where $0 < \delta < 1$, and δ is to be chosen in a moment.

At this point we use the uniform continuity of f. If ϵ is an arbitrary positive number, we can find δ small enough so that $|f(x-s)-f(x)| < \frac{1}{3}\epsilon$ if $|s| < \delta$, where the inequality holds for all x with the same δ. We use this inequality to estimate I_2:

$$|I_2| \le \frac{\epsilon c_n}{3} \int_{-\delta}^{\delta} (1-s^2)^n \, ds < \frac{\epsilon c_n}{3} \int_{-1}^{1} (1-s^2)^n \, ds = \frac{\epsilon}{3}.$$

Next, since f is continuous on a compact set, it is bounded, say $|f| \le M$. In I_3, we have $(1-s^2)^n \le (1-\delta^2)^n$, while

$$1/c_n = \int_{-1}^{1} (1-t^2)^n \, dt > \int_{-\delta/2}^{\delta/2} (1-t^2)^n \, dt > \delta(1-\tfrac{1}{4}\delta^2)^n.$$

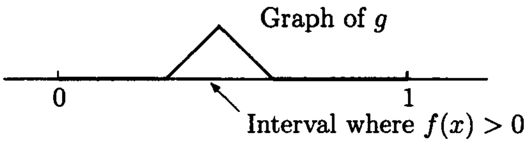
Hence we have

$$\begin{aligned} |I_3| &\le 2Mc_n \int_{\delta}^{1} (1-s^2)^n \, ds \\ &\le 2M\delta^{-1}(1-\delta)(1-\delta^2)^n(1-\tfrac{1}{4}\delta^2)^{-n}. \end{aligned}$$

Since $(1-\delta^2)/(1-\frac{1}{4}\delta^2) < 1$, this estimate shows that $I_3 \to 0$ as $n \to \infty$. Exactly similar reasoning applies to I_1. By taking n large enough, we then have $|I_1|$ and $|I_3|$ each less than $\frac{1}{3}\epsilon$, so

$$|I(x) - f(x)| \le |I_1| + |I_2| + |I_3| < \epsilon,$$

provided that n is large enough. Therefore the polynomial I furnishes the approximation f_3 if n is large enough.



The possibility of approximating a continuous function by polynomials has many applications. It was used in §10 in the proof of the existence of continuous nowhere differentiable functions. Here is another application.

Let f be continuous on the interval $[a, b]$. The quantities $\int_a^b f(x)x^n\,dx$ (for $n = 0, 1, 2, \ldots$) are called the *moments* of f. We shall show that a continuous function with a compact domain in $\mathbf{R}_1$ is determined by its moments; that is, *two continuous functions with the same sequence of moments are identical.*[2] (We do not say anything about how a continuous function can actually be calculated from its moments.) It is an equivalent statement that a continuous function all of whose moments are zero must vanish identically, and this we now prove. Suppose that all the moments of f are zero. Without loss of generality, we suppose that $a = 0$ and $b = 1$. If f is not identically zero, it is positive (or negative) in some interval, and we can construct a continuous function g that is zero outside this interval and that makes $\int_0^1 fg\,dx = 2h > 0$. (See the figure.) Let M be an upper bound for $|f|$ on the interval $[0, 1]$. Construct a polynomial P such that $|g(x) - P(x)| < h/M$ for all x. Then

$$\begin{aligned}\int_0^1 fP\,dx &= \int_0^1 fg\,dx - \int_0^1 f\cdot(g - P)\,dx \\ &\geq 2h - M\cdot\max|g(x) - P(x)| > h.\end{aligned}$$

But $\int_0^1 fP\,dx = 0$, since all the moments of f vanish. The contradiction can be avoided only by having f identically zero.

As a corollary of the theorem about moments, we see that the set of all continuous functions can be put into one-to-one correspondence with a class of sequences of real numbers, since different continuous functions have different sequences of moments. Since there are just as many sequences of real numbers as there are real numbers (Exercise 3.13), it follows that there are just as many continuous functions as there are real numbers. (A more direct way of seeing this is to observe that a continuous function is determined by its values at the rational points, that is, by a sequence of real numbers.)

NOTES

[1] The theorem that a continuous function can be uniformly approximated by polynomials on an interval is known as the Weierstrass approximation theorem. The proof given here was invented by E. Landau.

[2] For a more general result, see David Vernon Widder, *The Laplace Transform*, Princeton University Press, eighth printing, 1972, pp. 60–61 and Chapter III.

20. Linear functions. A function f whose domain is $\mathbf{R}_1$ is said to be *linear* if $f(x)+f(y) = f(x+y)$ for all x and y. (This is a more special use of the term "linear" than is often made: the function f defined by $f(x) = ax + b$ is not linear in our sense when $b \neq 0$.) Clearly if $f(x) = ax$ then f is linear, and we might expect that all linear functions would be of this form; but not all of them are. To exhibit one that is not, we should have to appeal to one of the more abstruse properties of the real number system, which depends on ideas that have not been introduced in this book.[1] What is also not immediately obvious, but will be established shortly, is that a discontinuous linear function is necessarily wildly discontinuous: it is, for example, unbounded in every interval, and indeed its graph must be everywhere dense in $\mathbf{R}_2$. It is only to be expected that no very simple construction will produce a function of this kind.

Let us consider a linear function f. For every x we have $f(2x) = f(x + x) = 2f(x)$, and so by induction $f(nx) = nf(x)$ for every positive integer n. Since $f(x) = f(x + 0) = f(x) + f(0)$, we have $f(0) = 0$. Since $0 = f(0) = f(x - x) = f(x) + f(-x)$, we have $f(-x) = -f(x)$. Therefore $f(nx) = nf(x)$ for every integer, positive or negative. Replacing x by x/n, we obtain $f(x) = nf(x/n)$, or $f(x/n) = n^{-1}f(x)$. Now replace x by mx, and we have $f(mx/n) = n^{-1}f(mx) = (m/n)f(x)$. In other words, $f(rx) = rf(x)$ for every rational number r. In particular (put $x = 1$), $f(r) = rf(1)$ for every rational number r. It follows that if f is continuous for all x, then $f(x) = xf(1)$ for all x.

We can easily strengthen this result by showing that $f(x) = xf(1)$ for all x provided merely that f is continuous at some point c. For then $f(c + \delta) - f(c) \to 0$ as $\delta \to 0$; but $f(c + \delta) - f(c) = f(\delta)$, so $f(\delta) \to 0$ as $\delta \to 0$. That is, f is continuous at 0. Now if x is any real number, then $f(x + \delta) - f(x) = f(\delta) \to 0$ as $\delta \to 0$, so f is continuous at x. Thus f is continuous throughout $\mathbf{R}_1$, and we know this implies that $f(x) = xf(1)$.

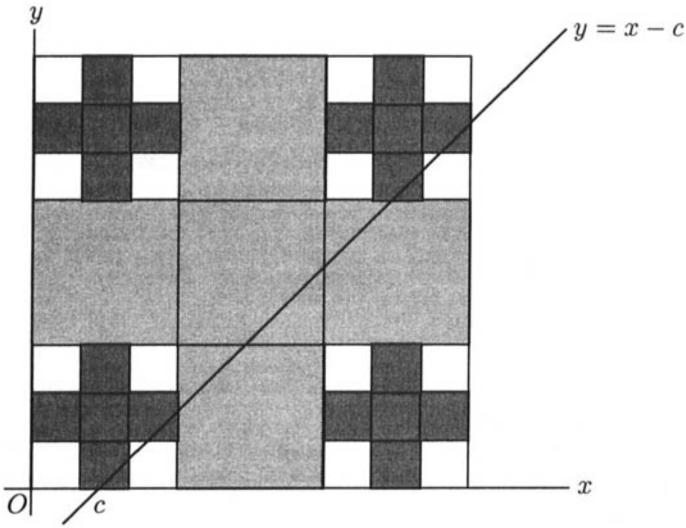
We can go still further in weakening the hypothesis and nevertheless being able to prove that a linear function is continuous. Suppose that f is merely bounded on some interval, or even on some set E having the following property: the set of all differences $x - y$ between points x and y of E contains a neighborhood of 0. That is, there is a positive δ such that if $|t| < \delta$, then there are points x and y in E such that $x - y = t$. Then we can still conclude[2] that f is continuous if it is linear, and so a linear f that is bounded on a set of the kind just described must be of the form $f(x) = cx$.

Suppose $|f(x)| \leq M$ on E. For numbers t that are differences between points of E, we can write $|f(t)| = |f(x-y)| = |f(x) - f(y)| \leq 2M$. Therefore if $|u| < \delta/n$, we have $|f(u)| = n^{-1}|f(nu)| \leq 2M/n$. Now let s be any real number, and let r be a rational number such that $|r - s| < \delta/n$. Then we have

$$\begin{aligned} |f(s) - sf(1)| &= |f(s-r) + (r-s)f(1)| \\ &\leq \frac{2M + \delta|f(1)|}{n}. \end{aligned}$$

Since n can be as large as we please, $f(s) - sf(1) = 0$.

There are many sets E, other than intervals, having the property used in this proof. They include the sets of positive Lebesgue measure (see §25), and some sets of measure zero. For example, the Cantor set (page 39) has the property. The proof of this fact can be given a very intuitive geometrical form.[3] Take the usual coordinate system in $\mathbf{R}_2$, and construct Cantor sets on the intervals $[0,1]$ of both the x and y axes, removing from the plane not only the middle thirds of intervals, but also all the points of the unit square that have one coordinate (at least) in a deleted interval, so that at each step we remove some cross-shaped sets. (See the figure.) Consider a line with equation $y = x - c$, where $0 \leq c \leq 1$. At each step, the line meets at least one of the squares that is not deleted in this step; these squares are closed, and nested, so their intersection contains some point (x, y) with $y = x - c$, and both x and y are points of the Cantor set.



To show[4] that $f(x) = cx$ for a linear function f whose graph is not everywhere dense in $\mathbf{R}_2$, we may appeal to an elementary fact from the theory of numbers. Let (x_1, x_2) and (y_1, y_2) be two pairs of real numbers that are not proportional; this means that $x_1y_2 \neq x_2y_1$, or in geometrical language that the points (x_1, x_2) and (y_1, y_2) of $\mathbf{R}_2$ are not on the same straight line through the origin. Then if a and b are any two real numbers whatsoever, we can find rational multipliers r_1 and r_2 so that $r_1x_1 + r_2x_2$ is as close as we like to a and simultaneously $r_1y_1 + r_2y_2$ is as close as we like to b. To prove this, we solve the equations $ux_1 + vx_2 = a$ and $uy_1 + vy_2 = b$ (which we can do since their determinant is not zero), and then choose r_1 and r_2 close to u and v, respectively.

Now suppose that f is linear and not of the form $f(x) =$

cx. The second hypothesis implies that we must be able to find points x_1 and x_2 such that $f(x_1)/x_1 \neq f(x_2)/x_2$. Then for any point (a, b) in $\mathbf{R}_2$, we can find rational numbers r_1 and r_2 such that $f(r_1x_1+r_2x_2) = r_1f(x_1)+r_2f(x_2)$ differs arbitrarily little from b, and at the same time $r_1x_1 + r_2x_2$ differs arbitrarily little from a. Thus there is a point of the graph of f as close as we please to the point (a, b) of $\mathbf{R}_2$.

Another proof that yields some additional insight can be given as follows. Suppose that f is linear but not of the form $f(x) = cx$, and we wish to show that the graph of f is dense in $\mathbf{R}_2$. Since $f(t+r) = f(t)+rf(1)$ for rational r, it is enough to show that f takes, arbitrarily close to the origin, values arbitrarily close to every real number, or, what is the same thing, arbitrarily close to every rational number. Let A be a positive rational number and $\epsilon < 1$ a small positive number that we use to specify closeness. We know that f is unbounded in $(0, \epsilon)$; suppose, for definiteness, that f takes arbitrarily large positive values there. Then there is an integer n exceeding A/ϵ such that for some s in $(0, \epsilon)$ we have $n + 1 \geq f(s) \geq n$. Since $f(rx) = rf(x)$ for all rational r and all x, we have $f(As/n) = (A/n)f(s)$, so $A + \epsilon > A(n + 1)/n \geq f(As/n) \geq A$, and As/n is a point of the interval $(0, \epsilon)$. We have therefore constructed a point arbitrarily close to 0 where f takes a value arbitrarily close to A, if $A > 0$. If $A < 0$, there is a point close to 0 where f takes a value close to $-A$, and since $f(-x) = -f(x)$, the same conclusion follows.

Here is an application in calculus of our theorems about linear functions.[5] *Suppose that the limit*

$$\lim_{R\to\infty} \frac{1}{2R} \int_{x-R}^{x+R} f(u)\, du$$

exists for every real x; denote it by $\varphi(x)$. We shall now show that $\varphi(x)$ *is necessarily of the form* $ax + b$. By the definition of $\varphi(x)$,

$$\varphi(x-h) = \lim_{R\to\infty} \frac{1}{2R} \int_{x-h-R}^{x-h+R} f(u)\,du,$$
$$\varphi(x+h) = \lim_{R\to\infty} \frac{1}{2R} \int_{x+h-R}^{x+h+R} f(u)\,du,$$

and

$$\begin{aligned}\varphi(x) &= \lim_{R\to\infty} \frac{1}{2R} \int_{x-(R+h)}^{x+(R+h)} f(u)\,du \\ &= \lim_{R\to\infty} \frac{1}{2R} \int_{x-(R-h)}^{x+(R-h)} f(u)\,du.\end{aligned}$$

Consequently, $\varphi(x-h) + \varphi(x+h) = 2\varphi(x)$. If we replace $x - h$ and $x + h$ by x and y, this says that $\varphi(x) + \varphi(y) = 2\varphi(\frac{1}{2}(x+y))$. Let $\varphi(x) - \varphi(0) = \psi(x)$; then we have

$$\begin{aligned}\psi(x) + \psi(y) &= \varphi(x) + \varphi(y) - 2\varphi(0) \\ &= 2\varphi(\tfrac{1}{2}(x+y)) - 2\varphi(0) \qquad (*) \\ &= 2\psi(\tfrac{1}{2}(x+y)).\end{aligned}$$

This is true for every y, and so in particular for $y = 0$; but $\psi(0) = 0$, so $\psi(x) = 2\psi(\frac{1}{2}x)$. Next replace x by $x + y$ to get $\psi(x+y) = 2\psi(\frac{1}{2}(x+y))$. However, $(*)$ above says that $2\psi(\frac{1}{2}(x+y)) = \psi(x) + \psi(y)$. Therefore $\psi(x) + \psi(y) = \psi(x+y)$; that is, ψ is linear. Now ψ is a limit of continuous functions, and so must have points of continuity (page 123). But we know that a linear function ψ that has a point of continuity has the form $\psi(x) = x\psi(1)$, which is to say that $\varphi(x) = \varphi(0) + x\psi(1)$, as asserted.

Exercise 20.1. Let $\varphi(x)$ denote[6]

$$\lim_{R\to\infty}\int_R^{x+R} f(u)\,du,$$

supposed to exist for every real x. Then $\varphi(x) = ax$.

NOTES

[1] What is needed is the existence of a Hamel basis for the real numbers: see K. Hrbáček and T. Jech, *Introduction to Set Theory*, Marcel Dekker, New York, 1978, pp. 144–147 (a second edition appeared in 1984). A discontinuous linear function is constructed in G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, second edition, Cambridge University Press, 1952, p. 96. Hamel's original paper is cited in note 4 below. See also G. S. Young, The linear functional equation, *American Mathematical Monthly* 65 (1958), 37–38. Additional history and references are given by J. W. Green and W. Gustin, Quasi-convex sets, *Canadian Journal of Mathematics* 2 (1950), 489–507.

[2] For this proof, and extensions, see H. Kestelman, On the functional equation $f(x + y) = f(x) + f(y)$, *Fundamenta Mathematicae* 34 (1947), 144–147. See also A. Wilansky, Additive functions, in *Lectures on Calculus*, K. O. May, ed., Holden-Day, San Francisco, 1967, pp. 97–124; and L. Reich, Über Approximation durch Werte additiver Funktionen, *Österreichische Akademie der Wissenschaften Mathematische-Naturwissenschaftliche Klasse. Sitzungsberichte. Abteilung II. Mathematische, Physikalische und Technische Wissenschaften* 201 (1992), no. 1–10, 169–181.

[3] A. Zygmund, *Trigonometric Series*, second edition, vol. 1, Cambridge University Press, reprinted 1977, p. 235. The property was discovered by H. Steinhaus. For a simple arithmetical proof see J. F. Randolph, Distances between points of the Cantor set, *American Mathematical Monthly* 47 (1940), 549–551. For related properties of Cantor sets, see N. C. Bose Majumder, On the distance set of the Cantor middle third set, III, *American Mathematical Monthly* 72 (1965), 725–729; J. M. Brown and K. W. Lee, The distance set of $C_\lambda \times C_\lambda$, *Journal of the London Mathematical Society* (2) 15 (1977), 551–556; Roger L. Kraft, What's the difference between Cantor sets?, *American Mathematical Monthly* 101 (1994), 640–650; Stephen Silverman, Intervals contained in arithmetic combinations of sets, *American Mathematical Monthly* 102 (1995), 351–353; Rodrigo Bamón, Sergio Plaza, and Jaime Vera, On central Cantor sets

with self-arithmetic difference of positive Lebesgue measure, *Journal of the London Mathematical Society (2)* 52 (1995), 137–146.

[4]G. Hamel, Eine Basis aller Zahlen und die unstetigen Lösungen der Funktionalgleichung $f(x+y) = f(x) + f(y)$, *Mathematische Annalen* 60 (1905), 459–462.

[5]M. Plancherel and G. Pólya, Sur les valeurs moyennes des fonctions réelles définies pour toutes les valeurs de la variable, *Commentarii Mathematici Helvetici* 3 (1931), 114–121; reprinted in *George Pólya: Collected Works*, Vol. III, Joseph Hersch and Gian-Carlo Rota, eds., MIT Press, Cambridge, MA, 1984, pp. 134–141.

[6]See the following papers by R. P. Agnew: Limits of integrals, *Duke Mathematical Journal* 9 (1942), 10–19; Mean values and Frullani integrals, *Proceedings of the American Mathematical Society* 2 (1951), 237–241; Frullani integrals and variants of the Egoroff theorem on essentially uniform convergence, *Acad. Serbe Sci. Publ. Inst. Math.* 6 (1954), 12–16.

21. Derivatives. We consider only functions whose domains are intervals in $\mathbf{R}_1$ and whose ranges are in $\mathbf{R}_1$. Along with the derivative[1] of a function f, which can be defined in the usual way, we shall consider some generalizations that have the advantage of applying to functions that are not necessarily differentiable in the usual sense. These are the four *Dini derivates*, for which we shall use the following notations and definitions:

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

$$f_+(x) = \liminf_{h \to 0^+} \frac{f(x+h) - f(x)}{h},$$

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

$$f_-(x) = \liminf_{h \to 0^-} \frac{f(x+h) - f(x)}{h};$$

the + and − refer to right and left, respectively, and their (upper or lower) positions refer to upper and lower limits. For each x, the four derivates exist, finite or infinite, for any function f at all.

It is common practice to use the phrase "the derivative of f" to mean, according to context, either the number $f'(x)$, that is, the derivative of f at the particular point x, or the function f' whose value at x is the number $f'(x)$. We shall use the same ambiguous terminology for the Dini derivates. If we are going to talk about them as functions, we have to extend our usual notion of function by considering functions whose values may include $+\infty$ or $-\infty$. We must be careful about such generalized functions; there are difficulties about forming sums or products, or (for example) about trying to differentiate them. It will be found that we do not in fact do anything ambiguous with derivates.

If $f^+(x) = f_+(x)$, we say that there is a *right-hand derivative* at x, and we denote it by $f'_+(x)$; similarly for the left-hand derivative $f'_-(x)$. Finally, the ordinary derivative $f'(x)$ exists (finite or infinite) if and only if all four derivates are equal.

Even when $f^+(x)$ and $g^+(x)$ are both finite, we do not necessarily have $(f+g)^+(x) = f^+(x) + g^+(x)$; but if $f'(x)$ exists (finite), we do have $(f+g)^+(x) = f'(x) + g^+(x)$ (compare page 106).

Exercise 21.1. If $f'_+(x)$ exists and is finite, then f is continuous on the right at x; if $f'(x)$ exists and is finite, then f is continuous at x.

Exercise 21.2. Show that f may be discontinuous at x when $f'(x)$ exists, but is infinite.

On the other hand, we have already seen (page 71) that a continuous function does not have to have a derivative anywhere (finite or infinite).

Exercise 21.3. Show that if $f'(a)$ exists (finite), we can write $f(x) - f(a) = (x-a)[f'(a) + \epsilon(x)]$, where $\lim_{x \to a} \epsilon(x) = 0$.

The so-called "chain rule" for differentiation says that if $f'(a)$ exists (finite), if $g(b) = a$, and if $g'(b)$ exists (finite), then for the function φ such that $\varphi(x) = f(g(x))$, the derivative $\varphi'(b)$ exists and is equal to $f'(a)g'(b)$. A fallacious proof proceeds as follows: as $h \to 0$,

$$\frac{\varphi(b+h) - \varphi(b)}{h} = \frac{f(g(b+h)) - f(g(b))}{g(b+h) - g(b)} \cdot \frac{g(b+h) - g(b)}{h}$$
$$\to f'(g(b))g'(b).$$

Exercise 21.4. Find the fallacy; give a correct proof by using Exercise 21.3.

Exercise 21.5. Show that *if* $f'(x) > 0$, *then* f *is increasing at* x, in the sense that there is an interval $(x-h, x+h)$ such that if s and t are in the interval and $s < x < t$, then $f(s) < f(x) < f(t)$. More generally, if $f_+(x) > 0$, then f is increasing on the right at x, in an obvious sense.

Exercise 21.6. A necessary and sufficient condition for $f'_+(x_0)$ to exist (finite or infinite) is that for every real number K, with at most one exception, $f(x) + Kx$ is monotonic on the right at x_0.[2]

Exercise 21.7. The only functions f for which $f(x) + Kx$ is monotonic on (a,b) for every real K, with at most one exception, are of the form $f(x) = px + q$ on (a,b).

Note, for comparison with the preceding exercise, that $f(x) = \sin x$ has $f(x) + Kx$ monotonic whenever $|K| > 1$.

We say that f has a *maximum* at x (an interior point of the domain of f) if there is a neighborhood N of x such that $f(y) \leq f(x)$ for all y in N; the maximum is *proper* if there is a neighborhood N' of x such that $f(y) < f(x)$ for y in N' and $y \neq x$.

Exercise 21.8. Show that if f has a maximum at x, then $f^+(x) \leq 0$ and $f_-(x) \geq 0$.

In particular, *if f has a maximum at x and $f'(x)$ exists, we must have* $f'(x) = 0$. There are of course similar results for minima.

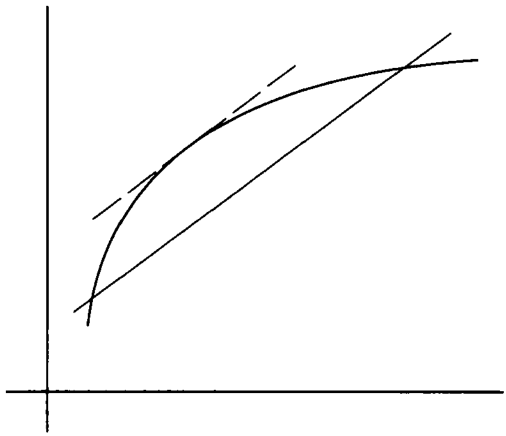
Every function has only countably many proper maxima.[3] To see this, assign to a proper maximum of f, occurring, say, at x, an interval (r_1, r_2) with rational endpoints, such that r_1 and r_2 are on opposite sides of x, and $f(y) < f(x)$ provided that $y \neq x$ and $r_1 < y < r_2$. This interval cannot also be assigned to some other proper maximum, occurring at z, since it would contain both z and x, and we should have to have both $f(z) < f(x)$ and $f(x) < f(z)$. Consequently, different proper maxima have different rational intervals assigned to them. Since there are only countably many intervals with rational endpoints, there are at most countably many proper maxima.

There can, however, be uncountably many improper maxima for a continuous function; for example, a constant function has improper maxima at all points. It can be shown that the values of any function at the points where its derivative is zero (or even where one of its Dini derivates is zero) form a set of measure zero.[4] In addition, *the ordinates of all maxima, proper or not, form a countable set.*[5] For we can assign to each such ordinate v an interval with rational endpoints on which the maximum value of f is v. Different values of v correspond to different intervals, and there are only countably many such intervals.

We have observed that when a derivative exists at a maximum or minimum inside the domain of the function (supposed here to be an interval), the derivative must be equal to 0 at that point; hence, if we want to show that a derivative takes the value 0, we frequently proceed by showing that the function from which it was derived has a maximum or minimum, not at an endpoint of its domain. If we want to show that f' assumes some other value c, we consider $g(x) = f(x) - cx$ and look for its maxima and minima. Any hypothesis that forces g to have a maximum or minimum in an interval (a, b) then guarantees that c is in the range of f'. Two hypotheses that do this for differentiable functions f are (A) that $f'(a) > c$ and $f'(b) < c$ (since then $g(x) > g(a)$ in a right-hand neighborhood of a, and so the largest value of g between a and b is not attained at a; similarly at b); or (B) that $g(a) = g(b)$ (since then g either is constant or has a proper maximum or minimum between a and b). Hypothesis (A) leads to the observation that derivatives have the intermediate value property:[6] *if a derivative takes two values, it takes every value between them.*

Exercise 21.9. Let f be a periodic differentiable function; let a be a given positive number; then there is a point x such that the tangent at x meets the graph again at a point a units farther along the x-axis (that is, $f(x + a) - f(x) = af'(x)$).

Hypothesis (B) leads to the *mean-value theorem* (also known as the law of the mean), which states that every difference quotient $[f(x) - f(y)]/(x - y)$ of a differentiable function f is in the range of f' (the usual formulation is—superficially—different). A proof is suggested by the diagram.[7] The function $g(x) = f(x) - \dfrac{x-a}{b-a}[f(b) - f(a)]$



takes the same value $f(a)$ at b and at a, so it has a maximum at some point c between a and b; at this point, the derivative of g is 0, and so $f'(c) = [f(b) - f(a)]/(b-a)$.

A less conventional procedure[8] is to start from $g(a) = g(b)$ and infer from the universal chord theorem (page 98) that there are intervals (x_n, y_n) in (a,b), each half as long as its predecessor, with $g(x_n) = g(y_n)$. These intervals are nested and hence converge to a point c, which will be in the open interval (a,b) if we pick the first two intervals to avoid a and b. Since we have a sequence of horizontal chords of g whose endpoints approach c, the tangent at c (assumed to exist) must be horizontal, that is, $g'(c) = 0$.

One should not overemphasize the existence of the intermediate point c, whose location is usually unknown; what is generally wanted in practice is that the difference quotient $[f(b) - f(a)]/(b - a)$ is between $\sup f'$ and $\inf f'$; this is actually an equivalent property because f' has the intermediate value property.[9] Another way of stating all this is to say that the range of f' is an interval which contains the range of the difference quotients $[f(x)-f(y)]/(x-y)$. The range of the difference quotients, however, does not necessarily contain the range of f'; in other words, the converse of the mean-value theorem may fail. An example is given by $f(x) = x^3$, where f' takes the value 0, whereas no difference quotient is 0. However, the least upper and greatest lower bounds of the set of values of difference quotients are the same as the corresponding bounds of f': for each value of f' is a limit of difference quotients, and therefore so are the least upper and greatest lower bounds of f'.

In the mean-value theorem we supposed that f is continuous in the closed interval $[a, b]$. We can, as a matter of fact, drop the continuity of f at the endpoints provided that we require continuity on the right at a and on the left at b in the case where the limits $f(a^+)$ and $f(b^-)$ exist, and otherwise require nothing at all at the endpoints. However, the greater generality so obtained is illusory, since if $f(a^+)$ does not exist and f' is finite near a, then f' assumes every finite value in every right-hand neighborhood of a, so that $(b - a)f'(c)$ can have any finite value we please.[10] For, if k is any number, then $f(x) - kx$ does not have a right-hand limit at a, and so cannot be monotonic in a right-hand neighborhood of a. It therefore has maxima and minima in every right-hand neighborhood of a, and its derivative is zero at such points x; then $f'(x) = k$.

As an application of the mean-value theorem, we now prove a theorem on the termwise differentiation of a sequence of functions. The elementary theorem on page 119 demands integrability of the derivatives, and its proof uses the theorem on the integration of a uniformly convergent sequence of functions; but it is possible to prove a more general theorem without using any integration at all. This is: *let the functions f_n have (finite) derivatives f_n' in an interval I; let the sequence $\{f_n(a)\}$ converge for some a in I, and let $\{f_n'\}$ converge uniformly, say to g. Then f_n converges to a limit f, uniformly on I if I is compact, otherwise uniformly on each compact subset of I; and $f'(x) = g(x)$ for all x in I.*

Exercise 21.10. Show, without using any integration, that a continuous function on an interval $[a, b]$ has an antiderivative.[11]

To prove the theorem on termwise differentiation, first apply the mean-value theorem to the difference $f_n - f_m$:

$$(f_n(x) - f_m(x)) - (f_n(a) - f_m(a)) \\ = (x - a)(f_n'(c) - f_m'(c)),$$

where c is between x and a (and, of course, may depend on m and n). The uniform convergence of $\{f_n'\}$ and the convergence of $\{f_n(a)\}$ thus make $\{f_n\}$ converge uniformly as long as $|x - a|$ is bounded. Let f be the limit of the f_n, and let ϵ be an arbitrary positive number. We have

$$\big|(f_n(x) - f_m(x)) - (f_n(a) - f_m(a))\big| \le |x - a|\epsilon$$

if n and m exceed some integer n_0. Letting $m \to \infty$, we see that

$$\big|(f_n(x) - f(x)) - (f_n(a) - f(a))\big| \le |x - a|\epsilon, \qquad n > n_0.$$

That is,

$$\left|\frac{f_n(x) - f_n(a)}{x - a} - \frac{f(x) - f(a)}{x - a}\right| \le \epsilon, \qquad n > n_0.$$

Also, $|f_n'(a) - g(a)| \le \epsilon$ if $n > n_1$ because of the convergence of the derivatives. Now fix an n that exceeds both n_0 and n_1. Then if $|x - a|$ is small enough,

$$\left|\frac{f_n(x) - f_n(a)}{x - a} - f_n'(a)\right| < \epsilon,$$

and so

$$\left|\frac{f(x) - f(a)}{x - a} - f_n'(a)\right| < 2\epsilon$$

if $|x - a|$ is small enough. But $|f_n'(a) - g(a)| \le \epsilon$, so

$$\left|\frac{f(x) - f(a)}{x - a} - g(a)\right| < 3\epsilon.$$

This inequality shows that $f'(a)$ exists and is equal to $g(a)$. Since we now know that $\{f_n\}$ converges everywhere, we can take a to be any point whatever in I, and the theorem follows.

Another application of the mean-value theorem yields the following theorem,[12] which has a transparent geometrical interpretation. *Suppose that f is differentiable in $[a, b]$ and that $f'(a) = f'(b)$; then there is a point c in (a, b) such that*

$$\frac{f(c) - f(a)}{c - a} = f'(c).$$

This says that if the graph of f has the same slope at a and at b, there must be a point c at which the tangent passes through the initial point $(a, f(a))$; a sketch will make this geometrically plausible.

In proving this, we may suppose that $f'(a) = f'(b) = 0$, since otherwise we would consider the function defined by

$f(x) - xf'(a)$. Consider the function g defined by

$$g(x) = \frac{f(x) - f(a)}{x - a}, \qquad a < x \le b; \quad g(a) = 0.$$

The function g is continuous in $[a, b]$ and differentiable in $(a, b]$. We have $g'(b) = -g(b)/(b - a)$. If $g(b) > 0$, then we have $g'(b) < 0$ and hence g decreasing at b (Exercise 21.5), while $g(a) = 0$, so g attains its maximum at a point c between a and b. Hence, there is a point c at which $g'(c) = 0$. A similar argument applies if $g(b) < 0$. If $g(b) = 0$, we have $g(a) = g(b) = 0$, and again $g'(c) = 0$ for some intermediate c. Since

$$g'(c) = \frac{f'(c)}{c - a} - \frac{f(c) - f(a)}{(c - a)^2},$$

our conclusion follows.

Still another application of the mean-value theorem justifies the intuitive idea that derivatives tend to behave worse than the functions from which they are derived.[13]

Exercise 21.11. If $f(x) > 0$, $f(0^+) = 0$, f is differentiable in $(0, 1]$, $h(x) \ge 0$, and $\int h(x)\,dx$ diverges at 0, then $h(f(x))f'(x)$ is unbounded as $x \to 0$; for example, $f'(x)/f(x)$ is unbounded, and so is $f'(x)/\{f(x) \log f(x)\}$.

We might hope to extend the mean-value theorem to cases where the derivative does not necessarily exist, but the most obvious generalization is certainly false. For example, if $f(x) = |x|$ we have $f'_+(x) = -1$ for $x < 0$ and $f'_+(x) = 1$ for $x \ge 0$, so that although $f(1) = f(-1)$, we do not have $f'_+(x) = 0$ for any x at all. Still less can we expect a mean-value theorem to hold for one of the Dini derivates. However, something analogous to the mean-value theorem does hold for Dini derivates and can substitute for the mean-value theorem in some applications.[14]

We shall establish the following result. *Let f be continuous in $[a, b]$. If C is any number that is larger than $[f(b) - f(a)]/(b - a)$, then at uncountably many points x in (a, b) we have $f^+(x) \leq C$.* Similarly $f_+(x) \geq c$ at uncountably many points x if $c < [f(b) - f(a)]/(b - a)$; not, in general, the same points in both cases. The left-hand derivates have the same property.

The proof of this proposition is much like the conventional proof of the mean-value theorem. Let C be any number larger than $[f(b) - f(a)]/(b - a)$, and consider the function g defined by

$$g(x) = f(x) - f(a) - C \cdot (x - a).$$

Then $g(a) = 0$, and $g(b) = f(b) - f(a) - C \cdot (b - a) < 0$. Let s be any number such that $0 = g(a) > s > g(b)$. The set of points x in $[a, b]$ such that $g(x) \geq s$ is the inverse image of a closed set and so is closed, since g is continuous. This set is also bounded, so it has a largest point, say x_s. Since g is continuous, we must have $g(x_s) = s$, while $g(x_s + h) < s$ when $0 < h < b - x_s$. Therefore $g^+(x_s) \leq 0$, whence $f^+(x_s) = g^+(x_s) + C \leq C$. Different values of s generate different values of x_s, and there are uncountably many values of s between 0 and $g(b)$. That is, there are uncountably many points x such that $f^+(x) \leq C$.

It is a familiar fact that *if f' exists and is nonnegative throughout an interval, then f is nondecreasing in that interval.* This follows from the mean-value theorem, since $f(y) - f(x) = (y - x)f'(c)$, with $x < c < y$. If $f'(c) \geq 0$, we infer that $f(y) \geq f(x)$ whenever $y \geq x$. We can use the theorem that we have just proved to establish a result that is stronger in two directions: we do not need to suppose that f' exists, and we can omit countably many points. More precisely, *if f is continuous, and one Dini derivate is nonnegative except perhaps for countably many points, it follows that f is nondecreasing.*[15]

For suppose that $f^+(x) \geq 0$ for $a \leq x \leq b$ except for countably many points. (The hypothesis $f_+(x) \geq 0$ implies this, and the proof for $f^-(x)$ is similar.) If f fails to be nondecreasing, there must be two points x and y such that $y > x$ and $f(y) < f(x)$. Our generalization of the mean-value theorem, with $[f(y) - f(x)]/(y - x) < C < 0$, then says that there are uncountably many points between x and y at which f^+ is negative, contradicting our hypothesis.

Exercise 21.12. The continuity of f is essential for the preceding theorem: construct a discontinuous function f for which $f^+(x) \geq 0$ for all x, yet $f(1) < f(-1)$.

We can now show that *if f is continuous, then all four Dini derivates have the same upper and lower bounds in an interval*; more generally, the collection of difference quotients $[f(x + h) - f(x)]/h$ has the same upper and lower bounds as the Dini derivates, provided, of course, that both x and $x + h$ belong to the interval in question. Suppose, for example, that m is a lower bound for f^+. Let $g(x) = f(x) - mx$. Since $g^+(x) \geq 0$, the function g is nondecreasing. If $h > 0$, we therefore have $g(x+h)-g(x) \geq 0$, or in other words

$$f(x + h) - f(x) - m(x + h) + mx \geq 0;$$

that is, $\dfrac{f(x + h) - f(x)}{h} \geq m$ for $h > 0$, so $f_+(x) \geq m$. Similarly, $f(x) - f(x - h) - mx + m(x - h) \geq 0$, that is, $\dfrac{f(x - h) - f(x)}{-h} \geq m$ for $h > 0$, so $f^-(x) \geq f_-(x) \geq m$.

Next, suppose that one derivate, say f^+, is continuous at x. This means that its upper and lower bounds are arbitrarily close to $f^+(x)$ in a sufficiently small neighborhood of x; the preceding theorem tells us that the same is true for the upper and lower bounds of the other three derivates. This means that all four derivates coincide (with $f^+(x)$) at the point x. That is, *if one derivate of a continuous function is continuous at a point, then there is a derivative at that point.*

A common error of students of calculus is to suppose that $f'(y)$ cannot exist if $\lim_{x \to y} f'(x)$ does not exist.

Exercise 21.13. The function f given by $f(x) = x^2 \sin(1/x)$ for $x \neq 0$ and $f(0) = 0$ is everywhere differentiable, but the limit $\lim_{x \to 0} f'(x)$ does not exist.

Exercise 21.14. Show that $f'(y)$ does exist if $\lim_{x \to y} f'(x)$ exists.

One reason for this misunderstanding, perhaps, is that a derivative, if discontinuous, is very discontinuous, so that functions with discontinuous derivatives are not commonly encountered in calculus. More precisely, *a derivative cannot have a simple jump.* This is to be interpreted in the following sense: if $f'(x)$ exists at every point x of an interval, and (for a point y of this interval) the limits $f'(y^+)$ and $f'(y^-)$ both exist, then both these limits are equal to $f'(y)$. On the other hand, the example $f(x) = |x|$ shows that the limits of f' from both sides can exist and be different at y if $f'(y)$ does not exist. The impossibility of a simple jump for a derivative is an immediate consequence of the intermediate value property of derivatives.

A continuous function cannot have a derivative that is everywhere infinite. Indeed, we can say much more: a continuous function must have $f^+(x) < +\infty$ on an uncountable set,[16] a fact that follows at once from the generalized mean-value theorem on page 149. Indeed, the generalized mean-value theorem says that $f^+(x) < C$ on an uncountable set if $C > [f(b) - f(a)]/(b - a)$.

It follows from a general theorem that we shall quote later (page 155) that whether f is continuous or not, it can have an infinite right-hand derivative f'_+ at most on a set of measure zero. On the other hand, if we do not require f to be continuous, we can have $f^+(x) = +\infty$ at every point x. An example of this phenomenon can be constructed as follows.[17] Let real numbers x in $[0, 1]$ be represented in base 3 as $0.a_1 a_2 \ldots$, where each a_n is 0, 1, or 2. If x has two representations, we choose the one that terminates. Then we put $f(x) = 0.b_1 b_2 \ldots$ (base 2), where $b_n = 1$ if $a_n = 2$ and otherwise $b_n = 0$. Now, since we excluded ternary representations ending in repeated 2's, the ternary representation of every x contains an infinite sequence of digits that are 0 or 1. Let one of these 0's or 1's occur at the rth ternary place. Let x' differ from x only by having 2 as its rth ternary digit; then $x' > x$, and in fact $x' - x = 3^{-r}$ or $2 \cdot 3^{-r}$. In either case, $f(x') - f(x) = 2^{-r}$. Hence

$$\frac{f(x') - f(x)}{x' - x} \geq \frac{3^r}{2^{r+1}}.$$

Since r can be arbitrarily large, it follows that $f^+(x) = +\infty$.

It can be shown that this function f is continuous except at the points that have terminating ternary expansions, and in fact is continuous on the right at these points, but discontinuous on the left.

Another interesting result about possible values of derivatives (of not necessarily continuous functions) is that if one level set of f' is dense, then every other level set of f' is of first category. That is to say, if $f'(x) = A$ (possibly infinite) on a dense set, then $f'(x)$ can exist and be different from A at most on a set of first category.[18]

It is enough to consider the set S where $f'(x) < A$, since the set where $f'(x) > A$ is the set where $(-f)'(x) < -A$. When A is finite, S is contained in the union of the sets $E_{n,m}$, where $x \in E_{n,m}$ provided that $|y - x| < 1/n$ implies

$$\frac{f(y) - f(x)}{y - x} < A - 1/m;$$

when $A = +\infty$, replace $A - 1/m$ by m. If we show that each $E_{n,m}$ is nowhere dense, we shall have proved our assertion.

Suppose then that some $E_{N,M}$ is dense in some interval I. Since the set of points where $f'(x) = A$ is dense in I, there is a point x_0 in I where $f'(x_0) = A$. Since $E_{N,M}$ is dense in I, the interval $(x_0 - 1/N, x_0 + 1/N)$ contains points of $E_{N,M}$; so we can choose points x_k in $E_{N,M}$ with $x_k \to x_0$. Thus

$$\frac{f(x_k) - f(x_0)}{x_k - x_0} < A - 1/M \qquad (\text{or} < M \text{ if } A = +\infty).$$

Letting $k \to \infty$, we get $f'(x_0) \leq A - 1/M$ (or $\leq M$), contradicting $f'(x_0) = A$ (or $+\infty$). Therefore $E_{n,m}$ is always nowhere dense.

This theorem can be generalized by using Dini derivates in the hypothesis; but the proof is more complicated. Since it is easy to show that at a point of discontinuity, at least one Dini derivate is infinite, it is then easy to get the following result:

If f is discontinuous at the points of an everywhere dense set and differentiable (with a finite derivative, and hence continuous) at the points of another everywhere dense set, then it must be continuous and not differentiable at the points of a set of second category.[19] We showed (page 125) that when f is continuous at the points of an everywhere dense set, its points of discontinuity form a set of first category. Thus the presence of an everywhere dense set of points of continuity means that there are only relatively few points of discontinuity; we now see that the presence of an everywhere dense set of points of discontinuity allows only relatively few points where a derivative exists.

A direct proof can be given as follows.

Let E_n be the set of points x such that $|y-x| < 1/n$ implies $|f(y)-f(x)|/|y-x| < n$. In the present context, "differentiable" means "having a finite derivative," so every point x where $f'(x)$ exists belongs to some E_n. To show that the set of such points is of first category, it is then enough to show that each E_n is nowhere dense.

Suppose that some E_N were dense in an open interval I. This interval contains a point w at which f is discontinuous, so there must be a positive h and a sequence $\{y_k\}$ converging to w such that $|f(y_k) - f(w)| \geq h$. Let k be so large that $|y_k - w| < 1/N$. Since E_N is dense in I, we can choose x_k in E_N so that x_k is between y_k and w; then $|x_k - w| < 1/N$ and $|y_k - x_k| < 1/N$. Since $x_k \in E_N$, the triangle inequality implies that

$$\begin{aligned}\frac{h}{|y_k - w|} &\leq \left|\frac{f(y_k) - f(x_k)}{y_k - w}\right| + \left|\frac{f(x_k) - f(w)}{y_k - w}\right| \\ &\leq \left|\frac{f(y_k) - f(x_k)}{y_k - x_k}\right| + \left|\frac{f(x_k) - f(w)}{x_k - w}\right| < 2N.\end{aligned}$$

Letting $y_k \to w$, we have a contradiction.

The ruler function (Exercise 13.1), which is continuous at the irrational numbers and discontinuous at the rational numbers, is differentiable nowhere, while its cube is differentiable at the points of a dense set.[20]

If one of the Dini derivates of a continuous function is zero everywhere in an interval, then the function is constant there: for we have shown that the function is both nonincreasing and nondecreasing. This implies that *two continuous functions having the same finite derivative throughout an interval differ only by a constant there.* On the other hand, it is possible for two continuous functions to have the same derivative, necessarily infinite at some points, throughout an interval, and not differ by a constant there (see page 164).

A considerable amount can be said about the derivates of a perfectly arbitrary function. We state the following facts without proof.[21] First, except at the points of a countable set, the upper derivate on one side is not less than the lower derivate on the other side. Next, if $f^+ = +\infty$ on a set E, then $f_- = -\infty$ on E except for a set of measure zero; similarly, if $f_+ = -\infty$ then $f^- = +\infty$ with the same possibility of exception. Finally, the set where f^+ and f_- are finite and different is of measure zero. Putting these facts together, we see that *except on a set of measure zero, there are only three possibilities*: (1) *there is a finite derivative*; (2) *the two upper derivates are* $+\infty$ *and the two lower derivates are* $-\infty$; (3) *the upper derivate on one side is* $+\infty$, *the lower derivate on the other side is* $-\infty$, *and the other two derivates are finite and equal.* Since only (1) is possible for a monotonic function, we see in particular that a monotonic function has a finite derivative almost everywhere; we shall give a direct proof of this in the next section. Another conclusion that we can draw from the general theorem is that if all derivates are bounded almost everywhere, the function has a derivative almost everywhere.

Derivatives are not as simple functions as one might think. For example, the product of two derivatives is not necessarily a derivative.[22]

NOTES

[1] For an extensive survey of theorems about derivatives, see Andrew Bruckner, *Differentiation of Real Functions*, American Mathematical Society, Providence, 1994.

[2] U. Dini, Grundlagen für eine Theorie der Functionen einer veränderlichen reellen Grösse, translated from the Italian edition of 1878, Teubner, Leipzig, 1892, pp. 279–280. Note that a function can be "increasing on the right at a point" without necessarily being increasing in any right-hand interval.

[3] W. Sierpiński: see S. Saks, *Theory of the Integral*, second revised

edition, English translation, Stechert, New York, 1937 (reissued by Dover, 1964), p. 261. The set of proper maxima can, however, be dense (see E. E. Posey and J. E. Vaughan, Functions with a proper local maximum in each interval, *American Mathematical Monthly* 90 (1983), 281–282), and indeed this is the case for most (in the sense of Baire category) continuous functions (see Vladimir Drobot and Michal Morayne, Continuous functions with a dense set of proper local maxima, *American Mathematical Monthly* 92 (1985), 209–211).

[4]S. Saks, book cited in the preceding note, pp. 271–272. See also D. E. Varberg, On absolutely continuous functions, *American Mathematical Monthly* 72 (1965), 831–841; E. P. Woodruff, Derivates of a function whose image is of Lebesgue measure zero, *Notices of the American Mathematical Society* 16 (1969), 666–667. For an elementary proof of a more elementary result, see G. Pólya and G. Szegő, *Problems and Theorems in Analysis*, vol. I, Springer, New York, 1972, pp. 81 and 269, problem 125 of part II.

[5]The countability of the set of ordinates of maxima was called to my attention by J. E. Vaughan.

[6]One could claim that this dates back to Galileo, who wrote in 1638 (Discorsi e dimostrazioni matematiche intorno a due nuove scienze, *Opera*, 1898, vol. 8, p. 283), "Plato perhaps had the idea that a moving object cannot pass from rest to a positive velocity without passing through all lower velocities." It seems likely that Galileo would have thought of velocity as being continuous, so he may merely have noticed the intermediate value property of continuous functions. The tentative attribution to Plato seems to be fanciful.

[7]A similar diagram once appeared on a Graduate Record Examination with directions something like: "Which of the following five theorems does the picture make you think of?"

[8]See A. K. Aziz and J. B. Diaz, On Pompeiu's proof of the mean-value theorem of the differential calculus of real-valued functions, *Contributions to Differential Equations*, vol. 1, pp. 467–481 (1963). See also H. Samelson, On Rolle's theorem, *American Mathematical Monthly* 86 (1979), 486.

[9]Compare J. Dieudonné, *Foundations of Modern Analysis*, second printing, Academic Press, 1969, pp. 158–160.

[10]E. W. Hobson, *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, vol. 1, third edition, Cambridge University Press, 1927, p. 363.

[11]This exercise was suggested by Robert B. Burckel.

[12]T. M. Flett, A mean value theorem, *Mathematical Gazette* 42 (1958), 38–39. See also S. G. Wayment, An integral mean value the-

orem, *Mathematical Gazette* 54 (1970), 300–301, and references given there; S. Reich, On mean value theorems, *American Mathematical Monthly* 76 (1969), 70–73; J. B. Diaz and R. Výborný, On some mean value theorems of the differential calculus, *Bulletin of the Australian Mathematical Society* 5 (1971), 227–238; S. Reich, Problem 5810, *American Mathematical Monthly* 78 (1971), 798.

[13] A special case is considered by L. J. Paige, A note on indeterminate forms, *American Mathematical Monthly* 61 (1954), 189–190.

[14] See A. P. Morse, Dini derivates of continuous functions, *Proceedings of the American Mathematical Society* 5 (1954), 126–130.

[15] For a generalization, see A. D. Miller and R. Výborný, Some remarks on functions with one-sided derivatives, *American Mathematical Monthly* 93 (1986), 471–475.

[16] For references see P. Erdős, Some remarks on set theory, *Annals of Mathematics* 44 (1943), 643–646.

[17] A. N. Singh, On infinite derivatives, *Fundamenta Mathematicae* 33 (1945), 106–107.

[18] F. M. Filipczak, On the derivative of a discontinuous function, *Colloquium Mathematicum* 13 (1964), 73–79; K. M. Garg, On bilateral derivatives and the derivative, *Transactions of the American Mathematical Society* 210 (1975), 295–329 (Proposition 3.9 and Corollary 5.2); R. P. Boas and G. T. Cargo, Level sets of derivatives, *Pacific Journal of Mathematics* 83 (1979), 37–44.

[19] This seems first to have been noticed explicitly by M. K. Fort, Jr. in 1951: A theorem concerning functions discontinuous on a dense set, *American Mathematical Monthly* 58 (1951), 408–410. It was found independently by S. Četković, Un théorème de la théorie des fonctions, *Comptes Rendus de l'Académie des Sciences* 245 (1957), 1692–1694; and by Solomon Markus [Marcus], Points of discontinuity and points of differentiability (in Russian), *Revue Roumaine de Mathématiques Pures et Appliquées* 2 (1957), 471–474. However, it is an easy consequence of an old theorem of W. H. Young (On the infinite derivates of a function of a single variable, *Arkiv för Matematik, Astronomi och Fysik* 1 (1903), 201–204). Young's theorem states that the set of points at which at least one Dini derivate is infinite is a countable intersection of open sets (hence, if it is dense, its complement is of first category). That Fort's theorem is a corollary of Young's was noticed by K. M. Garg, On the derivability of functions discontinuous at a dense set, *Revue Roumaine de Mathématiques Pures et Appliquées* 7 (1962), 175–179, and independently by Cargo (see Boas and Cargo, paper cited in the preceding note; that paper also contains a particularly simple proof of Young's theorem).

The proof of Fort's theorem in the text was suggested by A. C. Segal. Another simple proof is implicit in more general results obtained by E. M. Beesley, A. P. Morse, and D. C. Pfaff, Lipschitzian points, *American Mathematical Monthly* 79 (1972), 603–608.

[20]See Gerald J. Porter, On the differentiability of a certain well-known function, *American Mathematical Monthly* 69 (1962), 142; G. A. Heuer, Functions continuous at the irrationals and discontinuous at the rationals, *American Mathematical Monthly* 72 (1965), 370–373; J. E. Nymann, An application of Diophantine approximation, *American Mathematical Monthly* 76 (1969), 668–671; Richard B. Darst and Gerald D. Taylor, Differentiating powers of an old friend, *American Mathematical Monthly* 103 (1996), 415–416. A variation of the ruler function is considered in Alec Norton, Continued fractions and differentiability of functions, *American Mathematical Monthly* 95 (1988), 639–643.

[21]For a simple exposition, see F. Riesz and B. Sz.-Nagy, *Functional Analysis*, Ungar, New York, 1955, pp. 17–19.

[22]See A. M. Bruckner, Creating differentiability and destroying derivatives, *American Mathematical Monthly* 85 (1978), 554–562; M. W. Botsko, When is the product of two derivatives a derivative?, *Mathematics Magazine* 65 (1992), no. 3, 186–187.

22. Monotonic functions. A function f from an interval I in $\mathbf{R}_1$ into $\mathbf{R}_1$ is called *monotonic* if it is either nondecreasing or nonincreasing. That is, f is monotonic if either $f(y) \geq f(x)$ whenever $y > x$ in I, or else $f(y) \leq f(x)$ whenever $y > x$ in I. If one of these conditions holds with strict inequality throughout, we say that f is *strictly monotonic*. The familiar functions used in calculus are, if not monotonic, at least piecewise monotonic. Thus if $f(x) = x^2$, then f is decreasing when $x < 0$ and increasing when $x > 0$; if $f(x) = \cos x$, then f is alternately increasing and decreasing in the intervals $(-\pi, 0)$, $(0, \pi)$, and so on; if $f(x) = e^x$, then f is increasing throughout $\mathbf{R}_1$. All these functions are continuous. On the other hand, the function f defined by $f(x) = [x]$ (the greatest integer not exceeding x) is nondecreasing and is discontinuous at each integer x.

Exercise 22.1. A monotonic function is bounded on each compact subinterval of its domain.

Exercise 22.2. A monotonic function approaches a (finite) limit from each side at every interior point of its domain.

Exercise 22.3. The limit of a pointwise convergent sequence of monotonic functions is monotonic.

A function f is said to have a *jump* at a point x of its domain if f has limits from both sides at x, but f is not continuous at x. After Exercise 22.2, we can say that the only discontinuities of a monotonic function are jumps. The easiest monotonic functions to visualize are those with only a finite number of jumps, but a monotonic function can have a much more complicated structure than this. For example, if $f(x) = 2^{-n}$ in the interval $[1/(n+1), 1/n)$, then f is a nondecreasing function with jumps that have a limit point at 0.

A monotonic function can have at most countably many jumps, since the intervals from $f(x^-)$ to $f(x^+)$, if not empty, form a set of disjoint intervals in $\mathbf{R}_1$ (disjoint because f is monotonic), and such a set of intervals is countable (page 30). However, we shall show that the set of jumps of a monotonic function can be any countable set at all, even an everywhere dense one—for example, all the rational points in an interval. Let $\{x_n\}$ be a given countable set, and let j_n be positive numbers such that $\sum j_n < \infty$. We define functions f_n by putting $f_n(x) = 0$ for $x < x_n$ and $f_n(x) = j_n$ for $x \geq x_n$. Of course, the x_n will not in general be numbered in order of increasing magnitude. The series $\sum f_n$ converges uniformly (by the M-test, page 111), since $|f_n(x)| \leq j_n$ and $\sum j_n$ converges. If x_0 is not any of the x_n, then it is a point of continuity for all the f_n and hence is a point of continuity for f (page 112). On the other hand, if x_m is one of the x_n, then precisely one function f_n, namely f_m, is discontinuous at x_m. Consequently, $\sum_{n \neq m} f_n = f - f_m$ is continuous at x_m. Hence f, as the sum of a function that is continuous at x_m and a function that is discontinuous at x_m, is itself discontinuous at x_m. Indeed, f has a jump of amount j_m at x_m. We may reasonably call such an f a pure jump function. More generally, we call f a *pure jump function* if it is constructed similarly, but possibly with both right-hand and left-hand jumps, so that $f(x_m^-) \neq f(x_m) \neq f(x_m^+)$. If we construct a pure jump function whose right-hand and left-hand jumps are just those of a given nondecreasing function g, then $g - f$ will still be nondecreasing, and also continuous.

It may seem plausible that a pure jump function should have its derivative zero except at its jumps. This conjecture is almost, but not quite, true: *the derivative of a pure jump function is zero except on a set of measure zero*, but this set of measure zero may contain more points than the jumps.[1] We can get a better idea of the kind of thing that can happen by considering some special cases. Let f be the pure jump function that has jumps of amount 2^{-n} at the points 3^{-n} for $n = 1, 2, \ldots$; let g be the pure jump function that has jumps of amount 3^{-n} at the points 2^{-n}; let $f(0) = g(0) = 0$. Both f and g are continuous (on the right) at 0. However, we can easily show that $f'_+(0) = +\infty$ while $g'_+(0) = 0$. In fact, if $h > 0$, we have $[f(0 + h) - f(0)]/h = f(h)/h$; and if $3^{-m-1} \leq h < 3^{-m}$, then $f(h) = \sum_{k=m+1}^{\infty} 2^{-k} = 2^{-m}$, so that $f(h)/h > 3^m/2^m \to \infty$. Similarly, if $2^{-m-1} \leq h < 2^{-m}$, we have $g(h) = \sum_{k=m+1}^{\infty} 3^{-k} = \frac{1}{2} \cdot 3^{-m}$, so that $g(h)/h < 2^m/3^m \to 0$.

Exercise 22.4. Construct a monotonic pure jump function, having jumps with 0 as a limit point, such that $f'_+(0)$ is positive and finite.

There seems to be no essentially simpler way of proving that a pure jump function has a zero derivative almost everywhere than to appeal to the general theorem (which we shall prove presently) that every monotonic function has a finite derivative almost everywhere.

What is perhaps more surprising is that there can be a continuous monotonic function, not a constant, whose derivative is zero almost everywhere.[2] Functions with this property are called *singular* monotonic functions. We shall construct a singular monotonic function in some detail, since it can be used for various applications. One application (page 163) is the construction of a more complicated singular monotonic function that is constant in no interval.

We base the construction of a singular function on the Cantor set of §6; our example will be constant in each complementary interval of this set, and so its derivative will certainly be zero except perhaps at the points of the Cantor set, which is of measure zero. If x is any point of the interval $[0, 1]$, we write $x = 0.a_1a_2a_3\ldots$ (base 3), so that each a_k is 0, 1, or 2. The endpoints of the complementary intervals of the Cantor set are the numbers with terminating ternary expansions that can be written without using any 1's, for example $\frac{1}{3} = 0.1000\ldots = 0.0222\ldots$; $\frac{2}{3} = 0.2000\ldots$. If we halve all the digits of such an expansion and interpret the result as a number written in base 2, we get, for example, the numbers $0.0111\ldots$ (base 2) and $0.1000\ldots$ (base 2), which are the same. This happens for every pair of endpoints of the same complementary interval. Let us now define a function f by writing all the points x of the Cantor set in base 3, using no 1's, halving all the digits, and setting $f(x)$ equal to the resulting number, interpreted in base 2. This defines f on the Cantor set, and we have just seen that f has the same value at the two ends of each complementary interval. We can extend f to all of $[0, 1]$ by giving it the same value throughout each complementary interval that it has at the endpoints of that interval. We must now show that the *Cantor function* f is monotonic and continuous; it will also be interesting to investigate the derivates of f at the points of the Cantor set.

If x and y are two points of the Cantor set, not endpoints, and $x < y$, then the ternary expansions of x and y must be of the forms

$$x = 0.a_1 a_2 a_3 \ldots a_n a_{n+1} \cdots ,$$
$$y = 0.a_1 a_2 a_3 \ldots a_n b_{n+1} \cdots ,$$

with $b_{n+1} > a_{n+1}$. The binary expansions of $f(x)$ and $f(y)$ will then coincide through the nth digit, while the next digit of $f(y)$ will be 1 larger than the corresponding digit of $f(x)$. This means that $f(y) > f(x)$. Hence f is a nondecreasing function.

Since f is continuous in the intervals where it is constant, we have only to consider its continuity at points of the Cantor set. Let x be such a point, and consider a neighborhood of x of radius 3^{-n}. Every point y of the Cantor set that is in this neighborhood differs from x by a number whose ternary expansion starts with at least n zeros. The binary expansion of $f(y)$ then differs from that of $f(x)$ by a number whose binary expansion starts with at least n zeros, so that $f(y)$ differs from $f(x)$ by at most 2^{-n}. Since for a point y not in the Cantor set but in the same neighborhood of x the value of $f(y)$ is the same as its value at either endpoint of the complementary interval containing y, it follows that f is continuous at x.

We have therefore shown that f is continuous, monotonic, not constant, and singular.[3] We now investigate the differentiability of f at points of the Cantor set. At a left-hand endpoint of a complementary interval, the right-hand derivative f'_+ exists and is 0; and similarly $f'_-(x) = 0$ at a right-hand endpoint x.

Consider the derivatives on the other sides of endpoints of complementary intervals, say for definiteness a right-hand endpoint $x = 0.a_1a_2\ldots a_n20000\ldots$. If $m > n$ and h is between 3^{-m} and 3^{-m-1}, then $f(x+h)$ differs from $f(x)$ by something between 2^{-m} and 2^{-m-1}, and consequently $[f(x+h) - f(x)]/h$ must be between $2^{-m-1}/3^{-m}$ and $2^{-m}/3^{-m-1}$. Hence as $h \to 0$ (and so $m \to \infty$), this difference quotient becomes positively infinite. That is, at a right-hand endpoint x, we have $f'_+(x) = +\infty$ and $f'_-(x) = 0$. Similarly, $f'_-(x) = +\infty$ and $f'_+(x) = 0$ at a left-hand endpoint.

At limit points that are not endpoints, it can be shown that $f^+ = +\infty$, whereas f_+ can have any value between 0 and $+\infty$.

As a first application, we construct another singular function, one which is not constant on any interval.[4] Let C be the function just constructed, but extended to all of $\mathbf{R}_1$ by putting $C(x) = 0$ for $x \le 0$ and $C(x) = 1$ for $x \ge 1$.

Let $\{r_n\}$ be an enumeration of the rational numbers. Define

$$f(x) = \sum_{n=1}^{\infty} 2^{-n} C(2^n(x - r_n)).$$

Since the series is uniformly convergent (by the M-test), f is continuous. Each term in the sum is nondecreasing (since C is), so f is nondecreasing. In fact, f is strictly increasing: if $x < y$, then there is some rational number r_n between x and y, and $C(2^n(x-r_n)) = 0 < C(2^n(y-r_n))$, so $f(x) < f(y)$.

Finally, by Fubini's theorem on differentiation of series with nondecreasing terms (see page 171, below),

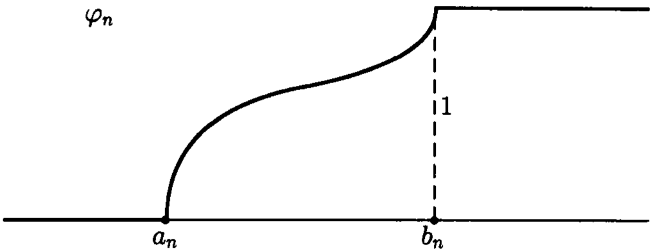
$$f'(x) = \sum_{n=1}^{\infty} C'(2^n(x-r_n)) = 0$$

for almost all x.

As another application of the singular function C, we can construct the example, mentioned on page 154, of two functions that have the same derivative (infinite at some points) throughout an interval, but do not differ by a constant. We need a function g that is continuous and nondecreasing with a finite derivative at each point not in the Cantor set and derivative $+\infty$ at each point of the Cantor set. Once we have such a g, we can put $h(x) = C(x) + g(x)$; then $h'(x) = g'(x) = +\infty$ at all points of the Cantor set (since all derivates of C are nonnegative); and $h'(x) = g'(x)$ at all points not in the Cantor set, since $C'(x) = 0$ at such points. However, g and h differ by C, which is not constant.

We proceed to construct g.[5] Let us enumerate the complementary intervals (a_n, b_n) of the Cantor set in order of decreasing length (the order among the finitely many intervals of each length is irrelevant). Let φ_n be a continuous nondecreasing function of the general character indicated in the figure: $\varphi_n(x) = 0$ for $x < a_n$, $\varphi_n(x) = 1$ for $x > b_n$, $(\varphi_n)'_+(a_n) = (\varphi_n)'_-(b_n) = +\infty$. (An explicit example is $\varphi_n(x) = (2/\pi)\tan^{-1}\{(x-a_n)^{1/2}(b_n-x)^{-1/2}\}$.)

Observe that the lengths of the intervals (a_n, b_n) are negative integral powers of 3, and let $h_n = (2/5)^m$ when-



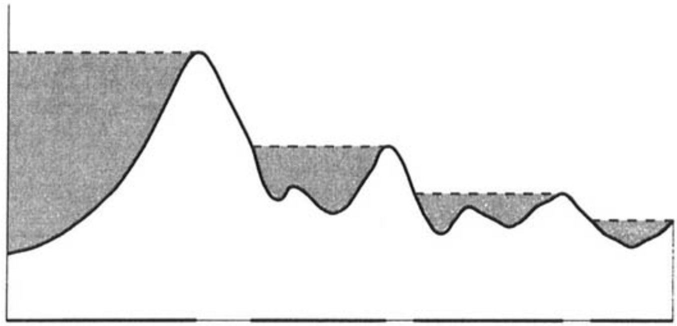
ever (a_n, b_n) has length 3^{-m}. Define $g(x) = \sum h_n \varphi_n(x)$ with the (infinite) sum extended over all n. In other words, the value $g(x)$ is the sum of the h_n over all intervals (a_n, b_n) to the left of x, plus $h_k \varphi_k(x)$ if x is in (a_k, b_k). Now there are 2^{m-1} intervals (a_n, b_n) of length 3^{-m}, and $h_n = (2/5)^m$ on each, so $\sum h_n$ converges. Therefore the series defining g converges uniformly, and so g is continuous. By construction, g is nondecreasing.

Let x be a point of the Cantor set, other than an a_n, and let $\delta > 0$. The numerator of the difference quotient

$$\Delta = \frac{g(x+\delta) - g(x)}{\delta} = \sum h_n \frac{\varphi_n(x+\delta) - \varphi_n(x)}{\delta}$$

will exceed h_k if the interval $(x, x+\delta)$ contains the complementary interval (a_k, b_k). It is easy to see that $(x, x+\delta)$ always contains a complementary interval of length $3^{-m} \geq \delta/9$. Then $\Delta \geq (2/5)^m/\delta \to \infty$, so $g'_+(x) = +\infty$. If, however, x is an a_n, then $g'_+(x) = +\infty$ by inspection. Similarly $g'_-(x) = +\infty$ at all x in the Cantor set. That is, $g'(x) = +\infty$ at all x in the Cantor set.

We now turn to the rather difficult proof that *a monotonic function has a finite derivative almost everywhere.*[6] The reader who is interested in seeing some applications of the theorem first may skip to page 170.



The proof depends on a lemma by F. Riesz, known as the "flowing water" or "rising sun" lemma. If g is a continuous function from an interval I into $\mathbf{R}_1$, if the graph of g is the cross section of the bed of a stream, and we consider the set E of points where water is flowing, it is intuitively clear that E consists of open intervals at whose ends g has the same value; if the graph is the profile of a mountain range, if the sun rises in the direction of the positive x-axis, and if E is the set of points that are in the shade, it is again intuitively clear that E consists of open intervals at whose ends g has the same value. (In both cases, there may be an exceptional interval at the extreme left, as in the picture.)

We now state the lemma in abstract terms and for a more general situation. *Let g be continuous on an interval I, except for jumps, and let G be defined by*

$$G(x) = \max(g(x^-), g(x), g(x^+)).$$

The set E of points x such that there is a $y > x$ with $g(y) > G(x)$ is an open set; and if (a, b) is any one of the intervals making up E, then $g(a^+) \leq G(b)$.

If we exchange left and right, and define E' as the set of points x such that there is a $y < x$ with $g(y) > G(x)$, then, similarly, if E' is the union of open intervals (a', b'), we have $G(a') \geq g(b'^-)$.

We first prove that E is open. Let $x_0 \in E$; then there is a $y > x_0$ with $g(y) > G(x_0)$. We have to show that this property holds for all x near x_0. If x varies slightly to the left of x_0, then $g(x)$ is near $g({x_0}^-)$; if x varies slightly to the right of x_0, then $g(x)$ is near $g({x_0}^+)$; in either case $G(x)$ can exceed $G(x_0)$ only slightly. Since $g(y) > G(x_0)$, we also have $g(y) > G(x)$ as long as $G(x)$ is not much greater than $G(x_0)$, which, as we have just seen, is the case when x is near x_0.

To prove the second statement of the lemma, let (a, b) be one of the intervals that make up E; then b is not a point of E. If it were the case that $g(a^+) > G(b)$, then there would exist a positive number ϵ and a point x between a and b such that $g(x) \geq G(b) + \epsilon$. Fix this x and this ϵ, and let z_1 denote the least upper bound of the numbers z in the interval $[x, b]$ such that $g(z) \geq G(b) + \epsilon$. Then $G(z_1) \geq G(b) + \epsilon$; in particular, $z_1 \neq b$. Thus, $z_1 \in E$, and so there is some $y > z_1$ for which $g(y) > G(z_1) \geq G(b) + \epsilon$. Since $b \notin E$, we must have $y \leq b$, but this contradicts the definition of z_1. This contradiction proves that $g(a^+)$ cannot exceed $G(b)$.

We are going to deduce the differentiability of a nondecreasing function from two consequences of Riesz's lemma. We lead up to these by showing first that our theorem will follow if we can prove that $f^+(x) < +\infty$ almost everywhere and $f^+(x) \leq f_-(x)$ almost everywhere. It will simplify the notation somewhat, and will do no harm, if we suppose that our interval (a, b) has its midpoint at 0, and that $f(0) = 0$. We can then reflect the graph of f around the origin to obtain the graph of a function f_0 whose value at a negative x is $-f(-x)$. This function f_0 is also nondecreasing. Its right-hand derivates at a negative x equal the left-hand derivates of f at $-x$, since

$$\frac{f_0(x+h)-f_0(x)}{h} = \frac{-f(-x-h)-(-f(-x))}{h} = \frac{f(-x+(-h))-f(-x)}{-h}.$$

If $f^+(x) \le f_-(x)$ holds almost everywhere for every nondecreasing f, then it also holds with f replaced by f_0 and x replaced by $-x$. That is, $f^-(x) \le f_+(x)$ for almost all x, so for almost all x we have

$$f^+(x) \le f_-(x) \le f^-(x) \le f_+(x) \le f^+(x).$$

If $f^+(x) < \infty$ for almost all x, then the preceding inequality shows that all four derivates are equal and finite for almost all x, that is, there is a finite derivative for almost all x.

We can reduce the problem still further: it is enough to show that the set on which $f_-(x) < r < R < f^+(x)$ has measure zero, whatever r and R may be. For, if $f^+(x) > f_-(x)$, then there are rational numbers r and R such that $f_-(x) < r < R < f^+(x)$. Since there are only countably many pairs of rational numbers, the set where $f^+(x) > f_-(x)$ is contained in the union of countably many sets of measure zero, and so is itself of measure zero.

The consequences of Riesz's lemma on which our proof depends are as follows:

(1) Let f be nondecreasing on $[a,b]$, and let E_R be the set of points x where f is continuous and $f^+(x) > R$. Then E_R can be covered by a countable set of disjoint intervals (a_k, b_k) whose total length $\sum(b_k - a_k)$ is at most

$$\sum[f(b_k{}^+) - f(a_k{}^+)]/R \le [f(b^-) - f(a^+)]/R.$$

(2) Let f be nondecreasing on $[a,b]$, and let E_r be the set of points x where f is continuous and $f_-(x) < r$. Then E_r can be covered by a countable set of disjoint intervals (a_k, b_k) such that $\sum[f(b_k{}^-) - f(a_k{}^+)] \le r\sum(b_k - a_k) \le r(b-a)$.

We defer the proofs of these two statements until we have shown how the theorem follows from them.

As a first step, we observe that (1) implies $f^+(x) < +\infty$ almost everywhere. For the set of jumps of f is countable, and hence of measure zero. Now if E is the set of points x where f is continuous and $f^+(x) = +\infty$, then we have the hypothesis of (1) for every positive R, so that E is covered by intervals (a_k, b_k) whose total length is at most $[f(b^-) - f(a^+)]/R$ for every R. That is, E is covered by a collection of intervals of arbitrarily small total length, whence E is of measure zero.

Next consider a set E of points x where f is continuous and $f_-(x) < r < R < f^+(x)$, so that the hypotheses of (1) and (2) are both satisfied. Apply (1) to each of the intervals (a_k, b_k) in (2). We find that the part of E in (a_k, b_k) is covered by intervals whose total length, L_k say, is at most $[f(b_k{}^-) - f(a_k{}^+)]/R$. Adding these inequalities for all (a_k, b_k), and applying (2), we find that

$$\sum L_k \le (1/R)\sum[f(b_k{}^-) - f(a_k{}^+)] \le (r/R)(b-a).$$

The same argument applies to any subinterval of (a,b); that is, the part of E in any interval (p,q) is covered by intervals of total length at most $(r/R)(q-p)$. Since E fills at most a fixed fraction $r/R < 1$ of any interval, E is of measure zero (Exercise 11.1).

We now turn to the proof of propositions (1) and (2).

(1) If $x \in E_R$, then $f^+(x) > R$, so there is a $y > x$ such that $[f(y) - f(x)]/(y-x) > R$, that is, $f(y) - Ry >$

$f(x) - Rx$. Apply Riesz's lemma to the function g such that $g(t) = f(t) - Rt$. Since f is continuous at x, so is g, and $G(x) = g(x)$. We conclude that E_R is covered by a countable set of disjoint intervals (a_k, b_k) such that $G(b_k) \geq g(a_k{}^+)$; in other words (since f is nondecreasing),

$$f(b_k{}^+) - Rb_k \geq f(a_k{}^+) - Ra_k,$$

that is,

$$f(b_k{}^+) - f(a_k{}^+) \geq R(b_k - a_k).$$

Adding these inequalities for the various values of k gives

$$R\sum(b_k - a_k) \leq \sum[f(b_k{}^+) - f(a_k{}^+)].$$

(2) If $x \in E_r$, then $f_-(x) < r$, so there is a $y < x$ such that $[f(y) - f(x)]/(y - x) < r$, that is (since $y < x$),

$$f(y) - f(x) > (y - x)r,$$
$$f(y) - ry > f(x) - rx.$$

Riesz's lemma, in the form with the inequality reversed in its hypothesis, can be applied to the function g defined by $g(t) = f(t) - rt$. We find that E_r is covered by a countable set of disjoint intervals (a_k, b_k) such that $g(b_k{}^-) \leq G(a_k)$; since f is nondecreasing, $G(a_k) = g(a_k^+)$, so

$$f(b_k{}^-) - rb_k \leq f(a_k{}^+) - ra_k,$$

that is,

$$f(b_k{}^-) - f(a_k{}^+) \leq r(b_k - a_k).$$

Summing over all the intervals, we obtain the conclusion of (2).

We now give some interesting applications of the theorem on differentiating monotonic functions. The existence of such applications helps to justify the amount of effort expended on the proof of the theorem.

(i) FUBINI'S THEOREM ON DIFFERENTIATION OF SERIES OF MONOTONIC FUNCTIONS. *Let $f_1 + f_2 + \cdots$ be a pointwise convergent series of nondecreasing functions on an interval $[a, b]$, with sum s. Then for almost every x in $[a, b]$, we have $f_1'(x) + f_2'(x) + \cdots = s'(x)$.*[7]

This is one example of a theorem that is more conveniently stated in terms of series than in terms of sequences.

If we exclude a set of measure zero, all the f_n have nonnegative derivatives, and so does s, since it too is a nondecreasing function. The series $f_1'(x) + f_2'(x) + \cdots$ has nonnegative terms, and so its partial sums $s_n'(x)$ form a nondecreasing sequence (for each x). The series therefore converges if its partial sums are bounded. But

$$\begin{aligned} \frac{s(x+h) - s(x)}{h} &= \frac{f_1(x+h) - f_1(x)}{h} + \cdots \\ &\geq \frac{f_1(x+h) - f_1(x)}{h} + \cdots \\ &\quad + \frac{f_n(x+h) - f_n(x)}{h}, \end{aligned}$$

since the f_n are nondecreasing. Letting $h \to 0$, we infer that $s'(x) \geq s_n'(x)$. Hence the differentiated series converges for almost all x, and we have only to verify that it has the right sum.

To identify the sum of the differentiated series, we first show that some subsequence of its partial sums converges to $s'(x)$ almost everywhere. Since $s(b) - s_n(b) \to 0$, there must be a subsequence of integers n for which the sum $\sum[s(b) - s_n(b)]$ converges (we can pick an n that makes the difference less than $\frac{1}{2}$, then a larger n that makes it less than $\frac{1}{4}$, and so on). For values of n in this subsequence, we have $s(x) - s_n(x) \leq s(b) - s_n(b)$, since $s(x) - s_n(x)$ is the "tail" of the series for $s(x)$ and so defines a nondecreasing function of x. Hence $\sum(s - s_n)$ (summed over the same values of n as before) is a pointwise convergent series of nondecreasing functions. By what we have already proved, the series $\sum[s'(x)-s'_n(x)]$ converges almost everywhere, and consequently its general term approaches zero almost everywhere. That is, we have found a subsequence of partial sums $s_n(x)$ such that $s'_n(x) \to s'(x)$ almost everywhere. Since the whole sequence of partial sums converges almost everywhere, it must converge to the limit of the subsequence, that is, to $s'(x)$.

(ii) DENSITY OF SETS. Let E be a set in $\mathbf{R}_1$. We say that a point x (in E or not) is a point of density for E if sufficiently small neighborhoods of x consist "mostly" of points of E. It is not quite easy to formulate this definition precisely. First let us cover E by a countable collection of disjoint open intervals. This can be done in various ways; we take the greatest lower bound of the total lengths of such coverings as giving a measure of the size of E, and call it the *outer measure* of E, denoted by $\mu(E)$. For an interval I, the outer measure $\mu(I)$ is the ordinary length of I. If E_1 and E_2 are sets lying in two disjoint intervals, then $\mu(E_1 \cup E_2) = \mu(E_1) + \mu(E_2)$. Now let N stand for a neighborhood of x; we say that x is a *point of density* for E if $\mu(E \cap N)/\mu(N) \to 1$ as $\mu(N) \to 0$; in other words, if the outer measure of the part of E in a small neighborhood of x is nearly equal to the length of the neighborhood. There is a similar definition for sets in any $\mathbf{R}_n$.

We shall now prove that *almost all points of E are points of density for E.*[8] This means that, generally speaking, a set nearly fills small neighborhoods of its points, and cannot, for example, occupy about half of every interval. (Compare Exercise 11.1. The same theorem also holds in any $\mathbf{R}_n$.)

We suppose that E is not of measure zero; otherwise the theorem has no content. We may also suppose that E is bounded, and hence lies in some compact interval I, since only small neighborhoods of points of E are relevant. Define a function λ by putting $\lambda(x)$ equal to the outer measure of the part of E that is to the left of x. Then λ is a nondecreasing function, and our theorem amounts to showing that $\lambda'(x) = 1$ for almost all x in E.

Let us first consider a function f analogous to λ, but obtained by using some fixed covering of E by a countable set of disjoint open intervals; that is, $f(x)$ is the total length of intervals belonging to the covering and lying to the left of x; if x is inside one of the intervals, we count only the part of that interval to the left of x. Now $f'(x) = 1$ when $x \in E$, since if $x \in E$ and h is small enough, then $x + h$ will be in the open interval covering x, and so $f(x+h) - f(x) = h$. Take a sequence of coverings of E whose total lengths μ_n approach $\mu(E)$ rapidly enough so that $\sum[\mu_n - \mu(E)]$ converges. Let f_n be the function f associated with the nth covering. Then $f_n(x) \to \lambda(x)$, and $f_n(x) - \lambda(x) \le \mu_n - \mu(E)$; moreover, $f_n - \lambda$ is a nondecreasing function. We have $\sum[f_n(x) - \lambda(x)]$ convergent, and by Fubini's theorem the differentiated series $\sum[f_n'(x) - \lambda'(x)]$ converges almost everywhere. Hence the terms of this series approach zero almost everywhere. Since $f_n'(x) = 1$ almost everywhere in E, we must also have $\lambda'(x) = 1$ almost everywhere in E.

(iii) THE MEASURE OF A LOCUS. Let F be a compact set in $\mathbf{R}_1$. If x is not in F, we know that there is a positive distance from x to F, and that it is attained for some point in F (Exercise 8.10). *Consider the set E_r of points that are at distance r from F, where $r > 0$. Then E_r is a set of measure zero.*[9] For, if not, E_r contains a point of density; let y be such a point. Let x be a point in F

at distance r from y. A neighborhood N of x of radius r cannot contain any point of E_r, since all points of N have distance less than r from x. Now any neighborhood I of y is half in N, and so $I \cap C(E_r)$ contains an interval at least half as long as I. The existence of such intervals contradicts the assumption that y is a point of density for E_r.

NOTES

[1] J. S. Lipiński, Sur la dérivée d'une fonction de sauts, *Colloquium Mathematicum* 4 (1957), 197–205; L. A. Rubel, Differentiability of monotonic functions, *Colloquium Mathematicum* 10 (1963), 277–279.

[2] This surprising behavior can be viewed either as typical or as atypical in the sense of Baire category, depending on what metric is used on the space of functions. Tudor Zamfirescu, Most monotone functions are singular, *American Mathematical Monthly* 88 (1981), 47–49; F. S. Cater, Most monotone functions are not singular, *American Mathematical Monthly* 89 (1982), 466–469.

[3] The Cantor function f is uniquely determined by the properties that f is monotonic, $f(0) = 0$, $f(x/3) = f(x)/2$ for all x in $[0,1]$, and $f(1-x) = 1 - f(x)$ for all x in $[0,1]$. Donald R. Chalice, A characterization of the Cantor function, *American Mathematical Monthly* 98 (1991), 255–258.

[4] This simple construction was given by G. Freilich, *American Mathematical Monthly* 80 (1973), 918–919. For other constructions, see Lajos Takács, An increasing continuous singular function, *American Mathematical Monthly* 85 (1978), 35–37, and its references.

[5] The construction given in the first edition of this book was fallacious. The present construction is modeled on that given by S. Saks (*Theory of the Integral*, second revised edition, English translation, Stechert, New York, 1937 (reissued by Dover, 1964), pp. 205–206) for the more general situation when the derivative is infinite on an arbitrary closed set of measure zero. Some interesting results in this connection are: *If g is continuous and has a (finite or infinite) derivative except on a countable set, this derivative being nonnegative almost everywhere, then g is nondecreasing* (G. Goldowsky; see the cited book of Saks); *if E is a countable set which is the intersection of countably many open sets, there is a (not necessarily continuous) function whose derivative is $+\infty$ on E and 0 outside E* (G. Piranian, The derivative of a monotonic discontinuous function, *Proceedings of the American Mathematical Society* 16 (1965), 243–244).

[6] The exposition follows F. Riesz and B. Sz.-Nagy, *Functional*

Analysis, Ungar, New York, 1955, with some simplifications given by H. Kestelman, *Modern theories of integration*, Clarendon Press, Oxford, 1937, pp. 199 ff. The theorem was originally proved by Lebesgue as a deduction from his entire theory of integration. A further simplification has been given by L. A. Rubel (paper cited in note 1), and a very short proof along different lines has been given by D. G. Austin, A geometric proof of the Lebesgue differentiation theorem, *Proceedings of the American Mathematical Society* 16 (1965), 220–221.

[7]This proof and that of the next theorem follow Riesz and Sz.-Nagy, book cited above. "Fubini's theorem" in the theory of integration is a different theorem.

[8]For another proof, see L. Zajíček, An elementary proof of the one-dimensional density theorem, *American Mathematical Monthly* 86 (1979), no. 4, 297–298.

[9]P. Erdős, Some remarks on the measurability of certain sets, *Bulletin of the American Mathematical Society* 51 (1945), 728–731 (proof by T. Radó).

23. Convex functions. We consider functions from an interval in $\mathbf{R}_1$ into $\mathbf{R}_1$. A function is usually called *convex* if the portion of its graph in every interval lies on or below its chord. We shall establish the rather striking fact that a continuous function is convex if we merely assume that the midpoint of each chord is above (or on) the graph of the function. Actually much milder hypotheses than continuity suffice; for example, it is enough to suppose that the function is bounded in some (small) interval.[1] Hence if a discontinuous function has the midpoint property, its graph must be rather wild. There exist such functions, for example, the discontinuous linear functions mentioned in §20.

The geometrical statements about chords can be formulated analytically as follows: to say that the midpoint of every chord lies on or above the graph means that

$$f\left(\frac{x_1+x_2}{2}\right) \leq \frac{f(x_1)+f(x_2)}{2}$$

for every x_1 and x_2 in the domain; to say that the entire chord lies on or above the graph means that

$$f(q_1x_1+q_2x_2) \leq q_1f(x_1)+q_2f(x_2) \qquad (*)$$

whenever $q_1 \geq 0$, $q_2 \geq 0$, and $q_1+q_2=1$. We are to show that when f is continuous, the first inequality (for all x_1 and x_2) implies the second.

If we write $\frac{1}{4}(x_1+x_2+x_3+x_4)$ as the average of the numbers $\frac{1}{2}(x_1+x_2)$ and $\frac{1}{2}(x_3+x_4)$ and then apply the midpoint inequality three times, we obtain

$$\begin{aligned} f\left(\tfrac{1}{4}(x_1+x_2+x_3+x_4)\right)& \\ &\leq \tfrac{1}{4}\left(f(x_1)+f(x_2)+f(x_3)+f(x_4)\right). \end{aligned}$$

By repeating this argument, we find that

$$f\left(\tfrac{1}{n}(x_1+\cdots+x_n)\right) \leq \tfrac{1}{n}\left(f(x_1)+\cdots+f(x_n)\right)$$

whenever the integer n is a power of 2. If we now take r of the x's equal to x_1 and the remaining $s=n-r$ of the x's equal to x_2, we get

$$f(\tfrac{r}{n}x_1+\tfrac{s}{n}x_2) \leq \tfrac{r}{n}f(x_1)+\tfrac{s}{n}f(x_2)$$

whenever n is a power of 2 and r and s are integers whose sum is n. Every real number can be approximated by rational numbers whose denominators are powers of 2 (by truncating the binary expansion). So if f is continuous, we may pass to the limit in the preceding inequality to obtain $(*)$.

From $(*)$, without any initial hypothesis about the continuity of f, we can deduce not only that f *is continuous at each point in the interior of its interval of convexity*, but that f *has finite right-hand and left-hand derivatives* at such points, these derivatives themselves being nondecreasing functions. The existence of the one-sided derivatives implies that f is continuous on the right and on the left at each point (Exercise 21.1), and so f is continuous (Exercise 13.2).

Exercise 23.1. Suppose f is a convex function on $(0, \infty)$, and let g be the function defined by $g(x) = f(x)/x$. Show that either g is monotonic, or else the graph of g consists of two monotonic pieces.[2]

To deduce our statements about derivatives from $(*)$, it is convenient to write this inequality in a somewhat different form. Supposing that $x_1 < x_2$, we introduce positive quantities g and h via $q_1x_1 + q_2x_2 = x_1 + g$ and $x_2 = x_1 + h$. Then $(*)$ takes the form

$$f(x_1 + g) \le \left(1 - \frac{g}{h}\right) f(x_1) + \frac{g}{h} f(x_1 + h),$$

and after dropping the subscript on x and rearranging, we get the equivalent inequality

$$\frac{f(x+g) - f(x)}{g} \le \frac{f(x+h) - f(x)}{h}, \qquad 0 < g < h.$$

This says that if we fix the left-hand endpoint of a chord, then the slope of the chord does not decrease as the other endpoint moves to the right. In other words, the difference quotient $[f(x+h)-f(x)]/h$ defines, for each x, a nondecreasing function of h. As $h \to 0^+$, this ratio therefore approaches a limit, which as far as we know up to now may be either finite or $-\infty$. That is, $f'_+(x)$ exists (finite or $-\infty$) for every x that is interior to our interval.

By working with negative h, we infer similarly that $f'_-(x)$ exists (finite or infinite).

From the midpoint inequality, applied with $x_1 = x - h$ and $x_2 = x + h$, we get $2f(x) \le f(x+h) + f(x-h)$. Equivalently, $[f(x) - f(x-h)]/h \le [f(x+h) - f(x)]/h$. When $h \to 0^+$, the right-hand side decreases to $f'_+(x)$, and the left-hand side increases to $f'_-(x)$. Therefore $f'_-(x) \le f'_+(x)$, and both sides of this inequality are finite.

Finally, we want to see that the one-sided derivatives are nondecreasing. Suppose, to the contrary, that there is a point x and a positive number g such that $f'_+(x) > f'_+(x+g)$. Since the right-hand derivative is the limit of difference quotients, we will have $[f(x+h) - f(x)]/h > [f(x+g+h) - f(x+g)]/h$ for all sufficiently small positive h. Equivalently, $f(x+h) + f(x+g) > f(x+g+h) + f(x)$. Now $(*)$ implies

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

(take $x_1 = x$, $x_2 = x + g + h$, $q_1 = g/(g+h)$, and $q_2 = h/(g+h)$), and similarly

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

Adding these two inequalities contradicts our supposition. Therefore f'_+ is a nondecreasing function, and for similar reasons f'_- is nondecreasing.

Let us note that the usual calculus criterion for convexity is indeed a sufficient condition. Suppose that $f''(x)$ exists and is nonnegative at every point of an interval. If h is a positive number such that both x and $x+2h$ are in this interval, then two applications of the mean-value theorem give $f(x+2h)-f(x+h)=hf'(c_1)$ and $f(x+h)-f(x)=hf'(c_2)$, with $x+h<c_1<x+2h$ and $x<c_2<x+h$. Apply the mean-value theorem once more to get

$$\frac{f(x+2h)-f(x+h)}{h}-\frac{f(x+h)-f(x)}{h} = (c_1-c_2)f''(c_3)\geq 0.$$

Consequently $f(x+h)\leq \frac{1}{2}(f(x+2h)+f(x))$. Now if $x_2>x_1$, put $x=x_1$ and $h=(x_2-x_1)/2$ to see that f is convex according to the midpoint definition.

The monotonicity of difference quotients that we noted above implies that

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

for every positive h. The right-hand side is the slope of an arbitrary chord of the graph of $y=f(x)$ with left-hand end at $(x,f(x))$; and the left-hand side is the slope of the right-hand tangent to the graph at x. Thus every chord through $(x,f(x))$ and going to the right lies above (or on) the (right-hand) tangent, which means that the entire curve, from x onward to the right, lies above (or on) the right-hand tangent. Similarly on the left; and since the right-hand tangent has larger slope than the left-hand tangent, the entire curve lies above (or on) the right-hand tangent line at x.

A straight line touching the curve and such that the entire curve lies above or on it is called a *supporting line*; the existence of a supporting line at each point can be (and often is) taken as the definition of convexity. We say that the function is *strictly convex* if each supporting line has just one point of contact with the graph.

We can now generalize the original definition of convex functions: not only is every chord of the graph above or on its arc, but if we put arbitrary positive weights at n points of the arc, their center of gravity will also be above or on the arc; and for a strictly convex function (and $n > 1$) the center of gravity will be strictly above the arc. Algebraically, this means that if $w_1, w_2, \ldots, w_n$ are positive weights whose sum is 1, then

$$f(w_1x_1 + \cdots + w_nx_n) \le w_1f(x_1) + \cdots + w_nf(x_n), \qquad (\dagger)$$

with strict inequality if f is strictly convex and at least two x_k are different. If f is concave (which means that $-f$ is convex), then the inequality is reversed.

Inequality ($\dagger$) is known as *Jensen's inequality.* To prove it, let $M = w_1x_1 + \cdots + w_nx_n$ and $w_1 + \cdots + w_n = 1$, so that M is the x-coordinate of the center of gravity of n weights w_k placed at the points $(x_k, f(x_k))$. Let us consider the supporting line determined by the right-hand tangent to the graph of f at $x = M$; the curve lies above this tangent. Let the equation of the tangent be $y = g(x) = ax + b$, where a and b are numbers that could be calculated but whose values are irrelevant.

Since the curve is above (or on) this supporting line through $(M, f(M))$, we have $ax + b \le f(x)$. Write this inequality for $x = x_1, x_2, \ldots, x_n$, multiply the kth inequality by w_k, and add the results:

$$w_1g(x_1) + \cdots + w_ng(x_n) \le w_1f(x_1) + \cdots + w_nf(x_n),$$

or, since the sum of the w's is 1,

$$b + a\sum_{k=1}^{n} w_kx_k \le \sum_{k=1}^{n} w_kf(x_k).$$

But $\sum_{k=1}^{n} w_k x_k = M$, so $g(M) \le \sum_{k=1}^{n} w_k f(x_k)$. Finally, the supporting line $y = g(x)$ goes through the point $(M, f(M))$, so $g(M) = f(M)$, and $f(M) \le \sum_{k=1}^{n} w_k f(x_k)$; this is (†).

Furthermore, if f is strictly convex and some x_k is different from M, then there was strict inequality in one of the inequalities that we added. Hence there is strict inequality in (†) except when all the x_k are equal to M.

In a similar way, we can get analogous inequalities for integrals instead of sums. Let w and x be continuous positive functions for $a \le t \le b$, with $\int_a^b w(t)\,dt = 1$. If we replace the number M by $\int_a^b w(t)x(t)\,dt$ and use the inequality $g(y) \le f(y)$ for every y between a and b, we get

$$f\left(\int_a^b w(t)x(t)\,dt\right) \le \int_a^b w(t)f(x(t))\,dt,$$

provided that f is convex, and the reversed inequality when f is concave. This is the integral form of Jensen's inequality.

The ordinary average, or more formally the *arithmetic mean*, of n numbers $x_1, \ldots, x_n$ is $(x_1 + \cdots + x_n)/n$; the *geometric mean* is $(x_1 x_2 \cdots x_n)^{1/n}$. For many purposes it is better to use weighted means,

$$w_1 x_1 + w_2 x_2 + \cdots + w_n x_n \quad \text{and} \quad x_1^{w_1} x_2^{w_2} \cdots x_n^{w_n},$$

where the w_k add to 1; the ordinary case has $w_k = 1/n$ for each k.

A famous (and very useful) theorem states that *the geometric mean of n positive numbers does not exceed their arithmetic mean*, and it is actually smaller except when all the numbers are equal to each other. This is just a consequence of the convexity of $-\log t$ for positive t. In fact, by (†) applied to this function,

$$-\log(w_1x_1 + \cdots + w_nx_n) \\ \leq -(w_1 \log x_1 + \cdots + w_n \log x_n),$$

that is,

$$\log(w_1x_1 + \cdots + w_nx_n) \geq \log(x_1)^{w_1} + \cdots + \log(x_n)^{w_n}.$$

If we exponentiate both sides, we get the inequality between the means; and there is equality if and only if all the x_k are equal.

There are many applications of Jensen's inequality in general and of the inequality between the geometric and arithmetic means in particular. We can solve many maximum and minimum problems for polynomials without using calculus, for example, to find the largest box that can be made from a square of paper a units on a side by cutting x by x squares out of the corners and folding up the resulting rectangles.[3] This problem asks us to maximize $x(a-2x)^2$, but it is just as satisfactory to maximize $\{4x(a-2x)^2\}^{1/3}$. Consider the numbers $4x$ and $a-2x$ with weights $\frac{1}{3}$ and $\frac{2}{3}$. Their weighted geometric mean $\{4x(a-2x)^2\}^{1/3}$ does not exceed the weighted arithmetic mean, namely $\frac{4}{3}x + \frac{2}{3}(a-2x) = \frac{2}{3}a$, and there is equality if and only if $4x = a - 2x$, that is, $x = a/6$. This says that the largest possible value for $\{4x(a-2x)^2\}^{1/3}$ is attained at $x = a/6$; and for this x we have $x(a-2x)^2 = 2a^3/27$. Notice that it was important to know when equality is attained in the inequality between the means.

Exercise 23.2. Show[4] that $(\sin x)^{\sin x} < (\cos x)^{\cos x}$ when $0 < x < \pi/4$.

As another application, we deduce a necessary condition for the convergence of a series $\sum a_n$ of positive terms.[5] Of course $a_n \to 0$ is a necessary condition, but $na_n \to 0$ is (in general) not. However, if the series is convergent, then n times the geometric mean of the first n terms must approach zero; that is, $n(a_1 a_2 \cdots a_n)^{1/n} \to 0$.

To see this, let $s_n = a_1 + a_2 + \cdots + a_n$, and $G_n = (a_1 a_2 \cdots a_n)^{1/n}$. We are assuming $s_n \to s$. It follows that the arithmetic mean of $s_1, s_2, \ldots, s_n$ approaches s.

Exercise 23.3. Prove the preceding statement.

Written in terms of the a_k, this says that

$$n^{-1}\{a_1 + (a_1 + a_2) + \cdots + (a_1 + a_2 + \cdots + a_n)\} \to s,$$

that is, $n^{-1}\{na_1 + (n-1)a_2 + \cdots + a_n\} \to s$, or equivalently

$$n^{-1}\{(n+1)s_n - a_1 - 2a_2 - 3a_3 - \cdots - na_n\} \to s.$$

Since $s_n \to s$, the preceding formula shows that

$$\frac{a_1 + 2a_2 + \cdots + na_n}{n} \to 0.$$

The left-hand side is the arithmetic mean of the n numbers displayed in the numerator, so their geometric mean also approaches 0. But this geometric mean is

$$\{n!\, a_1 a_2 \cdots a_n\}^{1/n} = (n!)^{1/n} G_n.$$

Since $(n!)^{1/n}/n \to 1/e$ (see Exercise 32.4), $nG_n \to 0$.

Important and useful inequalities[6] arise from the observation that the function $x \mapsto x^r$ defined on $(0, \infty)$ is convex when $r > 1$. (It is also convex when $r < 0$, but the resulting inequalities do not seem to be of much use; when $0 < r < 1$

it is concave.) We start by observing that[7] *if u and v are positive numbers, and $r > 1$, then the minimum value of the expression $q^{1-r}u^r + (1-q)^{1-r}v^r$ for $0 < q < 1$ equals $(u+v)^r$.* This can be verified by calculus: the indicated function of q blows up at both endpoints, so its minimum occurs at the interior point where the derivative is equal to zero, namely at $q = u/(u+v)$. However, it is more in the spirit of this section to argue, by convexity, that $\{qx + (1-q)y\}^r \le qx^r + (1-q)y^r$ for all positive numbers x and y, with equality if and only if $x = y$; putting $x = u/q$ and $y = v/(1-q)$ gives $(u+v)^r \le q^{1-r}u^r + (1-q)^{1-r}v^r$ for every value of q, with equality if and only if $u/q = v/(1-q)$, which is to say $q = u/(u+v)$.

To apply the observation, let a_k and b_k be positive numbers for $1 \le k \le n$. For each value of k and q we have $(a_k + b_k)^r \le q^{1-r}a_k^r + (1-q)^{1-r}b_k^r$, so (with all sums from 1 to n)

$$\sum (a_k + b_k)^r \le q^{1-r} \sum a_k^r + (1-q)^{1-r} \sum b_k^r.$$

Taking the minimum of the right-hand side in q with $u = (\sum a_k^r)^{1/r}$ and $v = (\sum b_k^r)^{1/r}$ gives, after taking rth roots of both sides,

$$\left(\sum (a_k + b_k)^r\right)^{1/r} \le \left(\sum a_k^r\right)^{1/r} + \left(\sum b_k^r\right)^{1/r}.$$

This is *Minkowski's inequality* for sums, which we used on page 24. The same argument, with sums replaced by integrals, yields Minkowski's inequality for integrals:

$$\left(\int |f+g|^r\right)^{1/r} \le \left(\int |f|^r\right)^{1/r} + \left(\int |g|^r\right)^{1/r}.$$

This is valid also when $r = 1$, as follows immediately from the triangle inequality.

By Jensen's inequality,

$$\left(\sum w_k x_k\right)^r \le \sum w_k x_k^r, \qquad \text{if } \sum w_k = 1,$$

for all positive numbers x_k. If the p_k are any positive numbers, then we may replace w_k by $p_k/\sum p_j$ to find

$$\left(\frac{\sum p_k x_k}{\sum p_k}\right)^r \le \frac{\sum p_k x_k^r}{\sum p_k},$$

that is,

$$\sum p_k x_k \le \left(\sum p_k\right)^{1-1/r} \left(\sum p_k x_k^r\right)^{1/r}.$$

Now let $p_k = y_k^{r/(r-1)}$ and $x_k = z_k y_k^{-1/(r-1)}$; then

$$\sum y_k z_k \le \left(\sum y_k^{r/(r-1)}\right)^{(r-1)/r} \left(\sum z_k^r\right)^{1/r},$$

which is *Hölder's inequality* for sums. Analogous reasoning gives Hölder's inequality for integrals:

$$\int |fg| \le \left(\int |f|^{r/(r-1)}\right)^{(r-1)/r} \left(\int |g|^r\right)^{1/r}.$$

The special case $r = 2$ says

$$\sum y_k z_k \le \left(\sum y_k^2\right)^{1/2} \left(\sum z_k^2\right)^{1/2},$$

$$\int |fg| \le \left(\int |f|^2\right)^{1/2} \left(\int |g|^2\right)^{1/2};$$

these inequalities are traditionally called *Cauchy's inequality* and *Schwarz's inequality*. They are also known by various combinations of the names Cauchy, Bunyakovsky, and

Schwarz, who found different versions, for sums and for Riemann integrals. Given Cauchy's inequality for sums, "any competent analyst" (as J. E. Littlewood used to put it) should have been able to find the appropriate generalization when it became needed.

NOTES

[1] The most general result is given by A. Ostrowski, Zur Theorie der konvexen Funktionen, *Commentarii Mathematici Helvetici* 1 (1929), 157–159.

[2] Marek Kuczma. See János Aczél and Che Tat Ng, A lemma on the angles between a fixed line and the lines connecting a fixed point on it with the points of a convex arc, in *General Inequalities, 6 (Oberwolfach, 1990)*, W. Walter, ed., International Series of Numerical Mathematics, vol. 103, Birkhäuser, Basel, 1992, pp. 463–464; Heinz König, On a property of convex functions, same book, p. 471.

[3] I learned of this application from M. Klamkin. Many similar problems are discussed by I. M. Niven, *Maxima and Minima Without Calculus*, Mathematical Association of America, Washington, DC, 1981.

[4] See the solution to problem 10261 in the *American Mathematical Monthly* 101 (1994), 690.

[5] J. R. Nurcombe, A necessary condition for convergence, *Mathematical Gazette* 59 (1975), 113–114.

[6] For the inequalities discussed here see G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, second edition, Cambridge University Press, 1952, chap. II and III. See also E. F. Beckenbach and R. Bellman, *Inequalities*, fourth printing, Springer-Verlag, 1983; D. S. Mitrinović, J. E. Pečarić, and A. M. Fink, *Classical and New Inequalities in Analysis*, Kluwer, Dordrecht, 1993.

[7] This lemma and its application to Minkowski's inequality are due to Heinz König, A simple proof of the Minkowski inequality, book cited in note 2, p. 469; I learned of this reference from Robert B. Burckel.

24. Infinitely differentiable functions. We next consider functions that can be differentiated more than once, or even infinitely often. For such functions there is a generalization of the mean-value theorem, called Taylor's theorem with remainder. We shall not go into the motivation for considering this particular formula, and we shall not try to obtain it under the most general hypotheses possible. However, we shall obtain the formula with one of the more useful forms of the remainder. Suppose f is a function whose domain contains the interval $[a, x]$, and the nth derivative $f^{(n)}$ exists and is continuous, or at least can be integrated to give $f^{(n-1)}$. We start from

$$f(x) = f(a) + \int_a^x f'(t)\,dt = f(a) - \int_a^x f'(t)\,d(x-t),$$

and integrate by parts (if $n \geq 2$), obtaining

$$f(x) = f(a) + (x-a)f'(a) + \int_a^x (x-t)f''(t)\,dt.$$

Repeating this process, we eventually get

$$\begin{aligned} f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2}f''(a) + \cdots \\ + \frac{(x-a)^{n-1}}{(n-1)!}f^{(n-1)}(a) + R_n(x), \end{aligned}$$

where

$$R_n(x) = \frac{1}{(n-1)!}\int_a^x f^{(n)}(t)(x-t)^{n-1}\,dt.$$

To illustrate one of the ways in which Taylor's theorem can be used, we prove the following theorem. *Suppose that f is a function on some interval $[x_0, \infty)$, that f'' is continuous, and that* $\lim_{x\to\infty} f(x) = \lim_{x\to\infty} f''(x) = 0$. *Then* $\lim_{x\to\infty} f'(x) = 0$.

To prove this, take Taylor's theorem with remainder of order 2, and write it in the form

$$f'(a) = \frac{f(x) - f(a)}{x - a} - \frac{1}{x - a}\int_a^x (x - t)f''(t)\,dt, \quad x > a.$$

Let x_1 be so large that both $|f(t)|$ and $|f''(t)|$ are less than ϵ when $t > x_1$. If $a > x_1$, then

$$\begin{aligned} |f'(a)| &\le \frac{2\epsilon}{x - a} + \frac{\epsilon}{x - a}\int_a^x (x - t)\,dt \\ &= \frac{2\epsilon}{x - a} + \frac{\epsilon(x - a)}{2} = \epsilon\left(\frac{2}{x - a} + \frac{x - a}{2}\right). \end{aligned}$$

Here x is at our disposal, as long as $x > a$; take $x = a + 2$. (The reason for this particular choice is that $2/(x - a) + (x - a)/2$ is smallest when $x = a + 2$.) Then the preceding inequality reduces to $|f'(a)| \le 2\epsilon$ for $a > x_1$, which is to say that $f'(a) \to 0$ as $a \to \infty$.

Slight modifications in the proof will yield considerably stronger results.[1] There is no need, for example, to suppose that $f(x) \to 0$; it is enough to have $f(x)$ bounded, as long as $f''(x) \to 0$. To see this, suppose that $|f(x)| \le M$, and put $\epsilon(x) = \max_{t \ge x} |f''(t)|$. Clearly ϵ is a nonincreasing function, and $\lim_{x \to \infty} \epsilon(x) = 0$. We have

$$\begin{aligned} |f'(a)| &\le \frac{2M}{x - a} + \frac{1}{x - a}\int_a^x (x - t)\epsilon(t)\,dt \\ &\le \frac{2M}{x - a} + \frac{\epsilon(a)}{2}(x - a). \end{aligned}$$

Now take x so that $x - a = \epsilon(a)^{-1/2}$. Then

$$|f'(a)| \le 2M\sqrt{\epsilon(a)} + \tfrac{1}{2}\sqrt{\epsilon(a)},$$

and again $f'(a) \to 0$.

Exercise 24.1. Suppose that f is a function on some interval $[x_0, \infty)$, that f'' is continuous and bounded, and that $\lim_{x\to\infty} f(x) = 0$. Then $\lim_{x\to\infty} f'(x) = 0$.

It is tempting to let $n \to \infty$ in Taylor's theorem to obtain an infinite series, the so-called Taylor series of f. If $R_n(x) \to 0$ (for a particular x), the series so obtained will converge, and in fact will converge to $f(x)$. However, we must not assume that this will always happen if f has derivatives of all orders (or, as we shall say, if f is infinitely differentiable), although it does happen for many of the simple functions that are considered in calculus.[2]

In the first place, the Taylor series might diverge; in the second place, it might converge, but to the wrong sum. We shall give examples of both possibilities.[3]

Even in elementary discussions, it is a commonplace that a Taylor series need not converge throughout the domain where the original function is infinitely differentiable; an example is given by $f(x) = 1/(1+x^2)$. Here the function is infinitely differentiable on the whole of $\mathbf{R}_1$, but its Taylor series (with center at 0) converges only for $|x| < 1$.

Much worse behavior can occur. We shall exhibit a function whose Taylor series diverges everywhere except, of course, at the point a itself. Consider the function f defined by the integral

$$f(x) = \int_0^\infty e^{-t}\cos(t^2 x)\, dt.$$

For even n we have

$$f^{(n)}(0) = \pm \int_0^\infty t^{2n} e^{-t}\, dt = \pm(2n)!\,,$$

and for odd n we have $f^{(n)}(0) = 0$, since the derivatives of odd order of the cosine are all 0 at 0. Hence the Taylor series of f with center at 0 has as its general term

$$\pm\frac{(2n)!}{n!}x^n, \qquad n \text{ even},$$

which does not approach zero except at $x = 0$, and therefore the series cannot converge except at $x = 0$.

It is possible to show that[4] if $\{M_k\}$ is any sequence of numbers, there is an infinitely differentiable function f such that $f^{(k)}(0) = M_k$ for every k. This shows that infinitely differentiable functions whose Taylor series about 0 diverge (except at 0) must exist in great profusion.

By a more complicated construction, it can be shown that there are infinitely differentiable functions whose Taylor series diverge no matter what point is taken as center.[5]

For an example of the other kind of failure of a Taylor series, consider the function f defined by

$$f(x) = e^{-1/x^2}, \quad x \neq 0; \qquad f(0) = 0.$$

We can show that $f^{(k)}(0) = 0$ for every k, so that the Taylor series of f about 0 has all its terms 0 and thus certainly converges to the wrong sum. Clearly $f^{(k)}(x)$ has the form $R(x)e^{-1/x^2}$ for $x \neq 0$, where R is a rational function. Now $x^{-n}e^{-1/x^2} \to 0$ as $x \to 0$ for every integer n (this is equivalent to the familiar fact that $x^n e^{-x^2} \to 0$ as $x \to \infty$), and so $f^{(k)}(x) \to 0$ as $x \to 0$. By Exercise 21.14, this implies that $f^{(k)}(0) = 0$ also. We may alternatively argue directly that $f^{(k)}(0) = \lim_{x\to 0} x^{-1}f^{(k-1)}(x) = 0$.

It is not difficult to show, by direct estimates of the remainders, that the Taylor series of familiar functions such as the sine and the exponential converge everywhere to the functions that gave rise to them. Similarly, the Taylor series for $(1 + x)^p$ about $x = 0$, for any real p, converges to the right value for $|x| < 1$. Since this Taylor series is just the series obtained by expanding $(1+x)^p$ by the binomial theorem, we obtain a proof of the binomial theorem for negative or fractional exponents.

A function whose Taylor series about a converges to the function in some neighborhood of a is called *analytic* at a. As a contrast to the preceding negative examples, we prove a striking positive theorem (S. Bernstein's theorem): *If f and all its derivatives are nonnegative in an interval I, then f is analytic in the interval.* (An example is given by $f(x) = e^x$.)

Suppose that $a < x < b$ and $[a, b] \subset I$. Write

$$R_n(x) = \frac{1}{(n-1)!} \int_a^x f^{(n)}(t)(b-t)^{n-1} \left(\frac{x-t}{b-t}\right)^{n-1} dt.$$

Since $(x-t)/(b-t)$ decreases (as a function of t), it is largest at $t = a$, and since $f^{(n)}(t) \geq 0$ we have

$$\begin{aligned} R_n(x) &\leq \left(\frac{x-a}{b-a}\right)^{n-1} \frac{1}{(n-1)!} \int_a^x f^{(n)}(t)(b-t)^{n-1}\,dt \\ &\leq \left(\frac{x-a}{b-a}\right)^{n-1} \frac{1}{(n-1)!} \int_a^b f^{(n)}(t)(b-t)^{n-1}\,dt \\ &= \left(\frac{x-a}{b-a}\right)^{n-1} R_n(b). \end{aligned}$$

But $R_n(b) \leq f(b)$ because all the terms of the Taylor series are nonnegative. Hence

$$R_n(x) \leq \left(\frac{x-a}{b-a}\right)^{n-1} f(b),$$

and since $0 < x - a < b - a$, this means that $R_n(x) \to 0$.[6]

It is also true, but harder to prove, that if, on a given interval, each derivative of f has a fixed sign (possibly differing from one derivative to another), then f is analytic on the interval.[7]

Although a nonanalytic function can have a divergent Taylor series about every point of an interval, the phenomenon of convergence of the Taylor series to the wrong value cannot occur throughout an interval. In fact, we can prove that *if the Taylor series of a function, about each point in an interval, has a positive radius of convergence, there must be a subinterval in which the function is analytic.* Repeated applications of this fact lead to the conclusion that (under the same hypothesis) the points about which the Taylor expansion fails can form at most a nowhere dense set.

The proof depends on a simple application of Baire's theorem. Let $\rho(a)$ denote the radius of convergence of the Taylor series of f, formed at the point a. By a familiar formula, $1/\rho(a) = \limsup_{n\to\infty} |f^{(n)}(a)/n!|^{1/n}$. Since $1/\rho(a)$ is finite for each a in the interval in question, the quantity $\mu(a) = \sup_n |f^{(n)}(a)/n!|^{1/n}$ is finite for each a. The sets E_k of points a where $\mu(a) < k$ (for $k = 1, 2, \dots$) exhaust the interval and by Baire's theorem cannot all be nowhere dense. Hence there is a subinterval in which we have $|f^{(n)}(a)/n!|^{1/n} \leq k$ (for $n = 0, 1, 2, \dots$), first on a dense set, and then, by the continuity of $f^{(n)}$, throughout. In this interval f is analytic, since the inequality shows that the remainder in the Taylor series about a approaches zero for points x such that $|x - a| < 1/k$.

It is now natural to ask what happens when $\rho(a)$ is not only positive but bounded away from zero: $\rho(a) \geq \delta > 0$ for every a in an interval. It can be shown that this condition does make f analytic throughout the interval.[8]

NOTES

[1] For references and generalizations, see R. P. Boas, Asymptotic

relations for derivatives, *Duke Mathematical Journal* 3 (1937), 637–646.

[2]The functions for which this happens are a set of first category in the space of infinitely differentiable functions equipped with a certain natural complete metric. See R. Darst, Most infinitely differentiable functions are nowhere analytic, *Canadian Mathematical Bulletin* 16 (1973), 597–598; F. S. Cater, Differentiable, nowhere analytic functions, *American Mathematical Monthly* 91 (1984), 618–624.

[3]For literature, see H. Salzmann and K. Zeller, Singularitäten unendlich oft differenzierbarer Funktionen, *Mathematische Zeitschrift* 62 (1955), 354–367.

[4]E. Borel. See A. Rosenthal, On functions with infinitely many derivatives, *Proceedings of the American Mathematical Society* 4 (1953), 600–602; Salzmann and Zeller, paper cited in the preceding note; H. Mirkil, Differentiable functions, formal power series, and moments, *Proceedings of the American Mathematical Society* 7 (1956), 650–652; Steven G. Krantz and Harold R. Parks, *A Primer of Real Analytic Functions*, Birkhäuser, Basel, 1992, §2.2.

[5]This was shown by D. Morgenstern, Unendlich oft differenzierbare nicht-analytische Funktionen, *Mathematische Nachrichten* 12 (1954), 74, by using Baire's theorem; for a more explicit construction see H. Cartan, Sur les classes de fonctions définies par des inégalités portant sur leurs dérivées successives, *Actualités Scientifiques et Industrielles*, no. 867 (1940), pp. 20–22.

[6]For a different proof, see Ray Redheffer, From center of gravity to Bernstein's theorem, *American Mathematical Monthly* 90 (1983), 130–131.

[7]For an elementary proof, see J. A. M. McHugh, A proof of Bernstein's theorem on regularly monotonic functions, *Proceedings of the American Mathematical Society* 47 (1975), 358–360.

[8]See R. P. Boas, When is a C^∞ function analytic?, *Mathematical Intelligencer* 11 (1989), no. 4, 34–37; and references cited there.

Chapter 3

Integration

25. Lebesgue measure. So far, we have mostly been dealing, directly or indirectly, with what used to be called differential calculus, and was clearly separated from integral calculus. The word "integral" has generally been used in two different senses. One is the process of undoing differentiation to find what have been variously described as antiderivatives, primitives, or indefinite integrals. This is the technique of finding explicit formulas for antiderivatives, an art which has lost much of its importance now that there are not only extensive tables of antiderivatives, but also computer programs that can find antiderivatives faster than people can find them by hand. The other meaning of "integral" is what we picture informally as the area between a curve and a coordinate axis, and more formally as the limit of sums that approximate this area. Of course, these two notions are connected by what is (or at least used to be) called the fundamental theorem of the calculus. This chapter is about the second meaning of "integral." We shall deal only with real-valued functions on subsets of the real line.

The simplest function we can consider, besides a constant function, is a function whose range consists of just two values, say 0 and 1. If f is a function that takes the value 1 on a set E and the value 0 elsewhere, then the "area" under the graph of f should be the "length" of the set E. If, however, E is an irregular set, such as the set of irrational decimals, then E does not have a length in the ordinary sense. We need to generalize the idea of length to arrive at some notion of the size, or *measure*, of an irregular set E.

The measure of an interval is just its length. Each of the intervals (a, b), $[a, b]$, $(a, b]$, and $[a, b)$ has measure $b - a$. To define the measure of other sets, we take as a fundamental principle that "the whole is equal to the sum of its parts."[1] Thus, the measure of a finite collection of disjoint intervals is the sum of their lengths.

An open set in $\mathbf{R}_1$ consists of a countable collection of disjoint open intervals (page 30), and the measure of an open set is the sum of the lengths of its constituent open intervals. For example, the measure of the open set that is the union of the intervals $(n, n + 2^{-n})$ for $n = 1, 2, 3, \ldots$ is 1, while the measure of the open set that is the union of the intervals $(n, n + n^{-1})$ is infinite (since the harmonic series $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots$ diverges).

We can determine the measure of a closed set by taking complements. If E is a closed subset of an open interval I, then we know the measure of $I \cap C(E)$, since this is an open set. The measure of the closed set E is the difference between the measure (that is, length) of I and the measure of $I \cap C(E)$. We used this idea in §11 to see that the Cantor set has measure zero: the complement of the Cantor set with respect to the interval $(0, 1)$ consists of a countable union of disjoint intervals of total length 1.

Exercise 25.1. If E is an open set of finite measure, and a positive ϵ is prescribed, then there is a closed set F contained in E whose measure is within ϵ of the measure of E.

Exercise 25.2. If E is a closed set of finite measure, and a positive ϵ is prescribed, then there is an open set G containing E whose measure is within ϵ of the measure of E.

Next we would like to assign measures to sets that are neither open nor closed, for example, the set of irrational numbers between 0 and 1, or the set of decimal expansions that contain 3's but not 7's. A natural way to try to do this is to approximate an irregular set E with sets whose measures we know, that is, with open sets and closed sets. We cannot expect to approximate E by open sets from within (neither of the examples of sets just mentioned contains any neighborhoods), but we can approximate by open sets from outside. Alternatively, we might approximate E from within by closed sets. A set E is called (Lebesgue) *measurable* if it can be arbitrarily well approximated both from outside by open sets and from inside by closed sets. More formally, E is measurable if, for every prescribed positive ϵ, there exist a closed set F contained in E and an open set G containing E such that the measures of G and F differ by less than ϵ, or, what amounts to the same thing, the measure of $G \cap C(F)$ is less than ϵ. (The second formulation is useful for determining measurability of sets having infinite measure.)

If E is measurable, then when $\epsilon \to 0$ the measures of the open set G and the closed set F approach a common limit, which is called the measure of E and is denoted $m(E)$. Thus the measure of a measurable set E is the common value of the infimum of the measures of open sets containing E (in §22, we called this the outer measure of E and denoted it by $\mu(E)$) and the supremum of the measures of closed sets contained in E (this is known as the *inner measure* of E).

Not all sets are measurable. There exists a nonmeasurable set E contained in the interval $[0, 1]$ that has outer measure 1 and inner measure 0.[2]

Exercise 25.3. If E is an arbitrary set, and Z is a set of measure zero, then E is measurable if and only if $E \cup Z$ is measurable.

Open sets and closed sets are measurable, and so are sets built up from open sets and closed sets by countably many operations of forming unions, intersections, and complements. Sets formed in this way are known collectively as *Borel sets.* The set of decimal expansions that contain 3's but not 7's is the intersection of an open set and a closed set, and so is a Borel set. It can be shown that every Lebesgue measurable set differs from a Borel set by a set of measure zero, so for all practical purposes we need deal only with Borel sets.

Measures of sets combine as you would expect: If E_1 and E_2 are disjoint measurable sets, then $m(E_1 \cup E_2) = m(E_1) + m(E_2)$. More generally, for a sequence $\{E_n\}$ of mutually disjoint measurable sets, we have $m(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} m(E_n)$; in words, measure is a countably additive set function. If E_1 and E_2 are not disjoint, then $m(E_1 \cup E_2) = m(E_1) + m(E_2) - m(E_1 \cap E_2)$.

Exercise 25.4. If E is measurable, then for every bounded set S, the outer measure of the part of E in S and the outer measure of the complement of E with respect to S add up to the outer measure of S.[3]

Since the length of an interval $[a, b]$ is the same as the length of the interval $[a + c, b + c]$, it follows that measures of sets are invariant under translation of the sets by a fixed amount c. Since the length of $[ac, bc]$ is c times the length of $[a, b]$ (when $c > 0$), it follows that dilating a set by a factor of c will multiply the measure of the set by c.

Exercise 25.5. Show that a measurable set E of positive measure has the following property[4] used in §20: the set of differences $x - y$ between points x and y of E contains a neighborhood of 0.

You may be wondering what a set that is not measurable could look like. Such sets are not easy to construct,[5] but we can deduce from §20 that they must exist.[6] Suppose that f is a linear function that does not have the form $f(x) = ax$. We know by §20 and Exercise 25.5 that f cannot be bounded on any measurable set having positive measure. On the other hand, obviously f is bounded on the set E_1 consisting of the points x for which $|f(x)| \leq 1$ (in other words, E_1 is the inverse image of the interval $[-1, 1]$). Therefore either E_1 has measure zero, or else E_1 is not measurable. Now when n is a positive integer, the linearity property of f implies that the set E_n of points x for which $|f(x)| \leq n$ is the same as the set of points x for which $|f(x/n)| \leq 1$. In other words, E_n is just E_1 dilated by a factor of n; so E_n has measure zero if E_1 does. Therefore, if E_1 has measure zero, then so does $\bigcup_{n=1}^{\infty} E_n$. This union, however, is all of $\mathbf{R}_1$, which certainly does not have measure zero. We have seen that E_1 cannot be a measurable set of positive measure, and E_1 cannot be a set of measure zero, so it must be the case that E_1 is not measurable.

We have discussed the notion of measure only for sets in $\mathbf{R}_1$. You can probably guess how to define measure in $\mathbf{R}_n$ by approximating sets with unions of boxes instead of with unions of intervals. According to the astonishing Banach-Tarski "paradox,"[7] a ball in $\mathbf{R}_3$ the size of a pea can be divided into a finite number of pieces which can then be reassembled (via rigid motions in space) into a ball the size of the sun. Evidently the pieces cannot all have a well-defined volume, that is, some of them must be nonmeasurable sets.

NOTES

[1]This addition to the "common notions" of Euclid's *Elements* is attributed to the Renaissance scholar Clavius. See Thomas L. Heath, *The Thirteen Books of Euclid's Elements*, second edition, vol. I, Dover, New York, 1956, p. 232.

[2]See, for example, Andrew Simoson, On two halves being two wholes, *American Mathematical Monthly* 91 (1984), 190–193.

[3]The converse is also valid. It is known as Carathéodory's criterion for measurability.

[4]This property of sets of positive measure was found by H. Steinhaus, Sur les distances des points des ensembles de mesure positive, *Fundamenta Mathematicae* 1 (1920), 93–104. For a short, elegant proof, see p. 561 of R. L. Cooke, The Cantor-Lebesgue theorem, *American Mathematical Monthly* 86 (1979), 558–565.

[5]The axiom of choice is required, as was shown by R. M. Solovay, A model of set-theory in which every set of reals is Lebesgue measurable, *Annals of Mathematics* 92 (1970), 1–56. See also James M. Briggs and Thomas Schaffter, Measure and cardinality, *American Mathematical Monthly* 86 (1979), 852–855.

[6]That a measurable linear function must be continuous was observed by Wacław Sierpiński, Sur l'équation fonctionnelle $f(x+y) = f(x)+f(y)$, *Fundamenta Mathematicae* 1 (1920), 116–122; and Stefan Banach, Sur l'équation fonctionnelle $f(x+y) = f(x)+f(y)$, *ibid.*, 123–124. Moreover, a measurable *convex* function (§23) must be continuous; see Wacław Sierpiński, Sur les fonctions convexes mesurables, *ibid.*, 125–129.

[7]For more on the Banach-Tarski paradox, see Robert M. French, The Banach-Tarski theorem, *Mathematical Intelligencer* 10 (1988), no. 4, 21–28; Karl Stromberg, The Banach-Tarski paradox, *American Mathematical Monthly* 86 (1979), 151–161; and Stan Wagon, *The Banach-Tarski Paradox*, Cambridge University Press, 1985. The original paper is S. Banach and A. Tarski, Sur la décomposition des

ensembles de points en parties respectivement congruentes, *Fundamenta Mathematicae* 6 (1924), 244–277.

26. Measurable functions. Now that we have measurable sets to work with, we can define measurable functions. If f is going to be a "nice" function, then the set of points where f takes values between, say, 0 and 1 should be a "nice" set. Accordingly, if f is a real-valued function whose domain is an interval, we say that f is (Lebesgue) measurable if the inverse images, under f, of all intervals are measurable sets.

Exercise 26.1. A monotonic function is measurable.

Exercise 26.2. A continuous function is measurable.

Exercise 26.3. If f is a measurable function, and g is a function such that $f(x) = g(x)$ for almost all x, then g is a measurable function.

Some other examples of measurable functions, besides the ones in the preceding exercises, are step functions; sums, differences, and products of measurable functions; limits, and upper or lower limits, of sequences of measurable functions; the absolute value of a measurable function; and the pointwise maximum and the pointwise minimum of two measurable functions.

Recall (page 125) that a Baire class of functions consists of the functions that can be obtained by taking pointwise limits of sequences of functions from the preceding Baire classes (the class at level 0 being the continuous functions). Evidently functions of any Baire class are measurable, since they are limits of sequences of measurable functions. In the setting of Lebesgue measure, we ought to consider limits that exist at all points except for a set of measure zero. In this context, all measurable functions arise from continuous functions by a single limiting process. That is, it can be shown that a function f on an interval $[a, b]$ is measurable if and only if there is a sequence $\{f_n\}$ of continuous functions such that $\lim_{n\to\infty} f_n(x) = f(x)$ for almost all x in the interval.

It could happen that on the exceptional set of measure zero where the sequence $\{f_n(x)\}$ does not converge to $f(x)$, this sequence does not converge at all. To put it another way, it is not always possible to change a measurable function on a set of measure zero to get a function of Baire class 1. It is possible, however, to change any measurable function on a set of measure zero to get a function of Baire class no greater than 2.[1]

Exercise 26.4. Every measurable function on an interval $[a, b]$ is the difference of two measurable functions each of which is the pointwise limit almost everywhere of a *monotonic* sequence of continuous functions.

The fact that a limit of a sequence of measurable functions remains measurable is one of the significant advantages of Lebesgue's theory of integration over Riemann's. (The limit of a sequence of Riemann integrable functions need not be Riemann integrable.) Moreover, sequences of measurable functions that converge on a finite interval $[a, b]$ do so "almost uniformly," meaning that the convergence is uniform if we discard from the domain of the functions a suitable set of arbitrarily small positive measure. This is known as Egoroff's (or Egorov's) theorem.[2]

Later on we will use Egoroff's theorem, but I omit its proof[3] on the principle that you can learn to drive an automobile without having to learn how the engine works. (I once heard Norbert Wiener admit that he had used the ergodic theorem extensively without having read a proof;

it is only fair to add that he later created his own proof.) In practice, you can use the Lebesgue integral for many applications without having worked through proofs of all of its properties, provided that you understand the properties. (How much do you really know about what is going on inside your computer?)

A special subclass of measurable functions on an interval $[a, b]$ that arises in the theory of integration is the class of functions of *bounded variation.* Roughly speaking, these are functions f with a finite amount of up and down oscillation. We consider subdividing $[a, b]$ into a finite number of subintervals $[x_k, x_{k+1}]$ that touch each other only at their endpoints. We form the sum $\sum_k |f(x_{k+1}) - f(x_k)|$, and then we find the least upper bound of such sums over all possible subdivisions of $[a, b]$. This quantity is the "total variation" of f, and if it is finite, then we say that f has bounded variation. For example, a monotonic function has bounded variation, because every such sum equals $|f(b) - f(a)|$; and the difference of two monotonic functions has bounded variation, because (by the triangle inequality) its total variation is no bigger than the sum of the total variations of the two functions.

It is interesting that the examples just mentioned are the only ones that exist: *a function f of bounded variation on an interval $[a, b]$ can always be written as the difference of two monotonic functions.* To see this, let $T(x)$ denote the total variation of f on the interval $[a, x]$. Clearly T is a nondecreasing function.[4] Since we can write f as the difference of T and $T - f$, we need only show that $T - f$ is a nondecreasing function. In other words, we need to show that if c and d are arbitrary points of the interval, and $c < d$, then $\{T(d) - f(d)\} - \{T(c) - f(c)\} \geq 0$, or equivalently, $T(d) - T(c) \geq f(d) - f(c)$. This last inequality is true because $T(d) - T(c)$ represents the total variation of f on the interval $[c, d]$, and hence it is at least as big as the net change $|f(d) - f(c)|$.

Exercise 26.5. A function of bounded variation is differentiable almost everywhere.

Exercise 26.6. If f is a continuous function on a finite interval $[a, b]$, and the derivative f' exists at all points of (a, b) and is bounded, then f has bounded variation.

Exercise 26.7. Give an example of a continuous function on the interval $[0, 1]$ that does not have bounded variation.

An important subclass of the functions of bounded variation consists of functions that have small variation on collections of intervals of small total length. We say that a function f on an interval $[a, b]$ is *absolutely continuous* if, given any positive ϵ, we can find a positive number δ such that for every finite collection of disjoint subintervals $\{(x_k, x_k + h_k)\}$ of total length less than δ, we have $\sum_k |f(x_k + h_k) - f(x_k)| < \epsilon$.

Exercise 26.8. An absolutely continuous function is, in particular, continuous.

Exercise 26.9. An absolutely continuous function is, in particular, a function of bounded variation.

We saw in §22 that there are singular monotonic functions, such as the Cantor function, that are not constant even though the derivative is equal to zero almost everywhere. The condition of absolute continuity rules out this sort of behavior: *if f is absolutely continuous on $[a, b]$, and if $f'(x) = 0$ for almost all x, then f is a constant function.* To see this, let c be an arbitrary point in $[a, b]$

different from a; we want to show that $f(c) = f(a)$. Fix a positive ϵ, and observe that if x is a point of $[a, c]$ for which $f'(x) = 0$, then there is an interval $(x - h_x, x + h_x)$ such that $|f(x) - f(y)| \leq \epsilon|x - y|$ when y is in this interval; denote the totality of such intervals for different values of x by J_1. Let δ be the number corresponding to ϵ in the definition of absolute continuity. We can cover the set of measure zero where $f'(x)$ is different from zero or fails to exist by a countable collection of open intervals of total length less than δ; denote this collection of intervals by J_2. Since $J_1 \cup J_2$ is a collection of open intervals that covers the compact set $[a, c]$, there is a covering by a finite number of these intervals. For this finite number of intervals, list the endpoints interior to $[a, c]$ in increasing order as $x_1, x_2, \ldots, x_{n-1}$, and put $x_0 = a$ and $x_n = c$. By the triangle inequality, $|f(c) - f(a)| \leq \sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)|$. Those terms of this sum corresponding to intervals that are contained in intervals from the collection J_2 have a total less than ϵ, by the definition of absolute continuity. Those terms of the sum corresponding to intervals that are contained in intervals from the collection J_1 have a total less than $2\epsilon(c - a)$. Thus $|f(c) - f(a)| \leq \epsilon(1 + 2(c - a))$. Since ϵ is arbitrary, we obtain the desired conclusion $f(c) = f(a)$.

NOTES

[1]G. Vitali, Una proprietà delle funzioni misurabili, *Rendiconti Reale Istituto Lombardo di scienze e lettere* (2) 38 (1905), 599–603.

[2]For a generalization to infinite intervals, see Robert G. Bartle, An extension of Egorov's theorem, *American Mathematical Monthly* 87 (1980), 628–633.

[3]You can find a proof of Egoroff's theorem in most treatises that cover Lebesgue integration, such as H. L. Royden, *Real Analysis*, third edition, Macmillan, New York, 1988; or F. Riesz and B. Sz.-Nagy, *Functional Analysis*, Ungar, New York, 1955.

[4] For more about T, see Frank N. Huggins, Some interesting properties of the variation function, *American Mathematical Monthly* 83 (1976), 538–546.

27. Definition of the Lebesgue integral. What you are most likely to have seen about integration, beyond the most elementary level, is Riemann's definition of an integral. The functions on a bounded interval $[a, b]$ that are integrable in Riemann's sense are precisely those bounded functions that are continuous at almost all points of the interval;[1] that is, the discontinuities form a set of measure zero. Let us quote a definition of the Riemann integral, so that we shall be able to compare it with the definitions of other kinds of integrals. We imagine dividing $[a, b]$ into a finite number of subintervals (x_k, x_{k+1}), forming the sums $\sum_k (x_{k+1} - x_k) f(y_k)$, where $x_k \leq y_k \leq x_{k+1}$, and finding the limit of these sums as the maximum difference $x_{k+1} - x_k$ approaches 0. This limit is $\int_a^b f$, or $\int_a^b f(x)\,dx$. (The conventional dx really belongs to the integral viewed as an antiderivative. We can normally dispense with it, except in cases such as $\int_a^b x^2\,dx$ when the function being integrated has no standard name—see page 81.)

There are good reasons for wanting to be able to integrate more general kinds of functions. We can't, for example, use Riemann's theory to integrate a function on $[0, 1]$ that is equal to 1 on a generalized Cantor set (page 74) of positive measure and equal to 0 on the complementary set, because this function has too many discontinuities. You might expect that if we were clever enough, we would be able to design an integral that would integrate all kinds of functions over all kinds of sets. However, it can be proved that there is no useful integral that will even integrate all functions over all "nice" sets, or integrate all "nice" functions over all sets.[2] The best we can do is to integrate a large number of useful functions (the measurable functions) over a large number of sets (the measurable sets).

Lebesgue was one of the many mathematicians who attempted to improve on the Riemann integral. It seems that he was mainly interested in finding an integral that would be better than the Riemann integral for studying trigonometric series. His definition turned out to have more potential than he perhaps imagined to begin with. What has made the Lebesgue integral so central in modern mathematical analysis is that it turned out to be the right integral to use in what we now call functional analysis, that is, the study of spaces whose elements are functions. This Primer is not a book about functional analysis, but the existence of that subject is a main reason for devoting so much space to the properties of integrals that make much of functional analysis possible. Consequently, the Lebesgue integral is not merely a case of one more generalization for the sake of generalizing.

The Lebesgue integral has become the standard integral for use in mathematical analysis generally, and especially in the theory of probability. There is, however, a persistent belief (I have seen it expressed in recent textbooks) that Lebesgue integration is a difficult and abstruse subject that is best avoided as long as possible. This may have been a valid objection in earlier days. In 1938, I heard G. H. Hardy describe how difficult it had been, thirty years earlier, for mathematicians of his generation to learn Lebesgue integration. My own generation took it as a matter of course, and by now it is more accessible (as I hope to show you), and essential for work in analysis.

We begin by defining the integral of a bounded function f on a bounded interval I. To do this, we first divide I, not just into subintervals, but into a finite number of measurable subsets S_k in such a way that any overlaps between different sets have measure zero. We can then form upper and lower sums $\sum_k m(S_k) \sup_{S_k} f$ and $\sum_k m(S_k) \inf_{S_k} f$, which we do not expect to be equal. We next find the infimum of all the upper sums, and the supremum of all the lower sums. If these turn out to be equal, their common value is defined to be the Lebesgue integral of f, denoted by $\int_I f$. This gives us a definition of the integral of a function over an interval. To integrate a function over a measurable set S, we form the function that has the values of f on S and is equal to 0 off S, and integrate that. It can be shown that the bounded functions f on I that are integrable in this sense are precisely the bounded measurable functions.

Exercise 27.1. Compute the Lebesgue integral of the function on the interval $[0, 1]$ that is equal to 0 at the rational numbers and equal to 1 at the irrational numbers.

Exercise 27.2. Compute the Lebesgue integral of the Cantor function (page 162).

This definition of the Lebesgue integral is superficially much like the definition of the Riemann integral.[3] One of the most important differences is the use of measurable sets instead of intervals. In Lebesgue's original definition of an integral, the range of the function was partitioned, instead of the domain. That often used to be thought (incorrectly) to be the essential difference between Lebesgue's and Riemann's definitions.

Exercise 27.3. Lebesgue's integral is more general than Riemann's in the sense that a Riemann integrable function is measurable, and its Riemann integral equals its Lebesgue integral;

however, there are Lebesgue integrable functions that are not Riemann integrable.

So far, we have defined Lebesgue integrals for bounded functions on bounded domains, but there are many interesting or useful functions that do not have these properties. We can define integrals of more general functions by taking limits.[4] First suppose that f is a nonnegative function that is unbounded or that has an unbounded domain S. To define $\int_S f$, we integrate the truncated function that is equal to f at the points where f has a value less than n, and is equal to n elsewhere, over the set $S \cap [-n, n]$, and then we take the limit as $n \to \infty$. We do not consider a function to be integrable if its integral turns out to be infinite. For a function f that is positive at some points and negative at others, we split f as the difference $f_1 - f_2$ of two nonnegative functions, where $f_1 = (|f| + f)/2$ and $f_2 = (|f| - f)/2$. We put $\int f = \int f_1 - \int f_2$, as long as this is not the meaningless quantity $\infty - \infty$, and we call f *integrable* if both $\int f_1$ and $\int f_2$ are finite. In other words, a measurable function f is integrable if and only if the nonnegative function $|f|$ is integrable; consequently, Lebesgue integrals are often said to be absolutely convergent. For example, $\int_0^\infty x^{-3/2} \sin x \, dx$ is well defined, since the integrals over the collections of intervals where $\sin x \geq 0$ and where $\sin x \leq 0$ are both finite; but $\int_0^\infty \sin x \, dx$ and even $\int_0^\infty x^{-1} \sin x \, dx$ do not exist as Lebesgue integrals.

Exercise 27.4. Verify that $x^{-1} \sin x$ is not Lebesgue integrable on $(0, \infty)$ even though $\lim_{t \to \infty} \int_0^t x^{-1} \sin x \, dx$ is finite.

From the point of view of Lebesgue integration, any set whose measure is zero (sometimes called a "null set") can be disregarded, for the value of $\int_S f$ is unaffected if we change the values of f on a subset of S that is of measure zero. This means that when we integrate functions, we can identify any two functions that differ only on a set of measure zero. A more formal way of saying this is that we do not actually integrate a function, but rather an equivalence class of functions, where "equivalent" means "differing only on a set of measure zero." From this point of view, an integrable function has a rather nebulous character. We should not say that it has any particular value at any particular point, because changing its value at a point does not affect an integral of the function. In practice, however, we usually visualize an integrable function as some particular representative of its equivalence class.

It is often convenient to allow functions to have the values $\pm\infty$, but only on sets of measure zero. If f is integrable (meaning that f is measurable, and the integral of $|f|$ is finite), then the value $f(x)$ is finite almost everywhere. Moreover, the measure of the set of points x for which $|f(x)| > M$ approaches 0 as $M \to \infty$. Informally, we may say that an integrable function is large only on a small set.

Exercise 27.5. If f is integrable on the measurable set S, and if $\epsilon > 0$, then there exists $\delta > 0$ such that $\int_E |f| < \epsilon$ for every subset E with measure less than δ.

NOTES

[1] See, for example, Michael W. Botsko, An elementary proof that a bounded a.e. continuous function is Riemann integrable, *American Mathematical Monthly* 95 (1988), 249–252.

[2] Assuming the continuum hypothesis, S. Ulam showed that the only finite measure defined on all subsets of the real numbers that assigns measure zero to each point is identically zero (Zur Masstheorie in der allgemeinen Mengenlehre, *Fundamenta Mathematicae* 16 (1930), 141–150).

[3]It was discovered in the 1960's that the Lebesgue integral can actually be constructed by a procedure entirely analogous to Riemann's. See Carus Mathematical Monograph number 20, *The Generalized Riemann Integral*, by Robert M. McLeod.

[4]By making suitable conventions about how to handle the symbol ∞, one can define the Lebesgue integral of a (perhaps) unbounded measurable function over a (perhaps) unbounded measurable set directly (without taking limits) by the same procedure used for bounded functions on bounded domains. This is often considered to be one of the advantages of the Lebesgue integral.

28. Properties of Lebesgue integrals. As you would guess, the Lebesgue integral shares many properties with the Riemann integral. For instance, integration is a linear operation: $\int(cf+g) = c\int f + \int g$ for all integrable functions f and g and every constant c. The integral is additive over sets: if S_1 and S_2 are disjoint measurable sets, then $\int_{S_1 \cup S_2} f = \int_{S_1} f + \int_{S_2} f$. Integration is a positive operation, in the sense that if $f(x) \geq 0$ for almost all x, then $\int f \geq 0$; and if in addition $f > 0$ on a set of positive measure, then $\int f > 0$. More generally, if f and g are integrable functions, and $f(x) \leq g(x)$ almost everywhere, then $\int f \leq \int g$. A common way to show that an unbounded measurable function f is integrable is to find an integrable function g that dominates f (this means that $|f(x)| \leq |g(x)|$ almost everywhere).

Exercise 28.1. If f denotes the function whose value at x is $x^{-1/2}\sin(\log x)$, then f is integrable on the interval $(0, 1)$.

Now we turn to some properties of Lebesgue integrals that Riemann integrals might have, but don't, or that are not meaningful for Riemann integrals.

(i) Integration of Sequences and Series. We cannot necessarily integrate a convergent sequence of functions term by term. For example, if f_n is the function defined by $f_n(x) = n$ for $0 < x < 1/n$, and $f_n(x) = 0$ elsewhere, we have $f_n(x) \to 0$ for each x, while $\int_0^1 f_n(x)\,dx = 1$, so that $\lim_n\{\int_0^1 f_n(x)\,dx\}$ and $\int_0^1\{\lim_n f_n(x)\}\,dx$ both exist but are different.

Exercise 28.2. Find an example of the same phenomenon with continuous functions f_n.

We do know (page 117) that with the Riemann integral we can integrate a uniformly convergent sequence of functions term by term: if $f_n \to f$ uniformly, then $\int f_n \to \int f$. With Lebesgue integration, we get the much more powerful "dominated convergence theorem." It says that if $\{f_n\}$ is a sequence of Lebesgue integrable functions converging pointwise almost everywhere to a limit function f, and if there exists a Lebesgue integrable function g such that $|f_n(x)| \le g(x)$ almost everywhere for each n, then f is Lebesgue integrable, and $\int f_n \to \int f$. In terms of series, the theorem says that if $\sum_{n=1}^{\infty} f_n(x)$ converges for almost all x, and if there exists an integrable function g such that $|\sum_{n=1}^{N} f_n(x)| \le g(x)$ almost everywhere for all N, then the sum of the series is automatically integrable, and we can correctly integrate term by term.

We can deduce the dominated convergence theorem by applying Egoroff's theorem (page 202). Suppose $\epsilon > 0$. Since g is integrable, there is a set S of finite measure such that the integral of g over the complement $C(S)$ is less than ϵ. By Exercise 27.5, there is a positive δ such that $\int_E g < \epsilon$ for every subset E of S with measure less than δ. By Egoroff's theorem, we can choose E so that $f_n \to f$ uniformly on the complement of E. This means that $|f_n(x) - f(x)| < \epsilon/m(S)$ for sufficiently large n and for $x \notin E$. Accordingly, $\int_S |f_n - f| < \int_E |f_n - f| + \epsilon$ for such n. Since g dominates each f_n, it also dominates the limit function f, and so $\int_E |f_n - f| \leq \int_E 2g < 2\epsilon$. Similarly, $\int_{C(S)} |f_n - f| < 2\epsilon$. Hence $\int |f_n - f| < 5\epsilon$ when n is large enough, which means that $\int f_n \to \int f$.

Exercise 28.3. If $\{f_n\}$ is a sequence of measurable functions on a finite interval $[a, b]$ such that $|f_n(x)| \leq M$ for all n and almost all x, and if $f_n(x) \to f(x)$ for almost all x, then $\int_a^b f_n \to \int_a^b f$. This is the "bounded convergence theorem."

As a consequence of the bounded convergence theorem (or of the dominated convergence theorem), we can get a proposition known as *Fatou's lemma*: If $\{f_n\}$ is a sequence of nonnegative integrable functions, and if $\lim_n f_n(x) = f(x)$ for almost all x in S, then $\int_S f \leq \liminf_n \int_S f_n$. (It is allowed for both sides to be infinite: if f is not integrable, then the integrals $\int f_n$ must tend to infinity.) Part of the point is that the integrability of f is a conclusion, and not a hypothesis. To see this, let $f^{(N)}$ denote the minimum of f and the function that is identically equal to N on the interval $(-N, N)$ and identically equal to 0 outside this interval; define $f_n^{(N)}$ similarly. By definition, $\int_S f = \lim_{N\to\infty} \int_S f^{(N)}$. For a fixed value of N, the bounded convergence theorem applies to the sequence $\{f_n^{(N)}\}$, so $\int_S f^{(N)} = \lim_{n\to\infty} \int_S f_n^{(N)}$. But $f_n^{(N)}(x) \leq f_n(x)$ for all x, so $\lim_{n\to\infty} \int_S f_n^{(N)} \leq \liminf_{n\to\infty} \int_S f_n$. Consequently $\int_S f^{(N)} \leq \liminf_{n\to\infty} \int_S f_n$, and we get the conclusion by letting $N \to \infty$.

Exercise 28.4. Prove that if f is a nondecreasing function on the interval $[a, b]$, then $\int_a^b f' \leq f(b) - f(a)$.

Exercise 28.5. Prove the "monotone convergence theorem": If the f_n are integrable functions, and the sequence $\{f_n(x)\}$ is nondecreasing for each x, so that $f_n(x) \to f(x)$ (the limit possibly being infinite for some x), then $\lim_n \int f_n = \int f$.

Rephrased in terms of series, the monotone convergence theorem of the preceding exercise says that if $\sum_n \int |f_n|$ converges, then $\sum_n f_n(x)$ converges almost everywhere, and $\sum_n \int f_n = \int \sum_n f_n$. This implies, for a series of non-negative terms, that if termwise integration gives a convergent result, it is necessarily correct.

(ii) DIFFERENTIATION AND INTEGRATION. In elementary calculus, integrals and derivatives are related by the "fundamental theorem," which says that integration and differentiation are inverse processes, if we put appropriate restrictions on the functions involved. This relation carries over to the Lebesgue theory with less severe restrictions on the functions.

Suppose that f is integrable on an interval $[a, b]$, and let F be the function whose value at each x is $\int_a^x f$. Without some additional hypothesis, we cannot expect to conclude that the derivative of F is f. For example, if f has a simple jump at an interior point c of the interval $[a, b]$, then the derivative $F'(c)$ does not even exist. (Think of a step function: the graph of its indefinite integral has a corner at the point where $x = c$.) For another example, suppose that $f(x) = 1$ for all x different from c, and $f(c) = 2$; then $F'(c)$ exists and equals 1, which is different from the value $f(c)$. In fact, since we can modify the values of a function f on a set of measure zero without changing the value of its integral, the most we can hope for is that for almost all x the derivative $F'(x)$ exists and equals $f(x)$. This is precisely what happens.

To see why, let us split the function f into a sum of its positive part f_1 and its negative part f_2, where $f_1(x) = \max(f(x), 0)$, and similarly $f_2(x) = \min(f(x), 0)$. Then $\int_a^x f_1$ and $\int_a^x f_2$ are monotonic functions of x, so they have derivatives almost everywhere by a theorem proved in §22. Hence their sum F has a derivative almost everywhere.

It remains to show that this derivative equals $f(x)$ for almost all x. First consider the special case that f is continuous at an interior point c of $[a, b]$. Then the average value $h^{-1}\int_c^{c+h} f$ approaches $f(c)$ when $h \to 0$. But $h^{-1}\int_c^{c+h} f = h^{-1}\{F(c+h) - F(c)\}$, so $F'(c) = f(c)$. In the general situation, we may assume, by Exercise 26.4, that f is the limit almost everywhere of a nondecreasing sequence $\{f_n\}$ of continuous functions. Set $F_n(x) = \int_a^x (f_{n+1} - f_n)$. From the special case just considered, we know that $F_n'(x) = f_{n+1}(x) - f_n(x)$ for all x. Because the integrand is nonnegative, F_n is a nondecreasing function. Because of the telescoping sum in the integrand, $\sum_{j=1}^n F_j(x) = \int_a^x (f_{n+1} - f_1)$. By the monotone convergence theorem, $\sum_{j=1}^n F_j(x)$ converges for all x to $F(x) - \int_a^x f_1$. By Fubini's theorem on derivatives (page 171), $\sum_{j=1}^n F_j'(x)$ converges to $F'(x) - f_1(x)$ for almost all x. But $\sum_{j=1}^n F_j'(x) = \sum_{j=1}^n (f_{j+1}(x) - f_j(x)) = f_{n+1}(x) - f_1(x)$, which converges to $f(x) - f_1(x)$ for almost all x. Thus $F'(x) = f(x)$ for almost all x.

Now let us consider the related question of whether the integral of a derivative gives back the original function. Again, we cannot expect this to be true in general. Indeed, even if f has a *finite* derivative everywhere, f' may fail to be integrable. To visualize an example of such an f, imagine the line segment $(0, 1)$ partitioned into infinitely many intervals that condense at 0, with each interval carrying a pair of teeth of height 1, one positive and the other negative. The integral of $|f'|$ over each interval equals 4 (the total variation of f on the interval), and since there are infinitely many intervals, f' is not Lebesgue integrable on $(0, 1)$. The corners of the graph of f can be rounded off to make f differentiable at all points of $(0, 1)$.[1]

If f' exists almost everywhere and is integrable, it need not be the case that f is an indefinite integral of f'. For example, the derivative of the Cantor function (page 162) is zero almost everywhere, and so the integral of the derivative is identically zero.[2]

For a function f to have any chance of being the integral of its derivative, f must possess any properties that indefinite integrals have in general. One such property is continuity. If g is integrable on $[a, b]$, then the function G such that $G(x) = \int_a^x g$ is a continuous function of x; this follows from Exercise 27.5. For the same reason, an indefinite integral has the stronger property of absolute continuity. This necessary property is also sufficient: *if f is absolutely continuous on $[a, b]$, then f has a derivative almost everywhere, and f is an indefinite integral of its derivative.*

Indeed, f is, in particular, of bounded variation (Exercise 26.9), and hence is the difference $f_1 - f_2$ of two nondecreasing functions f_1 and f_2; so $f'(x)$ exists for almost all x. Moreover, f_1' and f_2' are integrable by Exercise 28.4, so f' is too because $|f'(x)| \le f_1'(x) + f_2'(x)$. Let $F(x) = \int_a^x f'$. We know that F is absolutely continuous, and $F'(x) = f'(x)$ for almost all x. Thus $F - f$ is an absolutely continuous function whose derivative is zero almost everywhere. From page 204, we know that $F - f$ is constant; that is, f is an indefinite integral of its derivative.

NOTES

[1] There is an explicit formula in section 10.7 of E. C. Titchmarsh,

Theory of Functions, second edition, Oxford University Press, 1968; but a verbal description seems to be more informative than a formula.

[2]On the other hand, if f' exists and is finite everywhere and is integrable, then f is an indefinite integral of f'. See, for example, Casper Goffman, On functions with summable derivative, *American Mathematical Monthly* 78 (1971), 874–875; Walter Rudin, *Real and Complex Analysis*, third edition, McGraw-Hill, New York, 1987, Theorem 7.21, p. 149. Also, if f is continuous, the right-hand derivative of f exists and is finite except at countably many points, and the right-hand derivative is integrable, then f is an indefinite integral of its right-hand derivative (and hence of f'). See P. L. Walker, On Lebesgue integrable derivatives, *American Mathematical Monthly* 84 (1977), 287–288; Edwin Hewitt and Karl Stromberg, *Real and Abstract Analysis*, second printing corrected, Springer-Verlag, New York, 1969, (18.41)(d), p. 299.

29. Applications of the Lebesgue integral.

(i) COMPUTING A MAXIMUM BY INTEGRATION. When $p > 0$, the space of (equivalence classes) of measurable functions f for which $\int |f|^p$ is finite is denoted by L^p (or $L^p(S)$ if it is necessary to specify the domain of the functions). If S is a finite interval, we can show that $f \in L^p(S)$ implies $f \in L^q(S)$ if $q < p$, but not necessarily for any $q > p$. Indeed, if $q < p$, then $|f(x)|^q \le \max(1, |f(x)|^p)$ for all x, so $|f|^q$ is integrable if $|f|^p$ is, and in fact $\int_S |f|^q \le m(S) + \int_S |f|^p$. On the other hand, if $q > p$, think of $S = (0, 1/2)$, and let $f(x)$ be $1/\{x(\log x)^2\}^{1/p}$; then f^p is integrable but f^q is not.

Exercise 29.1. Give an example of a measurable function f such that $f \in L^q(0,1)$ if $q < p$, but $f \notin L^p(0,1)$.

When $p \geq 1$, the quantity $(\int |f|^p)^{1/p}$ is denoted by $\|f\|_p$ and is called the L^p *norm* of f. Minkowski's inequality (page 184) shows that it satisfies the triangle inequality. Consequently, we can view the function space L^p as a metric space, the distance between functions f and g being $\|f - g\|_p$.

Exercise 29.2. Compute $\|f\|_p$ for the function f on the interval $[0, 1]$ such that $f(x) = x$. What happens when $p \to \infty$?

Now we consider the question of finding the maximum of a nonnegative function f that is integrable on a finite interval $[a, b]$. If f is bounded, there is, of course, no reason to suppose that f attains a maximum. Even the notion of its supremum, or least upper bound, makes no sense in the context of Lebesgue integrable functions, since we could change the value of f at a single point and get a different value for the supremum. If we are going to think about a supremum at all, we shall need some property that is not affected by how the function behaves on sets of measure zero.

We therefore introduce a property that acts as much as possible like a supremum. We define the *essential supremum* of a nonnegative function f (denoted by $\|f\|_\infty$, or, in older works, by $\operatorname{ess\,sup} f$ or $\operatorname{vrai\,max} f$) to be the number M with the property that, for every positive ϵ, the inequality $f(x) < M + \epsilon$ holds for almost all x, whereas $f(x) > M - \epsilon$ on some set of positive measure. If there is no such number M, we define $\|f\|_\infty$ to be ∞.

Exercise 29.3. Determine $\|f\|_\infty$ when f is the ruler function (Exercise 13.1).

We are now going to prove that $\lim_{p\to\infty} \|f\|_p = \|f\|_\infty$, whether $\|f\|_\infty$ is finite (case 1) or infinite (case 2).

Case 1. Let $M = \|f\|_\infty$. If $M = 0$, then f is zero almost everywhere, so $\|f\|_p = 0$ for all p, and there is nothing to prove. Suppose, then, that $M > 0$, and let ϵ be an arbitrary positive number less than M. By the definition of essential supremum, $f(x) < M + \epsilon$ for almost all x. Then

$$\int_a^b f^p < (M + \epsilon)^p (b - a).$$

In particular, $f \in L^p$ for every p. Moreover

$$\limsup_{p \to \infty} \|f\|_p \le \limsup_{p \to \infty} (M + \epsilon)(b - a)^{1/p} = M + \epsilon.$$

On the other hand, $f(x) > M - \epsilon$ on a set S of positive measure. Then

$$\left(\int_a^b f^p\right)^{1/p} \ge \left(\int_S f^p\right)^{1/p} > \left(\int_S (M - \epsilon)^p\right)^{1/p}$$
$$= (M - \epsilon)(m(S))^{1/p}.$$

Hence $\liminf_{p \to \infty} \|f\|_p \ge M - \epsilon$. Since ϵ is arbitrary, we deduce that the limit $\lim_{p \to \infty} \|f\|_p$ exists and equals M.

Case 2. If there were a positive number p such that f does not belong to L^p, then f would not belong to L^q for any $q > p$, and the limit $\lim_{p \to \infty} \|f\|_p$ would be infinite. We may therefore suppose that f belongs to every L^p. Our assumption now is that there is no M such that $f(x) < M$ almost everywhere. Consequently, for each positive n, we have $f(x) > n$ on a set S_n of positive measure. Then

$$\left(\int_a^b f^p\right)^{1/p} \ge \left(\int_{S_n} f^p\right)^{1/p} > \left(\int_{S_n} n^p\right)^{1/p}$$
$$= n \cdot (m(S_n))^{1/p}.$$

Letting $p \to \infty$, we obtain $\liminf_{p \to \infty} \|f\|_p \ge n$. Since n is arbitrary, this means that $\lim_{p \to \infty} \|f\|_p = \infty$.

(ii) A CONNECTION BETWEEN SUMS AND INTEGRALS. The integral test for convergence of infinite series says that if f is a positive decreasing function on $(0, \infty)$, then $\int_1^\infty f$ and $\sum_{k=1}^\infty f(kx)$ (for any given positive number x) either both converge or both diverge.

Exercise 29.4. Prove this integral test.

Exercise 29.5. If f is not monotonic, the conclusion may fail.

If f is merely nonnegative and integrable, nothing can be said about the convergence of $\sum_k f(kx)$ in general. There is, however, a sort of average connection between $\int f$ and $\sum_k f(kx)$: if the integral converges, the sum converges for almost all x.

To prove this,[1] we show that

$$\int_1^\infty \left(x^{-1} \sum_{k=1}^\infty f(kx)\right) dx \le \int_1^\infty f.$$

We saw in §28 that sums and integrals of positive functions can be rearranged however we like, so

$$\begin{aligned}\int_1^\infty \left(x^{-1} \sum_{k=1}^\infty f(kx)\right) dx &= \sum_{k=1}^\infty \int_1^\infty x^{-1} f(kx)\, dx \\ = \sum_{k=1}^\infty \int_k^\infty x^{-1} f(x)\, dx &= \sum_{k=1}^\infty \sum_{n=k}^\infty \int_n^{n+1} x^{-1} f(x)\, dx \\ = \sum_{n=1}^\infty \sum_{k=1}^n \int_n^{n+1} x^{-1} f(x)\, dx &= \sum_{n=1}^\infty n \int_n^{n+1} x^{-1} f(x)\, dx \\ &\le \sum_{n=1}^\infty \int_n^{n+1} f(x)\, dx = \int_1^\infty f.\end{aligned}$$

Hence if $\int_1^\infty f$ is finite, then $\sum_{k=1}^\infty f(kx)$ is finite for almost all x greater than 1. If we know that $\int_\delta^\infty f$ is finite for every positive δ, then we get by rescaling that $\sum_{k=1}^\infty f(kx)$ is finite for almost all x greater than 0.

(iii) SERIES OF ORTHOGONAL FUNCTIONS. The space $L^2(a,b)$ of real-valued functions with integrable square on an interval $[a,b]$ is a metric space, the distance between functions f and g being $\|f-g\|_2$, the norm of their difference. We can introduce a notion of inner product (or dot product, or scalar product) in this space by defining $\langle f,g\rangle$ to be $\int_a^b f\cdot g$; the integral is finite by Schwarz's inequality (page 185). Notice that $\|f\|_2^2 = \langle f,f\rangle$. We say that two functions are *orthogonal* if their inner product is zero. A sequence of functions $\{\varphi_n\}$ is orthogonal if $\langle\varphi_n,\varphi_m\rangle = 0$ when $m\neq n$; and *orthonormal* if also $\|\varphi_n\|_2 = 1$ for each n. For example, the set of functions $\{\sqrt{2/\pi}\,\sin(nx)\}_{n=1}^\infty$ is orthonormal on the interval $[0,\pi]$.

Exercise 29.6. Verify the orthonormality of this sequence.

The formal expansion of a function f as a series in the orthonormal functions φ_n is $\sum_n c_n\varphi_n$, where the coefficients c_n are given by $c_n = \langle f,\varphi_n\rangle$. When the orthonormal functions are sines and cosines, this is usually called a *Fourier series*.

If we calculate the square of the distance between f and a partial sum of its formal expansion, we obtain

$$\begin{aligned}\left\|f-\sum_{n=1}^N c_n\varphi_n\right\|_2^2 &= \|f\|_2^2 - 2\left\langle f,\sum_{n=1}^N c_n\varphi_n\right\rangle + \left\|\sum_{n=1}^N c_n\varphi_n\right\|_2^2\\ &= \|f\|_2^2 - 2\sum_{n=1}^N c_n^2 + \sum_{m,n=1}^N c_nc_m\langle\varphi_n,\varphi_m\rangle\end{aligned}$$

$$= \|f\|_2^2 - \sum_{n=1}^{N} c_n^2.$$

The left-hand side is nonnegative for every value of N, so we must have

$$\sum_{n=1}^{\infty} c_n^2 \le \int f^2.$$

This is *Bessel's inequality.* It suggests the question of whether it has a converse, that is, whether numbers c_n such that $\sum_n c_n^2$ converges are necessarily the coefficients in the expansion of some function in a series of whatever set of orthonormal functions is being considered.

If integration is taken to be Riemann integration, then the answer is "no." It is one of the successes of Lebesgue integration that it does let us find a function of integrable square whose coefficients are a given sequence $\{c_n\}$, provided that $\sum_n c_n^2$ converges. This is the *Riesz-Fischer theorem* (proved independently by F. Riesz and E. Fischer).[2]

The most natural way to construct a function with prescribed coefficients c_n is to write the series $\sum_n c_n \varphi_n(x)$. Such a series does not necessarily converge for all x, but we shall see that it does converge in the L^2 distance. In other words, the partial sums of the series form a Cauchy sequence in the L^2 metric: if we let s_n denote $\sum_{k=1}^{n} c_k \varphi_k$, then $\|s_m - s_n\|_2 \to 0$ as m and $n \to \infty$. In fact, the orthonormality of the φ_n implies that if $m > n$, then $\|s_m - s_n\|_2^2 = \sum_{k=n+1}^{m} c_k^2$, and this sum goes to 0 as m and $n \to \infty$ because $\sum_{k=1}^{\infty} c_k^2$ converges. We might hope that if a series is Cauchy in L^2, then there will be an L^2 function to which it converges; in other words, L^2 is complete as a metric space. We would expect that the L^2 limit of the partial sums s_n will be a function that has the given coefficients c_n.

We are going to show that this is actually what happens. To see that L^2 is a complete space, consider any sequence $\{s_n\}$ that is a Cauchy sequence in the L^2 metric. Since $\|s_m - s_n\|_2 \to 0$, we can find an increasing sequence of numbers n_k for which $n_k \to \infty$ and $\|s_m - s_{n_k}\|_2 < 2^{-k}$ when $m \geq n_k$, and in particular $\|s_{n_{k+1}} - s_{n_k}\|_2 < 2^{-k}$. If we apply Schwarz's inequality to the function $|s_{n_{k+1}} - s_{n_k}|$ and the constant function with value 1, we obtain

$$\begin{aligned}\int_a^b |s_{n_{k+1}} - s_{n_k}| &\leq \left\{\int_a^b (s_{n_{k+1}} - s_{n_k})^2 \int_a^b 1^2\right\}^{1/2} \\ &\leq 2^{-k}(b-a)^{1/2}.\end{aligned}$$

Consequently the sum $\sum_{k=1}^{\infty} \int_a^b |s_{n_{k+1}} - s_{n_k}|$ is finite, and so the telescoping sum $\sum_{k=1}^{\infty}(s_{n_{k+1}}(x) - s_{n_k}(x))$ converges almost everywhere (page 214). This is equivalent to saying that the sequence $\{s_{n_k}(x)\}$ converges almost everywhere to some $f(x)$. This f is the function we are looking for. We still have to show that f is the limit, in L^2, of the original sequence $\{s_n\}$.

By Fatou's lemma,

$$\|s_{n_k} - f\|_2 \leq \liminf_{j\to\infty} \|s_{n_k} - s_{n_j}\|_2 \leq 2^{-k}.$$

Therefore f is in L^2, since s_{n_k} is, and the subsequence $\{s_{n_k}\} \to f$ in L^2 since $\lim_{k\to\infty} \|s_{n_k} - f\|_2 = 0$. Now Minkowski's inequality (page 184) gives us $\|f - s_n\|_2 \leq \|f - s_{n_k}\|_2 + \|s_{n_k} - s_n\|_2$, so $s_n \to f$ in L^2. Thus L^2 is a complete metric space.

Returning to the situation that s_n is a partial sum of the series $\sum_k c_k\varphi_k$, write $\langle f, \varphi_m\rangle = \langle f - s_n, \varphi_m\rangle + \langle s_n, \varphi_m\rangle$. By the Schwarz inequality, the absolute value of the first term is no greater than $\|f - s_n\|_2\|\varphi_m\|_2$, which goes to 0 as $n \to \infty$. By the orthonormality of the φ_k, the second term equals 0 if $n < m$, and c_m if $n \geq m$. Letting $n \to \infty$, we see that $\langle f, \varphi_m \rangle = c_m$, so f does have the required coefficients.

NOTES

[1]The proof follows the proof of Lemma 1 in Richard R. Goldberg and Richard S. Varga, Moebius inversion of Fourier transforms, *Duke Mathematical Journal* 23 (1956), 553–559.

[2]F. Riesz, Sur les systèmes orthogonaux de fonctions, *Comptes Rendus Acad. Sci. Paris* 144 (1907), 615–619, 734–736; E. Fischer, Sur la convergence en moyen, *ibid.*, 1022–1024, 1148–1151. The essential element is the completeness of the space L^2. Some people refer to the fact that the spaces L^p are complete as the Riesz-Fischer theorem.

30. Stieltjes integrals. We are now going to look at another generalization of Riemann integrals, namely Stieltjes integrals, or more precisely Riemann-Stieltjes integrals. (There are also Lebesgue-Stieltjes integrals, which we will not discuss here.) One reason for introducing Stieltjes integrals is that we often encounter pairs of similar formulas, one with sums and the other with integrals. Here are some examples from mathematics and mathematical physics:

1. (Mechanics) $\sum_n m_n x_n$ and $\int m(t)x(t)\,dt$, where m_n and $m(t)$ are masses or a distribution of masses on a line; x_n and $x(t)$ are the coordinates of the masses, either discretely or continuously distributed; in either case we speak of the sum or integral as the *moment* of the physical system.

2. (Probability) If numbers u_n or $u(x)$ are outcomes of an experiment, occurring with probabilities p_n or $p(x)$, then $\sum_n u_n p_n$ or $\int u(x)p(x)\,dx$ is the *expectation* of the event.

3. (Analysis) Power series $\sum_n a_n t^n$ and Laplace transforms $\int a(t)e^{-xt}\,dt$. Here the analogy between series and

integral is more obvious if we replace t by e^{-u} in the series and write the sum as $\sum_n a_n e^{-nu}$.

4. (Fourier analysis) Sums $\sum_n a_n e^{inx}$ and integrals $\int a(t)e^{ixt}\,dt$.

5. (Quantum mechanics) If you have studied quantum mechanics, you may have seen the notation $\mathop{\mathsf{S}}_{\mu} w_{k\mu} v_{\mu}(r)$, where S stands for either $\sum$ or $\int$, depending on what is wanted in a particular problem.

Whether the notations for sums and for integrals are quite similar or quite different depends on the history of the subjects that are being studied—the older the problems, the more likely they are to have very different notations and terminology. The most modern example (5) does not yet have a standard notation; the one quoted was introduced by L. Schiff in the 1940's.[1]

Of course, mathematical notation does change, but it changes very slowly. A convenient common notation for pairs like those quoted here has been available for about a century, but is still little used except by specialists, and has made little or no headway in undergraduate instruction. It seems possible that the Stieltjes integral might have been more widely used if the potential users had known how to pronounce "Stieltjes." (The name is Dutch; a fair approximation to the pronunciation is to treat the j as a y.)

A Stieltjes integral is written like this:

$$\int_a^b f(x)\,dg(x)$$

(read "the integral of f with respect to g"). To begin with, we assume that a and b are finite; infinite limits of integration can be dealt with in the same way as for Riemann integrals. A Stieltjes integral is defined by replacing the

Riemann sums

$$\sum_{k=1}^{n} f(y_k)[x_{k+1} - x_k]$$

by the sums

$$\sum_{k=1}^{n} f(y_k)[g(x_{k+1}) - g(x_k)],$$

where $x_k \leq y_k \leq x_{k+1}$, and finding the limit of these sums as the maximum difference $x_{k+1} - x_k$ approaches 0.

Exercise 30.1. If f is continuous and g has a continuous derivative, then the Stieltjes integral $\int_a^b f(x)\,dg(x)$ reduces to the Riemann integral $\int_a^b f(x)g'(x)\,dx$.

If f and g are too general, the sums will not have a limit. They will have a limit, for instance, if f is continuous and g is monotonic and bounded; or, more generally, if f is continuous and g is the difference of two bounded monotonic functions—that is, g has bounded variation.[2]

The Stieltjes integral $\int f(x)\,dg(x)$ does not exist if f and g have a common point of discontinuity.[3] For example, suppose that $f(x) = g(x) = 0$ when $x < 0$, $f(x) = g(x) = 1$ when $x > 0$. If we partition $[-1, 1]$ so that 0 is not a division point, then the approximating sum $\sum_{k=1}^{n} f(y_k)[g(x_{k+1}) - g(x_k)]$ reduces to one term, the one for which $x_k < 0 < x_{k+1}$. This term is either 0 or 1, depending on where y_k is, so the approximating sums will not approach a limit when we divide $[-1, 1]$ into thinner and thinner pieces.

A useful special case of the Stieltjes integral arises when g is a step function, that is, when the interval (a, b) can be decomposed into a finite number of intervals, on each of which g is constant. Steps of zero length (isolated points) are not allowed. For a step function, the Stieltjes integral reduces to a sum $\sum_k f(c_k)j_k$, where j_k is the size of the jump of g at c_k, that is, $j_k = g(c_k{}^+) - g(c_k{}^-)$.

We should be careful in interpreting the symbol $\int_a^b f\,dg$ if there happens to be a jump at a or b. If, for example, there is a jump at $x = a$, we write either $\int_{a^-}^b f(x)\,dg(x)$ or $\int_{a^+}^b f(x)\,dg(x)$, where $\int_{a^-}^b$ means the limit of $\int_{a-\epsilon}^b$ as $\epsilon \to 0^+$, and $\int_{a^+}^b$ is defined similarly.

Exercise 30.2. Suppose f is a continuous function on $(0, \infty)$. Show that $\int_{1^-}^{n^-} f(x)\,d[x] = \sum_{k=1}^{n-1} f(k)$, where, as usual, $[x]$ denotes the greatest integer not exceeding x.

The Stieltjes integral has the very convenient property that if $\int f\,dg$ exists, then so does $\int g\,df$. The two integrals are connected by the formula for integration by parts:

$$\int_a^b f(x)\,dg(x) = f(b)g(b) - f(a)g(a) - \int_a^b g(x)\,df(x).$$

(This is easy to prove by rearranging the sums that define one of the integrals.) It is sometimes written in the condensed form $\int f\,dg = fg - \int g\,df$, where it is understood that the same limits are to be supplied on all three terms.

Exercise 30.3. Omitting the limits of integration can lead to an apparent paradox. For example, write $\int dx = \int e^{-x}e^x\,dx = \int f(x)\,dg(x)$ with $g(x) = e^x$. Then the condensed formula says $\int dx = e^{-x}e^x - \int e^x\,d(e^{-x})$, or $x = 1 + \int e^x e^{-x}\,dx$, so $x = 1 + x$. Explain what happened.

It follows from the formula for integration by parts that not only does $\int f\,dg$ exist when f is continuous and g is of bounded variation, but it also exists if f is a function of bounded variation and g is continuous. When f and g have continuous derivatives, the formula reduces to the usual formula for integration by parts:

$$\int_a^b f(x)g'(x)\,dx = f(b)g(b) - f(a)g(a) - \int_a^b g(x)f'(x)dx.$$

One of the applications of Stieltjes integrals is to provide a mathematically correct definition of the physicists' δ-function. The function $\delta(x)$ is supposed to have the properties of being zero except at $x = 0$, but satisfying the equation $\int f(x)\delta(x)\,dx = f(0)$, where the integration is over any interval for which 0 is an interior point. There is, of course, no such function. However, if we introduce the function $H(x)$ that has the values 0 for $x \leq 0$ and 1 for $x > 0$, then $\int_a^b f(x)\,dH(x) = f(0)$ when f is a continuous function. This means that we could interpret an integral containing $\delta(x)$ as a Stieltjes integral with respect to H.

This definition does, however, have the drawback that there is no good way to define $\delta'(x)$ by Stieltjes integration. A more satisfactory way to define $\delta(x)$, and even its derivatives, can be given by using the theory of *distributions* (also called "generalized functions"). This subject is, however, beyond the scope of this book.[4]

NOTES

[1] See §22 of Leonard I. Schiff, *Quantum Mechanics*, McGraw-Hill, New York, 1949.

[2] A proof of the existence theorem can be found in most textbooks that discuss Stieltjes integrals, for instance David V. Widder, *Advanced Calculus*, second edition, Dover, New York, 1989, chap. 5, §7; or Tom M. Apostol, *Mathematical Analysis*, second edition, Addison-Wesley, Reading, MA, 1974, section 7.16.

[3]It is possible to redefine the Stieltjes integral to take account of discontinuities explicitly, in which case the integral will be defined even when f and g have common discontinuities. See Kenneth A. Ross, Another approach to Riemann-Stieltjes integrals, *American Mathematical Monthly* 87 (1980), 660–662; Kenneth A. Ross, *Elementary Analysis: The Theory of Calculus*, Springer, New York, 1980, §35; H. J. ter Horst, Riemann-Stieltjes and Lebesgue-Stieltjes integrability, *American Mathematical Monthly* 91 (1984), 551–559.

[4]A mathematically rigorous theory of distributions was developed in the 1940's by the French mathematician Laurent Schwartz, who should not be confused with the German mathematician Hermann Amandus Schwarz (1843–1921) of Schwarz's inequality. For an introduction to distribution theory, see, for example, Bent E. Petersen, Weak derivatives and integration by parts, *American Mathematical Monthly* 85 (1978), 190–191; and chap. 9 of Gerald B. Folland, *Fourier Analysis and Its Applications*, Wadsworth & Brooks/Cole, Pacific Grove, Calif., 1992.

31. Applications of the Stieltjes integral.

(i) APPROXIMATING AN INTEGRAL. In §21 we introduced the mean-value theorem for derivatives, which can be thought of as a way of approximating a difference quotient of a function f in terms of the derivative f'. This theorem says that if f' exists at every point of $[a, b]$, then

$$(b-a)\inf_t f'(t) \leq f(b) - f(a) \leq (b-a)\sup_t f'(t).$$

If we write F instead of f', and if F is integrable, we have

$$(b-a)\inf_t F(t) \leq \int_a^b F(x)\,dx \leq (b-a)\sup_t F(t),$$

which is one version of the *mean-value theorem for integrals.*

If g is a positive function, it is easy to derive the more general approximation

$$\{\inf_t F(t)\} \int_a^b g(x)\,dx \le \int_a^b F(x)g(x)\,dx$$
$$\le \{\sup_t F(t)\} \int_a^b g(x)\,dx.$$

Exercise 31.1. Derive the preceding formula.

If F is continuous, so that it takes all values between its maximum and minimum, there is a more compact version,

$$\int_a^b F(x)g(x)\,dx = F(c) \int_a^b g(x)\,dx,$$

where c is some number between a and b. This is the traditional version, which seems to be preferred because equations are thought to be nicer than inequalities. For Stieltjes integrals, there is a still more general version: if G is a nondecreasing function, then

$$\{\inf_t F(t)\}\{G(b) - G(a)\} \le \int_a^b F(x)\,dG(x)$$
$$\le \{\sup_t F(t)\}\{G(b) - G(a)\}.$$

We know two simple sufficient conditions for the existence of $\int F\,dG$: that F is continuous and G is a function of bounded variation; or that F is a function of bounded variation and G is continuous. The preceding inequality depends on the first condition; now let us look for an inequality that depends on the second condition. Let us suppose that F is a positive decreasing function and G is continuous. Then

$$F(a)\inf_t\{G(t)-G(a)\} \le \int_a^b F(x)\,dG(x) \\ \le F(a)\sup_t\{G(t)-G(a)\}.$$

Notice that the same factor $F(a)$ appears on both sides. (There is a similar inequality when F increases.) The continuity of G is needed only to ensure the existence of $\int F\,dG$. This is one version of what is known as the *second mean-value theorem*, or *Bonnet's theorem*; it has many applications, for example in the theory of Fourier series.

The proof is less straightforward than the proofs of the more familiar versions with which we started. We begin with

$$\int_a^b F(x)\,dG(x) = \int_a^b F(x)\,d[G(x)-G(a)]$$

and integrate by parts:

$$\int_a^b F(x)\,dG(x) \\ = F(b)[G(b)-G(a)] - \int_a^b [G(x)-G(a)]\,dF(x).$$

It is convenient to rewrite the right-hand side in the form

$$F(b)[G(b)-G(a)] + \int_a^b [G(x)-G(a)]\,d[-F(x)].$$

The first term does not exceed $F(b)\sup_t[G(t)-G(a)]$, and the second term does not exceed

$$\sup_t [G(t) - G(a)] \int_a^b d[-F(x)] = \sup_t [G(t) - G(a)][F(a) - F(b)].$$

The upper bound $F(a) \sup_t [G(t) - G(a)]$ for $\int_a^b F(x)\, dG(x)$ follows by combining the preceding two estimates. The lower bound on $\int_a^b F(x)\, dG(x)$ is derived in a similar way.

(ii) STIELTJES INTEGRALS APPLIED TO SERIES. Stieltjes integrals provide a convenient technique for manipulating series of numbers without getting mixed up in details about indices. Consider two sequences $\{a_j\}$ and $\{b_j\}$; let us denote the sums $\sum_{j=1}^k b_j$ by B_k and adopt the convention that $B_0 = 0$. The formula

$$\sum_{k=M}^{N} a_k b_k = a_N B_N - a_M B_{M-1} - \sum_{k=M}^{N-1} B_k (a_{k+1} - a_k)$$

is known as *summation by parts*, since it is an analogue of integration by parts for Riemann integrals. It can be obtained by rearranging the series, but it is not easy to remember which signs and indices go where. The derivation of the formula is simpler if we write the sums as Stieltjes integrals.

Let us think of $\sum_k a_k b_k$ as an integral $\int a(t)\, dB(t)$, where a and B are step functions. Let $B(t)$ jump from B_{k-1} to B_k when $t = k$. Then the function a must not have jumps at the same points; instead, we suppose that $a(t)$ has a jump of amount $a_{k+1} - a_k$ when $t = k + 1/2$.

Then we have

$$\sum_{k=M}^{N} a_k b_k = \int_{M^-}^{N^+} a(t)\, dB(t),$$

and the formula for Stieltjes integration by parts produces

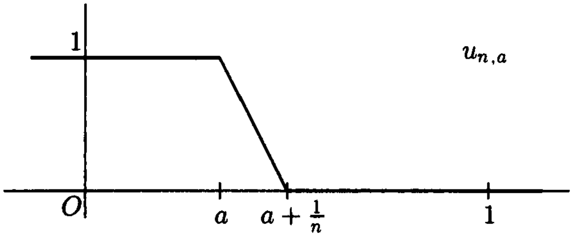
$$\sum_{k=M}^{N} a_k b_k = a(t)B(t)\Big|_{M^-}^{N^+} - \int_{M^-}^{N^+} B(t)\, da(t).$$

This, written out in full, is precisely the formula for summation by parts. Here the rearrangement of terms was taken care of in the derivation of the formula for integration by parts, and does not have to be repeated.

Exercise 31.2. Prove *Abel's test for convergence*: if $a_n \downarrow 0$ and $\sum_n b_n$ has bounded partial sums, then the series $\sum_n a_n b_n$ converges. (The "alternating series test" is the special case when $b_n = (-1)^n$.)

Exercise 31.3. Prove that if $f(x) = \sum_n a_n x^n$ for $|x| < R$, and $\sum_n a_n R^n$ has bounded partial sums, then f is bounded on the interval $(0, R)$.

(iii) RIESZ REPRESENTATION THEOREM. Let us consider real-valued functions whose domain is the space C of continuous functions on the interval $[0,1]$ (page 24). Examples are the function P such that $P(f) = f(1/2)$ for every element f of C, and the function I such that $I(f) = \int_0^1 f$. Functions like these, whose domain is itself a space of functions, are often called *functionals*. We will restrict attention to functionals L that are linear, so that $L(f+g) = L(f) + L(g)$ for all elements f and g of C, and $L(cf) = cL(f)$ for all constants c and all elements f of C. Moreover, we will only consider functionals L that are continuous, so that if $f_n \to f$ in the metric of C (which means that f_n converges uniformly to f), then $|L(f_n) - L(f)| \to 0$.



Exercise 31.4. The functionals P and I just defined are linear and continuous.

Exercise 31.5. A linear functional L is continuous if and only if there exists a constant M such that $|L(f)| \leq M \sup\{|f(x)| : 0 \leq x \leq 1\}$ for all elements f of C; one says that the functional L is "bounded."

One of the early successes of the Stieltjes integral was a theorem (proved by F. Riesz in 1909) that is now known as the Riesz representation theorem.[1] It says that every continuous linear functional L on the space C can be realized in the form $L(f) = \int_0^1 f\,dg$, where g is a suitable function of bounded variation.[2] For instance, the point evaluation functional P defined above can be realized in the form $P(f) = \int_0^1 f\,dH$, where H is a function that equals 0 on the interval $[0, 1/2]$ and 1 on the interval $(1/2, 1]$.

To define g, we introduce an auxiliary continuous function u that is equal to 1 when $x < 0$, is equal to 0 when $x > 1$, and decreases linearly when $0 \leq x \leq 1$. When n is a positive integer and $a \in [0,1]$, we write $u_{n,a}$ for the function defined by $u_{n,a}(x) = u(n(x-a))$. Then $u_{n,a}$ is equal to 1 when $x < a$, is equal to 0 when $x > a + n^{-1}$, and decreases linearly on the interval $[a, a+n^{-1}]$. If a is fixed, then the sequence $\{u_{n,a}(x)\}$ is nonincreasing for each x. We define a function g on $[0,1]$ via $g(0) = 0$ and $g(a) = \lim_{n\to\infty} L(u_{n,a})$ when $a > 0$.

Let us verify that this limit exists. It is equivalent to show that the telescoping sum

$$\sum_{k=1}^{n-1} \big(L(u_{k+1,a}) - L(u_{k,a})\big)$$

has a limit as $n \to \infty$, for this sum equals $L(u_{n,a}) - L(u_{1,a})$. Since an absolutely convergent infinite series is convergent, it will suffice to show that the sums

$$\sum_{k=1}^{n-1} \big|L(u_{k+1,a}) - L(u_{k,a})\big|$$

stay bounded as n increases. Because L is a linear functional, such a sum equals

$$L\left(\sum_{k=1}^{n-1} \pm(u_{k+1,a} - u_{k,a})\right)$$

for an appropriate choice of plus and minus signs. Now L is a continuous functional, so by Exercise 31.5 this expression is at most

$$M \cdot \sup_{0\le x\le 1} \left|\sum_{k=1}^{n-1} \pm\{u_{k+1,a}(x) - u_{k,a}(x)\}\right|.$$

By the triangle inequality, this is no larger than

$$\begin{aligned} M \cdot \sup \sum_{k=1}^{n-1} (u_{k,a} - u_{k+1,a}) &= M \cdot \sup(u_{1,a} - u_{n,a}) \\ &\le M. \end{aligned}$$

Accordingly, the limit defining g does exist.

Now we would like to see that $L(f) = \int f\,dg$ for every continuous function f. Consider a sum of the form $\sum_{k=1}^{N} f(y_k)\,(g(x_{k+1}) - g(x_k))$, where $x_1, x_2, \ldots, x_{N+1}$ is a partition of the interval $[0,1]$ and $x_k \le y_k \le x_{k+1}$. We want to show that such a sum is close to $L(f)$ when the maximum of the differences $x_{k+1} - x_k$ is small. Since f is automatically uniformly continuous (page 127), we may assume that this maximum difference is so small that $f(x)$ is close to $f(y_k)$ when $x_{k-1} \le x \le x_{k+1}$. By the definition of g, such a sum is close to the sum

$$f(y_1)L(u_{n,x_2}) + \sum_{k=2}^{N} f(y_k)\left(L(u_{n,x_{k+1}}) - L(u_{n,x_k})\right)$$

when n is large. By the linearity property of L, the latter sum is $L(f_n)$, where f_n denotes the continuous function $f(y_1)u_{n,x_2} + \sum_{k=2}^{N} f(y_k)\left(u_{n,x_{k+1}} - u_{n,x_k}\right)$. By construction, the function f_n is uniformly close to f.

Exercise 31.6. Verify the preceding statement.

Since the functional L is continuous, $L(f_n)$ is close to $L(f)$. Thus, the sums approximating $\int f\,dg$ are close to $L(f)$ when the partition is sufficiently fine. This means that $\int f\,dg$ exists and equals $L(f)$.

Exercise 31.7. The function g defined above is a function of bounded variation.

NOTES

[1] The name "Riesz representation theorem" is also used for theorems characterizing continuous linear functionals on the L^p spaces.

[2] If you look in a modern book, you will likely find the theorem stated as $L(f) = \int f\,d\mu$, where $d\mu$ is a "finite signed Borel (or Baire) measure."

32. Partial sums of infinite series. When a series is convergent, we can only rarely find its sum exactly. Usually we have to be satisfied with an approximation to the sum, and an estimate of the number of terms we would need in order to obtain a specified accuracy for it. On the other hand, if a series diverges, it may be interesting to find how rapidly its partial sums increase. Textbooks sometimes give students misleading ideas about series by devoting space to tests for convergence or divergence that may be of little practical value. In fact, if a series converges rapidly, almost any test will succeed; whereas a series that converges very slowly may be misleading.

Consider, for example, the series $\sum_{n=1}^{\infty} 1/n^2$. To get its sum with a remainder of less than 0.005 (roughly speaking, 2 decimal place accuracy) requires adding up 200 terms, and each additional digit requires 10 times as many terms as its predecessor. Clearly, even a fast computer will use up a lot of time before obtaining very high accuracy. It happens to be known that the sum is $\pi^2/6$, so that we could get very high accuracy now that π is known to several times more than 10^9 decimal places.[1] There are few other convergent series that are known to anywhere near such accuracy, unless they have been summed exactly.

The computer does let us sum many convergent series to considerable accuracy, and lets us find how fast the partial sums of divergent series increase. However, there are limits to what the computer can do. For example, the series $\sum_{n=3}^{\infty} 1/\{n(\log n)(\log\log n)\}$ diverges, but very slowly: although 56 terms will give a sum greater than 5, to get a sum of 6 would take more than 10^{19} terms.[2]

An example of a convergent series that has attracted a considerable amount of attention is obtained from the harmonic series $\sum_{n=1}^{\infty} 1/n$ by deleting all the terms whose denominators have one or more zeros in their decimal representations.[3] It is easy to show that the sum is less than 90; the sum of the first 9 terms (the ones with 1-digit denominators) is about 2.8; after 819 terms (the ones with denominators of 3 or fewer digits) the sum is about 6.7; after 66,429 terms (the ones with denominators of five or fewer digits) the sum is about 9.8; after 10^{18} terms the remainder still exceeds 1; and the sum is actually close to 23.

The number of terms required to compute the sum of a convergent series within a specified error, or the number of terms of a divergent series required to produce a sum exceeding a specified bound, often cannot be found by straightforward addition. However, there are ways of finding such numbers quite accurately if the law of formation of the terms is sufficiently regular. These methods are conveniently derived by using Stieltjes integration.

We are going to consider sums $\sum_{k=1}^{n} f(k)$, where f is continuous and has sufficiently many derivatives. We begin by writing the sum as a Stieltjes integral,

$$\sum_{k=1}^{n-1} f(k) = \int_{1^-}^{n^-} f(t)\, d[t],$$

where, as usual, $[t]$ denotes the largest integer that does not exceed t. We subtract $\int_1^n f(t)\, dt$ from both sides to get

$$\sum_{k=1}^{n-1} f(k) - \int_1^n f(t)\, dt = \int_{1^-}^{n^-} f(t)\, d\{[t] + \tfrac{1}{2} - t\}.$$

The point of the term $\frac{1}{2}$ that we have introduced (which does not change the value of the Stieltjes integral) is that the function $[t] + \frac{1}{2} - t$ is periodic with period 1 and is *antisymmetric* about the midpoint of each interval $[n, n+1]$. This function will occur so often that we give its negative a name of its own, $P_1(t) = -([t] + \frac{1}{2} - t)$.

We now have the difference between a sum and an integral expressed as a single integral, which we integrate by parts:

$$\sum_{k=1}^{n-1} f(k) - \int_1^n f(t)\,dt$$
$$= f(1)P_1(1^-) - f(n)P_1(n^-) + \int_1^n P_1(t)\,df(t).$$

(In places where there is no ambiguity, we can write n and 1 instead of n^- and 1^-.) Since P_1 has period 1, we have $P_1(n^-) = P_1(1^-) = \frac{1}{2}$. Consequently

$$\sum_{k=1}^{n-1} f(k) - \int_1^n f(t)\,dt = \frac{f(1) - f(n)}{2} + \int_1^n P_1(t)f'(t)\,dt.$$

We now add $f(n)$ to both sides to get

$$\sum_{k=1}^{n} f(k) - \int_1^n f(t)\,dt = \frac{f(1) + f(n)}{2} + \int_1^n P_1(t)f'(t)\,dt.$$

This is the *Euler-Maclaurin formula* in its simplest version.

As a side remark, we notice that this formula contains the integral test for convergence of series (compare §29(ii)). Indeed, we can show that *if* $\int_1^\infty |f'(t)|\,dt < \infty$, *then the integral* $\int_1^\infty f(t)\,dt$ *and the sum* $\sum_{n=1}^\infty f(n)$ *either both converge or both diverge.* (Of course, the lower limit 1 could be replaced with some other number; it is only the behavior as t and n tend to ∞ that is at issue.) In fact, $f(n)$ has a limit as $n \to \infty$ because $f(n) - f(1) = \int_1^n f'(t)\,dt$, and $\int_1^\infty f'(t)\,dt$ converges by hypothesis, while the integral on the right-hand side of the Euler-Maclaurin formula has a limit as $n \to \infty$ because $\int_1^\infty |P_1(t)f'(t)|\,dt \leq$

$\int_1^\infty |f'(t)|\,dt < \infty$. Consequently, if either of the two terms on the left-hand side of the Euler-Maclaurin formula has a limit as $n \to \infty$, so must the other term.

Exercise 32.1. The series $\sum_{n=1}^\infty n^{-1}\sin(n^{1/2})$ converges.

Exercise 32.2. If $f(t) \downarrow 0$ as $t \to \infty$, then $\int_1^\infty |f'(t)|\,dt$ converges, and the left-hand side of the Euler-Maclaurin formula approaches a limit as $n \to \infty$.

When $f(t) = 1/t$ in Exercise 32.2, the limit of the left-hand side of the Euler-Maclaurin formula is known as *Euler's constant* γ, that is,

$$\gamma = \lim_{n\to\infty}\left(-\log n + \sum_{k=1}^{n}\frac{1}{k}\right).$$

It is a long-standing question whether or not γ is rational. Since γ has been computed to thousands of decimal places, it is known that if γ is a rational number P/Q in lowest terms, then the denominator Q must have more than ten thousand decimal digits.

Exercise 32.3. When n is a positive integer, which of the quantities $\sum_{k=1}^n k^{-1}$ and $\int_1^n t^{-1}\,dt$ is bigger? When n increases, what happens to the difference of these quantities?

Exercise 32.4. By taking $f(t) = \log t$ in the Euler-Maclaurin formula, deduce that $\lim_{n\to\infty}(n!)^{1/n}/n = 1/e$.

We can improve the Euler-Maclaurin formula by repeated integration by parts. Since P_1 has period 1, it seems reasonable to use periodic integrals P_n. We set $Q_1(t) = t - \frac{1}{2}$ (on the interval $(0,1)$ this agrees with P_1),

$Q_2(t) = \frac{1}{2}t^2 - \frac{1}{2}t + \frac{1}{12}$, $Q_3(t) = \frac{1}{6}t^3 - \frac{1}{4}t^2 + \frac{1}{12}t$, and in general we choose the polynomial Q_k so that $Q'_k = Q_{k-1}$ and $Q_k(1) = Q_k(0)$ (however the last condition does not hold for $k = 1$). For $k > 1$, let P_k be the continuous, periodic continuation of Q_k from the interval $[0, 1]$ to the real line.

Exercise 32.5. Determine the polynomial Q_4.

For historical reasons, we define the so-called Bernoulli polynomials[4] $B_k(t) = k!\,Q_k(t)$ and the Bernoulli numbers $B_k = B_k(0) = k!\,Q_k(0)$. Several different notations for these Bernoulli numbers are in use; in the numbering system used here, we have $B_0 = 1$, $B_1 = -1/2$, $B_2 = 1/6$, $B_4 = -1/30$, $B_6 = 1/42$; $B_8 = -1/30$; $B_k = 0$ if k is odd and greater than 1. The extended formula that results from repeated integration by parts is "the" Euler-Maclaurin formula[5]

$$\begin{aligned}
&\sum_{k=1}^{n} f(k) - \int_1^n f(t)\,dt \\
&\qquad = \frac{f(1) + f(n)}{2} + \frac{B_2}{2!}\{f'(n) - f'(1)\} \\
&\qquad\quad + \cdots + \frac{B_{2m}}{(2m)!}\{f^{(2m-1)}(n) - f^{(2m-1)}(1)\} \\
&\qquad\quad + \int_1^n f^{(2m+1)}(t)P_{2m+1}(t)\,dt.
\end{aligned}$$

Let us see how this formula can be used to approximate the sum of a convergent series or the partial sums of a divergent series. We defined γ to be the limit as $n \to \infty$ of the difference $\sum_{k=1}^{n} f(k) - \int_1^n f(t)\,dt$ when $f(t) = 1/t$. We could use the same definition for any series (either convergent or divergent) for which this limit exists, although little attention has been paid to γ for any series except the harmonic series. If f is a function whose antiderivative we know, and if $f(t)$ and its derivatives tend to 0 as $t \to \infty$, we can easily compute to high accuracy both γ and $\sum_{k=1}^{n} f(k)$ (or $\sum_{k=1}^{\infty} f(k)$ if this sum converges).

To do this, we rewrite the Euler-Maclaurin formula as

$$\left(\sum_{k=1}^{n} f(k) - \int_{1}^{n} f(t)\,dt\right) - \gamma$$
$$= \tfrac{1}{2}f(n) + \frac{B_2}{2!}f'(n) + \cdots + \frac{B_{2m}}{(2m)!}f^{(2m-1)}(n) - R_m,$$

where the remainder $R_m = \int_n^\infty f^{(2m+1)}(t)P_{2m+1}(t)\,dt$. It can be shown[6] by using the periodicity and symmetry properties of P_{2m+1} and P_{2m+3} that $|R_m|$ is no larger than $|B_{2m+2}f^{(2m+1)}(n)/(2m+2)!|$ if the successive derivatives of f alternate in sign. Suppose $f^{(2m+1)}(t) \to 0$ fairly rapidly as $t \to \infty$ when m is not very large. Then we can compute γ fairly rapidly.[7] For example, taking $f(t) = 1/t$ (the harmonic series), $n = 10$ (ten terms of the series), and $m = 3$ (three Bernoulli numbers), we get

$$\left(\sum_{k=1}^{10} \frac{1}{k} - \int_{1}^{10} \frac{1}{t}\,dt\right) - \gamma$$
$$= \frac{1}{20} + \frac{B_2}{2!}\left(\frac{-1}{10^2}\right) + \frac{B_4}{4!}\left(\frac{-3!}{10^4}\right) + \frac{B_6}{6!}\left(\frac{-5!}{10^6}\right) - R_m.$$

The integral on the left-hand side evaluates to $\ln(10)$, and we have values for the Bernoulli numbers. If pressed, you could do the remaining arithmetic by hand with pencil and paper; with a simple hand-held calculator that displays nine digits after the decimal point, it works out to $\gamma \approx 0.577215664$. Since the remainder is no larger than

$|(B_8/8!)(7!/n^8)| = (1/240) \times 10^{-8} < 0.5 \times 10^{-10}$, we can be confident that at least the first eight digits are correct. (It happens that all nine digits are correct. However, the tenth digit of γ is a 9, so the value of γ *rounded* to nine places is 0.577215665.)

Once we have an approximate value for γ, we can apply the Euler-Maclaurin formula a second time, now with a much larger value of n, to approximate partial sums of a series. Consider, for example, the convergent series $\sum_{k=1}^{\infty} k^{-3}$. (The sum of this series is the value at $z = 3$ of a famous function known as the Riemann zeta function $\zeta(z)$.) Again taking $m = 3$ and $n = 10$, we can find an approximate value for the Euler constant γ of this series. It works out to be 0.702056903174, rounded to twelve decimal places. Since the remainder is at most $|(B_8/8!)(9!/2!)n^{-10}| = 0.15 \times 10^{-10}$, we can be sure that the first ten places are correct. For a convergent series, we can let $n \to \infty$, so that $\gamma = \sum_{k=1}^{\infty} k^{-3} - \int_1^{\infty} t^{-3}\,dt$. The integral evaluates to $1/2$, so we deduce that $\sum_1^{\infty} k^{-3} \approx 1.2020569031$, correct to ten decimal places. By taking $m = 3$ and $n = 10^4$, which would be a trivial computation on a computer, we can get γ and hence $\zeta(3)$ with an error no larger than 0.15×10^{-40}.

NOTES

[1] See note 1 to §3.

[2] There are tables of the number of terms required to compute some convergent series to a certain accuracy, and to make the sum of some divergent series greater than a certain number, in G. H. Hardy, *Orders of Infinity*, second edition, Cambridge University Press, 1924, pp. 68–69; and in R. P. Boas, Partial sums of infinite series, and how they grow, *American Mathematical Monthly* 84 (1977), 237–258.

[3] See, for example, A. D. Wadhwa, An interesting subseries of the harmonic series, *American Mathematical Monthly* 82 (1975), 931–933, and its references. More generally, a convergent series results from summing the reciprocals of those positive integers whose decimal representation omits the block of digits 170310 (or any specific

block of digits). For a short proof and an application to the irrationality of certain decimals, see Norbert Hegyvári, On some irrational decimal fractions, *American Mathematical Monthly* 100 (1993), 779–780.

[4]See D. H. Lehmer, A new approach to Bernoulli polynomials, *American Mathematical Monthly* 95 (1988), 905–911, for further information on the Bernoulli polynomials.

[5]There is a "second form" of the Euler-Maclaurin formula, in which the integration is carried out over an interval between half-integers instead of integers. See H. W. Gould and William Squire, Maclaurin's second formula and its generalization, *American Mathematical Monthly* 70 (1963), 44–52.

[6]R. P. Boas, paper cited in note 2.

[7]For the harmonic series, the numerical accuracy can be improved somewhat by treating $n + \frac{1}{2}$ as the basic quantity instead of n. See Duane W. DeTemple and Shun-Hwa Wang, Half integer approximations for the partial sums of the harmonic series, *Journal of Mathematical Analysis and Applications* 160 (1991), 149–156.

Answers to Exercises

1.1. The elements of E are precisely the points not in $C(E)$; so are the elements of $C(C(E))$.

1.2. (a) Every letter is either a consonant or a vowel, and all the vowels occur in "real functions."

(b) $C(E)$ consists of all the vowels.

(c) $C(F)$ consists only of consonants (in fact, not all of them).

(d) $F \cap E = \{r, l, f, n, c, t, s\}$, which contains no vowels.

1.3. Not very practical. (Repeat the preceding discussion with "bibliography" replacing "set.")

2.1. (a) All numbers greater than or equal to 1; all nonpositive numbers; 1; 0.

(b), (c), (d), (e): The same as (a).

(f) All nonnegative numbers; all nonpositive numbers; 0; 0.

(g), (h) No upper bound; all nonpositive numbers; $+\infty$; 0.

(i) All numbers greater than or equal to $5\pi/6$; all numbers less than or equal to $\pi/6$; $5\pi/6$; $\pi/6$.

2.2. Every nonempty set E that is bounded below has a greatest lower bound, denoted by $\inf E$, with the properties that every $x \in E$ satisfies $x \geq \inf E$, and if $A > \inf E$ there is at least one $x \in E$ such that $x < A$. If the least upper bound property is assumed and E is a set that is bounded below, let

F consist of all numbers x such that $-x \in E$. Since $x > M$ means $-x < -M$, the set F is bounded above and so has a least upper bound B. Then $-B$ is the greatest lower bound of E. For, if $x \in E$, then $-x \leq B$, so $x \geq -B$; if $A > -B$, then $-A < B$, and there is an $x \in E$ with $-x > -A$, so $x < A$.

2.3. $+\infty + (-\infty)$ would be $(a/0) + (b/0)$, where $a > 0$, and $b < 0$. For the rules of arithmetic to hold, this would have to equal $(a+b)/0$, which can be $+\infty$, $-\infty$, or undefined according as $a+b > 0$, $a+b < 0$, or $a+b = 0$. Then $0 \cdot (+\infty)$ is ambiguous too, since $0 \cdot (+\infty) = (+1 + (-1)) \cdot (+\infty) = +\infty + (-\infty)$. It is also impossible to attach a meaning to $(+\infty)/(+\infty)$.

2.4. If E contains the point x, then every upper bound of E is at least x, and every lower bound of E is at most x. The same statement is true with y in place of x if E also contains a point y. If $y > x$, then we have in particular $\sup E \geq y > x \geq \inf E$. Similarly if $y < x$.

3.1. Let the sets be E and F, and discard from F any elements it may have in common with E. If the reduced set (F_1, say) is finite, count F_1 and then E. If not, use the odd positive integers to label F_1 and the even positive integers to label E.

3.2. The elements of the kth set, E_k, may be denoted by $e_{k,1}, e_{k,2}, \ldots$. Associate $e_{k,n}$ with the lattice point (k, n).

3.3. Each disk contains a point with two rational coordinates that is not in any other disk; enumerate S by enumerating these points.

3.4. The linear polynomials $ax + b$ with integral coefficients are in one-to-one correspondence with the lattice points (a, b); the quadratic polynomials $ax^2 + bx + c$, with the three-dimensional lattice points (a, b, c); and so on.

3.5. If the real numbers x in (a, b) were countable, then the real numbers $x - a$ in $(0, b - a)$ would be countable, and so would the real numbers $(x - a)/(b - a)$ in $(0, 1)$.

3.6. The construction of the text applies word for word, since the number constructed contains no 3's.

3.7. The suggested procedure fails to count decimals such as $0.12345\ldots$ that have infinitely many nonzero digits (and most real numbers have this property).

3.8. If E with x_1 removed is finite, it is still finite after adjoining the one point x_1. If the process terminated at some stage, say after removing $x_1, \ldots, x_k$, then we would have a finite set left, and this set would still be finite after adjoining k points.

3.9. Let F be E with x_0 deleted. Select a countably infinite subset $\{x_1, x_2, \ldots\}$ of F. Let points of E other than $\{x_0, x_1, x_2, \ldots\}$ correspond to themselves (as points of F); let x_0 correspond to x_1, x_1 to x_2, etc.

3.10. The correspondence $x \leftrightarrow \tan x$ establishes a one-to-one correspondence between the real numbers in the interval $(-\pi/2, \pi/2)$ and the set of all real numbers. Geometrical arguments can also be used.

3.11. If "the set of all sets" is a set S, the aggregate of subsets of S is a set T that cannot be put into one-to-one correspondence with any subset of S. On the other hand, the subsets of S are sets and so elements of S, and their aggregate T is therefore in one-to-one correspondence with a subset of S, namely T itself. Contradiction.

3.12. If the functions could be put into one-to-one correspondence with the real numbers, each f could be labeled as f_x, where x is the real number that corresponds to the function f. Define a function g by $g(x) = f_x(x) + 1$. What real number could correspond to g?

3.13. Use the Schroeder-Bernstein theorem and the fact (Exercise 3.10) that there are just as many real numbers between 0 and 1 as there are real numbers. On the one hand, to each real number between 0 and 1 we can assign a sequence of real numbers, namely, the sequence of digits of the decimal that defines the real number. On the other hand, if we have a sequence of real numbers between 0 and 1, then we can prefix the digit 5 to each decimal expansion, write out the decimal expansions one below another, and traverse the resulting array along its diagonals (from upper right to lower left) to obtain the decimal expansion of a real number. Different sequences of real numbers generate different real numbers in this way, since the decimal we obtain cannot terminate.

4.1. (a) No, because $d(x, y)$ can be 0 without having $x = y$. In fact, $d((0, a), (0, b)) = 0$.

(b) No, because $d(x, y)$ is not always equal to $d(y, x)$; indeed, $d(x, y)$ is not always well-defined.

(c) No; $d(x, y)$ is not symmetric.

(d) Yes. Properties (1) and (2) are obvious. Write $x = e^u$, $y = e^v$, $z = e^w$, and write the triangle inequality in terms of u, v, w.

(e) No: $d(10^{-11}, 0) = 0$.

4.2. Properties (1) and (2) of both new metrics are obvious. For property (3), let the points denoted originally by x, y, z be (x_1, x_2), (y_1, y_2), (z_1, z_2). By the triangle inequality for $\mathbf{R}_1$,

$$|x_1 - z_1| \le |x_1 - y_1| + |y_1 - z_1|,$$
$$|x_2 - z_2| \le |x_2 - y_2| + |y_2 - z_2|.$$

Adding these two inequalities gives the triangle inequality for the first new metric. On the other hand, each line above is no larger than $\max(|x_1 - y_1|, |x_2 - y_2|) + \max(|y_1 - z_1|, |y_2 - z_2|)$, so the triangle inequality holds for the second new metric.

The triangle with vertices at $(0, 0)$, $(1, 0)$, and $(1, 1)$ has its hypotenuse equal to the sum of the other sides according to the first metric; for the second metric, consider the triangle with vertices at $(0, 0)$, $(1, 1)$, and $(2, 0)$. For the first metric, the points at unit distance from the origin are the points on the boundary of the square (or diamond) with vertices at $(0, 1)$, $(1, 0)$, $(0, -1)$, and $(-1, 0)$; for the second metric, the square with vertices at $(1, 1)$, $(1, -1)$, $(-1, -1)$, and $(-1, 1)$.

4.3. All the requirements for a metric space are still fulfilled.

4.4. $d(x, x) = 0$; $d(y, x) = d(x, y) = 1 > 0$ if $x \ne y$; if $x = z$, then $0 = d(x, z) \le d(x, y) + d(y, z)$; if $x \ne z$, then $1 = d(x, z) \le d(x, y) + d(y, z)$, since at least one of the terms on the right-hand side equals 1.

5.1. A neighborhood of the point x in C consists of all continuous functions y such that $|y(t) - x(t)| < r$ for all t in $[0, 1]$.

5.2. If $p = (m, n)$ is a point of the space, a neighborhood of p consists of all points with two integral coordinates at (ordinary) distance less than r from p. Each neighborhood contains only a finite number of points, and if $r < 1$, it contains only its center.

5.3. If N contains all points at distance less than r from x, then N contains all points at distance less than $r - d(x, y)$ from y (by the triangle inequality).

5.4. (i) Empty interior; all points are boundary points.

(ii) The set is its own interior. Its boundary consists of 1, $\frac{1}{2}$, $\frac{1}{3}$, ... , 0.

(iii) The set is its own interior. Its boundary consists of line segments of length 1 standing on the boundary points in (ii), together with two horizontal line segments of length 1, one at height 0 and the other at height 1.

(iv) Empty interior; the boundary is the whole set.

5.5. Neighborhoods of radius less than 1 contain only their centers (Exercise 5.2). Hence a sufficiently small neighborhood of any point of any nonempty set E contains no point of $C(E)$, so the boundary of E is empty.

5.6. The definition of boundary is symmetric in E and $C(E)$.

5.7. Let x be a boundary point of B, and let N be a neighborhood of x. Then N contains at least one point y of B, and by Exercise 5.3 contains a neighborhood of y. This neighborhood (and hence N) contains a point of E and a point of $C(E)$. Thus x is a boundary point of E, that is, $x \in B$. If E consists of the rational points of $\mathbf{R}_1$, its boundary is all of $\mathbf{R}_1$, so that the boundary of the boundary of E is empty.

5.8. (a) The boundary of N consists of the points at distance r from x.

(b) The boundary of N may be empty (Exercise 5.5), but if it has points they must be at distance r from x. For, if a point y of N is at distance less than r from x, so is a small neighborhood of y, so that y is an interior point. Similarly, a point y at distance greater than r from x is an interior point of $C(E)$.

5.9. Suppose $a < x < b$. A neighborhood of x is an interval of the form $(x - h, x + h)$, and if $h < \min(x - a, b - x)$, then this interval is inside (a, b). Hence x is an interior point.

The boundary points of $[a, b]$ are therefore the points a and b, and these are in $[a, b]$, so $[a, b]$ is closed.

5.10. The interval (a, b) is neither open nor closed in $\mathbf{R}_2$, since it has an empty interior but fails to contain the points a and b of its boundary. But $[a, b]$ is closed in $\mathbf{R}_2$, since now it contains all its boundary points.

5.11. The point 0 is not an interior point, while 1 is a boundary point that is not in the set.

5.12. By Exercise 5.3, every point of a neighborhood is an interior point of the neighborhood.

5.13. All points are interior points of the whole space, so the space is open. The boundary of the whole space is empty, and therefore contained in the space. Hence the whole space is both open and closed. The empty set contains its (empty) interior and its (empty) boundary.

5.14. The intervals $(n, n + \frac{1}{2})$ are sets of the required kind; so are unions of such sets.

5.15. Neither: its boundary is all of $\mathbf{R}_1$, and its interior is empty.

5.16. Any neighborhood is open. A neighborhood such as the set of points of the space at distance less than $\sqrt{2}$ from 0 is also closed, since its boundary is empty.

5.17. If E is any set in this space, neighborhoods of radius less than 1 of points of E belong to E, so E is open. The boundary of E is empty (Exercise 5.5), and therefore contained in E; hence E is closed.

5.18. If $x \subset G \subset E$ where G is open, then x is an interior point of G and therefore an interior point of E, since a neighborhood of x consisting exclusively of points of G also consists exclusively of points of E.

5.19. If E is open and $x \in E$, then some neighborhood of x contains only points of E, so not all neighborhoods of x can contain both points of E and points of $C(E)$. Hence no point of E is a boundary point.

Conversely, suppose that E contains none of its boundary points, and let $x \in E$. Then x is not a boundary point, whence some neighborhood of x fails to contain any points of $C(E)$ and so consists exclusively of points of E. Hence every point of E is an interior point.

5.20. E is open if and only if it contains none of its boundary points, therefore if and only if all the boundary points of E belong to $C(E)$; that is, if and only if $C(E)$ contains all the boundary points of $C(E)$ (since by Exercise 5.6 the boundary of E equals the boundary of $C(E)$); that is, if and only if $C(E)$ is closed.

5.21. Reformulation of Exercise 5.20: interchange E and $C(E)$.

5.22. A limit point of E cannot be an interior point of $C(E)$, for then it would have a neighborhood that does not intersect E. So a limit point of E is either an interior point of E or a boundary point of E. Every set contains its interior points, and a closed set also contains its boundary points. Hence a closed set E contains its limit points.

Conversely, every neighborhood of a boundary point x of a set E intersects E. If x is not already an element of E, then this intersection contains some point other than x, and so x is a limit point of E. Hence if E contains all its limit points, then E also contains all its boundary points, and so is closed.

5.23. Let x be a limit point of E, and let N_1 be a neighborhood of x. By hypothesis, N_1 contains a point y_1 of E such that $y_1 \neq x$. A neighborhood N_2 of x of radius less than $d(x, y_1)$ does not contain y_1, but does contain a point y_2 of E. And so on.

5.24. Let F be the set of limit points of E, and let x be a limit point of F. Then every neighborhood of x contains points of F, that is, limit points of E, and so contains a subneighborhood which contains points of E. Thus x is itself a limit point of E, and hence $x \in F$. That is, F contains all its limit points, and so is a closed set.

If x is a limit point of the boundary B of E, then every neighborhood of x contains a point of B, and hence contains a point of E and a point of $C(E)$. Therefore x is a boundary point of E; that is, B contains all its limit points, and so is a closed set.

5.25. (a) and (c): The points of $[0, 1]$. (b) The single point 0.

5.26. (i) The whole set.

(ii) The points $1, \frac{1}{2}, \frac{1}{3}, \ldots$ on each coordinate axis, together with $(0, 0)$.

(iii) Each radius $\theta = 1, \frac{1}{2}, \frac{1}{3}, \ldots, 0$, with $0 \leq r \leq 1$, consists of limit points.

5.27. If the set E has an empty boundary, then E is both open and closed (Exercise 5.19, definition of closed).

5.28. If x is a limit point of E, and $x \notin E$, then every neighborhood of x contains points of E and a point (namely, x) of $C(E)$; hence x is a boundary point of E. Conversely, if x is a boundary point of E but not an element of E, then every neighborhood of x contains a point of E (different from x, since $x \notin E$), so x is a limit point of E.

The complement of the closure of E consists of those points of $C(E)$ that are not boundary points of E, and hence (Exercise 5.6) are not boundary points of $C(E)$. Thus the complement of the closure of E contains none of its boundary points, and so is open; so the closure of E is closed.

5.29. No. For example, the set 1, 2, 3, ... , in $\mathbf{R}_1$.

5.30. (a) $[0,1]$; (b) $\{0, 1, \frac{1}{2}, \frac{1}{3}, \ldots\}$; (c) $[0,1]$.

5.31. By Exercise 5.28, the closure of the neighborhood N in question is the union of N and its boundary. The boundary is the set of points y such that $d(x, y) = r$.

The statement is not true in metric spaces in general. Consider, for example, the space consisting of $[0, \frac{1}{2}] \cup [1, \infty]$ with the $\mathbf{R}_1$ metric. Let N be the set of points at distance less than 1 from 0. The closure of N is the set $[0, \frac{1}{2}]$, while the set of points of the space at distance less than or equal to 1 from 0 is $[0, \frac{1}{2}] \cup \{1\}$.

5.32. Suppose we know that the union of n closed sets is closed. Let $F_1, F_2, \ldots, F_{n+1}$ be $n + 1$ closed sets. Then $F_1 \cup F_2 \cup \cdots \cup F_{n+1} = (F_1 \cup F_2 \cup \cdots \cup F_n) \cup F_{n+1}$, the union of two closed sets.

5.33. The proof that the intersection of two closed sets is closed extends practically word for word.

5.34. Consider the sets in $\mathbf{R}_1$, each consisting of a single rational number.

5.35. Unions of open sets are open: Consider a point x belonging to at least one of the sets of a collection of open sets. Then a neighborhood of x belongs to one of the open sets (because it is open) and therefore to the union of all the open sets.

Finite intersections of open sets are open: It is enough to consider two open sets (use induction for more). Let G_1 and G_2 be open, $x \in G_1 \cap G_2$. Then a neighborhood of x belongs to G_1, and a neighborhood of x belongs to G_2. The intersection of these neighborhoods contains a smaller neighborhood that belongs both to G_1 and G_2, and hence to their intersection.

Infinite intersections of open sets need not be open: Consider the intervals $(-1/n, 1/n)$ in $\mathbf{R}_1$. Their intersection is not an open set, since it contains the single point 0.

5.36. If N_2 has any boundary points, they must belong to the set of points y such that $d(x, y) \leq r/2$; hence $N_2 \subset N_1$. If the space consists of the integers with the $\mathbf{R}_1$ metric, and $r = 1$, we have $N_1 = N_2$ when $x = 0$.

6.1. A neighborhood in Ω is a single point if its radius is less than 1. Hence the only nowhere dense set in Ω is the empty set.

6.2. If a set fails to be nowhere dense, then its closure fills some neighborhood. If the set is also closed, then it coincides with its closure; hence it contains a neighborhood.

6.3. $\mathbf{R}_1$, considered as a subset of $\mathbf{R}_2$, is closed, and every point of $\mathbf{R}_1$ is a limit point of $\mathbf{R}_1$, so $\mathbf{R}_1$ is perfect. But $\mathbf{R}_1$ contains no neighborhood in $\mathbf{R}_2$.

6.4. The point $0.77245\ldots$ is greater than $\frac{2}{3}$, so not removed in the first step of the construction of the Cantor set; less than $\frac{7}{9}$, so not removed in the second step; and so on. At the fifth step we find an interval that contains the point and is removed. Therefore the point is not in the Cantor set.

6.5. The points of the Cantor set, excluding endpoints, can be written as ternary "decimals" that contain no 1's and do not end in all 0's or all 2's. Supposing them enumerated as $p_1, p_2, \ldots$, we form a new number t, whose ternary digits t_n are 0 or 2 according as the nth digit of p_n is 2 or 0. Since t differs from p_n in the nth digit, it cannot occur in our alleged enumeration. This construction would fail if it happened that the nth digit of p_n were always 0 (or always 2) from some n onward. This difficulty can be avoided by renumbering the alleged enumeration before starting the construction.

6.6. Yes, because there are only countably many rational points but uncountably many points in the Cantor set. The easiest way to find an explicit irrational point is to write down a base 3 "decimal" containing only the digits 0 and 2 that is nonterminating and nonrepeating. Any such expansion represents an irrational number in the Cantor set.

6.7. The points with two rational coordinates form a countable set that is everywhere dense in $\mathbf{R}_2$.

6.8. There are at least as many polynomials as there are constant terms for polynomials.

6.9. There are countably many rational constants; countably many first-degree polynomials with (two) rational coefficients; and so on (compare Exercise 3.4).

6.10. If p_n is a polynomial of degree n, then by approximating each coefficient within $\epsilon/(n+1)$ by a rational number, we approximate the polynomial on $[0,1]$ within ϵ. Now use Exercise 6.9.

6.11. There are uncountably many sequences of 0's and 1's (halve the elements of the ternary expansions considered in Exercise 6.5).

6.12. 1.

6.13. Let f_x be the function defined by $f_x(t) = 0$ for $0 \leq t < x$ and $f_x(t) = 1$ for $x \leq t \leq 1$. There are uncountably many such functions, since there is one for each x in $[0,1]$. The distance between any two f_x's is 1. The rest of the proof is like that for the space m.

7.1. Let f be a continuous function on a closed bounded set E, and suppose that f is not bounded. Then there are a point x_1 such that $|f(x_1)| > 1$; a point x_2 such that $|f(x_2)| > |f(x_1)| + 1 > 2$; and so on; generally, $|f(x_n)| > n$. The Bolzano-Weierstrass theorem yields a limit point x of the set $\{x_1, x_2, \dots\}$, and $x \in E$ since E is closed. Since f is continuous, we should have $f(y)$ close to $f(x)$ whenever y is close to x. But this is impossible, because there are x_n's arbitrarily close to x, and $|f(x_n)| \to \infty$.

7.2. The set E lies in some rectangle with sides parallel to the coordinate axes; divide the rectangle into quarters by lines bisecting all its sides, and proceed as in $\mathbf{R}_1$.

7.3. If some point x is in at least 3 intervals, say $E_1, \dots, E_k$, select one whose right-hand endpoint is largest and one whose left-hand endpoint is smallest. Since both intervals contain x, they overlap, and therefore cover the rest of the E_j, which we can then discard. Proceed similarly with another point x that is not in the two selected E_j, if there are any such points.

7.4. Suppose E is compact, F is closed, and $F \subset E$. Let F be covered by open sets G. The complement $C(F)$ of F is open, and $\{G\}$ and $C(F)$ together cover E. Since E is compact, a finite subcollection drawn from $\{G\}$ and $C(F)$ covers E, and so covers F. Discarding $C(F)$ still leaves F covered.

7.5. As for Exercise 7.2.

7.6. (a) The area of a triangle with vertices in S is a continuous function (on $\mathbf{R}_6$) of the 6 coordinates of the vertices. Since S is closed and bounded in $\mathbf{R}_2$, the set of coordinates of vertices is a compact subset of $\mathbf{R}_6$. Hence the area attains a maximum.

(b) No. For example, the largest triangle with vertices on a circumference is the inscribed equilateral triangle.

7.7. The set $E_1 = \{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots\}$ is covered by the intervals $(\frac{1}{2} - \frac{1}{8}, \frac{1}{2} + \frac{1}{8})$, $(\frac{1}{4} - \frac{1}{16}, \frac{1}{4} + \frac{1}{16})$, ... No finite number of these can cover E_1, since for any finite number there would be a smallest positive left-hand endpoint. $E_2 = \{1, 2, 3, \dots\}$ can be covered by the intervals $(\frac{1}{2}, \frac{3}{2})$, $(\frac{3}{2}, \frac{5}{2})$, ... Any finite number of these cover only a bounded part of $\mathbf{R}_1$, and E_2 is unbounded.

7.8. For any finite number of the intervals, let y be the smallest left-hand endpoint that occurs: then the point $y/2$ is not covered. Since E is not closed, the Heine-Borel theorem is not contradicted.

7.9. The set $[0, \infty)$ is closed but not bounded.

7.10. The Heine-Borel theorem gives sufficient conditions for a covering to be reducible to a finite covering, not necessary conditions.

7.11. Same as the preceding answer.

8.1. Let $s_n = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{n}$ be the nth partial sum of the harmonic series. Then $|s_n - s_{n+1}| = 1/(n+1) \to 0$. However, if n is a power of 2, say $n = 2^k$, and m is the next power 2^{k+1}, then $|s_n - s_m|$ is a sum of 2^k terms each at least as big as $1/(2^{k+1})$, so $|s_n - s_m| > \frac{1}{2}$. Hence the sequence $\{s_n\}$ is not a Cauchy sequence. The fourteenth-century scholar Nicole Oresme used essentially this argument to show that the harmonic series diverges.

8.2. Suppose that $\mathbf{R}_1$ is complete, and let E be a nonempty set that is bounded above. Let I_1 be a bounded interval whose right-hand endpoint is an upper bound for E, and whose left-hand endpoint is not an upper bound for E. Take I_2 to be either the left half or the right half of I_1, according as the midpoint of I_1 is or is not an upper bound for E. Similarly take I_3 to be either the left half or the right half of I_2, according as the midpoint of I_2 is or is not an upper bound for E; and so on. Each of these nested intervals is half the width of the preceding one, so the right-hand endpoints, all of which are upper bounds for E, form a Cauchy sequence, and so converge to a limit. The left-hand endpoints, none of which is an upper bound for E, similarly converge, necessarily to the same limit as the right-hand endpoints. This common limit is the least upper bound of E.

8.3. The *set* of all different s_n has a least upper bound L. Then, given $\epsilon > 0$, there is some s_n such that $L - s_n < \epsilon$. Since the s_n increase, $L - s_m < \epsilon$ for all $m \geq n$. Therefore the sequence $\{s_n\}$ converges to L.

8.4. For a Cauchy sequence $\{(x_n, y_n)\}$ in $\mathbf{R}_2$, $\{x_n\}$ and $\{y_n\}$ are Cauchy sequences in $\mathbf{R}_1$.

8.5. The rational points of $\mathbf{R}_1$.

8.6. If $s_n \to L$, every s_n from some index on is in any given neighborhood of L. If the s_n are not all equal to L from some index on, then every neighborhood of L contains one of them, other than L.

8.7. If $s_n \to L$ and $s_n \in F$, where F is a closed set, either $s_n = L$ from some point on and $L \in F$; or L is a limit point of the set of different s_n, and so $L \in F$ since F is closed.

8.8. E has a supremum L. There are points x_n of E (not necessarily all different) such that $L - x_n < 1/n$. Then $\{x_n\}$ has L as its limit. So $L \in E$ by Exercise 8.7.

8.9. If E has only a finite number of different elements, one of them occurs infinitely often, and its occurrences define the subsequence. Otherwise the elements of E have a limit point; apply the subsequence principle.

8.10. Let D be the distance in question, and take points x_n in F and y_n in G such that $d(x_n, y_n) \to D$. After discarding a finite number of terms, we may assume that $d(x_n, y_n) < D + 1$ for all n. Using Exercise 8.9, take a convergent subsequence from $\{x_n\}$; by renumbering, we may as well continue to call it $\{x_n\}$. Since F is closed, there is a point x in F such that $x_n \to x$. Again discarding a finite number of terms of the sequence, we may suppose that $d(x_n, x) < 1$ for all n. Then $\{y_n\}$ is a bounded sequence, because $d(y_n, x) < D + 2$ for all n. Consequently, $\{y_n\}$ has a convergent subsequence, which we may as well continue to call $\{y_n\}$, and $y_n \to y \in G$ since G is closed. Now $d(x, y) \leq d(x, x_n) + d(x_n, y_n) + d(y_n, y)$, and taking the limit as $n \to \infty$ proves that $d(x, y) \leq D$. Since $d(x, y)$ cannot be less than D, we must have $d(x, y) = D$.

8.11. (a) Yes. (b) Not necessarily. For (b), consider the integers in $\mathbf{R}_1$, with the $\mathbf{R}_1$ metric. The neighborhood of 0 for which $d(0, x) < \frac{1}{2}$ is the single point 0, and its diameter is 0.

8.12. Let the diameters of E and of its closure be δ and Δ, respectively; obviously $\Delta \geq \delta$. Choose x_n and y_n in the closure of E such that $d(x_n, y_n) > \Delta - 1/n$. By the definition of closure, there are points x_n' and y_n' in E within distance $1/n$ of x_n and y_n. Then $d(x_n', y_n') > \Delta - 3/n$, and hence $\delta \geq \Delta$. Therefore $\delta = \Delta$.

9.1. Consider the set $\{x\}$ containing only the single point x. It is nowhere dense if its closure, which is $\{x\}$ again, contains no neighborhood, that is, if $\{x\}$ is not a neighborhood. This will happen if the set of y with $d(y, x) < r$ contains points other than x for every positive r, and this is the case if x is a limit point of the space.

9.2. A nonempty perfect set in $\mathbf{R}_1$, considered as a metric space in itself, is of second category. Since its points are all limit points, a single point of it is nowhere dense, and therefore every countable subset of it is of first category. Hence no nonempty perfect set in $\mathbf{R}_1$ can be countable.

10.1. Let r be an arbitrary positive number less than 1, and let E_n be the set of points x in the interval $[0, r]$ at which f agrees with its nth partial sum. Since both f and its partial sums are continuous functions, E_n is a closed set. By hypothesis, the complete metric space $[0, r]$ is the union of the E_n, so by Baire's theorem, some E_n must contain some open subinterval of $[0, r]$. Thus f agrees with a polynomial on this subinterval. Continue exactly as in (ii), and then let $r \to 1$.

10.2. If E is the set in question, it is nowhere dense unless its closure (which is still E) contains a neighborhood (Exercise 6.2). If E contains a neighborhood, its complement is disjoint from this neighborhood and so is not everywhere dense.

10.3. Suppose that a closed interval in $\mathbf{R}_1$ is the union of a countably infinite number of disjoint, closed, nonempty sets E_n. By Baire's theorem, one of the sets is dense in some subinterval, and, being closed, contains that subinterval. Take the largest subinterval of this kind. Then repeat the process with what remains of the original interval. We obtain a countable collection of closed intervals I_n, each belonging to one E_n, with their union everywhere dense. If I_n and I_m have a common endpoint, this endpoint belongs both to E_n and E_m, which is impossible since the E_n are disjoint. If we remove the interiors of all the intervals I_n, the remaining set H is perfect. By applying Baire's theorem to H as in example (ii), we see that the part of H in some interval J belongs to a single E_n; in particular, so do all the endpoints of intervals I_n that are in J; hence these I_n all belong to the same E_n, and H is empty.

11.1. It is enough to show that $E \cap (a, b)$ is of measure zero for each (a, b), since $\mathbf{R}_1$ can be covered by countably many intervals. Cover $E \cap (a, b)$ by intervals (a_n, b_n) of total length at most $q(b - a)$. Then cover each $E \cap (a_n, b_n)$ similarly. We now have $E \cap (a, b)$ covered by intervals of total length at most $q(b_1 - a_1) + q(b_2 - a_2) + \cdots = q[(b_1 - a_1) + (b_2 - a_2) + \cdots] \leq q^2(b - a)$. Repeat the process; $q^n \to 0$.

12.1. (b) (domain the single point 0); (c) and (e) (domain all of $\mathbf{R}_1$).

12.2. Let x be a rational number p/q. If $m > q + 1$, then $m!\,x$ is an even integer, $\cos(m!\,\pi x) = 1$, the inner limit is 1, and so $f(x) = 1$. On the other hand, let x be irrational. Then $m!\,x$ is never an integer, $|\cos(m!\,\pi x)| < 1$, the inner limit is 0, and $f(x) = 0$.

13.1. If x_0 is rational, then $f(x_0) > 0$, but there are points x arbitrarily close to x_0 where $f(x) = 0$ (namely irrational x); so the definition of continuity is violated if we take ϵ to be $\frac{1}{2}f(x_0)$. Suppose, on the other hand, that x_0 is irrational, and that a positive ϵ is specified. There are only a finite number of rational numbers x in $[0, 1]$ for which $f(x) \geq \epsilon$; take δ to be the minimum distance of these rational numbers from x_0.

13.2. There is a positive δ_1 such that $|f(x) - f(x_0)| < \epsilon$ when $0 < x - x_0 < \delta_1$, and there is a positive δ_2 such that $|f(x) - f(x_0)| < \epsilon$ when $-\delta_2 < x - x_0 < 0$. Then $|f(x) - f(x_0)| < \epsilon$ when $|x - x_0| < \min(\delta_1, \delta_2)$.

13.3. $|f(x_0) - f(x)| = |d(x_0, y) - d(x, y)| \leq d(x_0, x)$ by the triangle inequality.

13.4. It is enough to show that $|D(x) - D(y)| \leq d(x, y)$. If p is any point in E, then $d(x, p) \leq d(x, y) + d(y, p)$ (triangle inequality). Since $D(x) \leq d(x, p)$ whatever the point p is, we infer that $D(x) \leq d(x, y) + d(y, p)$. The distance $d(y, p)$ can be made arbitrarily close to $D(y)$ by a proper choice of p. Hence $D(x) \leq d(x, y) + D(y)$, and so $D(x) - D(y) \leq d(x, y)$. By interchanging x and y, we also have $D(y) - D(x) \leq d(x, y)$ (by the symmetry of d).

13.5. For a constant function, the image of every nonempty set is a single point.

13.6. The inverse image of a neighborhood of $f(x_0)$ of radius $|f(x_0)|/2$ contains a neighborhood of x_0. For x in this neighborhood, $|f(x)| \geq |f(x_0)|/2 > 0$.

13.7. The set of real numbers of absolute value less than $1 + |f(x_0)|$ is open; hence its inverse image is open. This inverse image is not empty, since it contains x_0, and therefore it contains a neighborhood of x_0.

13.8. Negating the definition of continuity tells us that there is an $\epsilon > 0$ such that for every $\delta_n > 0$, there is an x_n for which $d(x_n, x_0) < \delta_n$ and $|f(x_n) - f(x_0)| \geq \epsilon$. Take $\delta_n = 1/n$.

14.1. If $f + g$ and f are continuous, so is $(f + g) - f = g$. If f is not continuous, neither is $-f$, and $f + (-f)$ is the constant function all of whose values are 0; the constant function is continuous.

14.2. Sum: Given a positive ϵ, we can find δ_1 such that $|f(x) - f(x_0)| < \epsilon/2$ when $d(x, x_0) < \delta_1$, and δ_2 such that $|g(x) - g(x_0)| < \epsilon/2$ when $d(x, x_0) < \delta_2$. Take $\delta = \min(\delta_1, \delta_2)$; then $|\{f(x) + g(x)\} - \{f(x_0) + g(x_0)\}| < \epsilon$ when $d(x, x_0) < \delta$.

Product: f and g are bounded in a neighborhood of x_0 (Exercise 13.7); let M be a common bound for the values $|f(x)|$ and $|g(x)|$. Then

$$\begin{aligned} |f(x)g(x) - f(x_0)g(x_0)| &\leq |\{f(x) - f(x_0)\}g(x)| \\ &\quad + |f(x_0)\{g(x) - g(x_0)\}| \\ &\leq M|f(x) - f(x_0)| + M|g(x) - g(x_0)|, \end{aligned}$$

and the right-hand side is small when $d(x, x_0)$ is small.

Quotient: It is enough to show that $1/g$ is continuous when g is continuous if $g(x_0) \neq 0$ (use the continuity of products and $f/g = f \cdot (1/g)$). By Exercise 13.6, we have $|g(x)| \geq m > 0$ in some neighborhood of x_0. Then for x in this neighborhood,

$$\left| \frac{1}{g(x)} - \frac{1}{g(x_0)} \right| = \frac{|g(x) - g(x_0)|}{|g(x)|\,|g(x_0)|} \leq m^{-2}|g(x) - g(x_0)|,$$

and the right-hand side is small when $d(x, x_0)$ is small.

14.3. Let G be an open set; then the complement $C(G)$ is closed. Hence if the images of closed sets are closed, the image of $C(G)$ is closed. Since f is univalent, the complement of the image of $C(G)$ is the image of G, and this complement is open (being the complement of a closed set). The converse is proved similarly.

14.4. If f does not take the value M, then the positive function g defined by $g(x) = 1/[M - f(x)]$ is continuous on the domain of f. Accordingly, g is bounded; let G be an upper bound for $g(x)$. Then $1/[M - f(x)] \leq G$. This implies that $f(x) \leq M - (1/G)$, whence M is not the least upper bound for the values of f after all.

14.5. Let I and L be the smallest and largest points of the range. These are both values assumed by the function; it then assumes every value of the interval $[I, L]$ by the intermediate value property.

14.6. If $f(x) \equiv 0$, there is nothing to prove. Otherwise $f(x_0) > 0$ for some x_0. Take $b > a$ so that $f(x) < \frac{1}{2}f(x_0)$ for $x \geq b$. Then $\max f(x)$ on $[a, \infty)$ equals $\max f(x)$ on $[a, b]$.

14.7. Let x_n be the last (that is, largest) point where f takes its maximum on $[n, \infty)$. (There is a last point since $f(x) \to 0$, and the set where $f(x) = f(x_n)$ is compact.) Clearly $x_{n+1} \geq x_n$, and $x_n \to \infty$.

14.8. The range of f is the same as the range of the restriction of f to $[0, p]$. The latter function is continuous on a compact set.

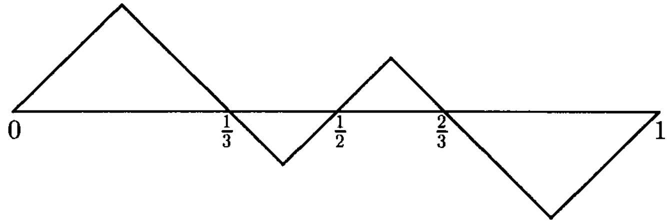
14.9. Same as the preceding answer.

14.10. Evidently

$$\int_a^{a+p} f(x)\,dx = \int_0^p f(x)\,dx + \int_p^{a+p} f(x)\,dx - \int_0^a f(x)\,dx.$$

A change of variables gives $\int_p^{a+p} f(x)\,dx = \int_0^a f(x+p)\,dx = \int_0^a f(x)\,dx$, since $f(x+p) = f(x)$. Alternatively,

$$\frac{d}{dt}\int_t^{t+p} f(x)\,dx = f(t+p) - f(t) = 0.$$



14.11. $\int_0^p [f(x+a) - 2f(x) + f(x-a)]\,dx = 0$.

14.12. Lévy's example was $f(x) = \sin^2(\pi x/b) - x\sin^2(\pi/b)$, where $b \neq 1/k$. P. R. Halmos has pointed out that, similarly, $f(x) = g(x) - x$ is an example whenever g is periodic, continuous, and of period b, with $g(0) = 0$ and $g(1) = 1$. The figure shows an example of a graph with a horizontal chord of length 1, but no horizontal chord of length b when $\frac{1}{3} < b < \frac{1}{2}$. More generally, a piecewise linear graph with alternating slopes of $+1$ and -1 and intercepts at $0, 1/(n+1), 1/n, 2/(n+1), 2/n, \dots, 1$ has no horizontal chord of length b when $1/(n+1) < b < 1/n$.

14.13. If $a = \frac{1}{2}$, our hypothesis says that f has a horizontal chord of length $\frac{1}{2}$. If $a = \frac{1}{3}$, then f has a horizontal chord either of length $\frac{1}{3}$ or of length $\frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$. And so on.

14.14. Consider the continuous function g such that $g(x) = f(x) - x$. We have $g(a) = f(a) - a \geq 0$, and $g(b) = f(b) - b \leq 0$. By the intermediate value property, $g(x) = 0$ for some x.

14.15. (a) Yes; (b) not necessarily. Let $T(t)$ be the indicated time on the erratic clock and $f(t) = t - T(t)$. Then $f(0) = 0$ and $f(24) = 0$, so f has a horizontal chord of length $1 = 24/24$, but not necessarily of any length $24/\lambda$ when λ is not an integer; 576 minutes is 2/5 of 24 hours. We must, however, modify our counterexample to the universal chord theorem to ensure that $T(t)$ is an increasing function. If λ is not an integer, we can take $T(t) = t + \epsilon(t\sin^2(\pi\lambda) - 24\sin^2(\pi\lambda t/24))$, where ϵ is small and positive. Then $T(t + (24/\lambda)) - T(t) \neq 24/\lambda$, and $T'(t) > 0$ if ϵ is small enough.

14.16. This is a limiting case of the theorem on simultaneous bisection of two areas (assuming that it has been proved for areas that are not convex). A direct proof is as follows. Take a point P on the curve, and find another point Q so that the two arcs joining P and Q have the same length. Let $f(P)$ be the part of the area inside the curve that lies to the right of the line joining P to Q. Then f is a continuous function with domain consisting of the points of the curve, since a small change in P produces a small change in Q. If P starts at P_0 and follows the curve, right and left have been interchanged when P reaches Q_0, so $f(P)$, if not originally half the area, is now on the other side of half the area, and so must have been equal to half the area for some position of P.

15.1. The numbers L_n are the elements of a nonincreasing sequence. Since they are bounded, they have a limit L. If n is so large that $L_n < L + \epsilon$, then $s_k < L + \epsilon$ for $k \geq n$. Since the L_n are decreasing toward L, we always have $L_n > L - \epsilon$, and so $s_k > L - \epsilon$ for some $k \geq n$. Taking larger and larger n's, we find an infinite number of s_k with the second property. Hence L has the properties defining $\limsup s_n$.

15.2. $l = \liminf s_n$ if, given any positive ϵ, we have $s_n \geq l - \epsilon$ if n is sufficiently large, and in addition an infinite number of s_n satisfy $s_n \leq l + \epsilon$. If $\{s_n\}$ is unbounded below, $\liminf s_n = -\infty$; if $\{s_n\}$ is bounded below, but l fails to exist, we write $\liminf s_n = +\infty$. In the examples, l is (i) -1; (ii) $+\infty$; (iii) $-\infty$; (iv) 0; (v) 0; (vi) $-\infty$.

15.3. We have $s_k \leq \frac{1}{2}\epsilon + \limsup s_n$ when k exceeds some n_1, and $t_k \leq \frac{1}{2}\epsilon + \limsup t_n$ when k exceeds some n_2; hence $s_k + t_k \leq \epsilon + \limsup s_n + \limsup t_n$ when k exceeds $\max(n_1, n_2)$. Therefore $\limsup(s_n + t_n)$ cannot be any larger number than $\limsup s_n + \limsup t_n$. The example $s_n = (-1)^n$ and $t_n = (-1)^{n+1}$ shows that strict inequality may occur. If $\lim t_n = T$, we have $s_k \geq \limsup s_n - \frac{1}{2}\epsilon$ for infinitely many k, and $t_k \geq T - \frac{1}{2}\epsilon$ for all large k, whence $s_k + t_k \geq \limsup s_n + T - \epsilon$ for infinitely many k.

15.4. If $\epsilon > 0$, we have $s_n \leq L + \epsilon$ when $n > n_1$, and $s_n \geq L - \epsilon$ when $n > n_2$; hence $|s_n - L| < \epsilon$ when $n > \max(n_1, n_2)$.

15.5. Same as the preceding answer.

16.1. (a) $s_n \to 0$ in C. For, let ϵ be a positive number; we have $|s_n(x)| < \epsilon$ if $1 - \epsilon < x \le 1$, independently of n, since $|x^n| \le 1$. If $0 \le x \le 1-\epsilon$, we have $|s_n(x)| \le (1-\epsilon)^n$. Therefore $\max_{0 \le x \le 1} |s_n(x)|$ does not exceed the larger of ϵ and $(1-\epsilon)^n$, and the second of these numbers is the smaller if n is large enough. "Divide and rule."

(b) $\{s_n\}$ does not converge in C. For, $s_n(x) \to 0$ for each x, but $s_n(1 - n^{-1}) = (1 - n^{-1})^n \to e^{-1}$, and $\max |s_n(x)|$ is no smaller than $s_n(1 - n^{-1})$.

16.2. $\sup_{x \in E} |s_n(x)| \le M$ by bounded convergence, and therefore $|\lim_{n\to\infty} s_n(x)| \le M$ for each $x \in E$.

17.1. If f is a discontinuous function, and $f_n = f$ for each n, then $f_n \to f$ uniformly. If g is discontinuous and bounded, and $g_n = n^{-1} g$ for each n, then $g_n \to 0$ uniformly.

17.2. The series is unchanged by termwise differentiation, so its sum $s(x)$ satisfies $s'(x) = s(x)$. Hence $s(x) = ce^x$, with $c = f(0) + f'(0) + \cdots$.

17.3. By the M-test (page 111), with E the set of positive integers, the series $\sum_n f_n(k)$ converges uniformly with respect to k. Hence we can make $|\sum_{n=N}^{\infty} f_n(k)| \le \epsilon$ by choosing N sufficiently large (not depending on k). Correspondingly, $|\sum_{n=N}^{\infty} L_n| \le \epsilon$. When $p(k)$ exceeds N, we have

$$\left|\sum_{n=1}^{p(k)} f_n(k) - \sum_{n=1}^{\infty} L_n\right| \le \left|\sum_{n=1}^{N} \{f_n(k) - L_n\}\right| + 2\epsilon.$$

With N fixed, let $k \to \infty$.

17.4. Take $M_n = |x|^n/n!$ and use the binomial theorem to write

$$\begin{aligned}(1 + x/k)^k &= \sum_{n=0}^{k} \binom{k}{n} \left(\frac{x}{k}\right)^n \\ &= 1 + \sum_{n=1}^{k} \left(1 - \frac{1}{k}\right)\left(1 - \frac{2}{k}\right) \cdots \left(1 - \frac{n-1}{k}\right) \frac{x^n}{n!}.\end{aligned}$$

18.1. For example, $f(x, y) = \sin 2\theta$, where $x = r\cos\theta$ and $y = r\sin\theta$.

19.1. No. Take $\epsilon = 1$, let δ be the corresponding number from the definition of uniform continuity, and let n be a positive integer such that $1/n < \delta$. The value of $f(x)$ for an arbitrary x differs by at most 1 from one of the finite number of values $f(1/n), f(2/n), \ldots, f(n/n)$.

19.2. Yes. There is a large N such that $|f(x) - L| < \frac{1}{2}\epsilon$ when $x > N$. By the triangle inequality, $|f(x) - f(y)| < \epsilon$ when x and y both exceed N. By uniform continuity on the compact interval $[0, N+1]$, there is a δ so that $|f(x) - f(y)| < \epsilon$ if x and y are in this interval and $|x - y| < \delta$. We may as well assume that $\delta < 1$; then $|f(x) - f(y)| < \epsilon$ whenever $|x - y| < \delta$.

19.3. No; a counterexample is $f(x) = \sin x$.

20.1. Since

$$\int_R^{x+y+R} f(u)\,du = \int_R^{y+R} f(u)\,du + \int_{y+R}^{x+y+R} f(u)\,du,$$

we get $\varphi(x+y) = \varphi(y) + \varphi(x)$. But φ is a limit of continuous functions, so it has the form $\varphi(x) = ax$.

21.1. If f'_+ exists (finite), $\lim_{h\to 0^+}[f(x+h) - f(x)]/h$ is finite, so $\lim_{h\to 0^+}[f(x+h) - f(x)] = 0$. For the statement about f', drop all superscript and subscript + signs in the preceding sentence.

21.2. If $f(x) = 0$ for $x < 0$ and $f(x) = 1$ for $x > 0$ and $f(0) = \frac{1}{2}$, then $f'(0) = +\infty$.

21.3. $\epsilon(x) = [f(x) - f(a)]/(x-a) - f'(a) \to 0$.

21.4. The suggested proof fails because $g(b+h) - g(b)$ might be zero for some values of h arbitrarily close to 0. From Exercise 21.3 we have

$$\begin{aligned}\varphi(b+h) - \varphi(b) &= f(g(b+h)) - f(g(b)) \\ &= [g(b+h) - g(b)][f'(g(b)) + \epsilon(g(b+h))].\end{aligned}$$

Divide by h and let $h \to 0$. Since $g'(b)$ exists, g is continuous at b, so $g(b+h) \to g(b)$, and $\epsilon(g(b+h)) \to 0$.

21.5. If $f_+(x) > 0$, then $\liminf_{h\to 0^+}[f(x+h) - f(x)]/h = \delta > 0$, so for all sufficiently small positive h we have $f(x+h) - f(x) > \frac{1}{2}h\delta > 0$.

21.6. (a) If $f'_+(x_0) = c$, finite, and if s is a given (small) positive number, we have

$$c - s < \frac{f(x) - f(x_0)}{x - x_0} < c + s \quad \text{for} \quad x_0 < x < x_0 + h,$$

provided that h is small enough (depending on s). Then

$$\begin{aligned}(c - s)(x - x_0) + K(x - x_0) &< [f(x) + Kx] - [f(x_0) + Kx_0] \\ &< (c + s)(x - x_0) + K(x - x_0)\end{aligned}$$

for $x_0 < x < x_0 + h$. Therefore $f(x) + Kx$ increases on the right at x_0 if $K > -c + s$ and decreases on the right if $K < -c - s$. Since s is arbitrary, $f(x) + Kx$ is monotonic on the right at x_0 if $K \neq -c$.

(b) If c is, for example, $+\infty$, then for a given (large) N we have

$$\frac{f(x) - f(x_0)}{x - x_0} > N \quad \text{for} \quad x_0 < x < x + h,$$

provided that h is small enough (depending on N). Then

$$[f(x) + Kx] - [f(x_0) + Kx_0] > (N + K)(x - x_0),$$

so $f(x) + Kx$ increases on the right at x_0 if $K > -N$; and N can be arbitrarily large.

(c) If $f_K(x) = f(x) + Kx$ is increasing (say) on the right at x_0, then there is a (small) positive h such that $f(x) - f(x_0) \geq -K(x - x_0)$ for $x_0 < x < x_0 + h$, whence $f_+(x_0) \geq -K$. Consequently, if f_K increases for every K, then $f_+(x_0) = f'_+(x_0) = +\infty$, while if f_K decreases for every K, then $f^+(x_0) = f'_+(x_0) = -\infty$. If f_K increases for some values of K and decreases for other values of K, then there must exist K_0 such that f_K increases for $K > K_0$ and decreases for $K < K_0$. Then on the one hand, $f_+(x_0) \geq -K_0$, and on the other hand, $f^+(x_0) \leq -K_0$, so $f'_+(x_0) = -K_0$.

21.7. As in part (c) of the preceding exercise, if f_K is sometimes increasing and sometimes decreasing, then it increases for

$K > K_0$ and decreases for $K < K_0$. So if $a < x < y < b$, then

$$\frac{f(y)-f(x)}{y-x} \geq -K, \qquad K > K_0,$$
$$\frac{f(y)-f(x)}{y-x} \leq -K, \qquad K < K_0,$$

and hence

$$\frac{f(y)-f(x)}{y-x} = -K_0.$$

Since K_0 is independent of x and y, we can fix y and obtain $f(x) = -K_0x + (f(y) + K_0y)$. If, however, f_K is (say) always increasing, then

$$\frac{f(y)-f(x)}{y-x} \geq -K$$

for every value of K (independent of x and y), which is impossible (let $K \to -\infty$).

21.8. $f(x+h) - f(x) \leq 0$ for $h > 0$, whence $f^+(x) \leq 0$; for $h < 0$, we have $f(x+h) - f(x) \leq 0$ and $[f(x+h) - f(x)]/h \geq 0$, whence $f_-(x) \geq 0$.

21.9. Let $g(x) = f(x+a) - f(x) - af'(x)$. Let c be a point where f attains its maximum. Then $f'(c) = 0$; therefore $g(c) = f(c+a) - f(c)$. Since $f(c+a)$ cannot exceed $f(c)$ (the largest value of f), we must have $g(c) \leq 0$. Let d be a point where f attains its minimum; in the same way, we get $g(d) \geq 0$. Since g is a derivative (because the continuous function f is the derivative of its own integral), g has the intermediate value property; therefore $g(x) = 0$ for some x.

21.10. By the Weierstrass approximation theorem from §19, we know that a continuous function g on a closed interval $[a, b]$ can be approximated uniformly by a sequence of polynomials. Polynomials have explicit antiderivatives (computed by a formal rule that uses no integration). By adjusting the constant terms of the antiderivatives, we can make them all equal to 0 at a. The theorem on termwise differentiation implies that these antiderivatives converge uniformly to a limit function that is an antiderivative of g.

21.11. The integral $g(x) = \int_{f(x)}^{f(1)} h(t)\,dt$ is unbounded as $x \to 0^+$, since $f(0^+) = 0$. But $g(1) - g(x) = (1-x)g'(c) = -(1-x)h(f(c))f'(c)$, by the mean-value theorem, with $0 < x < c$, and the left-hand side is unbounded as $x \to 0^+$. This proof requires fewer hypotheses than the following "change of variable" proof:

$$\int_0^1 h(f(x))f'(x)\,dx = \int_0^{f(1)} h(t)\,dt = +\infty;$$

the integrand is therefore unbounded.

Compare W. F. Osgood, Beweis der Existenz einer Lösung der Differentialgleichung $dy/dx = f(x,y)$ ohne Hinzunahme der Cauchy-Lipschitz'schen Bedingung, *Monatshefte für Mathematik und Physik* 9 (1898), 331–345; p. 344.

21.12. $f(x) = 1$ for $x < 0$, and $f(x) = 0$ for $x \geq 0$.

21.13. If $x \neq 0$, then $f'(x) = 2x\sin(1/x) - \cos(1/x)$, and this expression has no limit as $x \to 0$. However,

$$f'(0) = \lim_{x\to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0} x \sin\frac{1}{x} = 0.$$

21.14. We are supposing (tacitly) that f is continuous at y and that $f'(x)$ exists for all x in a neighborhood of y (with $x \neq y$). Then $[f(x) - f(y)]/(x-y) = f'(t)$ by the mean-value theorem, with t between x and y; the right-hand side is as close as we like to $\lim_{x\to y} f'(x)$ if x is sufficiently close to y, hence so is the left-hand side. This means that the derivative $f'(y)$ exists and is equal to $\lim_{x\to y} f'(x)$.

22.1. Let $[a,b]$ be the interval. If f is, say, nondecreasing, then $f(a) \leq f(x) \leq f(b)$ if $a \leq x \leq b$.

22.2. Let y be an interior point of the domain, and let $\{x_n\}$ be an increasing sequence with limit y. If f is, say, increasing, then $\{f(x_n)\}$ is an increasing bounded sequence which (Exercise 8.3) has a limit L. If $x_n < x < y$, we can find x_m so that

$x < x_m < y$, and then $f(x_n) \le f(x) \le f(x_m) \le L$. Since $f(x_n) \to L$ as $n \to \infty$, it follows that $f(x) \to L$ as $x \to y^-$.

22.3. If $f_n(x) \to f(x)$ for each x, and if $f_n(x) \le f_n(y)$ for all n when $x \le y$, then $f(x) \le f(y)$ when $x \le y$. The statement does not exclude the possibility that some f_n are increasing and others decreasing. If this happens, a finite number of exceptional f_n cause no difficulty. If there are infinitely many of both kinds, then f must be both nonincreasing and nondecreasing, and so constant.

22.4. Let f have jumps of amount $1/\{n(n+1)\}$ at the points $1/n$, and $f(0) = 0$. If $(m+1)^{-1} < h < m^{-1}$, then

$$f(h) = \sum_{k=m+1}^{\infty} \frac{1}{k(k+1)} = \frac{1}{m+1},$$

whence $f'_+(0) = 1$.

23.1. Put $y = q_1 x_1 + q_2 x_2$ in the convexity property $(*)$ and replace $f(x)$ by $xg(x)$ to obtain

$$\begin{aligned} yg(y) &\le q_1 x_1 g(x_1) + q_2 x_2 g(x_2) \\ &\le (q_1 x_1 + q_2 x_2) \max(g(x_1), g(x_2)); \end{aligned}$$

that is, $g(y) \le \max(g(x_1), g(x_2))$. This says that on each closed interval contained in $(0, \infty)$, the continuous function g attains its maximum at an endpoint.

Let E be the set of points x in $(0, \infty)$ such that there is a $y > x$ with $g(y) < g(x)$. By the rising sun lemma (page 166, applied to the function $-g$), the set E is open. If E is empty, then g is nondecreasing on $(0, \infty)$. Otherwise, let I be one of the open intervals of which E is composed, say $I = (a, b)$, where possibly $b = \infty$. (It will turn out that E consists of just one open interval.)

Now g must be nonincreasing on I, for if x_1 and x_2 are points of I such that $x_1 < x_2$ and $g(x_1) < g(x_2)$, then by taking a point $y > x_2$ such that $g(y) < g(x_2)$ we find that g does not attain its maximum on the interval $[x_1, y]$ at an endpoint. Because g is nonincreasing on I, the left endpoint a

must be 0; otherwise a would be contained in I by the definition of E.

If $b = \infty$, then g is nonincreasing on $(0, \infty)$. Suppose $b < \infty$. Since $b \notin E$, there is no point to the right of b at which g takes a value smaller than $g(b)$. It must be, then, that g is nondecreasing on (b, ∞), for if $b < x_1 < x_2$ and $g(x_2) < g(x_1)$, then g does not attain its maximum on the interval $[b, x_2]$ at an endpoint.

23.2. Apply the inequality between the arithmetic and geometric means of $\sin x$ and $\sin x + \cos x$ with weights $\tan x$ and $1 - \tan x$ to get $(\sin x)^{\tan x}(\sin x + \cos x)^{1-\tan x} \leq (\tan x)(\sin x) + (1 - \tan x)(\sin x + \cos x) = \cos x$. Since $\sin x + \cos x > 1$ for $0 < x < \pi/4$, the left-hand side exceeds $(\sin x)^{\tan x}$. Now raise each side to the power $\cos x$.

23.3. Given $s_n \to s$, to show that $(s_1 + \cdots + s_n)/n \to s$. By considering $s_n - s$ instead of s_n, we may assume that $s = 0$. Write

$$\frac{s_1 + \cdots + s_n}{n} = \frac{s_1 + \cdots + s_k}{n} + \frac{s_{k+1} + \cdots + s_n}{n},$$

and choose k so large that $|s_j| < \epsilon$ for $j > k$. Then

$$\frac{|s_{k+1} + \cdots + s_n|}{n} < \epsilon.$$

Now k is fixed, so $|s_1 + \cdots + s_k|/n < \epsilon$ if n is large enough, and so $|s_1 + \cdots + s_n|/n < 2\epsilon$.

24.1. Suppose that $|f''(x)| \leq M$ for all x. Let x_1 be so large that $|f(t)| < \epsilon$ for $t > x_1$. For $a > x_1$, Taylor's theorem with remainder of order 2 gives

$$\begin{aligned} |f'(a)| &\leq \frac{2\epsilon}{x-a} + \frac{1}{x-a}\int_a^x (x-t)M\,dt \\ &\leq \frac{2\epsilon}{x-a} + \frac{M(x-a)}{2} \end{aligned}$$

when $x > a$. Take $x - a = \sqrt{\epsilon}$ to get $|f'(a)| \leq 2\sqrt{\epsilon} + M\sqrt{\epsilon}/2$. Since ϵ is arbitrary, $f'(a) \to 0$ as $a \to \infty$.

25.1. The open set E is the union of countably many disjoint open intervals whose lengths form a convergent series. Discarding the tail of the series gives a finite number of open intervals whose total length is within $\epsilon/2$ of the measure of E. Shrinking each interval a little bit to a closed interval gives a finite number of closed intervals contained in E whose total length is within ϵ of the measure of E.

25.2. There is no harm in supposing that E is contained in some bounded open interval I, for it is enough to consider the part of E in each interval $[n, n+1]$. Now take complements and invoke the preceding exercise.

25.3. Let U be an open set containing Z such that $m(U) < \epsilon/2$. If E is measurable, let F be a closed subset of E and G an open set containing E such that $m(G \cap C(F)) < \epsilon/2$. Then $F \subset E \cup Z \subset G \cup U$, and $m((G \cup U) \cap C(F)) < \epsilon$. If $E \cup Z$ is measurable, let F be a closed subset of $E \cup Z$ and G an open set containing $E \cup Z$ such that $m(G \cap C(F)) < \epsilon/2$. Then $F \cap C(U) \subset E \subset G$, and $m(G \cap C(F \cap C(U))) < \epsilon$.

25.4. Notice that S is not assumed to be measurable. Suppose $F \subset E \subset G$, where F is closed, G is open, and the measure of $G \cap C(F)$ is less than ϵ. Cover S by countably many disjoint open intervals I_n whose total length is less than $\mu(S) + \epsilon$. The open sets $I_n \cap G$ cover $S \cap E$, so $\mu(S \cap E) \leq \sum \mu(I_n \cap G)$, and the open sets $I_n \cap C(F)$ cover $S \cap C(E)$, so $\mu(S \cap C(E)) \leq \sum \mu(I_n \cap C(F))$. Now $I_n \cap G$ and $I_n \cap C(F)$ cover I_n with an overlap equal to $I_n \cap (G \cap C(F))$. Since the I_n are disjoint, it follows that $\mu(S \cap E) + \mu(S \cap C(E)) \leq \sum \mu(I_n) + \mu(G \cap C(F)) < \mu(S) + 2\epsilon$. Consequently, $\mu(S \cap E) + \mu(S \cap C(E)) \leq \mu(S)$, and the reverse inequality always holds.

25.5. There is no harm in supposing that E has finite measure. Then we can find a countable collection of intervals $\{I_n\}$ covering E whose total length $\sum_n m(I_n) < (4/3)m(E)$. On the other hand, $m(E) \leq \sum_n m(E \cap I_n)$. Thus $\sum_n m(E \cap I_n) > (3/4)\sum_n m(I_n)$, and so there must be (at least) one n for which $m(E \cap I_n) > (3/4)m(I_n)$. Call this interval I_n simply I. Nothing essential changes if we make a dilation of $\mathbf{R}_1$, so we may as well assume that I is an interval of length 1; in this case, we have $m(E \cap I) > 3/4$.

I claim that every point in the interval $(-1/3, 1/3)$ can be written as the difference of two points of E. In the contrary case, there would be a number $x \in (-1/3, 1/3)$ for which E and its translate by x units would be disjoint. In particular, the measure of the union of $E \cap I$ and its translate by x units would be $2m(E \cap I) > 3/2$. On the other hand, this measure cannot be greater than the measure of the union of the whole of I and its translate by x units, which is at most $4/3$. Contradiction.

26.1. For a monotonic function, the inverse image of an interval is either an interval, a single point, or empty.

26.2. For a continuous function, the inverse image of an open interval is open (page 88), hence is a measurable set. The inverse image of a closed interval is closed, hence is measurable. A half-open interval is the intersection of an open interval and a closed interval, so its inverse image is the intersection of an open set and a closed set, hence is a Borel set, hence is a measurable set.

26.3. The inverse image of an interval under g differs from its inverse image under f by a set of measure zero, so it is a measurable set by Exercise 25.3.

26.4. Suppose $\{f_n\}$ is a sequence of continuous functions converging pointwise almost everywhere to f; this means that $f(x) = f_1(x) + \sum_{j=1}^{\infty}\{f_{j+1}(x) - f_j(x)\}$ for almost all x. Define nondecreasing sequences of continuous functions $\{g_n\}$ and $\{h_n\}$ by

$$g_n(x) = f_1(x) + \sum_{j=1}^{n} \max\{f_{j+1}(x) - f_j(x), 0\},$$

$$h_n(x) = -\sum_{j=1}^{n} \min\{f_{j+1}(x) - f_j(x), 0\}.$$

Since $g_n(x) - h_n(x) = f_1(x) + \sum_{j=1}^{n}\{f_{j+1}(x) - f_j(x)\} = f_{n+1}(x)$, we deduce that $g_n(x) - h_n(x) \to f(x)$ for almost all x. Therefore we are done if there are functions g and h such that $g_n \to g$ and $h_n \to h$.

However, it could happen for all x that $g_n(x)$ and $h_n(x)$ tend to ∞ as $n \to \infty$. We can avoid this difficulty by replacing the original sequence $\{f_n\}$ by a suitably chosen subsequence $\{f_{n_j}\}$ for which $\sum_{j=1}^{\infty} |f_{n_{j+1}}(x) - f_{n_j}(x)|$ is finite for almost all x. To choose the index n_j for a particular j, observe that almost all x belong to one of the nested sets S_k defined by $S_k = \{x : |f_m(x) - f(x)| < 2^{-j} \text{ when } m \geq k\}$. Choose n_j to be the first index k greater than n_{j-1} for which S_k has measure greater than $(b-a)(1-2^{-j})$.

26.5. Since a monotonic function is differentiable almost everywhere (page 165), so is the difference of two monotonic functions.

26.6. Suppose $|f'(x)| \leq M$ for all x in (a, b). By the mean-value theorem (page 143), each difference $|f(x_{k+1}) - f(x_k)| \leq M(x_{k+1} - x_k)$, so the total variation of f is no more than $M(b-a)$.

26.7. One example is the function f such that $f(0) = 0$ and $f(x) = x\sin(\pi/x)$ for $x \neq 0$; its variation on an interval of the form $[(n+1)^{-1}, n^{-1}]$ exceeds $2|f(1/(n+\frac{1}{2}))| > 2/(n+1)$, and $\sum_n (1/n)$ diverges. Other examples, in view of Exercise 26.5, are the continuous nowhere differentiable functions of §10.

26.8. The δ that works in the definition of absolute continuity for a collection of intervals also works in the definition of continuity for a single interval.

26.9. Let δ correspond to $\epsilon = 1$ in the definition of absolute continuity. In computing the total variation, we need only consider divisions of $[a, b]$ in which all subintervals have length less than δ, since subdividing intervals increases the computed variation. For any such division, collect the subintervals into groups, each having total length less than δ. This can be done so that there are no more than $2(b-a)/\delta$ such groups, and so the total variation is at most $2(b-a)/\delta$.

27.1. Since the rational numbers form a set of measure zero, every measurable set of positive measure must contain irrational numbers. Hence every upper sum equals 1. Evidently no lower sum exceeds 1. The lower sum for the partition of $[0, 1]$ into two pieces, S_1 being the rational numbers and S_2 being the irrational numbers, equals 1. Thus the supremum of lower sums equals 1, and the infimum of upper sums equals 1, and so this common value is the value of the integral.

27.2. The integral of the Cantor function is $1/2$. The Cantor function is monotonic, so it is measurable (Exercise 26.1). Therefore we need to show, for an arbitrary positive ϵ, that there is an upper sum less than $(1/2) + \epsilon$ and a lower sum greater than $(1/2) - \epsilon$. Choose an integer n so that the total length of the intervals removed in the first n steps of the construction of the Cantor set is greater than $1 - \epsilon$. Take for a partition of $[0, 1]$ these intervals together with one measurable set S_0 that is the totality of points not removed in the first n steps. Since $m(S_0) < \epsilon$, and the Cantor function is everywhere between 0 and 1, the contribution that S_0 makes to an upper or lower sum is between 0 and ϵ. The Cantor function is constant on each of the intervals in our partition, so these sets make a contribution to the upper or lower sum that equals

$$\frac{1}{3}\cdot\frac{1}{2}+\frac{1}{9}\cdot\left(\frac{1}{4}+\frac{3}{4}\right)+\frac{1}{27}\cdot\left(\frac{1}{8}+\frac{3}{8}+\frac{5}{8}+\frac{7}{8}\right)+\cdots$$
$$=\sum_{k=1}^{n}\frac{2^{k-2}}{3^k}.$$

As $n \to \infty$, this sum approaches $1/2$.

27.3. Suppose f is Riemann integrable on $[a, b]$; this means that f is bounded, and f is continuous at all points of $[a, b]$ except for a subset S of measure 0. (Notice that "continuous almost everywhere" is not the same as "equal almost everywhere to a continuous function"—the function that is equal to 0 on the rational numbers and 1 on the irrational numbers has the second property but not the first.) Let f_R denote the restriction of f to the complement $C(S)$ of S. Then f_R is a continuous function on its domain $C(S)$, so the inverse image under f_R of an open interval I is an open subset of $C(S)$, and in particular is a measurable subset of $[a, b]$. Now the inverse image of I under f differs from the inverse image of I under f_R

by a set of measure 0 (a subset of S), so it too is a measurable set. Thus f is Lebesgue measurable.

Let I denote the value of the Riemann integral of f on $[a, b]$. If a positive ϵ is given, then there is a partition of $[a, b]$ into a finite number of sub*intervals* S_k such that both $\sum_k m(S_k) \sup_{S_k} f$ and $\sum_k m(S_k) \inf_{S_k} f$ differ from I by less than ϵ. Since these are particular upper and lower Lebesgue sums, the infimum of the Lebesgue upper sums is at most $I + \epsilon$, and the supremum of the Lebesgue lower sums is at least $I - \epsilon$. Since ϵ is arbitrary, and f is Lebesgue integrable, the Lebesgue integral of f must equal I.

27.4. Since $x^{-1} \sin x$ is bounded near $x = 0$, we are only concerned with what happens when $x \to \infty$. Integrating by parts gives $\int_1^t x^{-1} \sin x \, dx = \cos 1 - t^{-1} \cos t - \int_1^t x^{-2} \cos x \, dx$. The integral on the right-hand side is absolutely convergent (because $|x^{-2} \cos x| \leq x^{-2}$), so $\lim_{t \to \infty} \int_1^t x^{-1} \sin x \, dx$ exists. However, the positive part and the negative part of $x^{-1} \sin x$ are not individually integrable on $(0, \infty)$. In fact, the graph of $\sin x$ on the interval $(2n\pi, 2n\pi + \pi)$ lies above a rectangle of width $\pi/2$ and height $1/\sqrt{2}$, which means that $\int_{2n\pi}^{(2n+1)\pi} x^{-1} \sin x \, dx > \{2\sqrt{2}(2n+1)\}^{-1}$; and $\sum_n (2n+1)^{-1}$ diverges.

27.5. Let S_k denote the set of points x for which $k \leq |f(x)| < k + 1$. Since $\sum_{k=1}^{n} k \cdot m(S_k) \leq \int_S |f|$ for every n, the series $\sum_{k=1}^{\infty} k \cdot m(S_k)$ converges. Consequently, there exists a large integer N such that $\sum_{k \geq N} k \cdot m(S_k) < \epsilon/2$, and so $\sum_{k \geq N} (k+1) \cdot m(S_k) < 3\epsilon/4$. Let $\delta = \epsilon/(4N)$. If E is a measurable subset of S with measure less than δ, then the integral of $|f|$ over the part of E where $|f(x)| < N$ is no more than $N\delta = \epsilon/4$, while the integral of $|f|$ over the remaining part of E is less than $3\epsilon/4$.

28.1. Since $|\sin(\log x)| \leq 1$ for all x, our function is dominated by g, where $g(x) = x^{-1/2}$; and we know that g is integrable on $(0, 1)$ with integral equal to $2\sqrt{x}\,\big]_0^1 = 2$.

28.2. Take $f_n(x) = 4n^2 x$ if $0 \leq x \leq 1/(2n)$, and $f_n(x) = 4n - 4n^2 x$ if $1/(2n) \leq x \leq 1/n$, and let $f_n(x)$ be zero elsewhere.

28.3. Apply the dominated convergence theorem with g being a constant function with value M.

28.4. Extend f to equal $f(b)$ for $x > b$, and define a function f_n by $f_n(x) = n\{f(x+n^{-1}) - f(x)\}$. The monotonicity of f implies that f_n is nonnegative. We know that f has a derivative almost everywhere, so $f_n(x) \to f'(x)$ almost everywhere. Applying Fatou's lemma gives $\int_a^b f' \le \liminf_{n\to\infty} \int_a^b f_n$. But $\int_a^b f_n = f(b) - n\int_a^{a+n^{-1}} f \le f(b) - f(a)$ because f is nondecreasing.

28.5. Nothing changes if we replace f_n by $f_n - f_1$ and replace f by $f - f_1$, so we may as well assume that $f_n \ge 0$ for all n. Then Fatou's lemma implies that $\int f \le \liminf_{n\to\infty} \int f_n$. On the other hand, $\int f_n \le \int f$ since $f_n \le f$, so $\limsup_{n\to\infty} \int f_n \le \int f$. Putting the two inequalities together gives what we want.

29.1. An example is $f(x) = x^{-1/p}$.

29.2. $\|f\|_p = (\int_0^1 x^p\,dx)^{1/p} = (p+1)^{-1/p}$, and this quantity tends to 1 when $p \to \infty$.

29.3. It is 0, since the rational numbers are a set of measure zero.

29.4. Evidently $\int_1^\infty f(t)\,dt$ and $\int_1^\infty f(tx)\,dt$ either both converge or both diverge, since the latter equals $x^{-1}\int_x^\infty f(t)\,dt$. Therefore, we may as well assume that $x = 1$. Now $\sum_{k=1}^\infty f(k)$ is the integral of a step function that equals $f(n)$ on the interval $(n, n+1)$, and since this step function is never smaller than f, we have $\sum_{k=1}^\infty f(k) \ge \int_1^\infty f$. On the other hand, $\sum_{k=2}^\infty f(k)$ is the integral of a step function that equals $f(n+1)$ on the interval $(n, n+1)$, and since this step function is never bigger than f, we have $\sum_{k=2}^\infty f(k) \le \int_1^\infty f$.

29.5. If, for example, $f(t) = \sin^2(\pi t)$, then $\int_1^\infty f$ diverges, while $\sum_{k=1}^\infty f(k) = 0$. On the other hand, if g is a function that has the value 1 at the integers and the value 0 everywhere else, then $\int_1^\infty g = 0$ while $\sum_{k=1}^\infty g(k)$ diverges.

29.6. Since $2(\sin nx)(\sin mx) = \cos(n-m)x - \cos(n+m)x$, the integral $\int_0^\pi (\sin nx)(\sin mx)\,dx$ evaluates when $n \ne m$ to

$$\tfrac{1}{2}\{(n-m)^{-1}\sin(n-m)x - (n+m)^{-1}\sin(n+m)x\}\Big|_0^\pi = 0.$$

When $n = m$, the same identity implies (for $n \geq 1$) that $\int_0^\pi (\sin nx)^2\,dx = \{(x/2) - (4n)^{-1}\sin 2x\}\big|_0^\pi = \pi/2$.

30.1. By the mean-value theorem (page 143), we can rewrite each sum $\sum_{k=1}^n f(y_k)[g(x_{k+1}) - g(x_k)]$ approximating the Stieltjes integral $\int_a^b f(x)\,dg(x)$ as a sum $\sum_{k=1}^n f(y_k)g'(z_k)[x_{k+1} - x_k]$, where $x_k \leq z_k \leq x_{k+1}$. Since g' is uniformly continuous on $[a,b]$, we will have for sufficiently fine partitions of $[a,b]$ that this sum is close to a sum $\sum_{k=1}^n f(y_k)g'(y_k)[x_{k+1} - x_k]$ approximating the Riemann integral $\int_a^b f(x)g'(x)\,dx$.

30.2. The greatest integer function is a step function with a jump of 1 at each integer.

30.3. If the integration limits are restored, the formula correctly says $b - a = (1+b) - (1+a)$. Omitting the limits means that the two sides could differ by a constant.

31.1. Evidently $F(x) \leq \sup_t F(t)$ for every x, so if $g(x) \geq 0$, then $F(x)g(x) \leq g(x)\sup_t F(t)$. Integrating both sides gives $\int_a^b F(x)g(x)\,dx \leq \{\sup_t F(t)\}\int_a^b g(x)\,dx$. The other inequality is derived analogously.

31.2. Suppose the partial sums $B_n = \sum_{j=1}^n b_j$ stay between $-K$ and $+K$ for some positive constant K. The partial summation formula says $\sum_{j=1}^n a_j b_j = a_n B_n + \sum_{j=1}^{n-1} B_j(a_j - a_{j+1})$. When $n \to \infty$, the first term on the right-hand side tends to 0 because $a_n \to 0$ and $|B_n| \leq K$. The second term also has a limit because the series $\sum_{j=1}^\infty |B_j(a_j - a_{j+1})|$ converges: it is dominated by K times $\sum_{j=1}^\infty (a_j - a_{j+1})$, which converges to a_1 (telescoping sum).

31.3. Apply the partial summation formula with $a_n R^n$ being the sequence with bounded partial sums B_n, and $(x/R)^n$ being the sequence tending monotonically to zero. I claim that if $|B_n| \leq K$ for all n, then $|f(x)| \leq 2K$ when $0 < x < R$. Indeed, $\sum_{j=0}^n a_j x^j = (x/R)^n B_n + \sum_{j=0}^{n-1} B_j\{(x/R)^j - (x/R)^{j+1}\}$. The first term on the right-hand side has absolute value less than K, and the sum on the right-hand side has absolute value less than $K\{1 - (x/R)^n\} < K$. Let $n \to \infty$.

31.4. Evaluation at a point is a linear operation, and so is integration. The functional P is continuous because a uniformly convergent sequence of functions converges pointwise. The functional I is continuous because a uniformly convergent sequence of functions can be integrated term by term (page 117).

31.5. A sequence $f_n \to f$ if and only if $f_n - f \to 0$, and since L is linear, $L(f_n) \to L(f)$ if and only if $L(f_n - f) \to 0$. Thus, a linear functional L is continuous on the space C if and only if it is continuous at the zero function in C.

Suppose $|L(f)| \leq M \sup_x |f(x)|$ for some constant M. Now $f_n \to 0$ in C means, by definition, that $\sup_x |f_n(x)| \to 0$, and hence $L(f_n) \to 0$. So L is continuous.

Conversely, if there is no such M, then there must be, for each n, a function f_n in C such that $|L(f_n)| > n \sup_x |f_n(x)|$. Let g_n be f_n divided by $n \sup_x |f_n(x)|$. Then $g_n \to 0$ in C, but $|L(g_n)| > 1$, so $L(g_n) \not\to 0$. Thus L is not continuous.

31.6. To see that $f_n(x)$ is uniformly close to $f(x)$, it is enough to consider x in a typical subinterval $[x_k, x_{k+1}]$. When the partition is sufficiently fine, the uniform continuity of f implies that $f(y_k)$, $f(y_{k-1})$, and $f(x)$ are close to each other for such x. We may assume n is so large that n^{-1} is much smaller than the width of any subinterval. For x between x_k and x_{k+1}, there are only two terms of the the sum defining $f_n(x)$ that can possibly be nonzero. One is $f(y_{k-1})\left(u_{n,x_k}(x) - u_{n,x_{k-1}}(x)\right)$, which equals $f(y_{k-1})u_{n,x_k}(x)$. The other term that can make a contribution to the sum is $f(y_k)\left(u_{n,x_{k+1}}(x) - u_{n,x_k}(x)\right)$, which equals $f(y_k)(1 - u_{n,x_k}(x))$. Therefore $f_n(x)$ is equal to $f(y_k) + u_{n,x_k}(x)\left(f(y_{k-1}) - f(y_k)\right)$. (In the special case that $k = 1$, we have $f_n(x) = f(y_1)$.) Since $f(y_{k-1})$ and $f(y_k)$ are close together, and $u_{n,x_k}(x)$ is between 0 and 1, this means that $f_n(x)$ is close to $f(y_k)$, and hence close to $f(x)$.

31.7. We need to find an upper bound on sums of the form $\sum_{k=1}^{N} |g(x_{k+1}) - g(x_k)|$, where $x_1, \ldots, x_{N+1}$ is a partition of $[0,1]$. By the definition of g, such a sum is well approximated by $|L(u_{n,x_2})| + \sum_{k=2}^{N} |L(u_{n,x_{k+1}}) - L(u_{n,x_k})|$ when n is large. This sum equals $L\left(\pm u_{n,x_2} + \sum_{k=2}^{N} \pm(u_{n,x_{k+1}} - u_{n,x_k})\right)$ for an appropriate choice of plus and minus signs, by the linearity of L.

The boundedness of L implies that this is at most M times the supremum of $|\pm u_{n,x_2} + \sum_{k=2}^{N} \pm(u_{n,x_{k+1}} - u_{n,x_k})|$. The latter supremum, by the triangle inequality, is at most the supremum of $u_{n,x_2} + \sum_{k=2}^{N}(u_{n,x_{k+1}} - u_{n,x_k})$, and this telescoping sum reduces to $u_{n,x_{N+1}}$, which has supremum 1. Thus, the total variation of g is at most M.

32.1. If $f(t) = t^{-1}\sin(t^{1/2})$, then $f'(t) = -t^{-2}\sin(t^{1/2}) + (1/2)t^{-3/2}\cos(t^{1/2})$. Evidently $|f'|$ is integrable on $(1,\infty)$, since $|f'(t)| \leq (3/2)t^{-3/2}$ on this interval. Consequently, we will know that the sum $\sum_{n=1}^{\infty} n^{-1}\sin(n^{1/2})$ converges if we show that the limit $\lim_{n\to\infty}\int_1^n t^{-1}\sin(t^{1/2})\,dt$ exists. The substitution $t = u^2$ changes this integral to twice $\int_1^{\sqrt{n}} u^{-1}\sin(u)\,du$, and we saw in Exercise 27.4 that this has a limit as $n \to \infty$.

32.2. Since $|P_1(t)| \leq 1$ for all t, we have $\int_1^n |P_1(t)f'(t)|\,dt \leq \int_1^n |f'(t)|\,dt = -\int_1^n f'(t)\,dt = f(1) - f(n) \leq f(1)$. Thus the integral on the right-hand side of the Euler-Maclaurin formula has a limit: the improper integral is even absolutely convergent. The term $\frac{1}{2}\{f(1)+f(n)\}$ has a limit since $f(n) \to 0$. Since the right-hand side of the Euler-Maclaurin formula has a limit, so does the left-hand side.

32.3. The sum is bigger than the integral, but this difference decreases with n. Indeed, since P_1 is antisymmetric about the midpoint of each interval $(m, m+1)$, and negative on the left half, while $f'(t) = -1/t^2$ is negative with decreasing magnitude, the integral of $P_1(t)f'(t)$ over each interval $(m, m+1)$ is positive. Therefore the right-hand side of the Euler-Maclaurin formula is positive. When n is replaced by $n+1$, this right-hand side changes by $\frac{1}{2}\{f(n+1) - f(n)\} + \int_n^{n+1} P_1(t)f'(t)\,dt$, which is the same as $\int_n^{n+1}(t-[t])f'(t)\,dt$. Since $t - [t]$ is positive on the interval $(n, n+1)$, and $f'(t) = -1/t^2$ is negative, this change is negative.

32.4. Since $\sum_{k=1}^{n}\log k = \log(n!)$ and $\int_1^n \log t\,dt = n\log n - n + 1$, the Euler-Maclaurin formula gives

$$\log(n!) - n\log n + n - 1 = \frac{1}{2}\log n + \int_1^n P_1(t)t^{-1}\,dt.$$

The integral has absolute value no more than $\frac{1}{2}\log n$, since $|P_1(t)| \le \frac{1}{2}$, so after dividing this equation by n we find that

$$\lim_{n\to\infty} \left(n^{-1}\log(n!) - \log n\right) = -1.$$

Now exponentiate.

32.5. Since $Q_4'(t) = Q_3(t) = \frac{1}{6}t^3 - \frac{1}{4}t^2 + \frac{1}{12}t$, we must have $Q_4(t) = \frac{1}{24}t^4 - \frac{1}{12}t^3 + \frac{1}{24}t^2 + k$ and then $Q_5(t) = \frac{1}{120}t^5 - \frac{1}{48}t^4 + \frac{1}{72}t^3 + kt + c$ for some constants k and c. We will have $Q_4(0) = Q_4(1)$ whatever value of k we take, but to ensure $Q_5(0) = Q_5(1)$ we must choose $k = -1/720$.

Index

A PRIMER OF REAL FUNCTIONS

An updated version of an MAA classic—ideal for a real analysis course or an advanced calculus course

This is a revised, updated and augmented edition of a classic Carus monograph on the theory of functions of a real variable. The principal change in this fourth edition is the addition of a new chapter on integration and its applications.

The book retains the informal chatty style of the previous editions, remaining accessible to readers with some mathematical sophistication and a background in calculus. The book is thus suitable either for self-study or for supplemental reading in a course on advanced calculus or real analysis.

Not intended as a systematic treatise, this book has more the character of a sequence of loosely connected lectures on a variety of interesting topics connected with real functions. Many of these topics are not commonly encountered in undergraduate textbooks: for example, the existence of continuous everywhere-oscillating functions (via the Baire category theorem); two functions having equal derivatives, yet not differing by a constant; application of Stieltjes integration to the speed of convergence of infinite series.

Earlier editions covered sets, metric spaces, continuous functions, and differentiable functions. The fourth edition adds sections on measurable sets and functions, the Lebesgue and Stieltjes integrals, and applications.

This book recaptures the sense of wonder that was associated with the subject in its early days.

ISBN 0-88385-029-X

9 780883 850299