Calculus on Manifolds

A MODERN APPROACH TO CLASSICAL THEOREMS OF ADVANCED CALCULUS

Michael Spivak
Brandeis University

Calculus on Manifolds

A MODERN APPROACH TO CLASSICAL THEOREMS OF ADVANCED CALCULUS

W. A. Benjamin, Inc.
New York *Amsterdam*
1965

Editors' Foreword

Mathematics has been expanding in all directions at a fabulous rate during the past half century. New fields have emerged, the diffusion into other disciplines has proceeded apace, and our knowledge of the classical areas has grown ever more profound. At the same time, one of the most striking trends in modern mathematics is the constantly increasing interrelationship between its various branches. Thus the present-day students of mathematics are faced with an immense mountain of material. In addition to the traditional areas of mathematics as presented in the traditional manner—and these presentations do abound—there are the new and often enlightening ways of looking at these traditional areas, and also the vast new areas teeming with potentialities. Much of this new material is scattered indigestibly throughout the research journals, and frequently coherently organized only in the minds or unpublished notes of the working mathematicians. And students desperately need to learn more and more of this material.

This series of brief topical booklets has been conceived as a possible means to tackle and hopefully to alleviate some of these pedagogical problems. They are being written by active research mathematicians, who can look at the latest developments, who can use these developments to clarify and condense the required material, who know what ideas to underscore and what techniques to stress. We hope that they will also serve to present to the able undergraduate an introduction to contemporary research and problems in mathematics, and that they will be sufficiently informal that the personal tastes and attitudes of the leaders in modern mathematics will shine through clearly to the readers.

The area of differential geometry is one in which recent developments have effected great changes. That part of differential geometry centered about Stokes' Theorem, sometimes called the fundamental theorem of multivariate calculus, is traditionally taught in advanced calculus courses (second or third year) and is essential in engineering and physics as well as in several current and important branches of mathematics. However, the teaching of this material has been relatively little affected by these modern developments; so the mathematicians must relearn the material in graduate school, and other scientists are frequently altogether deprived of it. Dr. Spivak's book should be a help to those who wish to see Stoke's Theorem as the modern working mathematician sees it. A student with a good course in calculus and linear algebra behind him should find this book quite accessible.

Robert Gunning
Hugo Rossi

Princeton, New Jersey
Waltham, Massachusetts
August 1965

Preface

This little book is especially concerned with those portions of "advanced calculus" in which the subtlety of the concepts and methods makes rigor difficult to attain at an elementary level. The approach taken here uses elementary versions of modern methods found in sophisticated mathematics. The formal prerequisites include only a term of linear algebra, a nodding acquaintance with the notation of set theory, and a respectable first-year calculus course (one which at least mentions the least upper bound (sup) and greatest lower bound (inf) of a set of real numbers). Beyond this a certain (perhaps latent) rapport with abstract mathematics will be found almost essential.

The first half of the book covers that simple part of advanced calculus which generalizes elementary calculus to higher dimensions. Chapter 1 contains preliminaries, and Chapters 2 and 3 treat differentiation and integration.

The remainder of the book is devoted to the study of curves, surfaces, and higher-dimensional analogues. Here the modern and classical treatments pursue quite different routes; there are, of course, many points of contact, and a significant encounter occurs in the last section. The very classical equation reproduced on the cover appears also as the last theorem of the book. This theorem (Stokes' Theorem) has had a curious history and has undergone a striking metamorphosis.

The first statement of the Theorem appears as a postscript to a letter, dated July 2, 1850, from Sir William Thomson (Lord Kelvin) to Stokes. It appeared publicly as question 8 on the Smith's Prize Examination for 1854. This competitive examination, which was taken annually by the best mathematics students at Cambridge University, was set from 1849 to 1882 by Professor Stokes; by the time of his death the result was known universally as Stokes' Theorem. At least three proofs were given by his contemporaries: Thomson published one, another appeared in Thomson and Tait's *Treatise on Natural Philosophy*, and Maxwell provided another in *Electricity and Magnetism* [13]. Since this time the name of Stokes has been applied to much more general results, which have figured so prominently in the development of certain parts of mathematics that Stokes' Theorem may be considered a case study in the value of generalization.

In this book there are three forms of Stokes' Theorem. The version known to Stokes appears in the last section, along with its inseparable companions, Green's Theorem and the Divergence Theorem. These three theorems, the classical theorems of the subtitle, are derived quite easily from a modern Stokes' Theorem which appears earlier in Chapter 5. What the classical theorems state for curves and surfaces, this theorem states for the higher-dimensional analogues (manifolds) which are studied thoroughly in the first part of Chapter 5. This study of manifolds, which could be justified solely on the basis of their importance in modern mathematics, actually involves no more effort than a careful study of curves and surfaces alone would require.

The reader probably suspects that the modern Stokes' Theorem is at least as difficult as the classical theorems derived from it. On the contrary, it is a very simple consequence of yet another version of Stokes' Theorem; this very abstract version is the final and main result of Chapter 4.

It is entirely reasonable to suppose that the difficulties so far avoided must be hidden here. Yet the proof of this theorem is, in the mathematician's sense, an utter triviality—a straightforward computation. On the other hand, even the statement of this triviality cannot be understood without a horde of difficult definitions from Chapter 4. There are good reasons why the theorems should all be easy and the definitions hard. As the evolution of Stokes' Theorem revealed, a single simple principle can masquerade as several difficult results; the proofs of many theorems involve merely stripping away the disguise. The definitions, on the other hand, serve a twofold purpose: they are rigorous replacements for vague notions, and machinery for elegant proofs. The first two sections of Chapter 4 define precisely, and prove the rules for manipulating, what are classically described as "expressions of the form" $P\,dx + Q\,dy + R\,dz$, or $P\,dx\,dy + Q\,dy\,dz + R\,dz\,dx$. Chains, defined in the third section, and partitions of unity (already introduced in Chapter 3) free our proofs from the necessity of chopping manifolds up into small pieces; they reduce questions about manifolds, where everything seems hard, to questions about Euclidean space, where everything is easy.

Concentrating the depth of a subject in the definitions is undeniably economical, but it is bound to produce some difficulties for the student. I hope the reader will be encouraged to learn Chapter 4 thoroughly by the assurance that the results will justify the effort: the classical theorems of the last section represent only a few, and by no means the most important, applications of Chapter 4; many others appear as problems, and further developments will be found by exploring the bibliography.

The problems and the bibliography both deserve a few words. Problems appear after every section and are numbered (like the theorems) within chapters. I have starred those problems whose results are used in the text, but this precaution should be unnecessary—the problems are the most important part of the book, and the reader should at least attempt them all. It was necessary to make the bibliography either very incomplete or unwieldy, since half the major branches of mathematics could legitimately be recommended as reasonable continuations of the material in the book. I have tried to make it incomplete but tempting.

Many criticisms and suggestions were offered during the writing of this book. I am particularly grateful to Richard Palais, Hugo Rossi, Robert Seeley, and Charles Stenard for their many helpful comments.

Michael Spivak

Waltham, Massachusetts
August 1965

Contents

1

Functions on Euclidean Space

NORM AND INNER PRODUCT

Euclidean n-space $\mathbf{R}^n$ is defined as the set of all n-tuples $(x^1, \ldots, x^n)$ of real numbers x^i (a "1-tuple of numbers" is just a number and $\mathbf{R}^1 = \mathbf{R}$, the set of all real numbers). An element of $\mathbf{R}^n$ is often called a point in $\mathbf{R}^n$, and $\mathbf{R}^1$, $\mathbf{R}^2$, $\mathbf{R}^3$ are often called the line, the plane, and space, respectively. If x denotes an element of $\mathbf{R}^n$, then x is an n-tuple of numbers, the ith one of which is denoted x^i; thus we can write

$$x = (x^1, \ldots, x^n).$$

A point in $\mathbf{R}^n$ is frequently also called a vector in $\mathbf{R}^n$, because $\mathbf{R}^n$, with $x + y = (x^1 + y^1, \ldots, x^n + y^n)$ and $ax = (ax^1, \ldots, ax^n)$, as operations, *is* a vector space (over the real numbers, of dimension n). In this vector space there is the notion of the length of a vector x, usually called the **norm** $|x|$ of x and defined by $|x| = \sqrt{(x^1)^2 + \cdots + (x^n)^2}$. If $n = 1$, then $|x|$ is the usual absolute value of x. The relation between the norm and the vector space structure of $\mathbf{R}^n$ is very important.

1-1 Theorem. *If $x,y \in \mathbf{R}^n$ and $a \in \mathbf{R}$, then*
(1) $|x| \geq 0$ *and* $|x| = 0$ *if and only if* $x = 0$.
(2) $|\Sigma_{i=1}^n x^i y^i| \leq |x| \cdot |y|$; *equality holds if and only if x and y are linearly dependent.*
(3) $|x + y| \leq |x| + |y|$.
(4) $|ax| = |a| \cdot |x|$.

Proof

(1) is left to the reader.
(2) If x and y are linearly dependent, equality clearly holds. If not, then $\lambda y - x \neq 0$ for all $\lambda \in \mathbf{R}$, so

$$0 < |\lambda y - x|^2 = \sum_{i=1}^n (\lambda y^i - x^i)^2$$

$$= \lambda^2 \sum_{i=1}^n (y^i)^2 - 2\lambda \sum_{i=1}^n x^i y^i + \sum_{i=1}^n (x^i)^2.$$

Therefore the right side is a quadratic equation in λ with no real solution, and its discriminant must be negative. Thus

$$4\left(\sum_{i=1}^n x^i y^i\right)^2 - 4\sum_{i=1}^n (x^i)^2 \cdot \sum_{i=1}^n (y^i)^2 < 0.$$

(3)
$$\begin{aligned} |x + y|^2 &= \Sigma_{i=1}^n (x^i + y^i)^2 \\ &= \Sigma_{i=1}^n (x^i)^2 + \Sigma_{i=1}^n (y^i)^2 + 2\Sigma_{i=1}^n x^i y^i \\ &\leq |x|^2 + |y|^2 + 2|x| \cdot |y| \qquad \text{by (2)} \\ &= (|x| + |y|)^2. \end{aligned}$$

(4) $|ax| = \sqrt{\Sigma_{i=1}^n (ax^i)^2} = \sqrt{a^2 \Sigma_{i=1}^n (x^i)^2} = |a| \cdot |x|$. ∎

The quantity $\Sigma_{i=1}^n x^i y^i$ which appears in (2) is called the **inner product** of x and y and denoted $\langle x,y \rangle$. The most important properties of the inner product are the following.

1-2 Theorem. *If x, x_1, x_2 and y, y_1, y_2 are vectors in $\mathbf{R}^n$ and $a \in \mathbf{R}$, then*
(1) $\langle x,y \rangle = \langle y,x \rangle$ (*symmetry*).

(2) $\langle ax,y\rangle = \langle x,ay\rangle = a\langle x,y\rangle$ *(bilinearity)*.
$\langle x_1 + x_2, y\rangle = \langle x_1,y\rangle + \langle x_2,y\rangle$
$\langle x, y_1 + y_2\rangle = \langle x,y_1\rangle + \langle x,y_2\rangle$

(3) $\langle x,x\rangle \geq 0$ *and* $\langle x,x\rangle = 0$ *if and only if* $x = 0$ *(positive definiteness)*.

(4) $|x| = \sqrt{\langle x,x\rangle}$.

(5) $\langle x,y\rangle = \dfrac{|x+y|^2 - |x-y|^2}{4}$ *(polarization identity)*.

Proof

(1) $\langle x,y\rangle = \Sigma_{i=1}^n x^i y^i = \Sigma_{i=1}^n y^i x^i = \langle y,x\rangle$.

(2) By (1) it suffices to prove

$$\langle ax,y\rangle = a\langle x,y\rangle,$$
$$\langle x_1 + x_2, y\rangle = \langle x_1,y\rangle + \langle x_2,y\rangle.$$

These follow from the equations

$$\langle ax,y\rangle = \sum_{i=1}^n (ax^i)y^i = a\sum_{i=1}^n x^i y^i = a\langle x,y\rangle,$$

$$\begin{aligned}\langle x_1 + x_2, y\rangle &= \sum_{i=1}^n (x_1{}^i + x_2{}^i)y^i = \sum_{i=1}^n x_1{}^i y^i + \sum_{i=1}^n x_2{}^i y^i \\ &= \langle x_1,y\rangle + \langle x_2,y\rangle.\end{aligned}$$

(3) and (4) are left to the reader.

(5) $$\begin{aligned}&\frac{|x+y|^2 - |x-y|^2}{4} \\ &= \tfrac{1}{4}[\langle x+y, x+y\rangle^2 - \langle x-y, x-y\rangle^2] \qquad \text{by (4)} \\ &= \tfrac{1}{4}[\langle x,x\rangle + 2\langle x,y\rangle + \langle y,y\rangle - (\langle x,x\rangle - 2\langle x,y\rangle + \langle y,y\rangle)] \\ &= \langle x,y\rangle. \quad \blacksquare\end{aligned}$$

We conclude this section with some important remarks about notation. The vector $(0, \ldots ,0)$ will usually be denoted simply 0. The **usual basis** of $\mathbf{R}^n$ is $e_1, \ldots ,e_n$, where $e_i = (0, \ldots ,1, \ldots ,0)$, with the 1 in the ith place. If $T: \mathbf{R}^n \to \mathbf{R}^m$ is a linear transformation, the matrix of T with respect to the usual bases of $\mathbf{R}^n$ and $\mathbf{R}^m$ is the $m \times n$ matrix $A = (a_{ij})$, where $T(e_i) = \Sigma_{j=1}^n a_{ji}e_j$ —the coefficients of $T(e_i)$

appear in the ith *column* of the matrix. If $S\colon \mathbf{R}^m \to \mathbf{R}^p$ has the $p \times m$ matrix B, then $S \circ T$ has the $p \times n$ matrix BA [here $S \circ T(x) = S(T(x))$; most books on linear algebra denote $S \circ T$ simply ST]. To find $T(x)$ one computes the $m \times 1$ matrix

$$\begin{pmatrix} y^1 \\ \cdot \\ \cdot \\ \cdot \\ y^m \end{pmatrix} = \begin{pmatrix} a_{11}, & \ldots & ,a_{1n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_{m1}, & \ldots & ,a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x^1 \\ \cdot \\ \cdot \\ \cdot \\ x^n \end{pmatrix};$$

then $T(x) = (y^1, \ldots, y^m)$. One notational convention greatly simplifies many formulas: if $x \in \mathbf{R}^n$ and $y \in \mathbf{R}^m$, then (x,y) denotes

$$(x^1, \ldots, x^n, y^1, \ldots, y^m) \in \mathbf{R}^{n+m}.$$

Problems. **1-1.*** Prove that $|x| \le \sum_{i=1}^n |x^i|$.

1-2. When does equality hold in Theorem 1-1(3)? *Hint:* Re-examine the proof; the answer is not "when x and y are linearly dependent."

1-3. Prove that $|x - y| \le |x| + |y|$. When does equality hold?

1-4. Prove that $\big| |x| - |y| \big| \le |x - y|$.

1-5. The quantity $|y - x|$ is called the **distance** between x and y. Prove and interpret geometrically the "triangle inequality": $|z - x| \le |z - y| + |y - x|$.

1-6. Let f and g be integrable on $[a,b]$.

(a) Prove that $\left|\int_a^b f \cdot g\right| \le \left(\int_a^b f^2\right)^{\frac{1}{2}} \cdot \left(\int_a^b g^2\right)^{\frac{1}{2}}$. *Hint:* Consider separately the cases $0 = \int_a^b (f - \lambda g)^2$ for some $\lambda \in \mathbf{R}$ and $0 < \int_a^b (f - \lambda g)^2$ for all $\lambda \in \mathbf{R}$.

(b) If equality holds, must $f = \lambda g$ for some $\lambda \in \mathbf{R}$? What if f and g are continuous?

(c) Show that Theorem 1-1(2) is a special case of (a).

1-7. A linear transformation $T\colon \mathbf{R}^n \to \mathbf{R}^n$ is **norm preserving** if $|T(x)| = |x|$, and **inner product preserving** if $\langle Tx, Ty \rangle = \langle x, y \rangle$.

(a) Prove that T is norm preserving if and only if T is inner-product preserving.

(b) Prove that such a linear transformation T is 1-1 and T^{-1} is of the same sort.

1-8. If $x,y \in \mathbf{R}^n$ are non-zero, the **angle** between x and y, denoted $\angle(x,y)$, is defined as $\arccos(\langle x,y \rangle / |x| \cdot |y|)$, which makes sense by Theorem 1-1(2). The linear transformation T is **angle preserving** if T is 1-1, and for $x,y \ne 0$ we have $\angle(Tx,Ty) = \angle(x,y)$.

(a) Prove that if T is norm preserving, then T is angle preserving.

(b) If there is a basis $x_1, \ldots, x_n$ of $\mathbf{R}^n$ and numbers $\lambda_1, \ldots, \lambda_n$ such that $Tx_i = \lambda_i x_i$, prove that T is angle preserving if and only if all λ_i are equal.

(c) What are all angle preserving $T: \mathbf{R}^n \to \mathbf{R}^n$?

1-9. If $0 \leq \theta < \pi$, let $T: \mathbf{R}^2 \to \mathbf{R}^2$ have the matrix $\begin{pmatrix} \sin\theta, & \cos\theta \\ -\cos\theta, & \sin\theta \end{pmatrix}$. Show that T is angle preserving and if $x \neq 0$, then $\angle(x,Tx) = \theta$.

1-10.* If $T: \mathbf{R}^m \to \mathbf{R}^n$ is a linear transformation, show that there is a number M such that $|T(h)| \leq M|h|$ for $h \in \mathbf{R}^m$. *Hint:* Estimate $|T(h)|$ in terms of $|h|$ and the entries in the matrix of T.

1-11. If $x,y \in \mathbf{R}^n$ and $z,w \in \mathbf{R}^m$, show that $\langle (x,z),(y,w)\rangle = \langle x,y\rangle + \langle z,w\rangle$ and $|(x,z)| = \sqrt{|x|^2 + |z|^2}$. Note that (x,z) and (y,w) denote points in $\mathbf{R}^{n+m}$.

1-12.* Let $(\mathbf{R}^n)^*$ denote the dual space of the vector space $\mathbf{R}^n$. If $x \in \mathbf{R}^n$, define $\varphi_x \in (\mathbf{R}^n)^*$ by $\varphi_x(y) = \langle x,y\rangle$. Define $T: \mathbf{R}^n \to (\mathbf{R}^n)^*$ by $T(x) = \varphi_x$. Show that T is a 1-1 linear transformation and conclude that every $\varphi \in (\mathbf{R}^n)^*$ is φ_x for a unique $x \in \mathbf{R}^n$.

1-13.* If $x,y \in \mathbf{R}^n$, then x and y are called **perpendicular** (or **orthogonal**) if $\langle x,y\rangle = 0$. If x and y are perpendicular, prove that $|x+y|^2 = |x|^2 + |y|^2$.

SUBSETS OF EUCLIDEAN SPACE

The closed interval $[a,b]$ has a natural analogue in $\mathbf{R}^2$. This is the **closed rectangle** $[a,b] \times [c,d]$, defined as the collection of all pairs (x,y) with $x \in [a,b]$ and $y \in [c,d]$. More generally, if $A \subset \mathbf{R}^m$ and $B \subset \mathbf{R}^n$, then $A \times B \subset \mathbf{R}^{m+n}$ is defined as the set of all $(x,y) \in \mathbf{R}^{m+n}$ with $x \in A$ and $y \in B$. In particular, $\mathbf{R}^{m+n} = \mathbf{R}^m \times \mathbf{R}^n$. If $A \subset \mathbf{R}^m$, $B \subset \mathbf{R}^n$, and $C \subset \mathbf{R}^p$, then $(A \times B) \times C = A \times (B \times C)$, and both of these are denoted simply $A \times B \times C$; this convention is extended to the product of any number of sets. The set $[a_1,b_1] \times \cdots \times [a_n,b_n] \subset \mathbf{R}^n$ is called a **closed rectangle** in $\mathbf{R}^n$, while the set $(a_1,b_1) \times \cdots \times (a_n,b_n) \subset \mathbf{R}^n$ is called an **open rectangle.** More generally a set $U \subset \mathbf{R}^n$ is called **open** (Figure 1-1) if for each $x \in U$ there is an open rectangle A such that $x \in A \subset U$.

A subset C of $\mathbf{R}^n$ is **closed** if $\mathbf{R}^n - C$ is open. For example, if C contains only finitely many points, then C is closed.

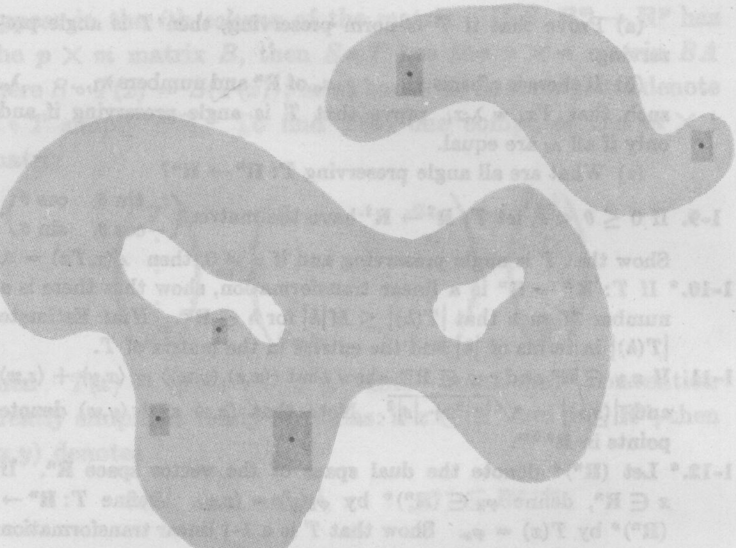

FIGURE 1-1

The reader should supply the proof that a closed rectangle in $\mathbf{R}^n$ is indeed a closed set.

If $A \subset \mathbf{R}^n$ and $x \in \mathbf{R}^n$, then one of three possibilities must hold (Figure 1-2):

1. There is an open rectangle B such that $x \in B \subset A$.
2. There is an open rectangle B such that $x \in B \subset \mathbf{R}^n - A$.
3. If B is any open rectangle with $x \in B$, then B contains points of both A and $\mathbf{R}^n - A$.

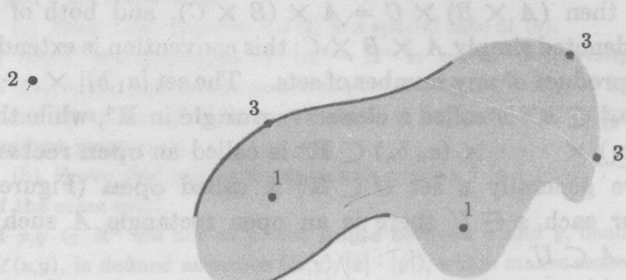

FIGURE 1-2

Those points satisfying (1) constitute the **interior** of A, those satisfying (2) the **exterior** of A, and those satisfying (3) the **boundary** of A. Problems 1-16 to 1-18 show that these terms may sometimes have unexpected meanings.

It is not hard to see that the interior of any set A is open, and the same is true for the exterior of A, which is, in fact, the interior of $\mathbf{R}^n - A$. Thus (Problem 1-14) their union is open, and what remains, the boundary, must be closed.

A collection $\mathcal{O}$ of open sets is an **open cover** of A (or, briefly, **covers** A) if every point $x \in A$ is in some open set in the collection $\mathcal{O}$. For example, if $\mathcal{O}$ is the collection of all open intervals $(a, a + 1)$ for $a \in \mathbf{R}$, then $\mathcal{O}$ is a cover of $\mathbf{R}$. Clearly no finite number of the open sets in $\mathcal{O}$ will cover $\mathbf{R}$ or, for that matter, any unbounded subset of $\mathbf{R}$. A similar situation can also occur for bounded sets. If $\mathcal{O}$ is the collection of all open intervals $(1/n, 1 - 1/n)$ for all positive integers n, then $\mathcal{O}$ is an open cover of $(0,1)$, but again no finite collection of sets in $\mathcal{O}$ will cover $(0,1)$. Although this phenomenon may not appear particularly scandalous, sets for which this state of affairs cannot occur are of such importance that they have received a special designation: a set A is called **compact** if every open cover $\mathcal{O}$ contains a finite subcollection of open sets which also covers A.

A set with only finitely many points is obviously compact and so is the infinite set A, which contains 0 and the numbers $1/n$ for all integers n (reason: if $\mathcal{O}$ is a cover, then $0 \in U$ for some open set U in $\mathcal{O}$; there are only finitely many other points of A not in U, each requiring at most one more open set).

Recognizing compact sets is greatly simplified by the following results, of which only the first has any depth (i.e., uses any facts about the real numbers).

1-3 Theorem (Heine-Borel). *The closed interval* $[a,b]$ *is compact.*

Proof. If $\mathcal{O}$ is an open cover of $[a,b]$, let

$$A = \{x: a \le x \le b \text{ and } [a,x] \text{ is covered by some finite number of open sets in } \mathcal{O}\}.$$