PEARSON NEW INTERNATIONAL EDITION

Introduction to Analysis

William R. Wade

Fourth Edition

ALWAYS LEARNING™

PEARSON®

Pearson New International Edition

Introduction to Analysis

William R. Wade

Fourth Edition

PEARSON®

Pearson Education Limited
Edinburgh Gate
Harlow
Essex CM20 2JE
England and Associated Companies throughout the world

Visit us on the World Wide Web at: www.pearsoned.co.uk

PEARSON®

ISBN 10: 1-292-03932-9
ISBN 13: 978-1-292-03932-9

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Printed in the United States of America

Table of Contents

CHAPTER 1

The Real Number System

You have already had several calculus courses in which you evaluated limits, differentiated functions, and computed integrals. You may even remember some of the major results of calculus, such as the Chain Rule, the Mean Value Theorem, and the Fundamental Theorem of Calculus. Although you are probably less familiar with multivariable calculus, you have taken partial derivatives, computed gradients, and evaluated certain line and surface integrals.

In view of all this, you must be asking: Why another course in calculus? The answer to this question is twofold. Although some proofs may have been presented in earlier courses, it is unlikely that the subtler points (e.g., completeness of the real numbers, uniform continuity, and uniform convergence) were covered. Moreover, the skills you acquired were mostly computational; you were rarely asked to prove anything yourself. This course develops the theory of calculus carefully and rigorously from basic principles and gives you a chance to learn how to construct your own proofs. It also serves as an introduction to analysis, an important branch of mathematics which provides a foundation for numerical analysis, functional analysis, harmonic analysis, differential equations, differential geometry, real analysis, complex analysis, and many other areas of specialization within mathematics.

1.1 INTRODUCTION

Every rigorous study of mathematics begins with certain undefined concepts, primitive notions on which the theory is based, and certain postulates, properties which are assumed to be true and given no proof. Our study will be based on the primitive notions of real numbers and sets, which will be discussed in this section.

We shall use standard notation for sets and real numbers. For example, **R** or $(-\infty, \infty)$ represents the set of *real numbers*, $\emptyset$ represents the *empty set* (the set with no elements), $a \in A$ means that a is an *element of* A, and $a \notin A$ means that a is *not* an element of A. We can represent a given finite set in two ways. We can list its elements directly, or we can describe it using sentences or equations. For example, the set of solutions to the equation $x^2 = 1$ can be written as

$$\{1, -1\} \quad \text{or} \quad \{x : x^2 = 1\}.$$

A set A is said to be a *subset* of a set B (notation: $A \subseteq B$) if and only if every element of A is also an element of B. If A is a subset of B but there is at least one element $b \in B$ that does not belong to A, we shall call A a *proper subset* of B (notation: $A \subset B$). Two sets A and B are said to be *equal* (notation: $A = B$)

From Chapter 1 of *Introduction to Analysis*, Fourth Edition. William R. Wade.

if and only if $A \subseteq B$ and $B \subseteq A$. If A and B are not equal, we write $A \neq B$. A set A is said to be *nonempty* if and only if $A \neq \emptyset$.

The *union* of two sets A and B (notation: $A \cup B$) is the set of elements x such that x belongs to A or B or both. The *intersection* of two sets A and B (notation: $A \cap B$) is the set of elements x such that x belongs to both A and B. The *complement* of B *relative to* A (notation: $A \setminus B$, sometimes B^c if A is understood) is the set of elements x such that x belongs to A but does not belong to B. For example,

$$\{-1, 0, 1\} \cup \{1, 2\} = \{-1, 0, 1, 2\}, \qquad \{-1, 0, 1\} \cap \{1, 2\} = \{1\},$$
$$\{1, 2\} \setminus \{-1, 0, 1\} = \{2\} \quad \text{and} \quad \{-1, 0, 1\} \setminus \{1, 2\} = \{-1, 0\}.$$

Let X and Y be sets. The *Cartesian product* of X and Y is the set of *ordered pairs* defined by

$$X \times Y := \{(x, y) : x \in X, y \in Y\}.$$

(The symbol := means "equal by definition" or "is defined to be.") Two points (x, y), $(z, w) \in X \times Y$ are said to be *equal* if and only if $x = z$ and $y = w$.

Let X and Y be sets. A *relation* on $X \times Y$ is any subset of $X \times Y$. Let $\mathcal{R}$ be a relation on $X \times Y$. The *domain* of $\mathcal{R}$ is the collection of $x \in X$ such that (x, y) belongs to $\mathcal{R}$ for some $y \in Y$. The *range* of $\mathcal{R}$ is the collection of $y \in Y$ such that (x, y) belongs to $\mathcal{R}$ for some $x \in X$. When $(x, y) \in \mathcal{R}$, we shall frequently write $x\mathcal{R}y$.

A *function* f from X into Y (notation: $f : X \to Y$) is a relation on $X \times Y$ whose domain is X (notation: $\text{Dom}(f) := X$) such that for each $x \in X$ there is a *unique* (one and only one) $y \in Y$ that satisfies $(x, y) \in f$. If $(x, y) \in f$, we shall call y the *value* of f at x (notation: $y = f(x)$ or $f : x \longmapsto y$) and call x a *preimage* of y under f. We said *a* preimage because, in general, a point in the range of f might have more than one preimage. For example, since $\sin(k\pi) = 0$ for $k = 0, \pm 1, \pm 2, \dots$, the value 0 has infinitely many preimages under $f(x) = \sin x$.

If f is a function from X into Y, we will say that f is *defined* on X and call Y the *codomain* of f. The *range* of f is the collection of all values of f; that is, the set $\text{Ran}(f) := \{y \in Y : f(x) = y \text{ for some } x \in X\}$. In general, then, the range of a function is a subset of its codomain and each $y \in \text{Ran}(f)$ has one or more preimages. If $\text{Ran}(f) = Y$ and each $y \in Y$ has exactly one preimage, $x \in X$, under f, then we shall say that $f : X \to Y$ *has an inverse*, and shall define the *inverse function* $f^{-1} : Y \to X$ by $f^{-1}(y) := x$, where $x \in X$ satisfies $f(x) = y$.

At this point it is important to notice a consequence of defining a function to be a set of ordered pairs. By the definition of equality of ordered pairs, two functions f and g are equal if and only if they have the same domain, and same values; that is, $f, g : X \to Y$, and $f(x) = g(x)$ for all $x \in X$. If they have different domains, they are different functions.

For example, let $f(x) = g(x) = x^2$. Then $f : [0, 1) \to [0, 1)$ and $g : (-1, 1) \to [0, 1)$ are two different functions. They both have the same range, $[0, 1)$, but each $y \in (0, 1)$ has exactly one preimage under f, namely $\sqrt{y}$, and two preimages under g, namely $\pm\sqrt{y}$. In particular, f has an inverse function, $f^{-1}(x) = \sqrt{x}$, but g does not. Making distinctions like this will actually make our life easier later in the course.

For the first half of this course, most of the concrete functions we consider will be *real-valued functions of a real variable* (i.e., functions whose domains and ranges are subsets of $\mathbf{R}$). We shall often call such functions simply *real functions*.

You are already familiar with many real functions.

1) The *exponential function* $e^x : \mathbf{R} \to (0, \infty)$ and its inverse function, the *natural logarithm*

$$\log x := \int_1^x \frac{dt}{t},$$

defined and real-valued for each $x \in (0, \infty)$. (Although this last function is denoted by $\ln x$ in elementary calculus texts, most analysts denote it, as we did just now, by $\log x$. We will follow this practice throughout this text. For a more constructive definition, see Exercise 4.5.5.)
2) The *trigonometric functions* (whose formulas are) represented by $\sin x$, $\cos x$, $\tan x$, $\cot x$, $\sec x$, and $\csc x$, and the inverse trigonometric functions $\arcsin x$, $\arccos x$, and $\arctan x$ whose ranges are, respectively, $[-\pi/2, \pi/2]$, $[0, \pi]$, and $(-\pi/2, \pi/2)$.
3) The *power functions* x^α, which can be defined constructively (see Appendix A.10 and Exercise 3.3.11) or by using the exponential function:

$$x^\alpha := e^{\alpha \log x}, \qquad x > 0, \quad \alpha \in \mathbf{R}.$$

We assume that you are familiar with the various algebraic laws and identities that these functions satisfy. A list of the most widely used trigonometric identities can be found in Appendix B. The most widely used properties of the power functions are $x^0 = 1$ for all $x \neq 0$; $x^n = x \cdot \ldots \cdot x$ (there are n factors here) when $n = 1, 2, \ldots$ and $x \in \mathbf{R}$; $x^\alpha > 0$, $x^\alpha \cdot x^\beta = x^{\alpha+\beta}$, and $(x^\alpha)^\beta = x^{\alpha \cdot \beta}$ for all $x > 0$ and $\alpha, \beta \in \mathbf{R}$; $x^\alpha = \sqrt[m]{x}$ when $\alpha = 1/m$ for some $m \in \mathbf{N}$ and the indicated root exists and is real; and $0^\alpha := 0$ for all $\alpha > 0$. (The symbol 0^0 is left undefined because it is indeterminate [see Example 4.31].)

We also assume that you can differentiate algebraic combinations of these functions using the basic formulas $(\sin x)' = \cos x$, $(\cos x)' = -\sin x$, and $(e^x)' = e^x$, for $x \in \mathbf{R}$; $(\log x)' = 1/x$ and $(x^\alpha)' = \alpha x^{\alpha-1}$, for $x > 0$ and $\alpha \in \mathbf{R}$; and

$$(\tan x)' = \sec^2 x \quad \text{for } x \neq \frac{(2n+1)\pi}{2}, \quad n \in \mathbf{Z}.$$

(You will have an opportunity to develop some of these rules in the exercises, e.g., see Exercises 4.2.9, 4.4.6, 4.5.3, 5.3.7, and 5.3.8.) Even with these assumptions, we shall repeat some material from elementary calculus.

We mentioned postulates in the opening paragraph. In the next two sections, we will introduce three postulates (containing a total of 13 different properties) which characterize the set of real numbers. Although you are probably already familiar with all but the last of these properties, we will use them to prove other equally familiar properties (e.g., in Example 1.4 we will prove that if $a \neq 0$, then $a^2 > 0$).

Why would we assume some properties and prove others? At one point, mathematicians thought that all laws about real numbers were of equal weight. Gradually, during the late 1800s, we discovered that many of the well-known laws satisfied by **R** are in fact consequences of others. The net result of this research is that the 13 properties listed below are considered to be fundamental properties describing **R**. All other laws satisfied by real numbers are secondary in the sense that they can be proved using these fundamental properties.

Why would we prove a law that is well known, perhaps even "obvious"? Why not just assume all known properties about **R** and proceed from there? We want this book to be reasonably self-contained, because this will make it easier for you to begin to construct your own proofs. We want the first proofs you see to be easily understood, because they deal with familiar properties that are unobscured by new concepts. But most importantly, we want to form a habit of proving all statements, even seemingly "obvious" statements.

The reason for this hard-headed approach is that some "obvious" statements are false. For example, divide an 8×8-inch square into triangles and trapezoids as shown on the left side of Figure 1.1. Since the 3-inch sides of the triangles perfectly match the 3-inch sides of the trapezoids, it is "obvious" that these triangles and trapezoids can be reassembled into a rectangle (see the right side of Figure 1.1). Or is it? The area of the square is $8 \times 8 = 64$ square inches but the area of the rectangle is $5 \times 13 = 65$ square inches. Since you cannot increase area by reassembling pieces, what looked right was in fact wrong. By computing slopes, you can verify that the rising diagonal on the right side of Figure 1.1 is, in fact, four distinct line segments that form a long narrow diamond which conceals that extra one square inch.

NOTE: *Reading a mathematics book is different from reading any other kind of book. When you see phrases like "you can verify" or "it is easy to see," you should use pencil and paper to do the calculations to be sure what we've said is correct.*

Here is another example. Grab a calculator and graph the functions $y = \log x$ and $y = \sqrt[100]{x}$. It is easy to see, using calculus, that $\log x$ and $\sqrt[100]{x}$ are both increasing and concave downward on $[0, \infty)$. Looking at the graphs (see Figure 1.2), it's "obvious" that $\log x$ is much larger than $\sqrt[100]{x}$ no matter how big x is. Or is it? Let's evaluate each function at e^{1000}: $\log(e^{1000}) = 1000 \log e = 1000$ is much smaller than $\sqrt[100]{e^{1000}} = e^{10} \approx 22,000$. Evidently, the graph of $y = \sqrt[100]{x}$

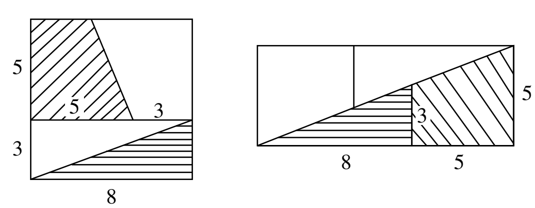

FIGURE 1.1

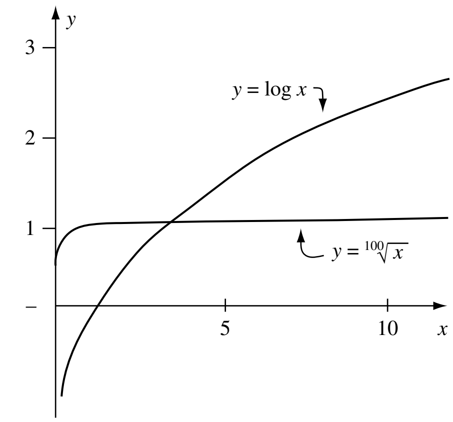

FIGURE 1.2

eventually crosses that of $y = \log x$. With a little calculus, you can prove that $\log x < \sqrt[100]{x}$ forever after that (see Exercise 4.4.6a).

What can be learned from these examples? We cannot always trust what we think we see. We must, as above, find some mathematical way of testing our perception, either verifying that it is correct, or rejecting it as wrong. This type of phenomenon is not a rare occurrence. You will soon encounter several other plausible statements that are, in fact, false. In particular, you must harbor a skepticism that demands proofs of all statements not assumed in postulates, even the "obvious" ones.

What, then, are you allowed to use when solving the exercises? You may use any property of real numbers (e.g., $2+3 = 5$, $2 < 7$, or $\sqrt{2}$ is irrational) without reference or proof. You may use any algebraic property of real numbers involving equal signs [e.g., $(x + y)^2 = x^2 + 2xy + y^2$ or $(x + y)(x - y) = x^2 - y^2$] and the techniques of calculus to find local maxima or minima of a given function without reference or proof. After completing the exercises in Section 1.2, you may also use any algebraic property of real numbers involving inequalities (e.g., $0 < a < b$ implies $0 < a^x < b^x$ for all $x > 0$) without reference or proof.

1.2 ORDERED FIELD AXIOMS

In this section we explore the algebraic structure of the real number system. We shall assume that the set of real numbers, **R**, is a field (i.e., that **R** satisfies the following postulate).

Postulate 1. [FIELD AXIOMS]. There are functions $+$ and $\cdot$, defined on $\mathbf{R}^2 := \mathbf{R} \times \mathbf{R}$, which satisfy the following properties for every $a, b, c \in \mathbf{R}$:

Closure Properties. $a + b$ and $a \cdot b$ belong to **R**.

Associative Properties. $a + (b + c) = (a + b) + c$ and $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

Commutative Properties. $a + b = b + a$ and $a \cdot b = b \cdot a$.

Distributive Law. $a \cdot (b + c) = a \cdot b + a \cdot c$.

Existence of the Additive Identity. There is a unique element $0 \in \mathbf{R}$ such that $0 + a = a$ for all $a \in \mathbf{R}$.

Existence of the Multiplicative Identity. There is a unique element $1 \in \mathbf{R}$ such that $1 \neq 0$ and $1 \cdot a = a$ for all $a \in \mathbf{R}$.

Existence of Additive Inverses. For every $x \in \mathbf{R}$ there is a unique element $-x \in \mathbf{R}$ such that

$$x + (-x) = 0.$$

Existence of Multiplicative Inverses. For every $x \in \mathbf{R} \setminus \{0\}$ there is a unique element $x^{-1} \in \mathbf{R}$ such that

$$x \cdot (x^{-1}) = 1.$$

We note in passing that the word *unique* can be dropped from the statements in Postulate 1 (see Appendix A).

We shall usually denote $a + (-b)$ by $a - b$, $a \cdot b$ by ab, a^{-1} by $\frac{1}{a}$ or $1/a$, and $a \cdot b^{-1}$ by $\frac{a}{b}$ or a/b. Notice that by the existence of additive and multiplicative inverses, the equation $x + a = 0$ can be solved for each $a \in \mathbf{R}$, and the equation $ax = 1$ can be solved for each $a \in \mathbf{R}$ provided that $a \neq 0$.

From these few properties (i.e., from Postulate 1), we can derive all the usual algebraic laws of real numbers, including the following:

$$(-1)^2 = 1, \tag{1}$$

$$0 \cdot a = 0, \quad -a = (-1) \cdot a, \quad -(-a) = a, \qquad a \in \mathbf{R}, \tag{2}$$

$$-(a - b) = b - a, \qquad a, b \in \mathbf{R}, \tag{3}$$

and

$$a, b \in \mathbf{R} \quad \text{and} \quad ab = 0 \quad \text{imply} \quad a = 0 \quad \text{or} \quad b = 0. \tag{4}$$

We want to keep our attention sharply focused on analysis. Since the proofs of algebraic laws like these lie more in algebra than analysis (see Appendix A), we will not present them here. In fact, with the exception of the absolute value and the Binomial Formula, we will assume all material usually presented in a high school algebra course (including the quadratic formula and graphs of the conic sections).

Postulate 1 is sufficient to derive all algebraic laws of **R**, but it does not completely describe the real number system. The set of real numbers also has an order relation (i.e., a concept of "less than").

Postulate 2. [Order Axioms]. There is a relation $<$ on $\mathbf{R} \times \mathbf{R}$ that has the following properties:

Trichotomy Property. Given $a, b \in \mathbf{R}$, one and only one of the following statements holds:

$$a < b, \quad b < a, \quad \text{or} \quad a = b.$$

Transitive Property. For $a, b, c \in \mathbf{R}$,

$$a < b \quad \text{and} \quad b < c \quad \text{imply} \quad a < c.$$

The Additive Property. For $a, b, c \in \mathbf{R}$,

$$a < b \quad \text{and} \quad c \in \mathbf{R} \quad \text{imply} \quad a + c < b + c.$$

The Multiplicative Properties. For $a, b, c \in \mathbf{R}$,

$$a < b \quad \text{and} \quad c > 0 \quad \text{imply} \quad ac < bc$$

and

$$a < b \quad \text{and} \quad c < 0 \quad \text{imply} \quad bc < ac.$$

By $b > a$ we shall mean $a < b$. By $a \le b$ and $b \ge a$ we shall mean $a < b$ or $a = b$. By $a < b < c$ we shall mean $a < b$ and $b < c$. In particular, $2 < x < 1$ makes no sense at all.

WARNING. *There are two Multiplicative Properties, so every time you multiply an inequality by an expression, you must carefully note the sign of that expression and adjust the inequality accordingly*. For example, $x < 1$ does NOT imply that $x^2 < x$ unless $x > 0$. If $x < 0$, then by the Second Multiplicative Property, $x < 1$ implies $x^2 > x$.

We shall call a number $a \in \mathbf{R}$ *nonnegative* if $a \ge 0$ and *positive* if $a > 0$. Postulate 2 has a slightly simpler formulation using the set of positive elements as a primitive concept (see Exercise 1.2.11). We have introduced Postulate 2 as above because these are the properties we use most often.

The real number system $\mathbf{R}$ contains certain special subsets: the set of *natural numbers*

$$\mathbf{N} := \{1, 2, \dots\},$$

obtained by beginning with 1 and successively adding 1s to form $2 := 1+1$, $3 := 2 + 1$, and so on; the set of *integers*

$$\mathbf{Z} := \{\dots, -2, -1, 0, 1, 2, \dots\}$$

(*Zahl* is German for number); the set of *rationals* (or fractions or quotients)

$$\mathbf{Q} := \left\{\frac{m}{n} : m, n \in \mathbf{Z} \text{ and } n \neq 0\right\};$$

and the set of *irrationals*

$$\mathbf{Q}^c = \mathbf{R} \setminus \mathbf{Q}.$$

Equality in $\mathbf{Q}$ is defined by

$$\frac{m}{n} = \frac{p}{q} \quad \text{if and only if} \quad mq = np.$$

Recall that each of the sets **N, Z, Q**, and **R** is a proper subset of the next; that is,

$$\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R}.$$

For example, every rational is a real number (because $m/n := mn^{-1}$ is a real number by Postulate 1), but $\sqrt{2}$ is an irrational.

Since we did not really define **N** and **Z**, we must make certain assumptions about them. If you are interested in the definitions and proofs, see Appendix A.

1.1 Remark. We will assume that the sets **N** and **Z** satisfy the following properties.

i) If $n, m \in \mathbf{Z}$, then $n + m$, $n - m$, and mn belong to **Z**.
ii) If $n \in \mathbf{Z}$, then $n \in \mathbf{N}$ if and only if $n \geq 1$.
iii) There is no $n \in \mathbf{Z}$ that satisfies $0 < n < 1$.

Using these properties, we can prove that **Q** satisfies Postulate 1 (see Exercise 1.2.9).

We notice in passing that none of the other special subsets of **R** satisfies Postulate 1. **N** satisfies all but three of the properties in Postulate 1: **N** has no additive identity (since $0 \notin \mathbf{N}$), **N** has no additive inverses (e.g., $-1 \notin \mathbf{N}$), and only one of the nonzero elements of **N** (namely, 1) has a multiplicative inverse. **Z** satisfies all but one of the properties in Postulate 1: Only two nonzero elements of **Z** have multiplicative inverses (namely, 1 and -1). $\mathbf{Q}^c$ satisfies all but four of the properties in Postulate 1: $\mathbf{Q}^c$ does not have an additive identity (since $0 \notin \mathbf{R} \setminus \mathbf{Q}$), does not have a multiplicative identity (since $1 \notin \mathbf{R} \setminus \mathbf{Q}$), and does not satisfy either closure property. Indeed, since $\sqrt{2}$ is irrational, the sum of irrationals may be rational ($\sqrt{2}+(-\sqrt{2}) = 0$) and the product of irrationals may be rational ($\sqrt{2} \cdot \sqrt{2} = 2$).

Notice that any subset of **R** satisfies Postulate 2. Thus **Q** satisfies both Postulates 1 and 2. The remaining postulate, introduced in Section 1.3, identifies a property that **Q** does not possess. In particular, Postulates 1 through 3 distinguish **R** from each of its special subsets **N, Z, Q**, and $\mathbf{Q}^c$. These postulates actually characterize **R**; that is, **R** is the only set that satisfies Postulates 1 through 3. (Such a set is called a *complete Archimedean ordered field*. We may as well admit a certain arbitrariness in choosing this approach. **R** has been developed axiomatically in at least five other ways [e.g., as a one-dimensional continuum or as a set of binary decimals with certain arithmetic operations]. The decision to present **R** using Postulates 1 through 3 is based partly on economy and partly on personal taste.)

Postulates 1 and 2 can be used to derive all identities and inequalities which are true for real numbers [e.g., see implications (5) through (9) below]. Since arguments based on inequalities are of fundamental importance to analysis, we begin to supply details of proofs at this stage.

What is a proof? Every mathematical result (for us this includes examples, remarks, lemmas, and theorems) has hypotheses and a conclusion. There are three main methods of proof: mathematical induction, direct deduction, and contradiction.

Mathematical induction, a special method for proving statements that depend on positive integers, will be covered in Section 1.4.

To construct a *deductive proof*, we assume the hypotheses to be true and proceed step by step to the conclusion. Each step is justified by a hypothesis, a definition, a postulate, or a mathematical result that has already been proved. (Actually, this is usually the way we write a proof. When constructing your own proofs, you may find it helpful to work forward from the hypotheses as far as you can and then work backward from the conclusion, trying to meet in the middle.)

To construct a *proof by contradiction*, we assume the hypotheses to be true, the conclusion to be false, and work step by step deductively until a *contradiction* occurs; that is, a statement that is obviously false or that is contrary to the assumptions made. At this point the proof by contradiction is complete. The phrase "suppose to the contrary" always indicates a proof by contradiction (e.g., see the proof of Theorem 1.9).

What about false statements? How do we "prove" that a statement is false? We can show that a statement is false by producing a single, concrete example (called a *counterexample*) that satisfies the hypotheses but not the conclusion of that statement. For example, to show that the statement "$x > 1$ implies $x^2 - x - 2 \neq 0$" is false, we need only observe that $x = 2$ is greater than 1 but $2^2 - 2 - 2 = 0$.

Here are some examples of deductive proofs. (*Note*: The symbol ■ indicates that the proof or solution is complete.)

1.2 *EXAMPLE.*

If $a \in \mathbf{R}$, prove that

$$a \neq 0 \quad \text{implies} \quad a^2 > 0. \tag{5}$$

In particular, $-1 < 0 < 1$.

Proof. Suppose that $a \neq 0$. By the Trichotomy Property, either $a > 0$ or $a < 0$.

Case 1. $a > 0$. Multiply both sides of this inequality by a, using the First Multiplicative Property. We obtain $a^2 = a \cdot a > 0 \cdot a$. Since (by (2)), $0 \cdot a = 0$ we conclude that $a^2 > 0$.

Case 2. $a < 0$. Multiply both sides of this inequality by a. Since $a < 0$, it follows from the Second Multiplicative Property that $a^2 = a \cdot a > 0 \cdot a = 0$. This proves that $a^2 > 0$ when $a \neq 0$.

Since $1 \neq 0$, it follows that $1 = 1^2 > 0$. Adding -1 to both sides of this inequality, we conclude that $0 = 1 - 1 > 0 - 1 = -1$. ■

1.3 *EXAMPLE.*

If $a \in \mathbf{R}$, prove that

$$0 < a < 1 \quad \text{implies} \quad 0 < a^2 < a \quad \text{and} \quad a > 1 \quad \text{implies} \quad a^2 > a. \tag{6}$$

Proof. Suppose that $0 < a < 1$. Multiply both sides of this inequality by a using the First Multiplicative Property. We obtain $0 = 0 \cdot a < a^2 < 1 \cdot a = a$. In particular, $0 < a^2 < a$.

On the other hand, if $a > 1$, then $a > 0$ by Example 1.2 and the Transitive Property. Multiplying $a > 1$ by a, we conclude that $a^2 = a \cdot a > 1 \cdot a = a$. ∎

Similarly (see Exercise 1.2.2), we can prove that

$$0 \le a < b \quad \text{and} \quad 0 \le c < d \quad \text{imply} \quad ac < bd, \tag{7}$$

$$0 \le a < b \quad \text{implies} \quad 0 \le a^2 < b^2 \quad \text{and} \quad 0 \le \sqrt{a} < \sqrt{b}, \tag{8}$$

and

$$0 < a < b \quad \text{implies} \quad \frac{1}{a} > \frac{1}{b} > 0. \tag{9}$$

Much of analysis deals with estimation (of error, of growth, of volume, etc.) in which these inequalities and the following concept play a central role.

1.4 Definition.

The *absolute value* of a number $a \in \mathbf{R}$ is the number

$$|a| := \begin{cases} a & a \ge 0 \\ -a & a < 0. \end{cases}$$

When proving results about the absolute value, we can always break the proof up into several cases, depending on when the parameters are positive, negative, or zero. Here is a typical example.

1.5 Remark. *The absolute value is multiplicative; that is*, $|ab| = |a|\,|b|$ *for all* $a, b \in \mathbf{R}$.

Proof. We consider four cases.

Case 1. $a = 0$ or $b = 0$. Then $ab = 0$, so by definition, $|ab| = 0 = |a|\,|b|$.

Case 2. $a > 0$ and $b > 0$. By the First Multiplicative Property, $ab > 0 \cdot b = 0$. Hence by definition, $|ab| = ab = |a|\,|b|$.

Case 3. $a > 0$ and $b < 0$, or, $b > 0$ and $a < 0$. By symmetry, we may suppose that $a > 0$ and $b < 0$. (That is, if we can prove it for $a > 0$ and $b < 0$, then by reversing the roles of a and b, we can prove it for $a < 0$ and $b > 0$.) By the Second Multiplicative Property, $ab < 0$. Hence by Definition 1.4, (2), associativity, and commutativity,

$$|ab| = -(ab) = (-1)(ab) = a((-1)b) = a(-b) = |a|\,|b|.$$

Case 4. $a < 0$ and $b < 0$. By the Second Multiplicative Property, $ab > 0$. Hence by Definition 1.4,

$$|ab| = ab = (-1)^2(ab) = (-a)(-b) = |a|\,|b|. \qquad \blacksquare$$

We shall soon see that there are more efficient ways to prove results about absolute values than breaking the argument into cases.

The following result is useful when solving inequalities involving absolute value signs.

1.6 Theorem. **[Fundamental Theorem of Absolute Values].** *Let $a \in \mathbf{R}$ and $M \geq 0$. Then $|a| \leq M$ if and only if $-M \leq a \leq M$.*

Proof. Suppose first that $|a| \leq M$. Multiplying by –1, we also have $-|a| \geq -M$.

Case 1. $a \geq 0$. By Definition 1.4, $|a| = a$. Thus by hypothesis,

$$-M \leq 0 \leq a = |a| \leq M.$$

Case 2. $a < 0$. By Definition 1.4, $|a| = -a$. Thus by hypothesis,

$$-M \leq -|a| = a < 0 \leq M.$$

This proves that $-M \leq a \leq M$ in either case.

Conversely, if $-M \leq a \leq M$, then $a \leq M$ and $-M \leq a$. Multiplying the second inequality by -1, we have $-a \leq M$. Consequently, $|a| = a \leq M$ if $a \geq 0$, and $|a| = -a \leq M$ if $a < 0$. $\blacksquare$

NOTE: In a similar way we can prove that $|a| < M$ if and only if $-M < a < M$.

Here is another useful result about absolute values.

1.7 Theorem. *The absolute value satisfies the following three properties.*

i) [Positive Definite] *For all $a \in \mathbf{R}$, $|a| \geq 0$ with $|a| = 0$ if and only if $a = 0$.*
ii) [Symmetric] *For all $a, b \in \mathbf{R}$, $|a - b| = |b - a|$.*
iii) [Triangle Inequalities] *For all $a, b \in \mathbf{R}$,*

$$|a + b| \leq |a| + |b| \quad \textit{and} \quad \big|\, |a| - |b| \,\big| \leq |a - b|.$$

Proof. i) If $a \geq 0$, then $|a| = a \geq 0$. If $a < 0$, then by Definition 1.4 and the Second Multiplicative Property, $|a| = -a = (-1)a > 0$. Thus $|a| \geq 0$ for all $a \in \mathbf{R}$.

If $|a| = 0$, then by definition $a = |a| = 0$ when $a \geq 0$ and $a = -|a| = 0$ when $a < 0$. Thus $|a| = 0$ implies that $a = 0$. Conversely, $|0| = 0$ by definition.

ii) By Remark 1.5, $|a - b| = |-1|\,|b - a| = |b - a|$.

iii) To prove the first inequality, notice that $|x| \le |x|$ holds for any $x \in \mathbf{R}$. Thus Theorem 1.6 implies that $-|a| \le a \le |a|$ and $-|b| \le b \le |b|$. Adding these inequalities (see Exercise 1.2.1), we obtain

$$-(|a| + |b|) \le a + b \le |a| + |b|.$$

Hence by Theorem 1.6 again, $|a + b| \le |a| + |b|$.

To prove the second inequality, apply the first inequality to $(a - b) + b$. We obtain

$$|a| - |b| = |a - b + b| - |b| \le |a - b| + |b| - |b| = |a - b|.$$

By reversing the roles of a and b and applying part ii), we also obtain

$$|b| - |a| \le |b - a| = |a - b|.$$

Multiplying this last inequality by -1 and combining it with the preceding one verifies

$$-|a - b| \le |a| - |b| \le |a - b|.$$

We conclude by Theorem 1.6 that $\big|\, |a| - |b| \,\big| \le |a - b|$. ■

Notice once and for all that this last inequality implies that $|a| - |b| \le |a - b|$ for all $a, b \in \mathbf{R}$. We will use this inequality several times.

WARNING. *Some students mistakenly mix absolute values and the Additive Property to conclude that $b < c$ implies $|a + b| < |a + c|$. It is important from the beginning to recognize that this implication is false unless both $a + b$ and $a + c$ are nonnegative*. For example, if $a = 1$, $b = -5$, and $c = -1$, then $b < c$ but $|a + b| = 4$ is not less than $|a + c| = 0$.

A correct way to estimate using absolute value signs usually involves one of the triangle inequalities.

1.8 *EXAMPLE.*

Prove that if $-2 < x < 1$, then $|x^2 - x| < 6$.

Proof. By hypothesis, $|x| < 2$. Hence by the triangle inequality and Remark 1.5,

$$|x^2 - x| \le |x|^2 + |x| < 4 + 2 = 6.$$ ■

The following result (which is equivalent to the Trichotomy Property) will be used many times in this and subsequent chapters.

1.9 Theorem. *Let $x, y, a \in \mathbf{R}$.*

i) *$x < y + \varepsilon$ for all $\varepsilon > 0$ if and only if $x \le y$.*
ii) *$x > y - \varepsilon$ for all $\varepsilon > 0$ if and only if $x \ge y$.*
iii) *$|a| < \varepsilon$ for all $\varepsilon > 0$ if and only if $a = 0$.*

Proof. i) Suppose to the contrary that $x < y + \varepsilon$ for all $\varepsilon > 0$ but $x > y$. Set $\varepsilon_0 = x - y > 0$ and observe that $y + \varepsilon_0 = x$. Hence by the Trichotomy Property, $y + \varepsilon_0$ cannot be greater than x. This contradicts the hypothesis for $\varepsilon = \varepsilon_0$. Thus $x \le y$.

Conversely, suppose that $x \le y$ and $\varepsilon > 0$ is given. Either $x < y$ or $x = y$. If $x < y$, then $x + 0 < y + 0 < y + \varepsilon$ by the Additive and Transitive Properties. If $x = y$, then $x < y + \varepsilon$ by the Additive Property. Thus $x < y + \varepsilon$ for all $\varepsilon > 0$ in either case. This completes the proof of part i).

ii) Suppose that $x > y - \varepsilon$ for all $\varepsilon > 0$. By the Second Multiplicative Property, this is equivalent to $-x < -y + \varepsilon$, hence by part i), equivalent to $-x \le -y$. By the Second Multiplicative Property, this is equivalent to $x \ge y$.

iii) Suppose that $|a| < \varepsilon = 0 + \varepsilon$ for all $\varepsilon > 0$. By part i), this is equivalent to $|a| \le 0$. Since it is always the case that $|a| \ge 0$, we conclude by the Trichotomy Property that $|a| = 0$. Therefore, $a = 0$ by Theorem 1.7i. ■

Let a and b be real numbers. A *closed interval* is a set of the form

$$[a, b] := \{x \in \mathbf{R} : a \le x \le b\}, \qquad [a, \infty) := \{x \in \mathbf{R} : a \le x\},$$
$$(-\infty, b] := \{x \in \mathbf{R} : x \le b\}, \quad \text{or} \quad (-\infty, \infty) := \mathbf{R},$$

and an *open interval* is a set of the form

$$(a, b) := \{x \in \mathbf{R} : a < x < b\}, \quad (a, \infty) := \{x \in \mathbf{R} : a < x\},$$
$$(-\infty, b) := \{x \in \mathbf{R} : x < b\}, \quad \text{or} \quad (-\infty, \infty) := \mathbf{R}.$$

By an *interval* we mean a closed interval, an open interval, or a set of the form

$$[a, b) := \{x \in \mathbf{R} : a \le x < b\} \quad \text{or} \quad (a, b] := \{x \in \mathbf{R} : a < x \le b\}.$$

Notice, then, that when $a < b$, the intervals $[a, b]$, $[a, b)$, $(a, b]$, and (a, b) correspond to line segments on the real line, but when $b < a$, these "intervals" are all the empty set.

An interval I is said to be *bounded* if and only if it has the form $[a, b]$, (a, b), $[a, b)$, or $(a, b]$ for some $-\infty < a \le b < \infty$, in which case the numbers a, b will be called the *endpoints* of I. All other intervals will be called *unbounded*. An interval with endpoints a, b is called *degenerate* if $a = b$ and *nondegenerate* if $a < b$. Thus a degenerate open interval is the empty set, and a degenerate closed interval is a point.

Analysis has a strong geometric flavor. Geometry enters the picture because the real number system can be identified with the real line in such a way that $a < b$ if and only if a lies to the left of b (see Figures 1.2, 2.1, and 2.2). This gives us a way of translating analytic results on $\mathbf{R}$ into geometric results on the number line, and vice versa. We close with several examples.

The absolute value is closely linked to the idea of length. The *length* of a bounded interval I with endpoints a, b is defined to be $|I| := |b - a|$, and the *distance* between any two points $a, b \in \mathbf{R}$ is defined by $|a - b|$.

Inequalities can be interpreted as statements about intervals. By Theorem 1.6, $|a| \le M$ if and only if a belongs to the closed interval $[-M, M]$; and by Theorem 1.9, a belongs to the open interval $(-\varepsilon, \varepsilon)$ for all $\varepsilon > 0$ if and only if $a = 0$.

We will use this point of view in Chapters 2 through 5 to give geometric interpretations to the calculus of functions defined on **R**, and in Chapters 11 through 13 to extend this calculus to functions defined on the Euclidean spaces $\mathbf{R}^n$.

EXERCISES

In each of the following exercises, verify the given statement carefully, proceeding step by step. Validate each step that involves an inequality by using some statement found in this section.

1.2.0 Let $a, b, c, d \in \mathbf{R}$ and consider each of the following statements. Decide which are true and which are false. Prove the true ones and give counterexamples to the false ones.

a) If $a < b$ and $c < d < 0$, then $ac > bd$.
b) If $a \le b$ and $c > 1$, then $|a + c| \le |b + c|$.
c) If $a \le b$ and $b \le a + c$, then $|a - b| \le c$.
d) If $a < b - \varepsilon$ for all $\varepsilon > 0$, then $a < 0$.

1.2.1. Suppose that $a, b, c \in \mathbf{R}$ and $a \le b$.

a) Prove that $a + c \le b + c$.
b) If $c \ge 0$, prove that $a \cdot c \le b \cdot c$.

1.2.2. Prove (7), (8), and (9). Show that each of these statements is false if the hypothesis $a \ge 0$ or $a > 0$ is removed.

1.2.3. **This exercise is used in Section 6.3**. The *positive part* of an $a \in \mathbf{R}$ is defined by

$$a^+ := \frac{|a| + a}{2}$$

and the *negative part* by

$$a^- := \frac{|a| - a}{2}.$$

a) Prove that $a = a^+ - a^-$ and $|a| = a^+ + a^-$.
b) Prove that

$$a^+ = \begin{cases} a & a \ge 0 \\ 0 & a \le 0 \end{cases} \quad \text{and} \quad a^- = \begin{cases} 0 & a \ge 0 \\ -a & a \le 0. \end{cases}$$

1.2.4. Solve each of the following inequalities for $x \in \mathbf{R}$.

a) $|2x + 1| < 7$
b) $|2 - x| < 2$

c) $|x^3 - 3x + 1| < x^3$

d) $\dfrac{x}{x-1} < 1$

e) $\dfrac{x^2}{4x^2 - 1} < \dfrac{1}{4}$

1.2.5. Let $a, b \in \mathbf{R}$.

a) Prove that if $a > 2$ and $b = 1 + \sqrt{a-1}$, then $2 < b < a$.
b) Prove that if $2 < a < 3$ and $b = 2 + \sqrt{a-2}$, then $0 < a < b$.
c) Prove that if $0 < a < 1$ and $b = 1 - \sqrt{1-a}$, then $0 < b < a$.
d) Prove that if $3 < a < 5$ and $b = 2 + \sqrt{a-2}$, then $3 < b < a$.

1.2.6. The *arithmetic mean* of $a, b \in \mathbf{R}$ is $A(a, b) = (a+b)/2$, and the *geometric mean* of $a, b \in [0, \infty)$ is $G(a, b) = \sqrt{ab}$. If $0 \le a \le b$, prove that $a \le G(a, b) \le A(a, b) \le b$. Prove that $G(a, b) = A(a, b)$ if and only if $a = b$.

1.2.7. Let $x \in \mathbf{R}$.

a) Prove that $|x| \le 2$ implies $|x^2 - 4| \le 4|x - 2|$.
b) Prove that $|x| \le 1$ implies $|x^2 + 2x - 3| \le 4|x - 1|$.
c) Prove that $-3 \le x \le 2$ implies $|x^2 + x - 6| \le 6|x - 2|$.
d) Prove that $-1 < x < 0$ implies $|x^3 - 2x + 1| < 1.26|x - 1|$.

1.2.8. For each of the following, find all values of $n \in \mathbf{N}$ that satisfy the given inequality.

a) $$\frac{1-n}{1-n^2} < 0.01$$

b) $$\frac{n^2 + 2n + 3}{2n^3 + 5n^2 + 8n + 3} < 0.025$$

c) $$\frac{n-1}{n^3 - n^2 + n - 1} < 0.002$$

1.2.9. a) Interpreting a rational m/n as $m \cdot n^{-1} \in \mathbf{R}$, use Postulate 1 to prove that

$$\frac{m}{n} + \frac{p}{q} = \frac{mq + np}{nq}, \quad \frac{m}{n} \cdot \frac{p}{q} = \frac{mp}{nq}, \quad -\frac{m}{n} = \frac{-m}{n}, \text{ and } \left(\frac{\ell}{n}\right)^{-1} = \frac{n}{\ell}$$

for $m, n, p, q, \ell \in \mathbf{Z}$ and $n, q, \ell \neq 0$.
b) Using Remark 1.1, Prove that Postulate 1 holds with $\mathbf{Q}$ in place of $\mathbf{R}$.
c) Prove that the sum of a rational and an irrational is always irrational. What can you say about the product of a rational and an irrational?
d) Let $m/n, p/q \in \mathbf{R}$ with $n, q > 0$. Prove that

$$\frac{m}{n} < \frac{p}{q} \quad \text{if and only if} \quad mq < np.$$

(Restricting this observation to $\mathbf{Q}$ gives a definition of "$<$" on $\mathbf{Q}$.)

1.2.10. Prove that

$$(ab + cd)^2 \leq (a^2 + c^2)(b^2 + d^2)$$

for all $a, b, c, d \in \mathbf{R}$.

1.2.11. a) Let $\mathbf{R}^+$ represent the collection of positive real numbers. Prove that $\mathbf{R}^+$ satisfies the following two properties.

i) For each $x \in \mathbf{R}$, one and only one of the following holds:

$$x \in \mathbf{R}^+, \quad -x \in \mathbf{R}^+, \quad \text{or} \quad x = 0.$$

ii) Given $x, y \in \mathbf{R}^+$, both $x + y$ and $x \cdot y$ belong to $\mathbf{R}^+$.

b) Suppose that $\mathbf{R}$ contains a subset $\mathbf{R}^+$ (not necessarily the set of positive numbers) which satisfies properties i) and ii). Define $x \prec y$ by $y - x \in \mathbf{R}^+$. Prove that Postulate 2 holds with $\prec$ in place of $<$.

1.3 COMPLETENESS AXIOM

In this section we introduce the last of three postulates that describe $\mathbf{R}$. To formulate this postulate, which distinguishes $\mathbf{Q}$ from $\mathbf{R}$, we need the following concepts.

1.10 Definition.

Let $E \subset \mathbf{R}$ be nonempty.

i) The set E is said to be *bounded above* if and only if there is an $M \in \mathbf{R}$ such that $a \leq M$ for all $a \in E$, in which case M is called an *upper bound* of E.

ii) A number s is called a *supremum* of the set E if and only if s is an upper bound of E and $s \leq M$ for all upper bounds M of E. (In this case we shall say that *E has a finite supremum s* and write $s = \sup E$.)

NOTE: Because French mathematicians (e.g., Borel, Jordan, and Lebesgue) did fundamental work on the connection between analysis and set theory, and *ensemble* is French for *set*, analysts frequently use E to represent a general set.

By Definition 1.10ii, a supremum of a set E (when it exists) is the smallest (or least) upper bound of E. By definition, then, in order to prove that $s = \sup E$ for some set $E \subset \mathbf{R}$, we must show two things: s is an upper bound, AND s is the smallest upper bound. Here is a typical example.

1.11 *EXAMPLE.*

If $E = [0, 1]$, prove that $\sup E = 1$.

Proof. By the definition of interval, 1 is an upper bound of E. Let M be any upper bound of E; that is, $M \geq x$ for all $x \in E$. Since $1 \in E$, it follows that $M \geq 1$. Thus 1 is the smallest upper bound of E. ∎

The following two remarks answer the question: How many upper bounds and suprema can a given set have?

1.12 Remark. *If a set has one upper bound, it has infinitely many upper bounds.*

Proof. If M_0 is an upper bound for a set E, then so is M for any $M > M_0$. ∎

1.13 Remark. *If a set has a supremum, then it has only one supremum.*

Proof. Let s_1 and s_2 be suprema of the same set E. Then both s_1 and s_2 are upper bounds of E, whence by Definition 1.10ii, $s_1 \leq s_2$ and $s_2 \leq s_1$. We conclude by the Trichotomy Property that $s_1 = s_2$. ∎

NOTE: *This proof illustrates a general principle. When asked to prove $a = b$, it is often easier to verify that $a \leq b$ and $b \leq a$ separately.*

The next result, a fundamental property of suprema, shows that the supremum of a set E can be approximated by a point in E (see Figure 1.3 for an illustration).

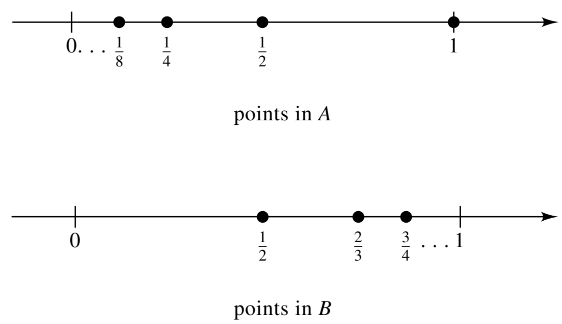

FIGURE 1.3

1.14 Theorem. [APPROXIMATION PROPERTY FOR SUPREMA].
If E has a finite supremum and $\varepsilon > 0$ is any positive number, then there is a point $a \in E$ such that

$$\sup E - \varepsilon < a \leq \sup E.$$

Proof. Suppose that the theorem is false. Then there is an $\varepsilon_0 > 0$ such that no element of E lies between $s_0 := \sup E - \varepsilon_0$ and $\sup E$. Since $\sup E$ is an upper bound for E, it follows that $a \leq s_0$ for all $a \in E$; that is, s_0 is an upper bound of E. Thus, by Definition 1.10ii, $\sup E \leq s_0 = \sup E - \varepsilon_0$. Adding $\varepsilon_0 - \sup E$ to both sides of this inequality, we conclude that $\varepsilon_0 \leq 0$, a contradiction. ∎

The Approximation Property can be used to show that the supremum of any subset of integers is itself an integer.

1.15 Theorem. *If $E \subset \mathbf{Z}$ has a supremum, then* $\sup E \in E$. *In particular, if the supremum of a set, which contains only integers, exists, that supremum must be an integer.*

Proof. Suppose that $s := \sup E$ and apply the Approximation Property to choose an $x_0 \in E$ such that $s - 1 < x_0 \leq s$. If $s = x_0$, then $s \in E$, as promised. Otherwise, $s - 1 < x_0 < s$ and we can apply the Approximation Property again to choose $x_1 \in E$ such that $x_0 < x_1 < s$.

Subtract x_0 from this last inequality to obtain $0 < x_1 - x_0 < s - x_0$. Since $-x_0 < 1-s$, it follows that $0 < x_1-x_0 < s+(1-s) = 1$. Thus $x_1-x_0 \in \mathbf{Z}\cap(0, 1)$, a contradiction by Remark 1.1iii. We conclude that $s \in E$. ■

The existence of suprema is the last assumption about **R** we make.

Postulate 3. [COMPLETENESS AXIOM]. If E is a nonempty subset of **R** that is bounded above, then E has a finite supremum.

We shall use Completeness Axiom many times. Our first two applications deal with the distribution of integers (Theorem 1.16) and rationals (Theorem 1.18) among real numbers.

1.16 Theorem. [ARCHIMEDEAN PRINCIPLE].
Given real numbers a and b, with $a > 0$, there is an integer $n \in \mathbf{N}$ such that $b < na$.

STRATEGY: The idea behind the proof is simple. By the Completeness Axiom and Theorem 1.15, any nonempty subset of integers that is bounded above has a "largest" integer. If k_0 is the largest integer that satisfies $k_0a \leq b$, then $n = (k_0 + 1)$ (which is larger than k_0) must satisfy $na > b$. In order to justify this application of the Completeness Axiom, we have two details to attend to: (1) Is the set $E := \{k \in \mathbf{N} : ka \leq b\}$ bounded above? (2) Is E nonempty? The answer to the second question depends on whether $b < a$ or not. Here are the details.

Proof. If $b < a$, set $n = 1$. If $a \leq b$, consider the set $E = \{k \in \mathbf{N} : ka \leq b\}$. E is nonempty since $1 \in E$. Let $k \in E$ (i.e., $ka \leq b$). Since $a > 0$, it follows from the First Multiplicative Property that $k \leq b/a$. This proves that E is bounded above by b/a. Thus, by the Completeness Axiom and Theorem 1.15, E has a finite supremum s that belongs to E, in particular, $s \in \mathbf{N}$.

Set $n = s + 1$. Then $n \in \mathbf{N}$ and (since n is larger than s), n cannot belong to E. Thus $na > b$. ■

Notice in Example 1.11 and Theorem 1.15 that the supremum of E belonged to E. The following result shows that this is not always the case.

1.17 *EXAMPLE.*

Let $A = \{1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots\}$ and $B = \{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots\}$. Prove that $\sup A = \sup B = 1$.

Proof. It is clear that 1 is an upper bound of both sets. It remains to see that 1 is the smallest upper bound of both sets. For A, this is trivial. Indeed, if M is any upper bound of A, then $M \geq 1$ (since $1 \in A$). On the other hand, if M is an upper bound for B, but $M < 1$, then $1 - M > 0$. In particular, $1/(1 - M) \in \mathbf{R}$.

Choose, by the Archimedean Principle, an $n \in \mathbf{N}$ such that $n > 1/(1-M)$. It follows (do the algebra) that $x_0 := 1-1/n > M$. Since $x_0 \in B$, this contradicts the assumption that M is an upper bound of B (see Figure 1.3). ■

The next proof shows how the Archimedean Principle is used to establish scale.

1.18 Theorem. [Density of Rationals].
If $a, b \in \mathbf{R}$ satisfy $a < b$, then there is a $q \in \mathbf{Q}$ such that $a < q < b$.

Strategy: To find a fraction $q = m/n$ such that $a < q < b$, we must specify both numerator m and denominator n. Let's suppose first that $a > 0$ and that the set $E := \{k \in \mathbf{N} : k/n \leq a\}$ has a supremum, k_0. Then $m := k_0 + 1$, being greater than the supremum of E, cannot belong to E. Thus $m/n > a$. Is this the fraction we look for? Is $m/n < b$? Not unless n is large enough. To see this, look at a concrete example: $a = 2/3$ and $b = 1$. If $n = 1$, then E has no supremum, When $n = 2,\ k_0 = 1$ and when $n = 3,\ k_0 = 2$. In both cases $(k_0 + 1)/n = 1$ is too big. However, when $n = 4,\ k_0 = 2$ so $(k_0 + 1)/4 = 3/4$ is smaller than b, as required.

How can we prove that for each fixed $a < b$ there always is an n large enough so that if k_0 is chosen as above, then $(k_0+1)/n < b$? By the choice of $k_0,\ k_0/n \leq a$. Let's look at the worst case scenario: $a = k_0/n$. Then $b > (k_0 + 1)/n$ means

$$b > \frac{k_0 + 1}{n} = \frac{k_0}{n} + \frac{1}{n} = a + \frac{1}{n}$$

(i.e., $b - a > 1/n$). Such an n can always be chosen by the Archimedean Principle.

What about the assumption that $\sup E$ exists? This requires that E be nonempty and bounded above. Once n is fixed, E will be bounded above by na. But the only way that E is nonempty is that at the very least, $1 \in E$ (i.e., that $1/n \leq a$). This requires a second restriction on n. We begin our formal proof at this point.

Proof. Suppose first that $a > 0$. Since $b - a > 0$, use the Archimedean Principle to choose an $n \in \mathbf{N}$ that satisfies

$$n > \max\left\{\frac{1}{a}, \frac{1}{b-a}\right\},$$

and observe that both $1/n < a$ and $1/n < b - a$.

Consider the set $E = \{k \in \mathbf{N} : k/n \leq a\}$. Since $1 \in E,\ E$ is nonempty. Since $n > 0,\ E$ is bounded above by na. Hence, by Theorem 1.15, $k_0 := \sup E$ exists and belongs to E, in particular, to **N**. Set $m = k_0 + 1$ and $q = m/n$. Since k_0 is the supremum of E, $m \notin E$. Thus $q > a$. On the other hand, since $k_0 \in E$, it follows from the choice of n that

$$b = a + (b - a) \geq \frac{k_0}{n} + (b - a) > \frac{k_0}{n} + \frac{1}{n} = \frac{m}{n} = q.$$

Now suppose that $a \leq 0$. Choose, by the Archimedean Principle, an integer $k \in \mathbf{N}$ such that $k > -a$. Then $0 < k + a < k + b$, and by the case already proved, there is an $r \in \mathbf{Q}$ such that $k + a < r < k + b$. Therefore, $q := r - k$ belongs to **Q** and satisfies the inequality $a < q < b$. ■

For some applications, we also need the following concepts.

1.19 Definition.

Let $E \subset \mathbf{R}$ be nonempty.

i) The set E is said to be *bounded below* if and only if there is an $m \in \mathbf{R}$ such that $a \geq m$ for all $a \in E$, in which case m is called a *lower bound* of the set E.
ii) A number t is called an *infimum* of the set E if and only if t is a lower bound of E and $t \geq m$ for all lower bounds m of E. In this case we shall say that *E has an infimum t* and write $t = \inf E$.
iii) E is said to be *bounded* if and only if it is bounded both above and below.

When a set E contains its supremum (respectively, its infimum) we shall frequently write $\max E$ for $\sup E$ (respectively, $\min E$ for $\inf E$).

(Some authors call the supremum the *least upper bound* and the infimum the *greatest lower bound*. We will not use this terminology because it is somewhat old fashioned and because it confuses some students, since the **least** upper bound of a given set is always greater than or equal to the **greatest** lower bound.)

To relate suprema to infima, we define the *reflection* of a set $E \subseteq \mathbf{R}$ by

$$-E := \{x : x = -a \text{ for some } a \in E\}.$$

For example, $-(1, 2] = [-2, -1)$.

The following result shows that the supremum of a set is the same as the negative of its reflection's infimum. This can be used to prove an Approximation Property and a Completeness Property for Infima (see Exercise 1.3.6).

1.20 Theorem. [Reflection Principle].
Let $E \subseteq \mathbf{R}$ be nonempty.

i) *E has a supremum if and only if $-E$ has an infimum, in which case*

$$\inf(-E) = -\sup E.$$

ii) *E has an infimum if and only if $-E$ has a supremum, in which case*

$$\sup(-E) = -\inf E.$$

Proof. The proofs of these statements are similar. We prove only the first statement.

Suppose that E has a supremum s and set $t = -s$. Since s is an upper bound for E, $s \geq a$ for all $a \in E$, so $-s \leq -a$ for all $a \in E$. Therefore, t is a lower bound of $-E$. Suppose that m is any lower bound of $-E$. Then $m \leq -a$ for all $a \in E$, so $-m$ is an upper bound of E. Since s is the supremum of E, it follows that $s \leq -m$ (i.e., $t = -s \geq m$). Thus t is the infimum of $-E$ and $\sup E = s = -t = -\inf(-E)$.

Conversely, suppose that $-E$ has an infimum t. By definition, $t \leq -a$ for all $a \in E$. Thus $-t$ is an upper bound for E. Since E is nonempty, E has a supremum by the Completeness Axiom. ■

Theorem 1.20 allows us to obtain information about infima from results about suprema, and vice versa (see the proof of the next theorem).

We shall use the following result many times.

1.21 Theorem. [MONOTONE PROPERTY].
Suppose that $A \subseteq B$ are nonempty subsets of $\mathbf{R}$.

i) *If B has a supremum, then* $\sup A \leq \sup B$.
ii) *If B has an infimum, then* $\inf A \geq \inf B$.

Proof. i) Since $A \subseteq B$, any upper bound of B is an upper bound of A. Therefore, $\sup B$ is an upper bound of A. It follows from the Completeness Axiom that $\sup A$ exists, and from Definition 1.10ii that $\sup A \leq \sup B$.

ii) Clearly, $-A \subseteq -B$. Thus by part i), Theorem 1.20, and the Second Multiplicative Property,

$$\inf A = -\sup(-A) \geq -\sup(-B) = \inf B. \quad ■$$

It is convenient to extend the definition of suprema and infima to all subsets of $\mathbf{R}$. To do this we expand the definition of $\mathbf{R}$ as follows. The set of *extended real numbers* is defined to be $\overline{\mathbf{R}} := \mathbf{R} \bigcup \{\pm\infty\}$. Thus x is an extended real number if and only if either $x \in \mathbf{R}$, $x = +\infty$, or $x = -\infty$.

Let $E \subseteq \mathbf{R}$ be nonempty. We shall define $\sup E = +\infty$ if E is unbounded above and $\inf E = -\infty$ if E is unbounded below. Finally, we define $\sup \emptyset = -\infty$ and $\inf \emptyset = +\infty$. Notice, then, that the supremum of a subset E of $\mathbf{R}$ (respectively, the infimum of E) is finite if and only if E is nonempty and bounded above (respectively, nonempty and bounded below). Moreover, under the convention $-\infty < a$ and $a < \infty$ for all $a \in \mathbf{R}$, the Monotone Property still holds for this extended definition; that is, if A and B are subsets of $\mathbf{R}$ and $A \subseteq B$, then $\sup A \leq \sup B$ and $\inf A \geq \inf B$, provided we use the convention that $-\infty < \infty$.

EXERCISES

1.3.0. Decide which of the following statements are true and which are false. Prove the true ones and give counterexamples to the false ones.

a) If A and B are nonempty, bounded subsets of **R**, then $\sup(A \cap B) \le \sup A$.
b) Let ε be a positive real number. If A is a nonempty, bounded subset of **R** and $B = \{\varepsilon x : x \in A\}$, then $\sup(B) = \varepsilon \sup(A)$.
c) If $A + B := \{a + b : a \in A \text{ and } b \in B\}$, where A and B are nonempty, bounded subsets of **R**, then $\sup(A + B) = \sup(A) + \sup(B)$.
d) If $A - B := \{a - b : a \in A \text{ and } b \in B\}$, where A and B are nonempty, bounded subsets of **R**, then $\sup(A - B) = \sup(A) - \sup(B)$

1.3.1. Find the infimum and supremum of each of the following sets.

a) $E = \{x \in \mathbf{R} : x^2 + 2x = 3\}$
b) $E = \{x \in \mathbf{R} : x^2 - 2x + 3 > x^2 \text{ and } x > 0\}$
c) $E = \{p/q \in \mathbf{Q} : p^2 < 5q^2 \text{ and } p, q > 0\}$
d) $E = \{x \in \mathbf{R} : x = 1 + (-1)^n/n \text{ for } n \in \mathbf{N}\}$
e) $E = \{x \in \mathbf{R} : x = 1/n + (-1)^n \text{ for } n \in \mathbf{N}\}$
f) $E = \{2 - (-1)^n/n^2 : n \in \mathbf{N}\}$

1.3.2. Prove that for each $a \in \mathbf{R}$ and each $n \in \mathbf{N}$ there exists a rational r_n such that $|a - r_n| < 1/n$.

1.3.3 . [Density of Irrationals] **This exercise is used in Section 3.3**. Prove that if $a < b$ are real numbers, then there is an irrational $\xi \in \mathbf{R}$ such that $a < \xi < b$.

1.3.4. Prove that a lower bound of a set need not be unique but the infimum of a given set E is unique.

1.3.5. Show that if E is a nonempty bounded subset of **Z**, then $\inf E$ exists and belongs to E.

1.3.6 . **This exercise is used in many sections, including 2.2 and 5.1**. Use the Reflection Principle and analogous results about suprema to prove the following results.

a) [Approximation Property for Infima] Prove that if a set $E \subset \mathbf{R}$ has a finite infimum and $\varepsilon > 0$ is any positive number, then there is a point $a \in E$ such that $\inf E + \varepsilon > a \ge \inf E$.
b) [Completeness Property for Infima] If $E \subseteq \mathbf{R}$ is nonempty and bounded below, then E has a (finite) infimum.

1.3.7. a) Prove that if x is an upper bound of a set $E \subseteq \mathbf{R}$ and $x \in E$, then x is the supremum of E.
b) Make and prove an analogous statement for the infimum of E.
c) Show by example that the converse of each of these statements is false.

1.3.8. Suppose that $E, A, B \subset \mathbf{R}$ and $E = A \cup B$. Prove that if E has a supremum and both A and B are nonempty, then $\sup A$ and $\sup B$ both exist, and $\sup E$ is one of the numbers $\sup A$ or $\sup B$.

1.3.9. A *dyadic rational* is a number of the form $k/2^n$ for some $k, n \in \mathbf{Z}$. Prove that if a and b are real numbers and $a < b$, then there exists a dyadic rational q such that $a < q < b$.

1.3.10. Let $x_n \in \mathbf{R}$ and suppose that there is an $M \in \mathbf{R}$ such that $|x_n| \le M$ for $n \in \mathbf{N}$. Prove that $s_n = \sup\{x_n, x_{n+1}, \dots\}$ defines a real number for each $n \in \mathbf{N}$ and that $s_1 \ge s_2 \ge \cdots$. Prove an analogous result about $t_n = \inf\{x_n, x_{n+1}, \dots\}$.

1.3.11. If $a, b \in \mathbf{R}$ and $b - a > 1$, then there is at least one $k \in \mathbf{Z}$ such that $a < k < b$.

1.4 MATHEMATICAL INDUCTION

In this section we introduce the method of Mathematical Induction and use it to prove the Binomial Formula, a result that shows how to expand powers of a binomial expression (i.e., an expression of the form $a + b$).

We begin by obtaining another consequence of the Completeness Axiom, the Well-Ordering Principle, which is a statement about the existence of least elements of subsets of **N**.

1.22 Theorem. [Well-Ordering Principle].
*If E is a nonempty subset of **N**, then E has a least element (i.e., E has a finite infimum and* $\inf E \in E$).

Proof. Suppose that $E \subseteq \mathbf{N}$ is nonempty. Then $-E$ is bounded above, by -1, so by the Completeness Axiom $\sup(-E)$ exists, and by Theorem 1.15, $\sup(-E) \in -E$. Hence by Theorem 1.20, $\inf E = -\sup(-E)$ exists, and $\inf E \in -(-E) = E$. ■

Our first application of the Well-Ordering Principle is called the *Principle of Mathematical Induction* or the *Axiom of Induction* (which, under mild assumptions, is equivalent to the Well-Ordering Principle—see Appendix A).

1.23 Theorem. *Suppose for each $n \in \mathbf{N}$ that $A(n)$ is a proposition (i.e., a verbal statement or formula) which satisfies the following two properties:*

i) $A(1)$ *is true.*
ii) *For every $n \in \mathbf{N}$ for which $A(n)$ is true, $A(n+1)$ is also true.*

Then $A(n)$ is true for all $n \in \mathbf{N}$.

Proof. Suppose that the theorem is false. Then the set $E = \{n \in \mathbf{N} : A(n)$ is false$\}$ is nonempty. Hence by the Well-Ordering Principle, E has a least element, say x.

Since $x \in E \subseteq \mathbf{N} \subset \mathbf{Z}$, we have by Remark 1.1ii that $x \ge 1$. Since $x \in E$, we have by hypothesis i) that $x \ne 1$. In particular, $x - 1 > 0$. Hence, by Remark 1.1i and iii, $x - 1 \ge 1$ and $x - 1 \in \mathbf{N}$.

Since $x - 1 < x$ and x is a least element of E, the statement $A(x-1)$ must be true. Applying hypothesis ii) to $n = x - 1$, we see that $A(x) = A(n+1)$ must also be true; that is, $x \notin E$, a contradiction. ■

Recall that if $x_0, x_1, \dots, x_n$ are real numbers and $0 \le j \le n$, then

$$\sum_{k=j}^{n} x_k := x_j + x_{j+1} + \cdots + x_n$$

denotes the sum of the x_k's as k ranges from j to n. The following examples illustrate the fact that the Principle of Mathematical Induction can be used to prove a variety of statements involving integers.

1.24 *EXAMPLE.*

Prove that

$$\sum_{k=1}^{n} (3k-1)(3k+2) = 3n^3 + 6n^2 + n$$

for $n \in \mathbf{N}$.

Proof. Let $A(n)$ represent the statement

$$\sum_{k=1}^{n} (3k-1)(3k+2) = 3n^3 + 6n^2 + n.$$

For $n = 1$ the left side of this equation is $2 \cdot 5$ and the right side is $3 + 6 + 1$. Therefore, $A(1)$ is true. Suppose that $A(n)$ is true for some $n \ge 1$. Then

$$\begin{aligned}\sum_{k=1}^{n+1} (3k-1)(3k+2) &= (3n+2)(3n+5) + \sum_{k=1}^{n} (3k-1)(3k+2) \\ &= (3n+2)(3n+5) + 3n^3 + 6n^2 + n \\ &= 3n^3 + 15n^2 + 22n + 10.\end{aligned}$$

On the other hand, a direct calculation reveals that

$$3(n+1)^3 + 6(n+1)^2 + (n+1) = 3n^3 + 15n^2 + 22n + 10.$$

Therefore, $A(n+1)$ is true when $A(n)$ is. We conclude by induction that $A(n)$ holds for all $n \in \mathbf{N}$. ∎

Two formulas encountered early in an algebra course are the perfect square and cube formulas:

$$(a+b)^2 = a^2 + 2ab + b^2 \quad \text{and} \quad (a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3.$$

Our next application of the Principle of Mathematical Induction generalizes these formulas from $n = 2$ and 3 to arbitrary $n \in \mathbf{N}$.

Recall that Pascal's triangle is the triangular array of integers whose rows begin and end with 1s with the property that an interior entry on any row is obtained by adding the two numbers in the preceding row immediately above that entry. Thus the first few rows of Pascal's triangle are as below.

$$\begin{array}{ccccccccccccc} & & & & & & 1 & & & & & & \\ & & & & & 1 & & 1 & & & & & \\ & & & & 1 & & 2 & & 1 & & & & \\ & & & 1 & & 3 & & 3 & & 1 & & & \\ & & 1 & & 4 & & 6 & & 4 & & 1 & & \\ & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\ 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 \end{array}$$

Notice that the third and fourth rows are precisely the coefficients that appeared in the perfect square and cube formulas above.

We can write down a formula for each entry in each row of the Pascal triangle. The first (and only) entry in the first row is

$$\binom{0}{0} := 1.$$

Using the notation $0! := 1$ and $n! := 1 \cdot 2 \cdots (n-1) \cdot n$ for $n \in \mathbf{N}$, define the *binomial coefficient* n choose k by

$$\binom{n}{k} := \frac{n!}{(n-k)!k!}$$

for $0 \le k \le n$ and $n = 0, 1, \ldots.$

Since $\binom{n}{0} = \binom{n}{n} = 1$ for all $n \in \mathbf{N}$, the following result shows that the binomial coefficient n over k does produce the $(k+1)$st entry in the $(n+1)$st row of Pascal's triangle.

1.25 ***Lemma.***
If $n, k \in \mathbf{N}$ and $1 \le k \le n$, then

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}.$$

Proof. By definition,

$$\begin{aligned} \binom{n}{k-1} + \binom{n}{k} &= \frac{n!\,k}{(n-k+1)!k!} + \frac{n!(n-k+1)}{(n-k+1)!k!} \\ &= \frac{n!(n+1)}{(n-k+1)!k!} = \binom{n+1}{k}. \end{aligned}$$

■

Binomial coefficients can be used to expand the nth power of a sum of two terms.

1.26 Theorem. [BINOMIAL FORMULA].
If $a, b \in \mathbf{R}$, $n \in \mathbf{N}$, and 0^0 is interpreted to be 1, then

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

Proof. The proof is by induction on n. The formula is obvious for $n = 1$. Suppose that the formula is true for some $n \in \mathbf{N}$. Then by the inductive hypothesis and Postulate 1,

$$\begin{aligned}(a+b)^{n+1} &= (a+b)(a+b)^n \\ &= (a+b)\left(\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k\right) \\ &= \left(\sum_{k=0}^{n}\binom{n}{k}a^{n-k+1}b^k\right) + \left(\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^{k+1}\right) \\ &= \left(a^{n+1} + \sum_{k=1}^{n}\binom{n}{k}a^{n-k+1}b^k\right) + \left(b^{n+1} + \sum_{k=0}^{n-1}\binom{n}{k}a^{n-k}b^{k+1}\right) \\ &= a^{n+1} + \sum_{k=1}^{n}\left(\binom{n}{k} + \binom{n}{k-1}\right)a^{n-k+1}b^k + b^{n+1}.\end{aligned}$$

Hence it follows from Lemma 1.25 that

$$(a+b)^{n+1} = a^{n+1} + \sum_{k=1}^{n}\binom{n+1}{k}a^{n+1-k}b^k + b^{n+1} = \sum_{k=0}^{n+1}\binom{n+1}{k}a^{n+1-k}b^k;$$

that is, the formula is true for $n+1$. We conclude by induction that the formula holds for all $n \in \mathbf{N}$. ■

We close this section with two optional, well-known results that further demonstrate the power of the Completeness Axiom and its consequences.

***1.27 Remark.** *If $x > 1$ and $x \notin \mathbf{N}$, then there is an $n \in \mathbf{N}$ such that $n < x < n+1$.*

Proof. By the Archimedean Principle, the set $E = \{m \in \mathbf{N} : x < m\}$ is nonempty. Hence by the Well-Ordering Principle, E has a least element, say m_0.

Set $n = m_0 - 1$. Since $m_0 \in E$, $n + 1 = m_0 > x$. Since m_0 is least, $n = m_0 - 1 \le x$. Since $x \notin \mathbf{N}$, we also have $n \ne x$. Therefore, $n < x < n+1$. ■