CHAPTER 1 BASIC PROPERTIES OF NUMBERS

The title of this chapter expresses in a few words the mathematical knowledge required to read this book. In fact, this short chapter is simply an explanation of what is meant by the "basic properties of numbers," all of which—addition and multiplication, subtraction and division, solutions of equations and inequalities, factoring and other algebraic manipulations—are already familiar to us. Nevertheless, this chapter is not a review. Despite the familiarity of the subject, the survey we are about to undertake will probably seem quite novel; it does not aim to present an extended review of old material, but to condense this knowledge into a few simple and obvious properties of numbers. Some may even seem too obvious to mention, but a surprising number of diverse and important facts turn out to be consequences of the ones we shall emphasize.

Of the twelve properties which we shall study in this chapter, the first nine are concerned with the fundamental operations of addition and multiplication. For the moment we consider only addition: this operation is performed on a pair of numbers—the sum $a + b$ exists for any two given numbers a and b (which may possibly be the same number, of course). It might seem reasonable to regard addition as an operation which can be performed on several numbers at once, and consider the sum $a_1 + \cdots + a_n$ of n numbers $a_1, \ldots, a_n$ as a basic concept. It is more convenient, however, to consider addition of pairs of numbers only, and to define other sums in terms of sums of this type. For the sum of three numbers a, b, and c, this may be done in two different ways. One can first add b and c, obtaining $b + c$, and then add a to this number, obtaining $a + (b + c)$; or one can first add a and b, and then add the sum $a + b$ to c, obtaining $(a + b) + c$. Of course, the two compound sums obtained are equal, and this fact is the very first property we shall list:

(P1) If a, b, and c are any numbers, then

$$a + (b + c) = (a + b) + c.$$

The statement of this property clearly renders a separate concept of the sum of three numbers superfluous; we simply agree that $a + b + c$ denotes the number $a + (b + c) = (a + b) + c$. Addition of four numbers requires similar, though slightly more involved, considerations. The symbol $a + b + c + d$ is defined to mean

$$
\begin{aligned}
&(1)\quad ((a + b) + c) + d,\\
\text{or }&(2)\quad (a + (b + c)) + d,\\
\text{or }&(3)\quad a + ((b + c) + d),\\
\text{or }&(4)\quad a + (b + (c + d)),\\
\text{or }&(5)\quad (a + b) + (c + d).
\end{aligned}
$$

This definition is unambiguous since these numbers are all equal. Fortunately, *this* fact need not be listed separately, since it follows from the property P1 already listed. For example, we know from P1 that

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

and it follows immediately that (1) and (2) are equal. The equality of (2) and (3) is a direct consequence of P1, although this may not be apparent at first sight (one must let $b + c$ play the role of b in P1, and d the role of c). The equalities (3) = (4) = (5) are also simple to prove.

It is probably obvious that an appeal to P1 will also suffice to prove the equality of the 14 possible ways of summing five numbers, but it may not be so clear how we can reasonably arrange a proof that this is so without actually listing these 14 sums. Such a procedure is feasible, but would soon cease to be if we considered collections of six, seven, or more numbers; it would be totally inadequate to prove the equality of all possible sums of an arbitrary finite collection of numbers $a_1, \ldots, a_n$. This fact may be taken for granted, but for those who would like to worry about the proof (and it is worth worrying about once) a reasonable approach is outlined in Problem 23. Henceforth, we shall usually make a tacit appeal to the results of this problem and write sums $a_1 + \cdots + a_n$ with a blithe disregard for the arrangement of parentheses.

The number 0 has one property so important that we list it next:

(P2) If a is any number, then

$$a + 0 = 0 + a = a.$$

An important role is also played by 0 in the third property of our list:

(P3) For every number a, there is a number $-a$ such that

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

Property P2 ought to represent a distinguishing characteristic of the number 0, and it is comforting to note that we are already in a position to prove this. Indeed, if a number x satisfies

$$a + x = a$$

for any one number a, then $x = 0$ (and consequently this equation also holds for all numbers a). The proof of this assertion involves nothing more than subtracting a from both sides of the equation, in other words, adding $-a$ to both sides; as the following detailed proof shows, all three properties P1–P3 must be used to justify this operation.

If	$a + x = a,$
then	$(-a) + (a + x) = (-a) + a = 0;$
hence	$((-a) + a) + x = 0;$
hence	$0 + x = 0;$
hence	$x = 0.$

As we have just hinted, it is convenient to regard subtraction as an operation derived from addition: we consider $a - b$ to be an abbreviation for $a + (-b)$. It is then possible to find the solution of certain simple equations by a series of steps (each justified by P1, P2, or P3) similar to the ones just presented for the equation $a + x = a$. For example:

$$\begin{array}{lr} \text{If} & x + 3 = 5, \\ \text{then} & (x + 3) + (-3) = 5 + (-3); \\ \text{hence} & x + (3 + (-3)) = 5 - 3 = 2; \\ \text{hence} & x + 0 = 2; \\ \text{hence} & x = 2. \end{array}$$

Naturally, such elaborate solutions are of interest only until you become convinced that they can always be supplied. In practice, it is usually just a waste of time to solve an equation by indicating so explicitly the reliance on properties P1, P2, and P3 (or any of the further properties we shall list).

Only one other property of addition remains to be listed. When considering the sums of three numbers a, b, and c, only two sums were mentioned: $(a + b) + c$ and $a + (b + c)$. Actually, several other arrangements are obtained if the order of a, b, and c is changed. That these sums are all equal depends on

(P4) If a and b are any numbers, then

$$a + b = b + a.$$

The statement of P4 is meant to emphasize that although the operation of addition of pairs of numbers might conceivably depend on the order of the two numbers, in fact it does not. It is helpful to remember that not all operations are so well behaved. For example, subtraction does not have this property: usually $a - b \neq b - a$. In passing we might ask just when $a - b$ does equal $b - a$, and it is amusing to discover how powerless we are if we rely only on properties P1–P4 to justify our manipulations. Algebra of the most elementary variety shows that $a - b = b - a$ only when $a = b$. Nevertheless, it is impossible to derive this fact from properties P1–P4; it is instructive to examine the elementary algebra carefully and determine which step(s) cannot be justified by P1–P4. We will indeed be able to justify all steps in detail when a few more properties are listed. Oddly enough, however, the crucial property involves multiplication.

The basic properties of multiplication are fortunately so similar to those for addition that little comment will be needed; both the meaning and the consequences should be clear. (As in elementary algebra, the product of a and b will be denoted by $a \cdot b$, or simply ab.)

(P5) If a, b, and c are any numbers, then

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c.$$

(P6) If a is any number, then

$$a \cdot 1 = 1 \cdot a = a.$$

Moreover, $1 \neq 0$.

(The assertion that $1 \neq 0$ may seem a strange fact to list, but we have to list it, because there is no way it could possibly be proved on the basis of the other properties listed—these properties would all hold if there were only one number, namely, 0.)

(P7) For every number $a \neq 0$, there is a number a^{-1} such that

$$a \cdot a^{-1} = a^{-1} \cdot a = 1.$$

(P8) If a and b are any numbers, then

$$a \cdot b = b \cdot a.$$

One detail which deserves emphasis is the appearance of the condition $a \neq 0$ in P7. This condition is quite necessary; since $0 \cdot b = 0$ for all numbers b, there is *no* number 0^{-1} satisfying $0 \cdot 0^{-1} = 1$. This restriction has an important consequence for division. Just as subtraction was defined in terms of addition, so division is defined in terms of multiplication: the symbol a/b means $a \cdot b^{-1}$. Since 0^{-1} is meaningless, $a/0$ is also meaningless—division by 0 is *always* undefined.

Property P7 has two important consequences. If $a \cdot b = a \cdot c$, it does not necessarily follow that $b = c$; for if $a = 0$, then both $a \cdot b$ and $a \cdot c$ are 0, no matter what b and c are. However, if $a \neq 0$, then $b = c$; this can be deduced from P7 as follows:

$$\begin{array}{lrl} \text{If} & a \cdot b &= a \cdot c \text{ and } a \neq 0, \\ \text{then} & a^{-1} \cdot (a \cdot b) &= a^{-1} \cdot (a \cdot c); \\ \text{hence} & (a^{-1} \cdot a) \cdot b &= (a^{-1} \cdot a) \cdot c; \\ \text{hence} & 1 \cdot b &= 1 \cdot c; \\ \text{hence} & b &= c. \end{array}$$

It is also a consequence of P7 that if $a \cdot b = 0$, then either $a = 0$ or $b = 0$. In fact,

$$\begin{array}{lrl} \text{if} & a \cdot b &= 0 \text{ and } a \neq 0, \\ \text{then} & a^{-1} \cdot (a \cdot b) &= 0; \\ \text{hence} & (a^{-1} \cdot a) \cdot b &= 0; \\ \text{hence} & 1 \cdot b &= 0; \\ \text{hence} & b &= 0. \end{array}$$

(It may happen that $a = 0$ and $b = 0$ are both true; this possibility is not excluded when we say "either $a = 0$ or $b = 0$"; in mathematics "or" is always used in the sense of "one or the other, or both.")

This latter consequence of P7 is constantly used in the solution of equations. Suppose, for example, that a number x is known to satisfy

$$(x - 1)(x - 2) = 0.$$

Then it follows that either $x - 1 = 0$ or $x - 2 = 0$; hence $x = 1$ or $x = 2$.

On the basis of the eight properties listed so far it is still possible to prove very little. Listing the next property, which combines the operations of addition and multiplication, will alter this situation drastically.

(P9) If a, b, and c are any numbers, then

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

(Notice that the equation $(b + c) \cdot a = b \cdot a + c \cdot a$ is also true, by P8.)

As an example of the usefulness of P9 we will now determine just when $a - b = b - a$:

$$\begin{array}{ll} \text{If} & a - b = b - a, \\ \text{then} & (a - b) + b = (b - a) + b = b + (b - a); \\ \text{hence} & a = b + b - a; \\ \text{hence} & a + a = (b + b - a) + a = b + b. \\ \text{Consequently} & a \cdot (1 + 1) = b \cdot (1 + 1), \\ \text{and therefore} & a = b. \end{array}$$

A second use of P9 is the justification of the assertion $a \cdot 0 = 0$ which we have already made, and even used in a proof on page 6 (can you find where?). This fact was not listed as one of the basic properties, even though no proof was offered when it was first mentioned. With P1–P8 alone a proof was not possible, since the number 0 appears only in P2 and P3, which concern addition, while the assertion in question involves multiplication. With P9 the proof is simple, though perhaps not obvious: We have

$$\begin{aligned} a \cdot 0 + a \cdot 0 &= a \cdot (0 + 0) \\ &= a \cdot 0; \end{aligned}$$

as we have already noted, this immediately implies (by adding $(a \cdot 0)$ to both sides) that $a \cdot 0 = 0$.

A series of further consequences of P9 may help explain the somewhat mysterious rule that the product of two negative numbers is positive. To begin with, we will establish the more easily acceptable assertion that $(-a) \cdot b = -(a \cdot b)$. To prove this, note that

$$\begin{aligned} (-a) \cdot b + a \cdot b &= [(-a) + a] \cdot b \\ &= 0 \cdot b \\ &= 0. \end{aligned}$$

It follows immediately (by adding $-(a \cdot b)$ to both sides) that $(-a) \cdot b = -(a \cdot b)$. Now note that

$$\begin{aligned} (-a) \cdot (-b) + [-(a \cdot b)] &= (-a) \cdot (-b) + (-a) \cdot b \\ &= (-a) \cdot [(-b) + b] \\ &= (-a) \cdot 0 \\ &= 0. \end{aligned}$$

Consequently, adding $(a \cdot b)$ to both sides, we obtain

$$(-a) \cdot (-b) = a \cdot b.$$

The fact that the product of two negative numbers is positive is thus a consequence of P1–P9. In other words, *if we want P1 to P9 to be true, the rule for the product of two negative numbers is forced upon us.*

The various consequences of P9 examined so far, although interesting and important, do not really indicate the significance of P9; after all we could have listed each of these properties separately. Actually, P9 is the justification for almost all algebraic manipulations. For example, although we have shown how to solve the equation

$$(x - 1)(x - 2) = 0,$$

we can hardly expect to be presented with an equation in this form. We are more likely to be confronted with the equation

$$x^2 - 3x + 2 = 0.$$

The "factorization" $x^2 - 3x + 2 = (x - 1)(x - 2)$ is really a triple use of P9:

$$\begin{aligned}(x - 1) \cdot (x - 2) &= x \cdot (x - 2) + (-1) \cdot (x - 2)\\ &= x \cdot x + x \cdot (-2) + (-1) \cdot x + (-1) \cdot (-2)\\ &= x^2 + x[(-2) + (-1)] + 2\\ &= x^2 - 3x + 2.\end{aligned}$$

A final illustration of the importance of P9 is the fact that this property is actually used every time one multiplies arabic numerals. For example, the calculation

$$\begin{array}{r} 13\\ \times 24\\ \hline 52\\ 26\\ \hline 312\end{array}$$

is a concise arrangement for the following equations:

$$\begin{aligned}13 \cdot 24 &= 13 \cdot (2 \cdot 10 + 4)\\ &= 13 \cdot 2 \cdot 10 + 13 \cdot 4\\ &= 26 \cdot 10 + 52.\end{aligned}$$

(Note that moving 26 to the left in the above calculation is the same as writing $26 \cdot 10$.) The multiplication $13 \cdot 4 = 52$ uses P9 also:

$$\begin{aligned}13 \cdot 4 &= (1 \cdot 10 + 3) \cdot 4\\ &= 1 \cdot 10 \cdot 4 + 3 \cdot 4\\ &= 4 \cdot 10 + 12\\ &= 4 \cdot 10 + 1 \cdot 10 + 2\\ &= (4 + 1) \cdot 10 + 2\\ &= 5 \cdot 10 + 2\\ &= 52.\end{aligned}$$

The properties P1–P9 have descriptive names which are not essential to remember, but which are often convenient for reference. We will take this opportunity to list properties P1–P9 together and indicate the names by which they are commonly designated.

(P1)	(Associative law for addition)	$a + (b + c) = (a + b) + c.$
(P2)	(Existence of an additive identity)	$a + 0 = 0 + a = a.$
(P3)	(Existence of additive inverses)	$a + (-a) = (-a) + a = 0.$
(P4)	(Commutative law for addition)	$a + b = b + a.$
(P5)	(Associative law for multiplication)	$a \cdot (b \cdot c) = (a \cdot b) \cdot c.$
(P6)	(Existence of a multiplicative identity)	$a \cdot 1 = 1 \cdot a = a; \quad 1 \neq 0.$
(P7)	(Existence of multiplicative inverses)	$a \cdot a^{-1} = a^{-1} \cdot a = 1$, for $a \neq 0.$
(P8)	(Commutative law for multiplication)	$a \cdot b = b \cdot a.$
(P9)	(Distributive law)	$a \cdot (b + c) = a \cdot b + a \cdot c.$

The three basic properties of numbers which remain to be listed are concerned with inequalities. Although inequalities occur rarely in elementary mathematics, they play a prominent role in calculus. The two notions of inequality, $a < b$ (a is less than b) and $a > b$ (a is greater than b), are intimately related: $a < b$ means the same as $b > a$ (thus $1 < 3$ and $3 > 1$ are merely two ways of writing the same assertion). The numbers a satisfying $a > 0$ are called **positive**, while those numbers a satisfying $a < 0$ are called **negative.** While positivity can thus be defined in terms of $<$, it is possible to reverse the procedure: $a < b$ can be defined to mean that $b - a$ is positive. In fact, it is convenient to consider the collection of all positive numbers, denoted by P, as the basic concept, and state all properties in terms of P:

(P10) (Trichotomy law) For every number a, one and only one of the following holds:

- (i) $a = 0$,
- (ii) a is in the collection P,
- (iii) $-a$ is in the collection P.

(P11) (Closure under addition) If a and b are in P, then $a + b$ is in P.

(P12) (Closure under multiplication) If a and b are in P, then $a \cdot b$ is in P.

These three properties should be complemented with the following definitions:

$$\begin{array}{lll} a > b & \text{if} & a - b \text{ is in } P; \\ a < b & \text{if} & b > a; \\ a \geq b & \text{if} & a > b \text{ or } a = b; \\ a \leq b & \text{if} & a < b \text{ or } a = b.^* \end{array}$$

Note, in particular, that $a > 0$ if and only if a is in P.

All the familiar facts about inequalities, however elementary they may seem, are consequences of P10–P12. For example, if a and b are any two numbers, then precisely one of the following holds:

(i) $a - b = 0$,
(ii) $a - b$ is in the collection P,
(iii) $-(a - b) = b - a$ is in the collection P.

Using the definitions just made, it follows that precisely one of the following holds:

(i) $a = b$,
(ii) $a > b$,
(iii) $b > a$.

A slightly more interesting fact results from the following manipulations. If $a < b$, so that $b - a$ is in P, then surely $(b + c) - (a + c)$ is in P; thus, if $a < b$, then $a + c < b + c$. Similarly, suppose $a < b$ and $b < c$. Then

$$\begin{array}{ll} & b - a \text{ is in } P, \\ \text{and} & c - b \text{ is in } P, \\ \text{so} & c - a = (c - b) + (b - a) \text{ is in } P. \end{array}$$

This shows that if $a < b$ and $b < c$, then $a < c$. (The two inequalities $a < b$ and $b < c$ are usually written in the abbreviated form $a < b < c$, which has the third inequality $a < c$ almost built in.)

The following assertion is somewhat less obvious: If $a < 0$ and $b < 0$, then $ab > 0$. The only difficulty presented by the proof is the unraveling of definitions. The symbol $a < 0$ means, by definition, $0 > a$, which means $0 - a = -a$ is in P. Similarly $-b$ is in P, and consequently, by P12, $(-a)(-b) = ab$ is in P. Thus $ab > 0$.

The fact that $ab > 0$ if $a > 0$, $b > 0$ and also if $a < 0$, $b < 0$ has one special consequence: $a^2 > 0$ if $a \neq 0$. Thus squares of nonzero numbers are

* There is one slightly perplexing feature of the symbols $\geq$ and $\leq$. The statements

$$\begin{array}{l} 1 + 1 \leq 3 \\ 1 + 1 \leq 2 \end{array}$$

are both true, even though we know that $\leq$ could be replaced by $<$ in the first, and by $=$ in the second. This sort of thing is bound to occur when $\leq$ is used with specific numbers; the usefulness of the symbol is revealed by a statement like Theorem 1—here equality holds for some values of a and b, while inequality holds for other values.

always positive, and in particular we have proved a result which might have seemed sufficiently elementary to be included in our list of properties: $1 > 0$ (since $1 = 1^2$).

The fact that $-a > 0$ if $a < 0$ is the basis of a concept which will play an extremely important role in this book. For any number a, we define the **absolute value** $|a|$ of a as follows:

$$|a| = \begin{cases} a, & a \geq 0 \\ -a, & a \leq 0. \end{cases}$$

Note that $|a|$ is always positive, except when $a = 0$. For example, we have $|-3| = 3$, $|7| = 7$, $|1 + \sqrt{2} - \sqrt{3}| = 1 + \sqrt{2} - \sqrt{3}$, and $|1 + \sqrt{2} - \sqrt{10}| = \sqrt{10} - \sqrt{2} - 1$. In general, the most straightforward approach to any problem involving absolute values requires treating several cases separately, since absolute values are defined by cases to begin with. This approach may be used to prove the following very important fact about absolute values.

THEOREM 1 For all numbers a and b, we have

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

PROOF We will consider 4 cases:

$$\begin{array}{lll} (1) & a \geq 0, & b \geq 0; \\ (2) & a \geq 0, & b \leq 0; \\ (3) & a \leq 0, & b \geq 0; \\ (4) & a \leq 0, & b \leq 0. \end{array}$$

In case (1) we also have $a + b \geq 0$, and the theorem is obvious; in fact,

$$|a + b| = a + b = |a| + |b|,$$

so that in this case equality holds.

In case (4) we have $a + b \leq 0$, and again equality holds:

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

In case (2), when $a \geq 0$ and $b \leq 0$, we must prove that

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

This case may therefore be divided into two subcases. If $a + b \geq 0$, then we must prove that

$$\begin{aligned} a + b &\leq a - b, \\ \text{i.e.,} \qquad b &\leq -b, \end{aligned}$$

which is certainly true since b is negative and $-b$ is positive. On the other hand, if $a + b \leq 0$, we must prove that

$$\begin{aligned} -a - b &\leq a - b, \\ \text{i.e.,} \qquad -a &\leq a, \end{aligned}$$

which is certainly true since a is positive and $-a$ is negative.

Finally, note that case (3) may be disposed of with no additional work, by applying case (2) with a and b interchanged. ▮

Although this method of treating absolute values (separate consideration of various cases) is sometimes the only approach available, there are often simpler methods which may be used. In fact, it is possible to give a much shorter proof of Theorem 1; this proof is motivated by the observation that

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

(Here, and throughout the book, $\sqrt{x}$ denotes the *positive* square root of x; this symbol is defined only when $x \geq 0$.) We may now observe that

$$\begin{aligned}(|a + b|)^2 = (a + b)^2 &= a^2 + 2ab + b^2 \\ &\leq a^2 + 2|a| \cdot |b| + b^2 \\ &= |a|^2 + 2|a| \cdot |b| + |b|^2 \\ &= (|a| + |b|)^2.\end{aligned}$$

From this we can conclude that $|a + b| \leq |a| + |b|$ because $x^2 < y^2$ implies $x < y$, provided that x and y are both non-negative; a proof of *this* fact is left to the reader (Problem 5).

One final observation may be made about the theorem we have just proved: a close examination of either proof offered shows that

$$|a + b| = |a| + |b|$$

if a and b have the same sign (i.e., are both positive or both negative), or if one of the two is 0, while

$$|a + b| < |a| + |b|$$

if a and b are of opposite signs.

We will conclude this chapter with a subtle point, neglected until now, whose inclusion is required in a conscientious survey of the properties of numbers. After stating property P9, we proved that $a - b = b - a$ implies $a = b$. The proof began by establishing that

$$a \cdot (1 + 1) = b \cdot (1 + 1),$$

from which we concluded that $a = b$. This result is obtained from the equation $a \cdot (1 + 1) = b \cdot (1 + 1)$ by dividing both sides by $1 + 1$. Division by 0 should be avoided scrupulously, and it must therefore be admitted that the validity of the argument depends on knowing that $1 + 1 \neq 0$. Problem 24 is designed to convince you that this fact cannot possibly be proved from properties P1–P9 alone! Once P10, P11, and P12 are available, however, the proof is very simple: We have already seen that $1 > 0$; it follows that $1 + 1 > 0$, and in particular $1 + 1 \neq 0$.

This last demonstration has perhaps only strengthened your feeling that it is absurd to bother proving such obvious facts, but an honest assessment of our present situation will help justify serious consideration of such details. In this chapter we have assumed that numbers are familiar objects, and that P1–P12 are merely explicit statements of obvious, well-known properties of numbers. It would be difficult, however, to justify this assumption. Although one learns how to "work with" numbers in school, just what numbers *are*, remains rather vague. A great deal of this book is devoted to elucidating the concept of numbers, and by the end of the book we will have become quite well acquainted with them. But it will be necessary to work with numbers throughout the book. It is therefore reasonable to admit frankly that we do not yet thoroughly understand numbers; we may still say that, in whatever way numbers are finally defined, they should certainly have properties P1–P12.

Most of this chapter has been an attempt to present convincing evidence that P1–P12 are indeed basic properties which we should assume in order to deduce other familiar properties of numbers. Some of the problems (which indicate the derivation of other facts about numbers from P1–P12) are offered as further evidence. It is still a crucial question whether P1–P12 actually account for *all* properties of numbers. As a matter of fact, we shall soon see that they do *not*. In the next chapter the deficiencies of properties P1–P12 will become quite clear, but the proper means for correcting these deficiencies is not so easily discovered. The crucial additional basic property of numbers which we are seeking is profound and subtle, quite unlike P1–P12. The discovery of this crucial property will require all the work of Part II of this book. In the remainder of Part I we will begin to see why some additional property is required; in order to investigate this we will have to consider a little more carefully what we mean by "numbers."

PROBLEMS

1. Prove the following:

(i) If $ax = a$ for some number $a \neq 0$, then $x = 1$.
(ii) $x^2 - y^2 = (x - y)(x + y)$.
(iii) If $x^2 = y^2$, then $x = y$ or $x = -y$.
(iv) $x^3 - y^3 = (x - y)(x^2 + xy + y^2)$.
(v) $x^n - y^n = (x - y)(x^{n-1} + x^{n-2}y + \cdots + xy^{n-2} + y^{n-1})$.
(vi) $x^3 + y^3 = (x + y)(x^2 - xy + y^2)$. (There is a particularly easy way to do this, using (iv), and it will show you how to find a factorization for $x^n + y^n$ whenever n is odd.)

2. What is wrong with the following "proof"? Let $x = y$. Then

$$\begin{aligned} x^2 &= xy, \\ x^2 - y^2 &= xy - y^2, \\ (x + y)(x - y) &= y(x - y), \\ x + y &= y, \\ 2y &= y, \\ 2 &= 1. \end{aligned}$$

3. Prove the following:

(i) $\frac{a}{b} = \frac{ac}{bc}$, if $b, c \neq 0$.

(ii) $\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$, if $b, d \neq 0$.

(iii) $(ab)^{-1} = a^{-1}b^{-1}$, if $a, b \neq 0$. (To do this you must remember the defining property of $(ab)^{-1}$.)

(iv) $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{db}$, if $b, d \neq 0$.

(v) $\frac{a}{b} \Big/ \frac{c}{d} = \frac{ad}{bc}$, if $b, c, d \neq 0$.

(vi) If $b, d \neq 0$, then $\frac{a}{b} = \frac{c}{d}$ if and only if $ad = bc$. Also determine when $\frac{a}{b} = \frac{b}{a}$.

4. Find all numbers x for which

(i) $4 - x < 3 - 2x$.
(ii) $5 - x^2 < 8$.
(iii) $5 - x^2 < -2$.
(iv) $(x - 1)(x - 3) > 0$. (When is a product of two numbers positive?)
(v) $x^2 - 2x + 2 > 0$.
(vi) $x^2 + x + 1 > 2$.
(vii) $x^2 - x + 10 > 16$.
(viii) $x^2 + x + 1 > 0$.
(ix) $(x - \pi)(x + 5)(x - 3) > 0$.
(x) $(x - \sqrt[3]{2})(x - \sqrt{2}) > 0$.
(xi) $2^x < 8$.
(xii) $x + 3^x < 4$.
(xiii) $\frac{1}{x} + \frac{1}{1 - x} > 0$.
(xiv) $\frac{x - 1}{x + 1} > 0$.

5. Prove the following:

(i) If $a < b$ and $c < d$, then $a + c < b + d$.
(ii) If $a < b$, then $-b < -a$.

(iii) If $a < b$ and $c > d$, then $a - c < b - d$.
(iv) If $a < b$ and $c > 0$, then $ac < bc$.
(v) If $a < b$ and $c < 0$, then $ac > bc$.
(vi) If $a > 1$, then $a^2 > a$.
(vii) If $0 < a < 1$, then $a^2 < a$.
(viii) If $0 \leq a < b$ and $0 \leq c < d$, then $ac < bd$.
(ix) If $0 \leq a < b$, then $a^2 < b^2$. (Use (viii).)
(x) If $a, b \geq 0$ and $a^2 < b^2$, then $a < b$. (Use (ix), backwards.)

6. Prove that if $0 < a < b$, then

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

Notice that the inequality $\sqrt{ab} \leq (a+b)/2$ holds for $a, b > 0$, without the additional assumption $a < b$. A generalization of this fact occurs in Problem 2-20.

***7.** (a) Prove that if $x^n = y^n$ and n is odd, then $x = y$. Hint: First explain why it suffices to consider only the case $x, y > 0$; then show that $x < y$ and $y > x$ are both impossible.
(b) Prove that if $x^n = y^n$ and n is even, then $x = y$ or $x = -y$.

***8.** Although the basic properties of inequalities were stated in terms of the collection P of all positive numbers, and $<$ was defined in terms of P, this procedure can be reversed. Suppose that P10–P12 are replaced by

(P′10) For any numbers a and b one, and only one, of the following holds:

(i) $a = b$,
(ii) $a < b$,
(iii) $b < a$.

(P′11) For any numbers a, b, and c, if $a < b$ and $b < c$, then $a < c$.
(P′12) For any numbers a, b, and c, if $a < b$, then $a + c < b + c$.
(P′13) For any numbers a, b, and c, if $a < b$ and $0 < c$, then $ac < bc$.

Show that P10–P12 can then be deduced as theorems.

9. Express each of the following with at least one less pair of absolute value signs.

(i) $|\sqrt{2} + \sqrt{3} - \sqrt{5} + \sqrt{7}|$.
(ii) $|(|a + b| - |a| - |b|)|$.
(iii) $|(|a + b| + |c| - |a + b + c|)|$.
(iv) $|x^2 - 2xy + y^2|$.
(v) $|(|\sqrt{2} + \sqrt{3}| - |\sqrt{5} - \sqrt{7}|)|$.

10. Express each of the following without absolute value signs, treating various cases separately when necessary.

(i) $|a + b| - |b|$.
(ii) $|(|x| - 1)|$.
(iii) $|x| - |x^2|$.
(iv) $a - |(a - |a|)|$.

11. Find all numbers x for which

(i) $|x - 3| = 8$.
(ii) $|x - 3| < 8$.
(iii) $|x + 4| < 2$.
(iv) $|x - 1| + |x - 2| > 1$.
(v) $|x - 1| + |x + 1| < 2$.
(vi) $|x - 1| + |x + 1| < 1$.
(vii) $|x - 1| \cdot |x + 1| = 0$.
(viii) $|x - 1| \cdot |x + 2| = 3$.

12. Prove the following:

(i) $|xy| = |x| \cdot |y|$.
(ii) $\left|\frac{1}{x}\right| = \frac{1}{|x|}$, if $x \neq 0$. (The best way to do this is to remember what $|x|^{-1}$ is.)
(iii) $\frac{|x|}{|y|} = \left|\frac{x}{y}\right|$, if $y \neq 0$.
(iv) $|x - y| \leq |x| + |y|$. (Give a very short proof.)
(v) $|x| - |y| \leq |x - y|$. (A very short proof is possible, if you write things in the right way.)
(vi) $|(|x| - |y|)| \leq |x - y|$. (Why does this follow immediately from (v)?)
(vii) $|x + y + z| \leq |x| + |y| + |z|$. Indicate when equality holds, and prove your statement.

13. The maximum of two numbers x and y is denoted by $\max(x, y)$. Thus $\max(-1, 3) = \max(3, 3) = 3$ and $\max(-1, -4) = \max(-4, -1) = -1$. The minimum of x and y is denoted by $\min(x, y)$. Prove that

$$\max(x, y) = \frac{x + y + |y - x|}{2},$$

$$\min(x, y) = \frac{x + y - |y - x|}{2}.$$

Derive a formula for $\max(x, y, z)$ and $\min(x, y, z)$, using, for example,

$$\max(x, y, z) = \max(x, \max(y, z)).$$

14. (a) Prove that $|a| = |-a|$. (The trick is not to become confused by too many cases. First prove the statement for $a \geq 0$. Why is it then obvious for $a \leq 0$?)

(b) Prove that $-b \leq a \leq b$ if and only if $|a| \leq b$. In particular, it follows that $-|a| \leq a \leq |a|$.

(c) Use this fact to give a new proof that $|a + b| \leq |a| + |b|$.

***15.** (a) Use Problem 1 and Problem 7 to prove that if x and y are not both 0, then

$$x^2 + xy + y^2 \neq 0,$$
$$x^4 + x^3y + x^2y^2 + xy^3 + y^4 \neq 0.$$

For every number $x \neq 0$, each of these expressions is positive for some positive number y and also for some negative y (namely, $y = \pm x$); it therefore seems reasonable that the $\neq$ signs can be replaced by $>$ signs. This maneuver is valid, but we are not yet in a position to prove this (see Problem 7-9). Parts (b) and (d) of this problem provide a direct demonstration that the $>$ signs hold.

(b) Using the fact that

$$x^2 + 2xy + y^2 = (x + y)^2 \geq 0,$$

show that the assumption $x^2 + xy + y^2 < 0$ leads to a contradiction.

(c) Show similarly that if x and y are not both 0, then

$$4x^2 + 6xy + 4y^2 > 0,$$
$$3x^2 + 5xy + 3y^2 > 0.$$

**(d) Show that if x and y are not both 0, then

$$x^4 + x^3y + x^2y^2 + xy^3 + y^4 > 0.$$

***16.** (a) Show that

$$(x + y)^2 = x^2 + y^2 \quad \text{only when } x = 0 \text{ or } y = 0,$$
$$(x + y)^3 = x^3 + y^3 \quad \text{only when } x = 0 \text{ or } y = 0 \text{ or } x = -y.$$

(b) Use Problem 15 to find out when $(x + y)^4 = x^4 + y^4$.

**(c) Find out when $(x + y)^5 = x^5 + y^5$. Hint: From the assumption $(x + y)^5 = x^5 + y^5$ you should be able to derive the equation $x^3 + 2x^2y + 2xy^2 + y^3 = 0$, if $xy \neq 0$. This implies that $(x + y)^3 = x^2y + xy^2 = xy(x + y)$.

You should now be able to make a good guess as to when $(x + y)^n = x^n + y^n$; the proof is contained in Problem 11-41.

17. (a) Suppose that $b^2 - 4c \geq 0$. Show that the numbers

$$\frac{-b + \sqrt{b^2 - 4c}}{2}, \qquad \frac{-b - \sqrt{b^2 - 4c}}{2}$$

both satisfy the equation $x^2 + bx + c = 0$.

(b) Suppose that $b^2 - 4c < 0$. Show that there are no numbers x satisfying $x^2 + bx + c = 0$; in fact, $x^2 + bx + c > 0$ for all x. Hint: "Complete the square," i.e., write $x^2 + bx + c = (x + b/2)^2 + ?$

(c) Use this fact to give another proof that if x and y are not both 0, then $x^2 + xy + y^2 > 0$.

(d) For which numbers α is it true that $x^2 + \alpha xy + y^2 > 0$ whenever x and y are not both 0?

(e) Find the smallest possible value of $x^2 + bx + c$ and of $ax^2 + bx + c$, for $a \neq 0$. (Use the trick in part (b).)

18. The fact that $a^2 \geq 0$ for all numbers a, elementary as it may seem, is nevertheless the fundamental idea upon which most important inequalities are ultimately based. The great-granddaddy of all inequalities is the *Schwartz inequality:*

$$x_1y_1 + x_2y_2 \leq \sqrt{x_1^2 + x_2^2}\sqrt{y_1^2 + y_2^2}.$$

The three proofs of the Schwartz inequality outlined below have only one thing in common—their reliance on the fact that $a^2 \geq 0$ for all a.

(a) Prove the Schwartz inequality by first proving that

$$(x_1^2 + x_2^2)(y_1^2 + y_2^2) = (x_1y_1 + x_2y_2)^2 + (x_1y_2 - x_2y_1)^2.$$

(b) Prove that if $x_1 = \lambda y_1$ and $x_2 = \lambda y_2$ for some number λ, then equality holds in the Schwartz inequality. Prove the same thing if $y_1 = y_2 = 0$. Now suppose that y_1 and y_2 are not both 0, and that there is no number λ such that $x_1 = \lambda y_1$ and $x_2 = \lambda y_2$. Then

$$\begin{aligned} 0 &< (\lambda y_1 - x_1)^2 + (\lambda y_2 - x_2)^2 \\ &= \lambda^2(y_1^2 + y_2^2) - 2\lambda(x_1y_1 + x_2y_2) + (x_1^2 + x_2^2). \end{aligned}$$

Using Problem 17, complete the proof of the Schwartz inequality.

(c) Prove the Schwartz inequality by using $2xy \leq x^2 + y^2$ (how is this derived?) with

$$x = \frac{x_i}{\sqrt{x_1^2 + x_2^2}}, \qquad y = \frac{y_i}{\sqrt{y_1^2 + y_2^2}},$$

first for $i = 1$ and then for $i = 2$.

(d) Deduce, from each of these three proofs, that equality holds only when $y_1 = y_2 = 0$ or when there is a number λ such that $x_1 = \lambda y_1$ and $x_2 = \lambda y_2$.

In our later work, three facts about inequalities will be crucial. Although proofs will be supplied at the appropriate point in the text, a personal assault on these problems is infinitely more enlightening than a perusal of a completely worked-out proof. The statements of these propositions involve some weird numbers, but their basic message is very simple: if x is close enough to x_0, and y is close enough to y_0, then $x + y$ will be close to $x_0 + y_0$, and xy will be close to x_0y_0, and $1/y$ will be close to $1/y_0$. The symbol "ε" which appears in these propositions is the fifth letter of the Greek alphabet ("epsilon"), and could just as well be replaced by a less intimidating Roman letter; however, tradition has made the use of ε almost sacrosanct in the contexts to which these theorems apply.

19. Prove that if

$$|x - x_0| < \frac{\varepsilon}{2} \quad \text{and} \quad |y - y_0| < \frac{\varepsilon}{2},$$

then

$$|(x + y) - (x_0 + y_0)| < \varepsilon,$$
$$|(x - y) - (x_0 - y_0)| < \varepsilon.$$

***20.** Prove that if

$$|x - x_0| < \min\left(\frac{\varepsilon}{2(|y_0| + 1)}, 1\right) \quad \text{and} \quad |y - y_0| < \frac{\varepsilon}{2(|x_0| + 1)},$$

then $|xy - x_0y_0| < \varepsilon$.
(The notation "min" was defined in Problem 13, but the formula provided by that problem is irrelevant at the moment; the first inequality in the hypothesis just means that

$$|x - x_0| < \frac{\varepsilon}{2(|y_0| + 1)} \quad \text{and} \quad |x - x_0| < 1;$$

at one point in the argument you will need the first inequality, and at another point you will need the second. One more word of advice: since the hypotheses only provide information about $x - x_0$ and $y - y_0$, it is almost a foregone conclusion that the proof will depend upon writing $xy - x_0y_0$ in a way that involves $x - x_0$ and $y - y_0$.)

***21.** Prove that if $y_0 \neq 0$ and

$$|y - y_0| < \min\left(\frac{|y_0|}{2}, \frac{\varepsilon|y_0|^2}{2}\right),$$

then $y \neq 0$ and

$$\left|\frac{1}{y} - \frac{1}{y_0}\right| < \varepsilon.$$

***22.** Replace the question marks in the following statement by expressions involving ε, x_0, and y_0 so that the conclusion will be true:

If $y_0 \neq 0$ and

$$|y - y_0| < ? \quad \text{and} \quad |x - x_0| < ?$$

then $y \neq 0$ and

$$\left|\frac{x}{y} - \frac{x_0}{y_0}\right| < \varepsilon.$$

This problem is trivial in the sense that its solution follows from Problems 20 and 21 with almost no work at all (notice that $x/y = x \cdot 1/y$). The crucial point is not to become confused; decide which of the two problems should be used first, and don't panic if your answer looks unlikely.

***23.** This problem shows that the actual placement of parentheses in a sum is irrelevant. The proofs involve "mathematical induction"; if you are not familiar with such proofs, but still want to tackle this problem, it can be saved until after Chapter 2, where proofs by induction are explained.

Let us agree, for definiteness, that $a_1 + \cdots + a_n$ will denote

$$a_1 + (a_2 + (a_3 + \cdots + (a_{n-2} + (a_{n-1} + a_n))) \cdots).$$

Thus $a_1 + a_2 + a_3$ denotes $a_1 + (a_2 + a_3)$, and $a_1 + a_2 + a_3 + a_4$ denotes $a_1 + (a_2 + (a_3 + a_4))$, etc.

(a) Prove that

$$(a_1 + \cdots + a_k) + a_{k+1} = a_1 + \cdots + a_{k+1}.$$

Hint: Use induction on k.

(b) Prove that if $n \geq k$, then

$$(a_1 + \cdots + a_k) + (a_{k+1} + \cdots + a_n) = a_1 + \cdots + a_n.$$

Hint: Use part (a) to give a proof by induction on k.

(c) Let $s(a_1, \ldots, a_k)$ be some sum formed from $a_1, \ldots, a_k$. Show that

$$s(a_1, \ldots, a_k) = a_1 + \cdots + a_k.$$

Hint: There must be two sums $s'(a_1, \ldots, a_l)$ and $s''(a_{l+1}, \ldots, a_k)$ such that

$$s(a_1, \ldots, a_k) = s'(a_1, \ldots, a_l) + s''(a_{l+1}, \ldots, a_k).$$

24. Suppose that we interpret "number" to mean either 0 or 1, and $+$ and $\cdot$ to be the operations defined by the following two tables.

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Check that properties P1–P9 all hold, even though $1 + 1 = 0$.

CHAPTER 2 NUMBERS OF VARIOUS SORTS

In Chapter 1 we used the word "number" very loosely, despite our concern with the basic properties of numbers. It will now be necessary to distinguish carefully various kinds of numbers.

The simplest numbers are the "counting numbers"

$$1, 2, 3, \ldots .$$

The fundamental significance of this collection of numbers is emphasized by its symbol **N** (for **natural numbers**). A brief glance at P1–P12 will show that our basic properties of "numbers" do not apply to **N**—for example, P2 and P3 do not make sense for **N**. From this point of view the system **N** has many deficiencies. Nevertheless, **N** is sufficiently important to deserve several comments before we consider larger collections of numbers.

The most basic property of **N** is the principle of "mathematical induction." Suppose $P(x)$ means that the property P holds for the number x. Then the principle of mathematical induction states that $P(x)$ is true for all natural numbers x provided that

(1) $P(1)$ is true.
(2) Whenever $P(k)$ is true, $P(k+1)$ is true.

Note that condition (2) merely asserts the truth of $P(k+1)$ under the assumption that $P(k)$ is true; this suffices to ensure the truth of $P(x)$ for all x, if condition (1) also holds. In fact, if $P(1)$ is true, then it follows that $P(2)$ is true (by using (2) in the special case $k = 1$). Now, since $P(2)$ is true it follows that $P(3)$ is true (using (2) in the special case $k = 2$). It is clear that each number will eventually be reached by a series of steps of this sort, so that $P(k)$ is true for all numbers k.

A favorite illustration of the reasoning behind mathematical induction envisions an infinite line of people,

person number 1, person number 2, person number 3, ...

If each person has been instructed to tell any secret he hears to the person behind him (the one with the next largest number) and a secret is told to person number 1, then clearly every person will eventually learn the secret. If $P(x)$ is the assertion that person number x will learn the secret, then the instructions given (to tell all secrets learned to the next person) assures that condition (2) is true, and telling the secret to person number 1 makes (1) true. The following example is a less facetious use of mathematical induction. There is a useful and striking formula which expresses the sum of the first n numbers in a simple way:

$$1 + \cdots + n = \frac{n(n+1)}{2}.$$

To prove this formula, note first that it is clearly true for $n = 1$. Now ***assume*** that for some integer k we have

$$1 + \cdots + k = \frac{k(k+1)}{2}.$$

Then

$$\begin{aligned} 1 + \cdots + k + (k+1) &= \frac{k(k+1)}{2} + k + 1 \\ &= \frac{k(k+1) + 2k + 2}{2} \\ &= \frac{k^2 + 3k + 2}{2} \\ &= \frac{(k+1)(k+2)}{2}, \end{aligned}$$

so the formula is also true for $k + 1$. By the principle of induction this proves the formula for all natural numbers n. This particular example illustrates a phenomenon that frequently occurs, especially in connection with formulas like the one just proved. Although the proof by induction is often quite straightforward, the method by which the formula was discovered remains a mystery. Problems 4 and 5 indicate how some formulas of this type may be derived.

The principle of mathematical induction may be formulated in an equivalent way without speaking of "properties" of a number, a term which is sufficiently vague to be eschewed in a mathematical discussion. A more precise formulation states that if A is any collection (or "set"—a synonymous mathematical term) of natural numbers and

(1) 1 is in A,
(2) $k + 1$ is in A whenever k is in A,

then A is the set of all natural numbers. It should be clear that this formulation adequately replaces the less formal one given previously—we just consider the set A of natural numbers x which satisfy $P(x)$. For example, suppose A is the set of natural numbers n for which it is true that

$$1 + \cdots + n = \frac{n(n+1)}{2}.$$

Our previous proof of this formula showed that A contains 1, and that $k + 1$ is in A, if k is. It follows that A is the set of all natural numbers, i.e., that the formula holds for all natural numbers n.

There is yet another rigorous formulation of the principle of mathematical induction, which looks quite different. If A is any collection of natural numbers, it is tempting to say that A must have a smallest member. Actually, this statement can fail to be true in a rather subtle way. A particularly important set of natural numbers is the collection A that contains no natural numbers at all, the "empty collection" or "null set,"* denoted by $\emptyset$. The null set $\emptyset$ is a collection of natural numbers that has no smallest member—in fact, it has no members at all. This is the only possible exception, however; if A is a nonnull set of natural numbers, then A has a least member. This "intuitively obvious" statement, known as the "well-ordering principle," can be proved from the principle of induction as follows. Suppose that the set A has no least member. Let B be the set of natural numbers n such that $1, \ldots, n$ are all *not* in A. Clearly 1 is in B (because if 1 were in A, then A would have 1 as smallest member). Moreover, if $1, \ldots, k$ are not in A, surely $k + 1$ is not in A (otherwise $k + 1$ would be the smallest member of A), so $1, \ldots, k + 1$ are all not in A. This shows that if k is in B, then $k + 1$ is in B. It follows that every number n is in B, i.e., the numbers $1, \ldots, n$ are *not* in A for any natural number n. Thus $A = \emptyset$, which completes the proof.

It is also possible to prove the principle of induction from the well-ordering principle (Problem 9). Either principle may be considered as a basic assumption about the natural numbers.

There is still another form of induction which should be mentioned. It sometimes happens that in order to prove $P(k + 1)$ we must assume not only $P(k)$, but also $P(l)$ for all natural numbers $l \leq k$. In this case we rely on the "principle of complete induction": If A is a set of natural numbers and

(1) 1 is in A,
(2) $k + 1$ is in A if $1, \ldots, k$ are in A,

then A is the set of all natural numbers.

Although the principle of complete induction may appear much stronger than the ordinary principle of induction, it is actually a consequence of that principle. The proof of this fact is left to the reader, with a hint (Problem 10). Applications will be found in Problems 6, 16, 19, and 20.

Closely related to proofs by induction are "recursive definitions." For example, the number $n!$ (read "n factorial") is defined as the product of all the natural numbers less than or equal to n:

$$n! = 1 \cdot 2 \cdot \ldots \cdot (n - 1) \cdot n.$$

This can be expressed more precisely as follows:

(1) $1! = 1$,
(2) $n! = n \cdot (n - 1)!$.

This form of the definition exhibits the relationship between $n!$ and $(n - 1)!$

* Although it may not strike you as a collection, in the ordinary sense of the word, the null set arises quite naturally in many contexts. We frequently consider the set A, consisting of all x satisfying some property P; often we have no guarantee that P is satisfied by *any* number, so that A might be $\emptyset$—in fact often one proves that P is always false by showing that $A = \emptyset$.

in an explicit way that is ideally suited for proofs by induction. Problem 21 reviews a definition already familiar to you, which may be expressed more succinctly as a recursive definition; as this problem shows, the recursive definition is really necessary for a rigorous proof of some of the basic properties of the definition.

One definition which may not be familiar involves some convenient notation which we will constantly be using. Instead of writing

$$a_1 + \cdots + a_n,$$

we will usually employ the Greek letter Σ (capital sigma, for "sum") and write

$$\sum_{i=1}^{n} a_i.$$

In other words, $\sum_{i=1}^{n} a_i$ denotes the sum of the numbers obtained by letting $i = 1, 2, \ldots, n$. Thus

$$\sum_{i=1}^{n} i = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

Notice that the letter i really has nothing to do with the number denoted by $\sum_{i=1}^{n} i$, and can be replaced by any convenient symbol (except n, of course!):

$$\sum_{j=1}^{n} j = \frac{n(n+1)}{2},$$

$$\sum_{j=1}^{i} j = \frac{i(i+1)}{2},$$

$$\sum_{n=1}^{j} n = \frac{j(j+1)}{2}.$$

To define $\sum_{i=1}^{n} a_i$ precisely really requires a recursive definition:

$$(1)\quad \sum_{i=1}^{1} a_i = a_1,$$

$$(2)\quad \sum_{i=1}^{n} a_i = \sum_{i=1}^{n-1} a_i + a_n.$$

But only purveyors of mathematical austerity would insist too strongly on such precision. In practice, all sorts of modifications of this symbolism are used, and no one ever considers it necessary to add any words of explanation.

The symbol

$$\sum_{\substack{i=1\\ i\neq 4}}^{n} a_i,$$

for example, is an obvious way of writing

$$a_1 + a_2 + a_3 + a_5 + a_6 + \cdots + a_n,$$

or more precisely,

$$\sum_{i=1}^{3} a_i + \sum_{i=5}^{n} a_i.$$

The deficiencies of the natural numbers which we discovered at the beginning of this chapter may be partially remedied by extending this system to the set of **integers**

$$\ldots, -2, -1, 0, 1, 2, \ldots.$$

This set is denoted by **Z** (from German "Zahl," number). Of properties P1–P12, only P7 fails for **Z**.

A still larger system of numbers is obtained by taking quotients m/n of integers (with $n \neq 0$). These numbers are called **rational numbers,** and the set of all rational numbers is denoted by **Q** (for "quotients"). In this system of numbers all of P1–P12 are true. It is tempting to conclude that the "properties of numbers," which we studied in some detail in Chapter 1, refer to just one set of numbers, namely, **Q**. There is, however, a still larger collection of numbers to which properties P1–P12 apply—the set of all **real numbers,** denoted by **R**. The real numbers include not only the rational numbers, but other numbers as well (the **irrational numbers**) which can be represented by infinite decimals; π and $\sqrt{2}$ are both examples of irrational numbers. The proof that π is irrational is not easy—we shall devote all of Chapter 16 of Part III to a proof of this fact. The irrationality of $\sqrt{2}$, on the other hand, is quite simple, and was known to the Greeks. (Since the Pythagorean theorem shows that an isosceles right triangle, with sides of length 1, has a hypotenuse of length $\sqrt{2}$, it is not surprising that the Greeks should have investigated this question.) The proof depends on a few observations about the natural numbers. Every natural number n can be written either in the form $2k$ for some integer k, or else in the form $2k + 1$ for some integer k (this "obvious" fact has a simple proof by induction (Problem 7)). Those natural numbers of the form $2k$ are called **even;** those of the form $2k + 1$ are called **odd.** Note that even numbers have even squares, and odd numbers have odd squares:

$$(2k)^2 = 4k^2 = 2 \cdot (2k^2),$$
$$(2k+1)^2 = 4k^2 + 4k + 1 = 2 \cdot (2k^2 + 2k) + 1.$$

In particular it follows that the converse must also hold: if n^2 is even, then n is even; if n^2 is odd, then n is odd. The proof that $\sqrt{2}$ is irrational is now quite simple. Suppose that $\sqrt{2}$ were rational; that is, suppose there were natural numbers p and q such that

$$\left(\frac{p}{q}\right)^2 = 2.$$

We can assume that p and q have no common divisor (since all common divisors could be divided out to begin with). Now we have

$$p^2 = 2q^2.$$

This shows that p^2 is even, and consequently p must be even; that is, $p = 2k$ for some natural number k. Then

$$p^2 = 4k^2 = 2q^2,$$

so

$$2k^2 = q^2.$$

This shows that q^2 is even, and consequently that q is even. Thus both p and q are even, contradicting the fact that p and q have no common divisor. This contradiction completes the proof.

It is important to understand precisely what this proof shows. We have demonstrated that there is no rational number x such that $x^2 = 2$. This assertion is often expressed more briefly by saying that $\sqrt{2}$ is irrational. Note, however, that the use of the symbol $\sqrt{2}$ implies the existence of *some* number (necessarily irrational) whose square is 2. We have not proved that such a number exists and we can assert confidently that, at present, a proof is *impossible* for us. Any proof at this stage would have to be based on P1–P12 (the only properties of **R** we have mentioned); since P1–P12 are also true for **Q** the exact same argument would show that there is a rational number whose square is 2, and this we know is false. (Note that the reverse argument will not work—our proof that there is no rational number whose square is 2 cannot be used to show that there is no real number whose square is 2, because our proof used not only P1–P12 but also a special property of **Q**, the fact that every number in **Q** can be written p/q for integers p and q.)

This particular deficiency in our list of properties of the real numbers could, of course, be corrected by adding a new property which asserts the existence of square roots of positive numbers. Resorting to such a measure is, however, neither aesthetically pleasing nor mathematically satisfactory; we would still not know that every number has an nth root if n is odd, and that every positive number has an nth root if n is even. Even if we assumed this, we could not prove the existence of a number x satisfying $x^5 + x + 1 = 0$ (even though there does happen to be one), since we do not know how to write the solution of the equation in terms of nth roots (in fact, it is known that the solution cannot be written in this form). And, of course, we certainly do not wish to assume that all equations have solutions, since this is false (no real number x satisfies $x^2 + 1 = 0$, for example). In fact, this direction of investigation is not a fruitful one. The most useful hints about the property distinguishing **R** from **Q**, the most compelling evidence for the necessity of elucidating this property, do not come from the study of numbers alone. In order to study the properties of the real numbers in a more profound way, we must study more than the real numbers. At this point we must begin with the foundations of calculus, in particular the fundamental concept on which calculus is based—functions.

PROBLEMS

1. Prove the following formulas by induction.

(i) $1^2 + \cdots + n^2 = \dfrac{n(n+1)(2n+1)}{6}$.

(ii) $1^3 + \cdots + n^3 = (1 + \cdots + n)^2$.

2. Find a formula for

(i) $\sum_{i=1}^{n} (2i - 1) = 1 + 3 + 5 + \cdots + (2n - 1)$.

(ii) $\sum_{i=1}^{n} (2i - 1)^2 = 1^2 + 3^2 + 5^2 + \cdots + (2n - 1)^2$.

Hint: What do these expressions have to do with $1 + 2 + 3 + \cdots + 2n$ and $1^2 + 2^2 + 3^2 + \cdots + (2n)^2$?

3. If $0 \le k \le n$, the "binomial coefficient" $\binom{n}{k}$ is defined by

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k!}, \text{ if } k \neq 0, n$$

$\binom{n}{0} = \binom{n}{n} = 1$. (This becomes a special case of the first formula if we define $0! = 1$.)

(a) Prove that

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

(The proof does not require an induction argument.)

This relation gives rise to the following configuration, known as "Pascal's triangle"—a number not on one of the sides is the sum of the two numbers above it; the binomial coefficient $\binom{n}{k}$ is the kth number in the nth row.

```
          1
        1   1
      1   2   1
    1   3   3   1
  1   4   6   4   1
1   5  10  10   5   1
         . . .
```

(b) Notice that all the numbers in Pascal's triangle are natural numbers. Use part (a) to prove by induction that $\binom{n}{k}$ is always a natural number. (Your entire proof by induction will, in a sense, be summed up in a glance by Pascal's triangle.)

(c) Give another proof that $\binom{n}{k}$ is a natural number by showing that $\binom{n}{k}$ is the number of sets of exactly k integers each chosen from $1, \ldots, n$.

(d) Prove the "binomial theorem": If a and b are any numbers and n is a natural number, then

$$\begin{aligned}(a+b)^n &= a^n + \binom{n}{1}a^{n-1}b + \binom{n}{2}a^{n-2}b^2 + \cdots + \binom{n}{n-1}ab^{n-1} + b^n \\ &= \sum_{j=0}^{n}\binom{n}{j}a^{n-j}b^j.\end{aligned}$$

(e) Prove that

(i) $$\sum_{j=0}^{n}\binom{n}{j} = \binom{n}{0} + \cdots + \binom{n}{n} = 2^n.$$

(ii) $$\sum_{j=0}^{n}(-1)^j\binom{n}{j} = \binom{n}{0} - \binom{n}{1} + \cdots \pm \binom{n}{n} = 0.$$

4. (a) Prove by induction on n that

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

if $r \neq 1$ (if $r = 1$, evaluating the sum certainly presents no problem).

(b) Derive this result by setting $S = 1 + r + \cdots + r^n$, multiplying this equation by r, and solving the two equations for S.

5. The formula for $1^2 + \cdots + n^2$ may be derived as follows. We begin with the formula

$$(k+1)^3 - k^3 = 3k^2 + 3k + 1.$$

Writing this formula for $k = 1, \ldots, n$ and adding, we obtain

$$\begin{aligned}2^3 - 1^3 &= 3 \cdot 1^2 + 3 \cdot 1 + 1 \\ 3^3 - 2^3 &= 3 \cdot 2^2 + 3 \cdot 2 + 1 \\ &\vdots \\ (n+1)^3 - n^3 &= 3 \cdot n^2 + 3 \cdot n + 1 \\ \hline (n+1)^3 - 1 &= 3[1^2 + \cdots + n^2] + 3[1 + \cdots + n] + n.\end{aligned}$$

Thus we can find $\sum_{k=1}^{n} k^2$ if we already know $\sum_{k=1}^{n} k$ (which could have been found in a similar way). Use this method to find

(i) $1^3 + \cdots + n^3$.
(ii) $1^4 + \cdots + n^4$.
(iii) $\dfrac{1}{1 \cdot 2} + \dfrac{1}{2 \cdot 3} + \cdots + \dfrac{1}{n(n+1)}$.
(iv) $\dfrac{3}{1^2 \cdot 2^2} + \dfrac{5}{2^2 \cdot 3^2} + \cdots + \dfrac{2n+1}{n^2(n+1)^2}$.

***6.** Use the method of Problem 5 to show that $\sum_{k=1}^{n} k^p$ can always be written in the form

$$\frac{n^{p+1}}{p+1} + An^p + Bn^{p-1} + Cn^{p-2} + \cdots .$$

(The first 10 such expressions are

$$\sum_{k=1}^{n} k = \tfrac{1}{2}n^2 + \tfrac{1}{2}n$$

$$\sum_{k=1}^{n} k^2 = \tfrac{1}{3}n^3 + \tfrac{1}{2}n^2 + \tfrac{1}{6}n$$

$$\sum_{k=1}^{n} k^3 = \tfrac{1}{4}n^4 + \tfrac{1}{2}n^3 + \tfrac{1}{4}n^2$$

$$\sum_{k=1}^{n} k^4 = \tfrac{1}{5}n^5 + \tfrac{1}{2}n^4 + \tfrac{1}{3}n^3 - \tfrac{1}{30}n$$

$$\sum_{k=1}^{n} k^5 = \tfrac{1}{6}n^6 + \tfrac{1}{2}n^5 + \tfrac{5}{12}n^4 - \tfrac{1}{12}n^2$$

$$\sum_{k=1}^{n} k^6 = \tfrac{1}{7}n^7 + \tfrac{1}{2}n^6 + \tfrac{1}{2}n^5 - \tfrac{1}{6}n^3 + \tfrac{1}{42}n$$

$$\sum_{k=1}^{n} k^7 = \tfrac{1}{8}n^8 + \tfrac{1}{2}n^7 + \tfrac{7}{12}n^6 - \tfrac{7}{24}n^4 + \tfrac{1}{12}n^2$$

$$\sum_{k=1}^{n} k^8 = \tfrac{1}{9}n^9 + \tfrac{1}{2}n^8 + \tfrac{2}{3}n^7 - \tfrac{7}{15}n^5 + \tfrac{2}{9}n^3 - \tfrac{1}{30}n$$

$$\sum_{k=1}^{n} k^9 = \tfrac{1}{10}n^{10} + \tfrac{1}{2}n^9 + \tfrac{3}{4}n^8 - \tfrac{7}{10}n^6 + \tfrac{1}{2}n^4 - \tfrac{1}{12}n^2$$

$$\sum_{k=1}^{n} k^{10} = \tfrac{1}{11}n^{11} + \tfrac{1}{2}n^{10} + \tfrac{5}{6}n^9 - 1n^7 + 1n^5 - \tfrac{1}{2}n^3 + \tfrac{5}{66}n.$$

Notice that the coefficients in the second column are always $\frac{1}{2}$, and that after the third column the powers of n with nonzero coefficients decrease by 2 until n^2 or n is reached. The coefficients in all but the first two columns seem to be rather haphazard, but there actually is some sort of pattern; finding it may be regarded as a super-perspicacity test. See Problem 26-16 for the complete story.)

7. Prove that every natural number is either even or odd.

8. Prove that if a set A of natural numbers contains n_0 and contains $k + 1$ whenever it contains k, then A contains all natural numbers $\geq n_0$.

9. Prove the principle of mathematical induction from the well-ordering principle.

10. Prove the principle of complete induction from the ordinary principle of induction. Hint: If A contains 1 and A contains $n + 1$ whenever it contains $1, \ldots, n$, consider the set B of all k such that $1, \ldots, k$ are all in A.

11. (a) If a is rational and b is irrational, is $a + b$ necessarily irrational? What if a and b are both irrational?
(b) If a is rational and b is irrational, is ab necessarily irrational? (Careful!)
(c) Is there a number a such that a^2 is irrational, but a^4 is rational?
(d) Are there two irrational numbers whose sum and product are both rational?

12. (a) Prove that $\sqrt{3}$, $\sqrt{5}$, and $\sqrt{6}$ are irrational. Hint: To treat $\sqrt{3}$, for example, use the fact that every integer is of the form $3n$ or $3n + 1$ or $3n + 2$. Why doesn't this proof work for $\sqrt{4}$?
(b) Prove that $\sqrt[3]{2}$ and $\sqrt[3]{3}$ are irrational.

***13.** Prove that

(a) $\sqrt{2} + \sqrt{3}$ is irrational.
(b) $\sqrt{6} - \sqrt{2} - \sqrt{3}$ is irrational.

14. (a) Prove that if $x = p + \sqrt{q}$ where p and q are rational, and m is a natural number, then $x^m = a + b\sqrt{q}$ for some rational a and b.
(b) Prove also that $(p - \sqrt{q})^m = a - b\sqrt{q}$.

15. (a) Prove that if m and n are natural numbers and $m^2/n^2 < 2$, then $(m + 2n)^2/(m + n)^2 > 2$; show, moreover, that

$$\frac{(m + 2n)^2}{(m + n)^2} - 2 < 2 - \frac{m^2}{n^2}.$$

(b) Prove the same results with all inequality signs reversed.
(c) Prove that if $m/n < \sqrt{2}$, then there is another rational number m'/n' with $m/n < m'/n' < \sqrt{2}$.

***16.** It seems likely that $\sqrt{n}$ is irrational whenever the natural number n is not the square of another natural number. Although the method of Problem 12 may actually be used to treat any particular case, it is not clear in advance that it will always work, and a proof for the general case requires some extra information. A natural number p is called a **prime number** if it is impossible to write $p = ab$ for natural numbers a and b unless one of these is p, and the other 1; for convenience we also agree that 1 is *not* a prime number. The first few prime numbers are 2, 3, 5, 7, 11, 13, 17, 19. If $n > 1$ is not a prime, then $n = ab$, with a and b both $< n$; if either a or b is not a prime it can be factored similarly; continuing in this way proves that we can write n as a product of primes. For example, $28 = 4 \cdot 7 = 2 \cdot 2 \cdot 7$.

(a) Turn this argument into a rigorous proof by complete induction. (To be sure, any reasonable mathematician would accept the informal argument, but this is partly because it would be obvious to him how to state it rigorously.)

A fundamental theorem about integers, which we will not prove here, states that this factorization is unique, except for the order of the factors. Thus, for example, 28 can never be written as a product of primes one of which is 3, nor can it be written in a way that involves 2 only once (now you should appreciate why 1 is not allowed as a prime).

(b) Using this fact, prove that $\sqrt{n}$ is irrational unless $n = m^2$ for some natural number m.
(c) Prove more generally that $\sqrt[k]{n}$ is irrational unless $n = m^k$.
(d) No discussion of prime numbers should fail to allude to Euclid's beautiful proof that there are infinitely many of them. Prove that there cannot be only finitely many prime numbers $p_1, p_2, p_3, \ldots, p_n$ by considering $p_1 \cdot p_2 \cdot \ldots \cdot p_n + 1$.

***17.** (a) Prove that if x satisfies

$$x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0,$$

for some integers $a_{n-1}, \ldots, a_0$, then x is irrational unless x is an integer. (Why is this a generalization of Problem 16?)
(b) Prove that $\sqrt{2} + \sqrt[3]{2}$ is irrational. Hint: Start by working out the first 6 powers of this number.

18. Prove Bernoulli's inequality: If $h > -1$, then

$$(1 + h)^n \geq 1 + nh.$$

Why is this trivial if $h > 0$?

19. The Fibonacci sequence $a_1, a_2, a_3, \ldots$ is defined as follows:

$$\begin{aligned} a_1 &= 1, \\ a_2 &= 1, \\ a_n &= a_{n-1} + a_{n-2} \qquad \text{for } n \geq 3. \end{aligned}$$

This sequence, which begins 1, 1, 2, 3, 5, 8, . . . , was discovered by Fibonacci (circa 1175–1250), in connection with a problem about rabbits. Fibonacci assumed that an initial pair of rabbits gave birth to one new pair of rabbits per month, and that after two months each new pair behaved similarly. The number a_n of pairs born in the nth month is $a_{n-1} + a_{n-2}$, because a pair of rabbits is born for each pair born the previous month, and moreover each pair born two months ago now gives birth to another pair. The number of interesting results about this sequence is truly amazing—there is even a Fibonacci Association which publishes a journal, *The Fibonacci Quarterly*. Prove that

$$a_n = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n}{\sqrt{5}}.$$

One way of deriving this astonishing formula is presented in Problem 23-8.

20. The result in Problem 1-6 has an important generalization: If $a_1, \ldots, a_n \geq 0$, then

$$\sqrt[n]{a_1 \cdot \ldots \cdot a_n} \leq \frac{a_1 + \cdots + a_n}{n}.$$

(a) Why is this true if $a_1 = \cdots = a_n$? Suppose not all a_i are equal, say $a_i \neq a_j$. If a_i and a_j are both replaced by $(a_i + a_j)/2$ what happens to the "arithmetic mean" $A_n = (a_1 + \cdots + a_n)/n$? What happens to the "geometric mean" $G_n = \sqrt[n]{a_1 \cdot \ldots \cdot a_n}$? Why does repeating this process enough times eventually prove that $G_n \leq A_n$? (This is another place where it is a good exercise to provide a formal proof by induction, as well as informal reason.)

The reasoning in the previous proof is closely related to another interesting proof.

(b) Using the fact that $G_n \leq A_n$ when $n = 2$, prove, by induction on k, that $G_n \leq A_n$ for $n = 2^k$.

(c) For a general n, let $2^m > n$. Apply part (b) to the 2^m numbers

$$a_1, \ldots, a_n, \underbrace{A_n, \ldots, A_n}_{2^m - n \text{ times}}$$

to prove that $G_n \leq A_n$.

21. The following is a recursive definition of a^n:

$$a^1 = a,$$
$$a^{n+1} = a^n \cdot a.$$

Prove, by induction, that

$$a^{n+m} = a^n \cdot a^m,$$
$$(a^n)^m = a^{nm}.$$

(Don't try to be fancy: use either induction on n or induction on m, not both at once.)

22. Suppose we know properties P1 and P4 for the natural numbers, but that multiplication has never been mentioned. Then the following can be used as a recursive definition of multiplication:

$$1 \cdot b = b,$$
$$(a + 1) \cdot b = a \cdot b + b.$$

Prove the following (in the order suggested!):

$a \cdot (b + c) = a \cdot b + a \cdot c$ (use induction on a),
$a \cdot 1 = a$,
$a \cdot b = b \cdot a$ (you just finished proving the case $b = 1$).

23. In this chapter we began with the natural numbers and gradually built up to the real numbers. A completely rigorous discussion of this process requires a little book in itself (see Part V). No one has ever figured out how to get to the real numbers without going through this process, but if we do accept the real numbers as given, then the natural numbers can be *defined* as the real numbers of the form 1, $1 + 1$, $1 + 1 + 1$, etc. The whole point of this problem is to show that there is a rigorous mathematical way of saying "etc."

(a) A set A of real numbers is called **inductive** if

(1) 1 is in A,
(2) $k + 1$ is in A whenever k is in A.

Prove that

(i) **R** is inductive.
(ii) The set of positive real numbers is inductive.
(iii) The set of positive real numbers unequal to $\frac{1}{2}$ is inductive.
(iv) The set of positive real numbers unequal to 5 is not inductive.
(v) If A and B are inductive, then the set C of real numbers which are in both A and B is also inductive.

(b) A real number n will be called a **natural number** if n is in *every* inductive set.

(i) Prove that 1 is a natural number.
(ii) Prove that $k + 1$ is a natural number if k is a natural number.