# Appendix A*

## Mathematical Notation and Methods of Proof

*Please, sir, I want some more.*
—Oliver
Charles Dickens, *Oliver Twist*

*Anyone who understands algebraic notation reads at a glance in an equation results reached arithmetically only with great labour and pains.*
—Augustin Cournot
*Researches into the Mathematical Principles of the Theory of Wealth*
Translated by Nathaniel T. Bacon
Macmillan, 1897, p. 4

In this book, an effort has been made to use "mathematical English" as much as possible, keeping mathematical notation to a minimum. However, mathematical notation is a convenient shorthand that can greatly simplify the amount of writing we have to do. Moreover, it is commonly used in every branch of mathematics, so the ability to read and write mathematical notation is an essential ingredient of mathematical understanding. Finally, there are some theorems whose proofs become "obvious" if the right notation is used.

Proving theorems in mathematics is as much an art as a science. For the beginner, it is often hard to know what approach to use in proving a theorem; there are many approaches, any one of which might turn out to be the best. To become proficient at proofs, it is important to study as many examples as possible and to get plenty of practice.

This appendix summarizes basic mathematical notation applied to sets. Summation notation, a useful shorthand for dealing with sums, is also discussed. Finally, some approaches to proofs are illustrated with generic examples.

### Set Notation

A ***set*** is a collection of objects, called the ***elements*** (or ***members***) of the set. Examples of sets include the set of all words in this text, the set of all books in your college library, the set of positive integers, and the set of all vectors in the plane whose equation is $2x + 3y - z = 0$.

It is often possible to list the elements of a set, in which case it is conventional to enclose the list within braces. For example, we have

$$\{1, 2, 3\}, \quad \{a, t, x, z\}, \quad \{2, 4, 6, \ldots, 100\}, \quad \left\{\frac{\pi}{4}, \frac{2\pi}{5}, \frac{\pi}{2}, \frac{4\pi}{7}, \ldots, \frac{5\pi}{6}\right\}$$

Note that ellipses (. . .) denote elements omitted when a pattern is present. (What is the pattern in the last two examples?) Infinite sets are often expressed using ellipses. For example, the set of positive integers is usually denoted by $\mathbb{N}$ or $\mathbb{Z}^+$, so

$$\mathbb{N} = \mathbb{Z}^+ = \{1, 2, 3, \ldots\}$$

The set of all integers is denoted by $\mathbb{Z}$, so

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$$

Two sets are considered to be ***equal*** if they contain exactly the same elements. The *order* in which elements are listed does not matter, and repetitions are not counted. Thus,

*Exercises and selected odd-numbered answers for this appendix can be found on the student companion website.

$$\{1, 2, 3\} = \{2, 1, 3\} = \{1, 3, 2, 1\}$$

The symbol $\in$ means "is an element of" or "is in," and the symbol $\notin$ denotes the negation—that is, "is not an element of" or "is not in." For example,

$$5 \in \mathbb{Z}^+ \quad \text{but} \quad 0 \notin \mathbb{Z}^+$$

It is often more convenient to describe a set in terms of a rule satisfied by all of its elements. In such cases, ***set builder notation*** is appropriate. The format is

$$\{x : x \text{ satisfies } P\}$$

where $P$ represents a property or a collection of properties that the element $x$ must satisfy. The colon is pronounced "such that." For example,

$$\{n : n \in \mathbb{Z}, n > 0\}$$

is read as "the set of all $n$ such that $n$ is an integer and $n$ is greater than zero." This is just another way of describing the positive integers $\mathbb{Z}^+$. (We could also write $\mathbb{Z}^+ = \{n \in \mathbb{Z} : n > 0\}$.)

The ***empty set*** is the set with no elements. It is denoted by either $\varnothing$ or $\{\ \}$.

## Example A.1

Describe in words the following sets:

(a) $A = \{n : n = 2k, k \in \mathbb{Z}\}$
(b) $B = \{m/n : m, n \in \mathbb{Z}, n \neq 0\}$
(c) $C = \{x \in \mathbb{R} : 4x^2 - 4x - 3 = 0\}$
(d) $D = \{x \in \mathbb{Z} : 4x^2 - 4x - 3 = 0\}$

**Solution** (a) $A$ is the set of numbers $n$ that are integer multiples of 2. Therefore, $A$ is the set of all even integers.

(b) $B$ is the set of all expressions of the form $m/n$, where $m$ and $n$ are integers and $n$ is nonzero. This is the set of *rational numbers*, usually denoted by $\mathbb{Q}$. (Note that this way of describing $\mathbb{Q}$ produces many repetitions; however, our convention, as noted above, is that we include only one occurrence of each element. Thus, this expression precisely describes the set of all rational numbers.)

(c) $C$ is the set of all real solutions of the equation $4x^2 - 4x - 3 = 0$. By factoring or using the quadratic formula, we find that the roots of this equation are $-\frac{1}{2}$ and $\frac{3}{2}$. (Verify this.) Therefore,

$$C = \{-\tfrac{1}{2}, \tfrac{3}{2}\}$$

(d) From the solution to (c) we see that there are *no* solutions to $4x^2 - 4x - 3 = 0$ in $\mathbb{R}$ that are integers. Therefore, $D$ is the empty set, which we can express by writing $D = \varnothing$.

John Venn (1834–1923) was an English mathematician who studied at Cambridge University and later lectured there. He worked primarily in mathematical logic and is best known for inventing Venn diagrams.

If every element of a set $A$ is also an element of a set $B$, then $A$ is called a ***subset*** of $B$, denoted $A \subseteq B$. We can represent this situation schematically using a ***Venn diagram***, as shown in Figure A.1. (The rectangle represents the *universal set*, a set large enough to contain all of the other sets in question—in this case, $A$ and $B$.)

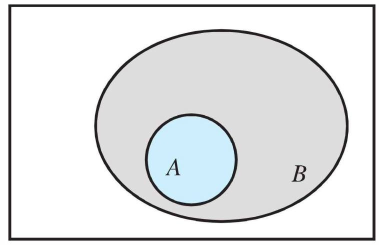



**Figure A.1**
$A \subseteq B$

**Example A.2**

(a) $\{1, 2, 3\} \subseteq \{1, 2, 3, 4, 5\}$

(b) $\mathbb{Z}^+ \subseteq \mathbb{Z} \subseteq \mathbb{R}$

(c) Let $A$ be the set of all positive integers whose last two digits are 24 and let $B$ be the set of all positive integers that are evenly divisible by 4. Then if $n$ is in $A$, it is of the form

$$n = 100k + 24$$

for some integer $k$. (For example, $36{,}524 = 100 \cdot 365 + 24$.) But then

$$n = 100k + 24 = 4(25k + 6)$$

so $n/4 = 25k + 6$, which is an integer. Hence, $n$ is evenly divisible by 4, so it is in $B$. Therefore, $A \subseteq B$.

We can show that two sets $A$ and $B$ are equal by showing that each is a subset of the other. This strategy is particularly useful if the sets are defined abstractly or if it is not easy to list and compare their elements.

**Example A.3**

Let $A$ be the set of all positive integers whose last two digits form a number that is evenly divisible by 4. In the case of a one-digit number, we take its tens digit to be 0. Let $B$ be the set of all positive integers that are evenly divisible by 4. Show that $A = B$.

**Solution** As in Example A.2(c), it is easy to see that $A \subseteq B$. If $n$ is in $A$, then we can split off the number $m$ formed by its last two digits by writing

$$n = 100k + m$$

for some integer $k$. But, since $m$ is divisible by 4, we have $m = 4r$ for some integer $r$. Therefore,

$$n = 100k + m = 100k + 4r = 4(25k + r)$$

so $n$ is also evenly divisible by 4. Hence, $A \subseteq B$.

To show that $B \subseteq A$, let $n$ be in $B$. That is, $n$ is evenly divisible by 4. Let's say that $n = 4s$, where $s$ is an integer. If $m$ is the number formed by the last two digits of $n$, then, as above, $n = 100k + m$ for some integer $k$. But now

$$m = n - 100k = 4s - 100k = 4(s - 25k)$$

which implies that $m$ is evenly divisible by 4, since $s - 25k$ is an integer. Therefore, $n$ is in $A$, and we have shown that $B \subseteq A$.

Since $A \subseteq B$ and $B \subseteq A$, we must have $A = B$.

The ***intersection*** of sets $A$ and $B$ is denoted by $A \cap B$ and consists of the elements that $A$ and $B$ have in common. That is,

$$A \cap B = \{x : x \in A \quad \text{and} \quad x \in B\}$$

Figure A.2 shows a Venn diagram of this case. The ***union*** of $A$ and $B$ is denoted by $A \cup B$ and consists of the elements that are in either $A$ or $B$ (or both). That is,

$$A \cup B = \{x : x \in A \quad \text{or} \quad x \in B\}$$

See Figure A.3.

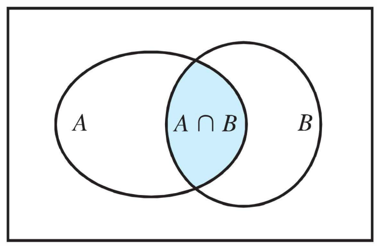



**Figure A.2**
$A \cap B$

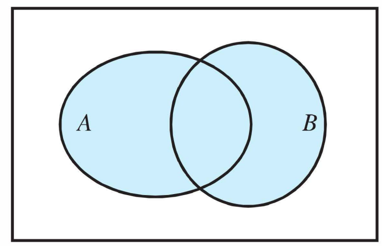



**Figure A.3**
$A \cup B$

### Example A.4

Let $A = \{n^2 : n \in \mathbb{Z}^+, 1 \leq n \leq 4\}$ and let $B = \{n \in \mathbb{Z}^+ : n \leq 10 \text{ and } n \text{ is odd}\}$. Find $A \cap B$ and $A \cup B$.

**Solution** We see that

$$A = \{1^2, 2^2, 3^2, 4^2\} = \{1, 4, 9, 16\} \quad \text{and} \quad B = \{1, 3, 5, 7, 9\}$$

Therefore, $A \cap B = \{1, 9\}$ and $A \cup B = \{1, 3, 4, 5, 7, 9, 16\}$.

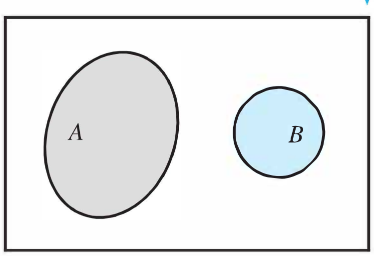



**Figure A.4**
Disjoint sets

If $A \cap B = \varnothing$, then $A$ and $B$ are called ***disjoint sets.*** (See Figure A.4.) For example, the set of even integers and the set of odd integers are disjoint.

## Summation Notation

Summation notation is a convenient shorthand to use to write out a sum such as

$$1 + 2 + 3 + \cdots + 100$$

where we want to leave out all but a few terms. As in set notation, ellipses (. . .) convey that we have established a pattern and have simply left out some intermediate terms. In the above example, readers are expected to recognize that we are summing all of the positive integers from 1 to 100. However, ellipses can be ambiguous. For example, what would one make of the following sum?

$$1 + 2 + \cdots + 64$$

Is this the sum of all positive integers from 1 to 64 or just the powers of two, $1 + 2 + 4 + 8 + 16 + 32 + 64$? It is often clearer (and shorter) to use ***summation notation*** (or ***sigma notation***).

$\Sigma$ is the capital Greek letter *sigma,* corresponding to S (for "sum"). Summation notation was introduced by Fourier in 1820 and was quickly adopted by the mathematical community.

We can abbreviate a sum of the form

$$a_1 + a_2 + \cdots + a_n \tag{1}$$

as

$$\sum_{k=1}^{n} a_k \tag{2}$$

which tells us to sum the terms $a_k$ over all integers $k$ ranging from 1 to $n$. An alternative version of this expression is

$$\sum_{1 \le k \le n} a_k$$

The subscript $k$ is called the ***index of summation.*** It is a "dummy variable" in the sense that it does not appear in the actual sum in expression (1). Therefore, we can use any letter we like as the index of summation (as long as it doesn't already appear somewhere else in the expressions we are summing). Thus, expression (2) can also be written as

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

The index of summation need not start at 1. The sum $a_3 + a_4 + \cdots + a_{99}$ becomes

$$\sum_{k=3}^{99} a_k$$

although we can arrange for the index to begin at 1 by rewriting the expression as

$$\sum_{k=1}^{97} a_{k+2}.$$

The key to using summation notation effectively is being able to recognize patterns.

**Example A.5**

Write the following sums using summation notation.

(a) $1 + 2 + 4 + \cdots + 64$ (b) $1 + 3 + 5 + \cdots + 99$ (c) $3 + 8 + 15 + \cdots + 99$

**Solution** (a) We recognize this expression as a sum of powers of 2:

$$1 + 2 + 4 + \cdots + 64 = 2^0 + 2^1 + 2^2 + \cdots + 2^6$$

Therefore, the index of summation appears as the exponent, and we have $\sum_{k=0}^{6} 2^k$.

(b) This expression is the sum of all the odd integers from 1 to 99. Every odd integer is of the form $2k + 1$, so the sum is

$$\begin{aligned} &1 + 3 + 5 + \cdots + 99 \\ &= (2 \cdot 0 + 1) + (2 \cdot 1 + 1) + (2 \cdot 2 + 1) + \cdots + (2 \cdot 49 + 1) \\ &= \sum_{k=0}^{49} (2k + 1) \end{aligned}$$

(c) The pattern here is less clear, but a little reflection reveals that each term is 1 less than a perfect square:

$$\begin{aligned} &3 + 8 + 15 + \cdots + 99 \\ &= (2^2 - 1) + (3^2 - 1) + (4^2 - 1) + \cdots + (10^2 - 1) \\ &= \sum_{k=2}^{10} (k^2 - 1) \end{aligned}$$

**Example A.6**

Rewrite each of the sums in Example A.5 so that the index of summation starts at 1.

**Solution** (a) If we use the change of variable $i = k + 1$, then, as $k$ goes from 0 to 6, $i$ goes from 1 to 7. Since $k = i - 1$, we obtain

$$\sum_{k=0}^{6} 2^k = \sum_{i=1}^{7} 2^{i-1}$$

(b) Using the same substitution as in part (a), we get

$$\sum_{k=0}^{49} (2k + 1) = \sum_{i=1}^{50} (2(i - 1) + 1) = \sum_{i=1}^{50} (2i - 1)$$

(c) The substitution $i = k - 2$ will work (try it), but it is easier just to add a term corresponding to $k = 1$, since $1^2 - 1 = 0$. Therefore,

$$\sum_{k=2}^{10} (k^2 - 1) = \sum_{k=1}^{10} (k^2 - 1)$$

Multiple summations arise when there is more than one index of summation, as there is with a matrix. The notation

$$\sum_{i,j=1}^{n} a_{ij} \tag{3}$$

means to sum the terms $a_{ij}$ as $i$ and $j$ each range independently from 1 to $n$. The sum in expression (3) is equivalent to either

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}$$

where we sum first over $j$ and then over $i$ (we always work from the inside out), or

$$\sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij}$$

where the order of summation is reversed.

**Example A.7**

Write out $\sum_{i,j=1}^{3} i^j$ using both possible orders of summation.

**Solution**

$$\begin{aligned} \sum_{i=1}^{3} \sum_{j=1}^{3} i^j &= \sum_{i=1}^{n} (i^1 + i^2 + i^3) \\ &= (1^1 + 1^2 + 1^3) + (2^1 + 2^2 + 2^3) + (3^1 + 3^2 + 3^3) \\ &= (1 + 1 + 1) + (2 + 4 + 8) + (3 + 9 + 27) = 56 \end{aligned}$$

and

$$\sum_{j=1}^{3}\sum_{i=1}^{3} i^j = \sum_{j=1}^{n}(1^j + 2^j + 3^j)$$
$$= (1^1 + 2^1 + 3^1) + (1^2 + 2^2 + 3^2) + (1^3 + 2^3 + 3^3)$$
$$= (1 + 2 + 3) + (1 + 4 + 9) + (1 + 8 + 27) = 56$$

**Remark** Of course, the value of the sum in Example A.7 is the same no matter which order of summation we choose, because the sum is *finite*. It is also possible to consider *infinite sums* (known as *infinite series* in calculus), but such sums do not always have a value and great care must be taken when rearranging or manipulating their terms. For example, suppose we let

$$S = \sum_{k=0}^{\infty} 2^k$$

Then
$$S = 1 + 2 + 4 + 8 + \cdots$$
$$= 1 + 2(1 + 2 + 4 + \cdots)$$
$$= 1 + 2S$$

from which it follows that $S = -1$. This is clearly nonsense, since $S$ is a sum of *non-negative* terms! (Where is the error?)

*How to Solve It* is the title of a book by the mathematician George Pólya (1887–1985). Since its publication in 1945, *How to Solve It* has sold over a million copies and has been translated into 17 languages. Pólya was born in Hungary, but because of the political situation in Europe, he moved to the United States in 1940. He subsequently taught at Brown and Stanford Universities, where he did mathematical research and developed a well-deserved reputation as an outstanding teacher. The Pólya Prize is awarded annually by the Society for Industrial and Applied Mathematics for major contributions to areas of mathematics close to those on which Pólya worked. The Mathematical Association of America annually awards Pólya Lectureships to mathematicians demonstrating the high-quality exposition for which Pólya was known.

## Methods of Proof

The notion of proof is at the very heart of mathematics. It is one thing to know *what* is true; it is quite another to know *why* it is true and to be able to demonstrate its truth by means of a logically connected sequence of statements. The intention here is not to try to teach you how to do proofs; you will become better at doing proofs by studying examples and by practicing—something you should do often as you work through this text. The intention of this brief section is simply to provide a few elementary examples of some types of proofs. The proofs of theorems in the text will provide further illustrations of "how to solve it."

Roughly speaking, mathematical proofs fall into two categories: ***direct proofs*** and ***indirect proofs***. Many theorems have the structure "if $P$, then $Q$," where $P$ (the *hypothesis*, or *premise*) and $Q$ (the *conclusion*) are statements that are either true or false. We denote such an implication by $P \Rightarrow Q$. A direct proof proceeds by establishing a chain of implications

$$P \Rightarrow P_1 \Rightarrow P_2 \Rightarrow \cdots \Rightarrow P_n \Rightarrow Q$$

leading directly from $P$ to $Q$.

**Example A.8**

Prove that any two consecutive perfect squares differ by an odd number. This instruction can be rephrased as "Prove that if $a$ and $b$ are consecutive perfect squares, then $a - b$ is odd." Hence, it has the form $P \Rightarrow Q$, with $P$ being "$a$ and $b$ are consecutive perfect squares" and $Q$ being "$a - b$ is odd."

**Solution** Assume that $a$ and $b$ are consecutive perfect squares, with $a > b$. Then

$$a = (n + 1)^2 \quad \text{and} \quad b = n^2$$

for some integer $n$. But now

$$\begin{aligned} a - b &= (n + 1)^2 - n^2 \\ &= n^2 + 2n + 1 - n^2 \\ &= 2n + 1 \end{aligned}$$

so $a - b$ is odd.

There are two types of indirect proofs that can be used to establish a conditional statement of the form $P \Rightarrow Q$. A ***proof by contradiction*** assumes that the hypothesis $P$ is true, just as in a direct proof, but then supposes that the conclusion $Q$ is *false*. The strategy then is to show that this is not possible (i.e., to rule out the possibility that the conclusion is false) by finding a contradiction to the truth of $P$. It then follows that $Q$ must be true.

## Example A.9

Let $n$ be a positive integer. Prove that if $n^2$ is even, so is $n$. (Take a few minutes to try to find a direct proof of this assertion; it will help you to appreciate the indirect proof that follows.)

**Solution** Assume that $n$ is a positive integer such that $n^2$ is even. Now suppose that $n$ is not even. Then $n$ is odd, so

$$n = 2k + 1$$

for some integer $k$. But if so, we have

$$n^2 = (2k + 1)^2 = 4k^2 + 4k + 1$$

so $n^2$ is odd, since it is 1 more than the even number $4k^2 + 4k$. This contradicts our hypothesis that $n^2$ is even. We conclude that our supposition that $n$ was *not* even must have been false; in other words, $n$ must be even.

Closely related to the method of proof by contradiction is ***proof by contrapositive***. The *negative* of a statement $P$ is the statement "it is not the case that $P$," abbreviated symbolically as $\neg P$ and pronounced "not $P$." For example, if $P$ is "$n$ is even," then $\neg P$ is "it is not the case that $n$ is even"—in other words, "$n$ is odd."

The *contrapositive* of the statement $P \Rightarrow Q$ is the statement $\neg Q \Rightarrow \neg P$. A conditional statement $P \Rightarrow Q$ and its contrapositive $\neg Q \Rightarrow \neg P$ are logically equivalent in the sense that they are either both true or both false. (For example, if $P \Rightarrow Q$ is a theorem, then so is $\neg Q \Rightarrow \neg P$. To see this, note that if the hypothesis $\neg Q$ is true, then $Q$ is false. The conclusion $\neg P$ cannot be false, for if it were, then $P$ would be true and our known theorem $P \Rightarrow Q$ would imply the truth of $Q$, giving us a contradiction. It follows that $\neg P$ is true and we have proved $\neg Q \Rightarrow \neg P$.) Here is a contrapositive proof of the assertion in Example A.9.

**Example A.10**

Let $n$ be a positive integer. Prove that if $n^2$ is even, so is $n$.

**Solution** The contrapositive of the given statement is

"If $n$ is not even, then $n^2$ is not even" or "If $n$ is odd, so is $n^2$"

To prove this contrapositive, assume that $n$ is odd. Then $n = 2k + 1$ for some integer $k$. As before, this means that $n^2 = (2k + 1)^2 = 4k^2 + 4k + 1$ is odd, which completes the proof of the contrapositive. Since the contrapositive is true, so is the original statement.

Although we do not require a new method of proof to handle it, we will briefly consider how to prove an "if and only if" theorem. A statement of the form "$P$ if and only if $Q$" signals a *double implication,* which we denote by $P \Leftrightarrow Q$. To prove such a statement, we must prove $P \Rightarrow Q$ and $Q \Rightarrow P$. To do so, we can use the techniques described earlier, where appropriate. It is important to notice that the "*if* " part of $P \Leftrightarrow Q$ is "$P$ if $Q$," which is $Q \Rightarrow P$; the "only if" part of $P \Leftrightarrow Q$ is "$P$ only if $Q$," meaning $P \Rightarrow Q$. The implication $P \Rightarrow Q$ is sometimes read as "$P$ is sufficient for $Q$" or "$Q$ is necessary for $P$"; $Q \Rightarrow P$ is read "$Q$ is sufficient for $P$" or "$P$ is necessary for $Q$." Taken together, they are $P \Leftrightarrow Q$, or "$P$ is necessary and sufficient for $Q$" and vice versa.

**Example A.11**

A pawn is placed on a chessboard and is allowed to move one square at a time, either horizontally or vertically. A *pawn's tour* of a chessboard is a path taken by a pawn, moving as described, that visits each square exactly once, starting and ending on the same square. Prove that there is a pawn's tour of an $n \times n$ chessboard if and only if $n$ is even.

**Solution** [ $\Leftarrow$ ] ("if") Assume that $n$ is even. It is easy to see that the strategy illustrated in Figure A.5 for a $6 \times 6$ chessboard will always give a pawn's tour.

[ $\Rightarrow$ ] ("only if") Suppose that there is a pawn's tour of an $n \times n$ chessboard. We will give a proof by contradiction that $n$ must be even. To this end, let's assume that $n$ is odd. At each move, the pawn moves to a square of a different color. The total number of moves in its tour is $n^2$, which is also an odd number, according to the proof in Example A.10. Therefore, the pawn must end up on a square of the opposite color from that of the square on which it started. (Why?) This is impossible, since the pawn ends where it started, so we have a contradiction. It follows that $n$ cannot be odd; hence, $n$ is even and the proof is complete.

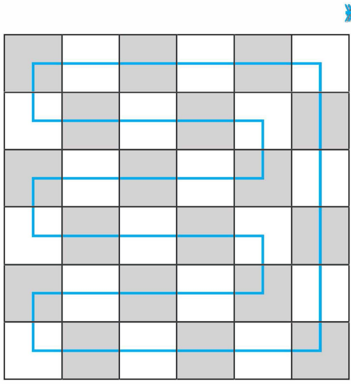


Figure A.5

Some theorems assert that several statements are *equivalent*. This means that each is true if and only if all of the others are true. Showing that $n$ statements are equivalent requires $\binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n^2 - n}{2}$ "if and only if" proofs. In practice, however, it is often easier to establish a "ring" of $n$ implications that links all of the statements. The proof of the Fundamental Theorem of Invertible Matrices provides an excellent example of this approach.

# Appendix B*

## Mathematical Induction

The ability to spot patterns is one of the keys to success in mathematical problem solving. Consider the following pattern:

$$1 = 1$$
$$1 + 3 = 4$$
$$1 + 3 + 5 = 9$$
$$1 + 3 + 5 + 7 = 16$$
$$1 + 3 + 5 + 7 + 9 = 25$$

*Great fleas have little fleas*
*upon their backs to bite 'em,*
*And little fleas have lesser fleas,*
*and so* ad infinitum.
—Augustus De Morgan
*A Budget of Paradoxes*
Longmans, Green, and Company, 1872, p. 377

The sums are all perfect squares: $1^2, 2^2, 3^2, 4^2, 5^2$. It seems reasonable to conjecture that this pattern will continue to hold; that is, the sum of consecutive odd numbers, starting at 1, will always be a perfect square. Let's try to be more precise. If the sum is $n^2$, then the last odd number in the sum is $2n - 1$. (Check this in the five cases above.) In symbols, our conjecture becomes

$$1 + 3 + 5 + \cdots + (2n - 1) = n^2 \quad \text{for all } n \geq 1 \tag{1}$$

Notice that Equation (1) is really an *infinite* collection of statements, one for each value of $n \geq 1$. Although our conjecture seems reasonable, we cannot assume that the pattern continues—we need to prove it. This is where ***mathematical induction*** comes in.

### First Principle of Mathematical Induction

Let $S(n)$ be a statement about the positive integer $n$. If

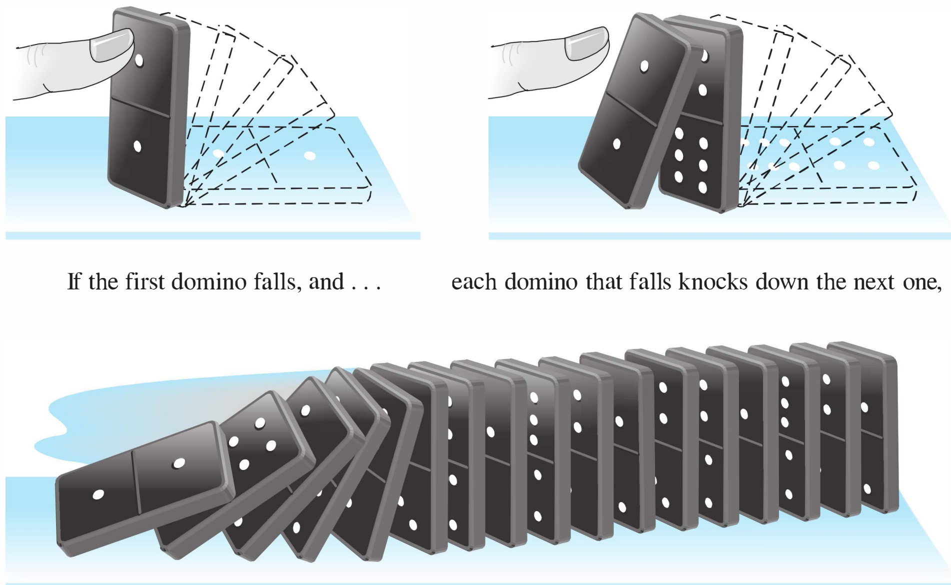
1. $S(1)$ is true and
2. for all $k \geq 1$, the truth of $S(k)$ implies the truth of $S(k + 1)$

then $S(n)$ is true for all $n \geq 1$.

Verifying that $S(1)$ is true is called the ***basis step.*** The assumption that $S(k)$ is true for some $k \geq 1$ is called the ***induction hypothesis.*** Using the induction hypothesis to prove that $S(k + 1)$ is then true is called the ***induction step.*** Mathematical induction has been referred to as the *domino principle* because it is analogous to showing that a line of dominoes will fall down if (1) the first domino can be knocked down (the basis step) and (2) knocking down any domino (the induction hypothesis) will knock over the next domino (the induction step). See Figure B.1.

We now use the principle of mathematical induction to prove Equation (1).

*Exercises and selected odd-numbered answers for this appendix can be found on the student companion website.



If the first domino falls, and . . . each domino that falls knocks down the next one, . . . then all the dominoes can be made to fall by pushing over the first one.

**Figure B.1**

**Example B.1** Use mathematical induction to prove that

$$1 + 3 + 5 + \cdots + (2n - 1) = n^2$$

for all $n \geq 1$.

**Solution** For $n = 1$, the sum on the left-hand side is just 1, while the right-hand side is $1^2$. Since $1 = 1^2$, this completes the basis step.

Now assume that the formula is true for some integer $k \geq 1$. That is, assume that

$$1 + 3 + 5 + \cdots + (2k - 1) = k^2$$

(This is the induction hypothesis.) The induction step consists of proving that the formula is true when $n = k + 1$. We see that when $n = k + 1$, the left-hand side of formula (1) is

$$\begin{aligned} 1 + 3 + 5 + \cdots + (2(k+1) - 1) &= 1 + 3 + 5 + \cdots + (2k + 1) \\ &= \underbrace{1 + 3 + 5 + \cdots + (2k - 1)}_{k^2} + (2k + 1) \\ &= \quad k^2 \quad + 2k + 1 \quad \leftarrow \text{by the induction hypothesis} \\ &= (k + 1)^2 \end{aligned}$$

which is the right-hand side of Equation (1) when $n = k + 1$.

This completes the induction step, and we conclude that Equation (1) is true for all $n \geq 1$, by the principle of mathematical induction.

The next example gives a proof of a useful formula for the sum of the first $n$ positive integers. The formula appears several times in the text; for example, see the solution to Exercise 51 in Section 2.4.

**Example B.2** Prove that

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

for all $n \geq 1$.

**Solution** The formula is true for $n = 1$, since

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

Assume that the formula is true for $n = k$; that is,

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

We need to show that the formula is true when $n = k + 1$; that is, we must prove that

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

But we see that

$$\begin{aligned} 1 + 2 + \cdots + (k + 1) &= (1 + 2 + \cdots + k) + (k + 1) \\ &= \frac{k(k + 1)}{2} + (k + 1) \qquad \text{by the induction hypothesis} \\ &= \frac{k(k + 1) + 2(k + 1)}{2} \\ &= \frac{k^2 + 3k + 2}{2} \\ &= \frac{(k + 1)(k + 2)}{2} \\ &= \frac{(k + 1)[(k + 1) + 1]}{2} \end{aligned}$$

which is what we needed to show.

This completes the induction step, and we conclude that the formula is true for all $n \geq 1$, by the principle of mathematical induction.

In a similar vein, we can prove that the sum of the squares of the first $n$ positive integers satisfies the formula

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n + 1)(2n + 1)}{6}$$

for all $n \geq 1$. (Verify this for yourself.)

The basis step need not be for $n = 1$, as the next two examples illustrate.

**Example B.3** Prove that $n! > 2^n$ for all integers $n \geq 4$.

**Solution** The basis step here is when $n = 4$. The inequality is clearly true in this case, since

$$4! = 24 > 16 = 2^4$$

Assume that $k! > 2^k$ for some integer $k \geq 4$. Then

$$\begin{aligned}(k+1)! &= (k+1)k! \\ &> (k+1)2^k && \text{by the induction hypothesis} \\ &\geq 5 \cdot 2^k && \text{since } k \geq 4 \\ &> 2 \cdot 2^k = 2^{k+1}\end{aligned}$$

which verifies the inequality for $n = k + 1$ and completes the induction step.

We conclude that $n! > 2^n$ for all integers $n \geq 4$, by the principle of mathematical induction.

If $a$ is a nonzero real number and $n \geq 0$ is an integer, we can give a recursive definition of the power $a^n$ that is compatible with mathematical induction. We define $a^0 = 1$ and, for $n \geq 0$,

$$a^{n+1} = a^n a$$

(This form avoids the ellipses used in the version $a^n = \overbrace{aa \cdots a}^{n \text{ times}}$.) We can now use mathematical induction to verify a familiar property of exponents.

**Example B.4** Let $a$ be a nonzero real number. Prove that $a^m a^n = a^{m+n}$ for all integers $m, n \geq 0$.

**Solution** At first glance, it is not clear how to proceed, since there are *two* variables, $m$ and $n$. But we simply need to keep one of them fixed and perform our induction using the other. So, let $m \geq 0$ be a fixed integer. When $n = 0$, we have

$$a^m a^0 = a^m \cdot 1 = a^m = a^{m+0}$$

using the definition $a^0 = 1$. Hence, the basis step is true.

Now assume that the formula holds when $n = k$, where $k \geq 0$. Then $a^m a^k = a^{m+k}$. For $n = k + 1$, using our recursive definition and the fact that addition and multiplication are associative, we see that

$$\begin{aligned}a^m a^{k+1} &= a^m(a^k a) && \text{by definition} \\ &= (a^m a^k)a \\ &= a^{m+k}a && \text{by the induction hypothesis} \\ &= a^{(m+k)+1} && \text{by definition} \\ &= a^{m+(k+1)}\end{aligned}$$

Therefore, the formula is true for $n = k + 1$, and the induction step is complete.

We conclude that $a^m a^n = a^{m+n}$ for all integers $m, n \geq 0$, by the principle of mathematical induction.

In Examples B.1 through B.4, the use of the induction hypothesis during the induction step is relatively straightforward. However, this is not always the case. An alternative version of the principle of mathematical induction is often more useful.

### Second Principle of Mathematical Induction

Let $S(n)$ be a statement about the positive integer $n$. If

1. $S(1)$ is true and
2. the truth of $S(1), S(2), \ldots, S(k)$ implies the truth of $S(k + 1)$

then $S(n)$ is true for all $n \geq 1$.

The only difference between the two principles of mathematical induction is in the induction hypothesis: The first version assumes that $S(k)$ is true, whereas the second version assumes that all of $S(1), S(2), \ldots, S(k)$ are true. This makes the second principle seem weaker than the first, since we need to assume more in order to prove $S(k + 1)$ (although, paradoxically, the second principle is sometimes called *strong* induction). In fact, however, the two principles are logically equivalent: Each one implies the other. (Can you see why?)

The next example presents an instance in which the second principle of mathematical induction is easier to use than the first. Recall that a prime number is a positive integer whose only positive integer factors are 1 and itself.

**Example B.5** Prove that every positive integer $n \geq 2$ either is prime or can be factored into a product of primes.

**Solution** The result is clearly true when $n = 2$, since 2 is prime. Now assume that for all integers $n$ between 2 and $k$, $n$ either is prime or can be factored into a product of primes. Let $n = k + 1$. If $k + 1$ is prime, we are done. Otherwise, it must factor into a product of two smaller integers—say,

$$k + 1 = ab$$

Since $2 \leq a, b \leq k$ (why?), the induction hypothesis applies to $a$ and $b$. Therefore,

$$a = p_1 \cdots p_r \quad \text{and} \quad b = q_1 \cdots q_s$$

where the $p$'s and $q$'s are all prime. Then

$$ab = p_1 \cdots p_r q_1 \cdots q_s$$

gives a factorization of $ab$ into primes, completing the induction step.

We conclude that the result is true for all integers $n \geq 2$, by the second principle of mathematical induction.

Do you see why the first principle of mathematical induction would have been difficult to use here?

We conclude with a highly nontrivial example that involves a combination of induction and *backward* induction. The result is the Arithmetic Mean–Geometric Mean Inequality, discussed in Chapter 7 in Exploration: Geometric Inequalities and Optimization Problems. The clever proof in Example B.6 is due to Cauchy.

**Example B.6**

Let $x_1, \ldots, x_n$ be nonnegative real numbers. Prove that

$$\sqrt[n]{x_1x_2\cdots x_n} \leq \frac{x_1 + x_2 + \cdots + x_n}{n}$$

for all integers $n \geq 2$.

**Solution** For $n = 2$, the inequality becomes $\sqrt{xy} \leq (x + y)/2$. You are asked to verify this in Problems 1 and 2 of the Exploration mentioned above.

If $S(n)$ is the stated inequality, we will prove that $S(k)$ implies $S(2k)$. Assume that $S(k)$ is true; that is,

$$\sqrt[k]{x_1x_2\cdots x_k} \leq \frac{x_1 + x_2 + \cdots + x_k}{k}$$

for all nonnegative real numbers $x_1, \ldots, x_k$. Let

$$x_1 = \frac{y_1 + y_2}{2}, \quad x_2 = \frac{y_3 + y_4}{2}, \quad \ldots, \quad x_k = \frac{y_{2k-1} + y_{2k}}{2}$$

Then

$$\begin{aligned}
\sqrt[2k]{y_1\cdots y_{2k}} &= \sqrt[k]{\sqrt{y_1\cdots y_{2k}}} = \sqrt[k]{\sqrt{y_1y_2}\cdots\sqrt{y_{2k-1}y_{2k}}} \\
&\leq {}^{k}\left(\frac{y_1 + y_2}{2}\right)\cdots\left(\frac{y_{2k-1} + y_{2k}}{2}\right) && \text{by } S(2) \\
&= \sqrt[k]{x_1\cdots x_k} \\
&\leq \frac{x_1 + x_2 + \cdots + x_k}{k} && \text{by } S(k) \\
&= \frac{\left(\frac{y_1 + y_2}{2}\right) + \cdots + \left(\frac{y_{2k-1} + y_{2k}}{2}\right)}{k} \\
&= \frac{y_1 + \cdots + y_{2k}}{2k}
\end{aligned}$$

which verifies $S(2k)$.

Thus, the Arithmetic Mean–Geometric Mean Inequality is true for $n = 2, 4, 8, \ldots$—the powers of 2. In order to complete the proof, we need to "fill in the gaps." We will use backward induction to prove that $S(k)$ implies $S(k - 1)$. Assuming $S(k)$ is true, let

$$x_k = \frac{x_1 + x_2 + \cdots + x_{k-1}}{k - 1}$$

Then

$$\sqrt[k]{x_1x_2\cdots x_{k-1}\left(\frac{x_1 + x_2 + \cdots + x_{k-1}}{k-1}\right)} \leq \frac{x_1 + x_2 + \cdots + \left(\frac{x_1 + x_2 + \cdots + x_{k-1}}{k-1}\right)}{k}$$

$$= \frac{kx_1 + kx_2 + \cdots + kx_{k-1}}{k(k-1)}$$

$$= \frac{x_1 + x_2 + \cdots + x_{k-1}}{k-1}$$

Equivalently,

$$x_1x_2\cdots x_{k-1}\left(\frac{x_1 + x_2 + \cdots + x_{k-1}}{k-1}\right) \leq \left(\frac{x_1 + x_2 + \cdots + x_{k-1}}{k-1}\right)^k$$

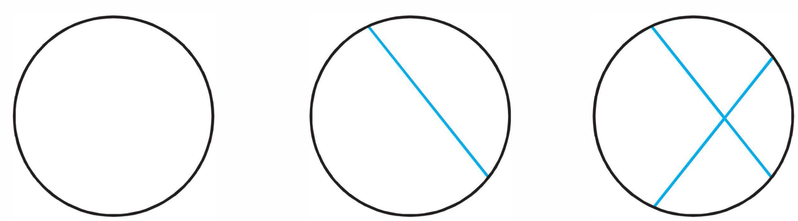
or

$$x_1x_2\cdots x_{k-1} \leq \left(\frac{x_1 + x_2 + \cdots + x_{k-1}}{k-1}\right)^{k-1}$$

Taking the $(k-1)$th root of both sides yields $S(k-1)$.

The two inductions, taken together, show that the Arithmetic Mean–Geometric Mean Inequality is true for all $n \geq 2$.

**Remark** Although mathematical induction is a powerful and indispensable tool, it cannot work miracles. That is, it cannot prove that a pattern or formula holds if it does not. Consider the diagrams in Figure B.2, which show the maximum number of regions $R(n)$ into which a circle can be subdivided by $n$ straight lines.



$R(0) = 1 = 2^0$ $R(1) = 2 = 2^1$ $R(2) = 4 = 2^2$

**Figure B.2**

Based on the evidence in Figure B.2, we might conjecture that $R(n) = 2^n$ for $n \geq 0$ and try to prove this conjecture using mathematical induction. We would not succeed, since this formula is not correct! If we had considered one more case, we would have discovered that $R(3) = 7 \neq 8 = 2^3$, thereby demolishing our conjecture. In fact, the correct formula turns out to be

$$R(n) = \frac{n^2 + n + 2}{2}$$

which *can* be verified by induction. (Can you do it?)

For other examples in which a pattern appears to be true, only to disappear when enough cases are considered, see Richard K. Guy's delightful article "The Strong Law of Small Numbers" in the *American Mathematical Monthly*, Vol. 95 (1988), pp. 697–712.

# Appendix C*

## Complex Numbers

A ***complex number*** is a number of the form $a + bi$, where $a$ and $b$ are real numbers and $i$ is a symbol with the property that $i^2 = -1$. The real number $a$ is considered to be a special type of complex number, since $a = a + 0i$. If $z = a + bi$ is a complex number, then the ***real part*** of $z$, denoted by Re $z$, is $a$, and the ***imaginary part*** of $z$, denoted by Im $z$, is $b$. Two complex numbers $a + bi$ and $c + di$ are ***equal*** if their real parts are equal and their imaginary parts are equal—that is, if $a = c$ and $b = d$. A complex number $a + bi$ can be identified with the point $(a, b)$ and plotted in the plane (called the ***complex plane,*** or the ***Argand plane***), as shown in Figure C.1. In the complex plane, the horizontal axis is called the ***real axis*** and the vertical axis is called the ***imaginary axis.***

*[The] extension of the number concept to include the irrational, and we will at once add, the imaginary, is the greatest forward step which pure mathematics has ever taken.*
—Hermann Hankel
*Theorie der Complexen Zahlensysteme*
Leipzig, 1867, p. 60

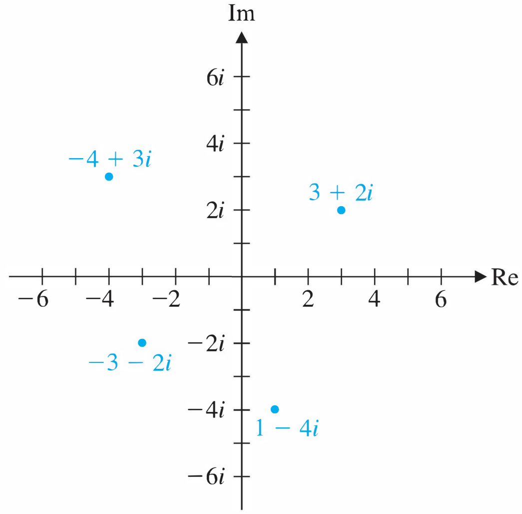



**Figure C.1**
The complex plane

There is nothing “imaginary” about complex numbers—they are just as “real” as the real numbers. The term *imaginary* arose from the study of polynomial equations such as $x^2 + 1 = 0$, whose solutions are not “real” (i.e., real numbers). It is worth remembering that at one time negative numbers were thought of as “imaginary” too.

### Operations on Complex Numbers

The ***sum*** of the complex numbers $a + bi$ and $c + di$ is defined as

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

Notice that, with the identification of $a + bi$ with $(a, b)$, $c + di$ with $(c, d)$, and $(a + c) + (b + d)i$ with $(a + c, b + d)$, addition of complex numbers is the same as vector addition. The ***product*** of $a + bi$ and $c + di$ is

$$\begin{aligned}(a + bi)(c + di) &= a(c + di) + bi(c + di)\\ &= ac + adi + bci + bdi^2\end{aligned}$$

Jean-Robert Argand (1768–1822) was a French accountant and amateur mathematician. His geometric interpretation of complex numbers appeared in 1806 in a book that he published privately. He was not, however, the first to give such an interpretation. The Norwegian-Danish surveyor Caspar Wessel (1745–1818) gave the same version of the complex plane in 1787, but his paper was not noticed by the mathematical community until after his death.

*Exercises and selected odd-numbered answers for this appendix can be found on the student companion website.

Since $i^2 = -1$, this expression simplifies to $(ac - bd) + (ad + bc)i$. Thus, we have

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i$$

Observe that, as a special case, $a(c + di) = ac + adi$, so the ***negative*** of $c + di$ is $-(c + di) = (-1)(c + di) = -c - di$. This fact allows us to compute the ***difference*** of $a + bi$ and $c + di$ as

$$\begin{aligned}(a + bi) - (c + di) &= (a + bi) + (-1)(c + di) \\ &= (a + (-c)) + (b + (-d))i \\ &= (a - c) + (b - d)i\end{aligned}$$

**Example C.1**

Find the sum, difference, and product of $3 - 4i$ and $-1 + 2i$.

**Solution** The sum is

$$(3 - 4i) + (-1 + 2i) = (3 - 1) + (-4 + 2)i = 2 - 2i$$

The difference is

$$(3 - 4i) - (-1 + 2i) = (3 - (-1)) + (-4 - 2)i = 4 - 6i$$

The product is

$$\begin{aligned}(3 - 4i)(-1 + 2i) &= -3 + 6i + 4i - 8i^2 \\ &= -3 + 10i - 8(-1) = 5 + 10i\end{aligned}$$

The ***conjugate*** of $z = a + bi$ is the complex number

$$\bar{z} = a - bi$$

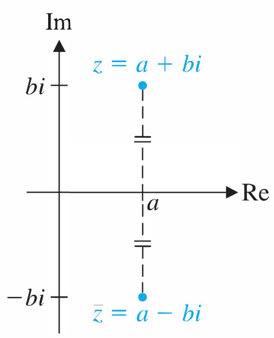



**Figure C.2**
Complex conjugates

($\bar{z}$ is pronounced "$z$ bar.") Figure C.2 gives the geometric interpretation of the conjugate.

To find the quotient of two complex numbers, we multiply the numerator and the denominator by the conjugate of the denominator.

**Example C.2**

Express $\dfrac{-1 + 2i}{3 + 4i}$ in the form $a + bi$.

**Solution** We multiply the numerator and denominator by $\overline{3 + 4i} = 3 - 4i$. Using Example C.1, we obtain

$$\frac{-1 + 2i}{3 + 4i} = \frac{-1 + 2i}{3 + 4i} \cdot \frac{3 - 4i}{3 - 4i} = \frac{5 + 10i}{3^2 + 4^2} = \frac{5 + 10i}{25} = \frac{1}{5} + \frac{2}{5}i$$

On the following page is a summary of some of the properties of conjugates. The proofs follow from the definition of conjugate; you should verify them for yourself.

1. $\overline{\overline{z}} = z$
2. $\overline{z + w} = \overline{z} + \overline{w}$
3. $\overline{zw} = \overline{z}\,\overline{w}$
4. If $z \neq 0$, then $\overline{(w/z)} = \overline{w}/\overline{z}$.
5. $z$ is real if and only if $\overline{z} = z$.

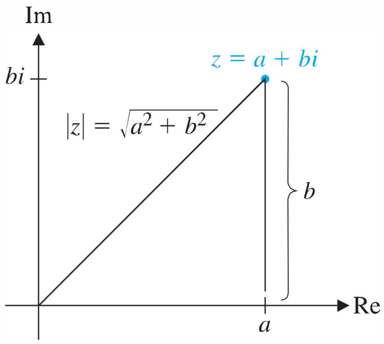


**Figure C.3**

The ***absolute value*** (or ***modulus***) $|z|$ of a complex number $z = a + bi$ is its distance from the origin. As Figure C.3 shows, Pythagoras' Theorem gives

$$|z| = |a + bi| = \sqrt{a^2 + b^2}$$

Observe that

$$z\overline{z} = (a + bi)(a - bi) = a^2 - abi + bai - b^2i^2 = a^2 + b^2$$

Hence,

$$z\overline{z} = |z|^2$$

This gives us an alternative way of describing the division process for the quotient of two complex numbers. If $w$ and $z \neq 0$ are two complex numbers, then

$$\frac{w}{z} = \frac{w}{z}\cdot\frac{\overline{z}}{\overline{z}} = \frac{w\overline{z}}{z\overline{z}} = \frac{w\overline{z}}{|z|^2}$$

Below is a summary of some of the properties of absolute value. (You should try to prove these using the definition of absolute value and other properties of complex numbers.)

1. $|z| = 0$ if and only if $z = 0$.
2. $|z| = |\overline{z}|$
3. $|zw| = |z||w|$
4. If $z \neq 0$, then $\left|\frac{1}{z}\right| = \frac{1}{|z|}$.
5. $|z + w| \leq |z| + |w|$

## Polar Form

As you have seen, the complex number $z = a + bi$ can be represented geometrically by the point $(a, b)$. This point can also be expressed in terms of ***polar coordinates*** $(r, \theta)$, where $r \geq 0$, as shown in Figure C.4. We have

$$a = r\cos\theta \quad \text{and} \quad b = r\sin\theta$$

so

$$z = a + bi = r\cos\theta + (r\sin\theta)i$$

Thus, any complex number can be written in the ***polar form***

$$z = r(\cos\theta + i\sin\theta)$$

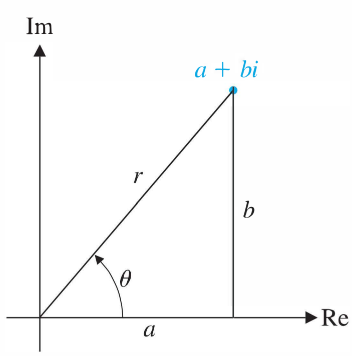


**Figure C.4**

where $r = |z| = \sqrt{a^2 + b^2}$ and $\tan \theta = b/a$. The angle $\theta$ is called an ***argument*** of $z$ and is denoted by arg $z$. Observe that arg $z$ is not unique: Adding or subtracting any integer multiple of $2\pi$ gives another argument of $z$. However, there is only one argument $\theta$ that satisfies

$$-\pi < \theta \leq \pi$$

This is called the ***principal argument*** of $z$ and is denoted by Arg $z$.

**Example C.3**

Write the following complex numbers in polar form using their principal arguments:

(a) $z = 1 + i$ (b) $w = 1 - \sqrt{3}i$

**Solution** (a) We compute

$$r = |z| = \sqrt{1^2 + 1^2} = \sqrt{2} \quad \text{and} \quad \tan \theta = \frac{1}{1} = 1$$

Therefore, Arg $z = \theta = \dfrac{\pi}{4}$ (= 45°), and we have

$$z = \sqrt{2}\left(\cos \frac{\pi}{4} + i \sin \frac{\pi}{4}\right)$$

as shown in Figure C.5.

(b) We have

$$r = |w| = \sqrt{1^2 + (-\sqrt{3})^2} = \sqrt{4} = 2 \quad \text{and} \quad \tan \theta = \frac{-\sqrt{3}}{1} = -\sqrt{3}$$

Since $w$ lies in the fourth quadrant, we must have Arg $z = \theta = -\dfrac{\pi}{3}$ (= −60°). Therefore,

$$w = 2\left(\cos\left(-\frac{\pi}{3}\right) + i \sin\left(-\frac{\pi}{3}\right)\right)$$

See Figure C.5.

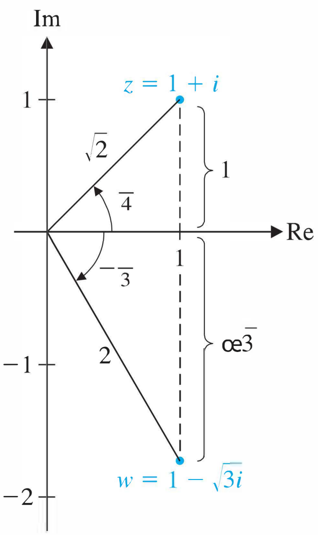


**Figure C.5**

The polar form of complex numbers can be used to give geometric interpretations of multiplication and division. Let

$$z_1 = r_1(\cos \theta_1 + i \sin \theta_1) \quad \text{and} \quad z_2 = r_2(\cos \theta_2 + i \sin \theta_2)$$

Multiplying, we obtain

$$\begin{aligned} z_1 z_2 &= r_1 r_2 (\cos \theta_1 + i \sin \theta_1)(\cos \theta_2 + i \sin \theta_2) \\ &= r_1 r_2 [(\cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2) + i(\sin \theta_1 \cos \theta_2 + \cos \theta_1 \sin \theta_2)] \end{aligned}$$

Using the trigonometric identities

$$\begin{aligned} \cos(\theta_1 + \theta_2) &= \cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2 \\ \sin(\theta_1 + \theta_2) &= \sin \theta_1 \cos \theta_2 + \cos \theta_1 \sin \theta_2 \end{aligned}$$

we obtain

$$z_1z_2 = r_1r_2[\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)] \qquad (1)$$

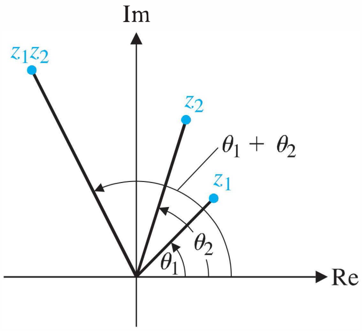



**Figure C.6**

which is the polar form of a complex number with absolute value $r_1r_2$ and argument $\theta_1 + \theta_2$. This shows that

$$|z_1z_2| = |z_1||z_2| \quad \text{and} \quad \arg(z_1z_2) = \arg z_1 + \arg z_2$$

Equation (1) says that *to multiply two complex numbers, we multiply their absolute values and add their arguments*. See Figure C.6.

Similarly, using the subtraction identities for sine and cosine, we can show that

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}[\cos(\theta_1 - \theta_2) + i\sin(\theta_1 - \theta_2)] \quad \text{if } z \neq 0$$

(Verify this.) Therefore,

$$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|} \quad \text{and} \quad \arg\left(\frac{z_1}{z_2}\right) = \arg z_1 - \arg z_2$$

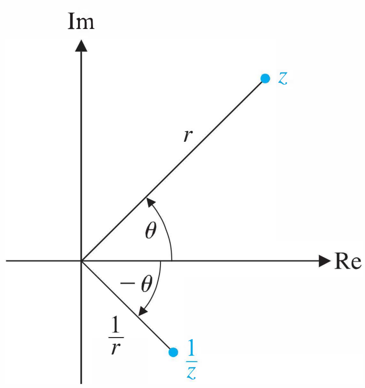



**Figure C.7**

and we see that *to divide two complex numbers, we divide their absolute values and subtract their arguments*.

As a special case of the last result, we obtain a formula for the reciprocal of a complex number in polar form. Setting $z_1 = 1$ (and therefore $\theta_1 = 0$) and $z_2 = z$ (and therefore $\theta_2 = \theta$), we obtain the following:

> If $z = r(\cos\theta + i\sin\theta)$ is nonzero, then
>
> $$\frac{1}{z} = \frac{1}{r}(\cos\theta - i\sin\theta)$$

See Figure C.7.

## Example C.4

Find the product of $1 + i$ and $1 - \sqrt{3}i$ in polar form.

**Solution** From Example C.3, we have

$$1 + i = \sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right) \quad \text{and} \quad 1 - \sqrt{3}i = 2\left(\cos\left(-\frac{\pi}{3}\right) + i\sin\left(-\frac{\pi}{3}\right)\right)$$

Therefore,

$$\begin{aligned}(1 + i)(1 - \sqrt{3}i) &= 2\sqrt{2}\left[\cos\left(\frac{\pi}{4} - \frac{\pi}{3}\right) + i\sin\left(\frac{\pi}{4} - \frac{\pi}{3}\right)\right] \\ &= 2\sqrt{2}\left[\cos\left(-\frac{\pi}{12}\right) + i\sin\left(-\frac{\pi}{12}\right)\right]\end{aligned}$$

See Figure C.8.

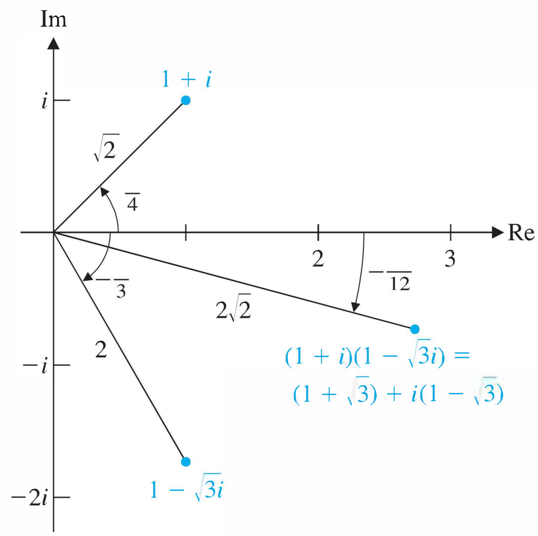



**Figure C.8**

**Remark** Since $(1+i)(1-\sqrt{3}i) = (1+\sqrt{3}) + i(1-\sqrt{3})$ (check this), we must have

$$1+\sqrt{3} = 2\sqrt{2}\cos\left(-\frac{\pi}{12}\right) = -2\sqrt{2}\cos\left(\frac{\pi}{12}\right)$$

and

$$1-\sqrt{3} = 2\sqrt{2}\sin\left(-\frac{\pi}{12}\right) = -2\sqrt{2}\sin\left(\frac{\pi}{12}\right)$$

(Why?) This implies that

$$\cos\left(\frac{\pi}{12}\right) = \frac{1+\sqrt{3}}{2\sqrt{2}} \quad \text{and} \quad \sin\left(\frac{\pi}{12}\right) = \frac{\sqrt{3}-1}{2\sqrt{2}}$$

We therefore have a method for finding the sine and cosine of an angle such as $\pi/12$ that is not a special angle but that can be obtained as a sum or difference of special angles.

## De Moivre's Theorem

If $n$ is a positive integer and $z = r(\cos\theta + i\sin\theta)$, then repeated use of Equation (1) yields formulas for the powers of $z$:

$$z^2 = r^2(\cos 2\theta + i\sin 2\theta)$$
$$z^3 = zz^2 = r^3(\cos 3\theta + i\sin 3\theta)$$
$$z^4 = zz^3 = r^4(\cos 4\theta + i\sin 4\theta)$$
$$\vdots$$

Abraham De Moivre (1667–1754) was a French mathematician who made important contributions to trigonometry, analytic geometry, probability, and statistics.

In general, we have the following result, known as ***De Moivre's Theorem.***

**Theorem C.1** **De Moivre's Theorem**

If $z = r(\cos\theta + i\sin\theta)$ and $n$ is a positive integer, then

$$z^n = r^n(\cos n\theta + i\sin n\theta)$$

Stated differently, we have

$$|z^n| = |z|^n \quad \text{and} \quad \arg(z^n) = n \arg z$$

In words, De Moivre's Theorem says that *to take the* $n$th *power of a complex number, we take the* $n$th *power of its absolute value and multiply its argument by* $n$.

**Example C.5** Find $(1 + i)^6$.

**Solution** From Example C.3(a), we have

$$1 + i = \sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right)$$

Hence, De Moivre's Theorem gives

$$\begin{aligned}(1 + i)^6 &= (\sqrt{2})^6\left(\cos\frac{6\pi}{4} + i\sin\frac{6\pi}{4}\right)\\ &= 8\left(\cos\frac{3\pi}{2} + i\sin\frac{3\pi}{2}\right)\\ &= 8(0 + i(-1)) = -8i\end{aligned}$$

See Figure C.9, which shows $1 + i, (1 + i)^2, (1 + i)^3, \ldots, (1 + i)^6$.

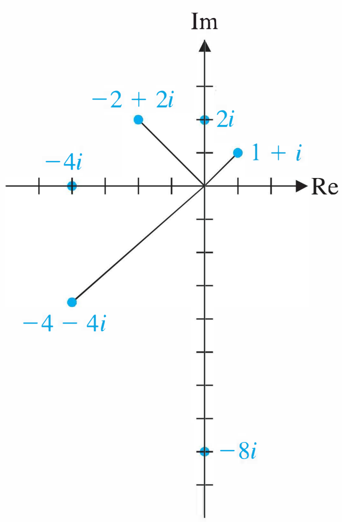



**Figure C.9**
Powers of $1 + i$

We can also use De Moivre's Theorem to find $n$th roots of complex numbers. An $n$th root of the complex number $z$ is any complex number $w$ such that

$$w^n = z$$

In polar form, we have

$$w = s(\cos\varphi + i\sin\varphi) \quad \text{and} \quad z = r(\cos\theta + i\sin\theta)$$

so, by De Moivre's Theorem,

$$s^n(\cos n\varphi + i\sin n\varphi) = r(\cos\theta + i\sin\theta)$$

Equating the absolute values, we see that

$$s^n = r \quad \text{or} \quad s = r^{1/n} = \sqrt[n]{r}$$

We must also have

$$\cos n\varphi = \cos\theta \quad \text{and} \quad \sin n\varphi = \sin\theta$$

(Why?) Since the sine and cosine functions each have period $2\pi$, these equations imply that $n\varphi$ and $\theta$ differ by an integer multiple of $2\pi$; that is,

$$n\varphi = \theta + 2k\pi \quad \text{or} \quad \varphi = \frac{\theta + 2k\pi}{n}$$

where $k$ is an integer. Therefore,

$$w = r^{1/n}\left[\cos\left(\frac{\theta + 2k\pi}{n}\right) + i\sin\left(\frac{\theta + 2k\pi}{n}\right)\right]$$

describes the possible $n$th roots of $z$ as $k$ ranges over the integers. It is not hard to show that $k = 0, 1, 2, \ldots, n - 1$ produce distinct values of $w$, so there are exactly $n$ different $n$th roots of $z = r(\cos\theta + i\sin\theta)$. We summarize this result as follows:

Let $z = r(\cos\theta + i\sin\theta)$ and let $n$ be a positive integer. Then $z$ has exactly $n$ distinct $n$th roots given by

$$r^{1/n}\left[\cos\left(\frac{\theta + 2k\pi}{n}\right) + i\sin\left(\frac{\theta + 2k\pi}{n}\right)\right] \quad (2)$$

for $k = 0, 1, 2, \ldots, n - 1$.

## Example C.6

Find the three cube roots of $-27$.

**Solution** In polar form, $-27 = 27(\cos\pi + i\sin\pi)$. It follows that the cube roots of $-27$ are given by

$$(-27)^{1/3} = 27^{1/3}\left[\cos\left(\frac{\pi + 2k\pi}{3}\right) + i\sin\left(\frac{\pi + 2k\pi}{3}\right)\right] \quad \text{for } k = 0, 1, 2$$

Using formula (2) with $n = 3$, we obtain

$$27^{1/3}\left[\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right] = 3\left(\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) = \frac{3}{2} + \frac{3\sqrt{3}}{2}i$$

$$27^{1/3}\left[\cos\left(\frac{\pi + 2\pi}{3}\right) + i\sin\left(\frac{\pi + 2\pi}{3}\right)\right] = 3(\cos\pi + i\sin\pi) = -3$$

$$27^{1/3}\left[\cos\left(\frac{\pi + 4\pi}{3}\right) + i\sin\left(\frac{\pi + 4\pi}{3}\right)\right] = 3\left(\cos\frac{5\pi}{3} + i\sin\frac{5\pi}{3}\right)$$

$$= 3\left(\frac{1}{2} - \frac{\sqrt{3}}{2}i\right) = \frac{3}{2} - \frac{3\sqrt{3}}{2}i$$

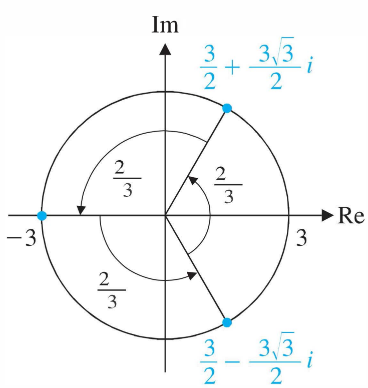
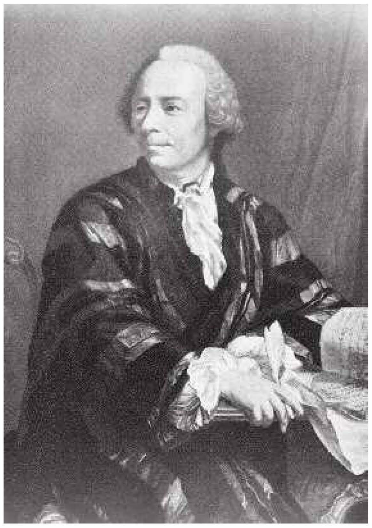



**Figure C.10**
The cube roots of $-27$

As Figure C.10 shows, the three cube roots of $-27$ are equally spaced $2\pi/3$ radians (120°) apart around a circle of radius 3 centered at the origin.

In general, formula (2) implies that the $n$th roots of $z = r(\cos\theta + i\sin\theta)$ will lie on a circle of radius $r^{1/n}$ centered at the origin. Moreover, they will be equally spaced $2\pi/n$ radians $(360/n)°$ apart. (Verify this.) Thus, if we can find one $n$th root of $z$, the remaining $n$th roots of $z$ can be obtained by rotating the first root through successive increments of $2\pi/n$ radians. Had we known this in Example C.6, we could have used the fact that the real cube root of $-27$ is $-3$ and then rotated it twice through an angle of $2\pi/3$ radians (120°) to get the other two cube roots.





Leonhard Euler (1707–1783) was the most prolific mathematician of all time. He has over 900 publications to his name, and his collected works fill over 70 volumes. There are so many results attributed to him that "Euler's formula" or "Euler's Theorem" can mean many different things, depending on the context.

Euler worked in so many areas of mathematics, it is difficult to list them all. His contributions to calculus and analysis, differential equations, number theory, geometry, topology, mechanics, and other areas of applied mathematics continue to be influential. He also introduced much of the notation we currently use, including $\pi$, $e$, $i$, $\Sigma$ for summation, $\Delta$ for difference, and $f(x)$ for a function, and was the first to treat sine and cosine as functions.

Euler was born in Switzerland but spent most of his mathematical life in Russia and Germany. In 1727, he joined the St. Petersburg Academy of Sciences, which had been founded by Catherine I, the wife of Peter the Great. He went to Berlin in 1741 at the invitation of Frederick the Great, but returned in 1766 to St. Petersburg, where he remained until his death. When he was young, he lost the vision in one eye as the result of an illness, and by 1776 he had lost the vision in the other eye and was totally blind. Remarkably, his mathematical output did not diminish, and he continued to be productive until the day he died.

## Euler's Formula

In calculus, you learn that the function $e^z$ has a power series expansion

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots$$

that converges for every real number $z$. It can be shown that this expansion also works when $z$ is a complex number and that the complex exponential function $e^z$ obeys the usual rules for exponents. The sine and cosine functions also have power series expansions:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + - \cdots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + - \cdots$$

If we let $z = ix$, where $x$ is a real number, then we have

$$e^z = e^{ix} = 1 + ix + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \cdots$$

Using the fact that $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, $i^5 = i$, and so on, repeating in a cycle of length 4, we see that

$$\begin{aligned} e^{ix} &= 1 + ix - \frac{x^2}{2!} - \frac{ix^3}{3!} + \frac{x^4}{4!} + \frac{ix^5}{5!} - \frac{x^6}{6!} - \frac{ix^7}{7!} + + - - \cdots \\ &= \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + - \cdots\right) + i\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + - \cdots\right) \\ &= \cos x + i \sin x \end{aligned}$$

This remarkable result is known as ***Euler's formula.***

**Theorem C.2** **Euler's Formula**

For any real number $x$,

$$e^{ix} = \cos x + i \sin x$$

Using Euler's formula, we see that the polar form of a complex number can be written more compactly as

$$z = r(\cos\theta + i\sin\theta) = re^{i\theta}$$

For example, from Example C.3(a), we have

$$1 + i = \sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right) = \sqrt{2}e^{i\pi/4}$$

We can also go in the other direction and convert a complex exponential back into polar or standard form.

**Example C.7** Write the following in the form $a + bi$:

(a) $e^{i\pi}$ (b) $e^{2+i\pi/4}$

**Solution** (a) Using Euler's formula, we have

$$e^{i\pi} = \cos\pi + i\sin\pi = -1 + i\cdot 0 = -1$$

(If we write this equation as $e^{i\pi} + 1 = 0$, we obtain what is surely one of the most remarkable equations in mathematics. It contains the fundamental operations of addition, multiplication, and exponentiation; the additive identity 0 and the multiplicative identity 1; the two most important transcendental numbers, $\pi$ and $e$; and the complex unit $i$—all in one equation!)

(b) Using rules for exponents together with Euler's formula, we obtain

$$e^{2+i\pi/4} = e^2 e^{i\pi/4} = e^2\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right) = e^2\left(\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}\right)$$

$$= \frac{e^2\sqrt{2}}{2} + \frac{e^2\sqrt{2}}{2}i$$

If $z = re^{i\theta} = r(\cos\theta + i\sin\theta)$, then

$$\bar{z} = r(\cos\theta - i\sin\theta) \tag{3}$$

The trigonometric identities

$$\cos(-\theta) = \cos\theta \text{ and } \sin(-\theta) = -\sin\theta$$

allow us to rewrite Equation (3) as

$$\bar{z} = r(\cos(-\theta) + i\sin(-\theta)) = re^{i(-\theta)}$$

This gives the following useful formula for the conjugate:

If $z = re^{i\theta}$, then

$$\bar{z} = re^{-i\theta}$$

**Note** Euler's formula gives a quick, one-line proof of De Moivre's Theorem:

$$[r(\cos\theta + i\sin\theta)]^n = (re^{i\theta})^n = r^n e^{in\theta} = r^n(\cos n\theta + i\sin n\theta)$$

# Appendix D*

## Polynomials

A ***polynomial*** is a function $p$ of a single variable $x$ that can be written in the form

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \tag{1}$$

*Euler gave the most algebraic of the proofs of the existence of the roots of [a polynomial] equation. . . . I regard it as unjust to ascribe this proof exclusively to Gauss, who merely added the finishing touches.*
—Georg Frobenius, 1907
Quoted on the MacTutor History of Mathematics archive, http://www-history.mcs.st-and.ac.uk/history/

where $a_0, a_1, \ldots, a_n$ are constants ($a_n \neq 0$), called the ***coefficients*** of $p$. With the convention that $x^0 = 1$, we can use summation notation to write $p$ as

$$p(x) = \sum_{k=0}^{n} a_kx^k$$

The integer $n$ is called the ***degree*** of $p$, which is denoted by writing $\deg p = n$. A polynomial of degree zero is called a ***constant polynomial.***

### Example D.1

Which of the following are polynomials?

(a) $2 - \frac{1}{3}x + \sqrt{2}x^2$ (b) $2 - \dfrac{1}{3x^2}$ (c) $\sqrt{2x^2}$

(d) $\ln\left(\dfrac{2e^{5x^3}}{e^{3x}}\right)$ (e) $\dfrac{x^2 - 5x + 6}{x - 2}$ (f) $\sqrt{x}$

(g) $\cos(2\cos^{-1}x)$ (h) $e^x$

**Solution** (a) This is the only one that is obviously a polynomial.

(b) A polynomial of the form shown in Equation (1) cannot become infinite as $x$ approaches a finite value [$\lim_{x \to c} p(x) \neq \pm\infty$], whereas $2 - 1/3x^2$ approaches $-\infty$ as $x$ approaches zero. Hence, it is not a polynomial.

(c) We have

$$\sqrt{2x^2} = \sqrt{2}\sqrt{x^2} = \sqrt{2}|x|$$

which is equal to $\sqrt{2}x$ when $x \geq 0$ and to $-\sqrt{2}x$ when $x < 0$. Therefore, this expression is formed by "splicing together" two polynomials (a *piecewise polynomial*), but it is not a polynomial itself.

*Exercises and selected odd-numbered answers for this appendix can be found on the student companion website.

(d) Using properties of exponents and logarithms, we have

$$\ln\left(\frac{2e^{5x^3}}{e^{3x}}\right) = \ln(2e^{5x^3 - 3x}) = \ln 2 + \ln(e^{5x^3 - 3x})$$

$$= \ln 2 + 5x^3 - 3x = \ln 2 - 3x + 5x^3$$

so this expression is a polynomial.

(e) The domain of this function consists of all real numbers $x \neq 2$. For these values of $x$, the function simplifies to

$$\frac{x^2 - 5x + 6}{x - 2} = \frac{(x - 2)(x - 3)}{x - 2} = x - 3$$

so we can say that it is a polynomial *on its domain*.

(f) We see that this function cannot be a polynomial (even on its domain $x \geq 0$), since repeated differentiation of a polynomial of the form shown in Equation (1) eventually results in zero and $\sqrt{x}$ does not have this property. (Verify this.)

(g) The domain of this expression is $-1 \leq x \leq 1$. Let $\theta = \cos^{-1} x$ so that $\cos\theta = x$. Using a trigonometric identity, we see that

$$\cos(2\cos^{-1} x) = \cos 2\theta = 2\cos^2\theta - 1 = 2x^2 - 1$$

so this expression is a polynomial on its domain.

(h) Analyzing this expression as we did the one in (f), we conclude that it is not a polynomial.

Two polynomials are ***equal*** if the coefficients of corresponding powers of $x$ are all equal. In particular, equal polynomials must have the same degree. The ***sum*** of two polynomials is obtained by adding together the coefficients of corresponding powers of $x$.

**Example D.2**

Find the sum of $2 - 4x + x^2$ and $1 + 2x - x^2 + 3x^3$.

**Solution** We compute

$$(2 - 4x + x^2) + (1 + 2x - x^2 + 3x^3) = (2 + 1) + (-4 + 2)x$$
$$+ (1 + (-1))x^2 + (0 + 3)x^3$$
$$= 3 - 2x + 3x^3$$

where we have "padded" the first polynomial by giving it an $x^3$ coefficient of zero.

We define the ***difference*** of two polynomials analogously, subtracting coefficients instead of adding them. The ***product*** of two polynomials is obtained by repeatedly using the distributive law and then gathering together corresponding powers of $x$.

**Example D.3** Find the product of $2 - 4x + x^2$ and $1 + 2x - x^2 + 3x^3$.

**Solution** We obtain

$$\begin{aligned}(2 - 4x + x^2)&(1 + 2x - x^2 + 3x^3)\\ &= 2(1 + 2x - x^2 + 3x^3) - 4x(1 + 2x - x^2 + 3x^3)\\ &\quad + x^2(1 + 2x - x^2 + 3x^3)\\ &= (2 + 4x - 2x^2 + 6x^3) + (-4x - 8x^2 + 4x^3 - 12x^4)\\ &\quad + (x^2 + 2x^3 - x^4 + 3x^5)\\ &= 2 + (4x - 4x) + (-2x^2 - 8x^2 + x^2) + (6x^3 + 4x^3 + 2x^3)\\ &\quad + (-12x^4 - x^4) + 3x^5\\ &= 2 - 9x^2 + 12x^3 - 13x^4 + 3x^5\end{aligned}$$

Observe that for two polynomials $p$ and $q$, we have

$$\deg(pq) = \deg p + \deg q$$

If $p$ and $q$ are polynomials with $\deg q \le \deg p$, we can divide $q$ into $p$, using long division to obtain the quotient $p/q$. The next example illustrates the procedure, which is the same as for long division of one integer into another. Just as the quotient of two integers is not, in general, an integer, the quotient of two polynomials is not, in general, another polynomial.

**Example D.4** Compute $\dfrac{1 + 2x - x^2 + 3x^3}{2 - 4x + x^2}$.

**Solution** We will perform long division. It is helpful to write each polynomial with *decreasing* powers of $x$. Accordingly, we have

$$x^2 - 4x + 2\,\overline{)\,3x^3 - x^2 + 2x + 1}$$

We begin by dividing $x^2$ into $3x^3$ to obtain the partial quotient $3x$. We then multiply $3x$ by the divisor $x^2 - 4x + 2$, subtract the result, and bring down the next term from the dividend $(3x^3 - x^2 + 2x + 1)$:

$$\begin{array}{r} 3x\phantom{ + 2x + 1} \\ x^2 - 4x + 2\,\overline{)\,3x^3 - \phantom{1}x^2 + 2x + 1} \\ \underline{3x^3 - 12x^2 + 6x\phantom{ + 1}} \\ 11x^2 - 4x + 1 \end{array}$$

Then we repeat the process with $11x^2$, multiplying 11 by $x^2 - 4x + 2$ and subtracting the result from $11x^2 - 4x + 1$. We obtain

$$\begin{array}{r} 3x + 11 \\ x^2 - 4x + 2\,\overline{)\,3x^3 - \phantom{1}x^2 + \phantom{4}2x + \phantom{2}1} \\ \underline{3x^3 - 12x^2 + \phantom{4}6x\phantom{ + 21}} \\ 11x^2 - \phantom{4}4x + \phantom{2}1 \\ \underline{11x^2 - 44x + 22} \\ 40x - 21 \end{array}$$

We now have a remainder $40x - 21$. Its degree is less than that of the divisor $x^2 - 4x + 2$, so the process stops, and we have found that

$$3x^3 - x^2 + 2x + 1 = (x^2 - 4x + 2)(3x + 11) + (40x - 21)$$

or

$$\frac{3x^3 - x^2 + 2x + 1}{x^2 - 4x + 2} = 3x + 11 + \frac{40x - 21}{x^2 - 4x + 2}$$

Example D.4 can be generalized to give the following result, known as the ***division algorithm.***

**Theorem D.1** **The Division Algorithm**

If $f$ and $g$ are polynomials with $\deg g \leq \deg f$, then there are polynomials $q$ and $r$ such that

$$f(x) = g(x)q(x) + r(x)$$

where either $r = 0$ or $\deg r < \deg g$.

In Example D.4,

$$f(x) = 3x^3 - x^2 + 2x + 1, \quad g(x) = x^2 - 4x + 2, \quad q(x) = 3x + 11,$$

$$\text{and} \quad r(x) = 40x - 21$$

In the division algorithm, if the remainder is zero, then

$$f(x) = g(x)q(x)$$

and we say that $g$ is a ***factor*** of $f$. (Notice that $q$ is also a factor of $f$.) There is a close connection between the factors of a polynomial and its zeros. A ***zero*** of a polynomial $f$ is a number $a$ such that $f(a) = 0$. [The number $a$ is also called a ***root*** of the polynomial equation $f(x) = 0$.] The following result, known as the ***Factor Theorem,*** establishes the connection between factors of a polynomial and its zeros.

**Theorem D.2** **The Factor Theorem**

Let $f$ be a polynomial and let $a$ be a constant. Then $a$ is a zero of $f$ if and only if $x - a$ is a factor of $f(x)$.

**Proof** By the division algorithm,

$$f(x) = (x - a)q(x) + r(x)$$

where either $r(x) = 0$ or $\deg r < \deg(x - a) = 1$. Thus, in either case, $r(x) = r$ is a constant. Now,

$$f(a) = (a - a)q(a) + r = r$$

so $f(a) = 0$ if and only if $r = 0$, which is equivalent to

$$f(x) = (x - a)q(x)$$

as we needed to prove.

There is no method that is guaranteed to find the zeros of a given polynomial. However, there are some guidelines that are useful in special cases. The case of a polynomial with *integer* coefficients is particularly interesting. The following result, known as the ***Rational Roots Theorem,*** gives criteria for a zero of such a polynomial to be a *rational* number.

**Theorem D.3** **The Rational Roots Theorem**

Let

$$f(x) = a_0 + a_1x + \cdots + a_nx^n$$

be a polynomial with integer coefficients and let $a/b$ be a rational number written in lowest terms. If $a/b$ is a zero of $f$, then $a_0$ is a multiple of $a$ and $a_n$ is a multiple of $b$.

**Proof** If $a/b$ is a zero of $f$, then

$$a_0 + a_1\left(\frac{a}{b}\right) + \cdots + a_{n-1}\left(\frac{a}{b}\right)^{n-1} + a_n\left(\frac{a}{b}\right)^n = 0$$

Multiplying through by $b^n$, we have

$$a_0b^n + a_1ab^{n-1} + \cdots + a_{n-1}a^{n-1}b + a_na = 0 \qquad (1)$$

which implies that

$$a_0b^n + a_1ab^{n-1} + \cdots + a_{n-1}a^{n-1}b = -a_na^n \qquad (2)$$

The left-hand side of Equation (2) is a multiple of $b$, so $a_na^n$ must be a multiple of $b$ also. Since $a/b$ is in lowest terms, $a$ and $b$ have no common factors greater than 1. Therefore, $a_n$ must be a multiple of $b$.

We can also write Equation (1) as

$$-a_0b^n = a_1ab^{n-1} + \cdots + a_{n-1}a^{n-1}b + a_na^n$$

and a similar argument shows that $a_0$ must be a multiple of $a$. (Show this.)

**Example D.5**

Find all the rational roots of the equation

$$6x^3 + 13x^2 - 4 = 0 \qquad (3)$$

**Solution** If $a/b$ is a root of this equation, then 6 is a multiple of $b$ and $-4$ is a multiple of $a$, by the Rational Roots Theorem. Therefore,

$$a \in \{\pm1, \pm2, \pm4\} \quad \text{and} \quad b \in \{\pm1, \pm2, \pm3, \pm6\}$$

Forming all possible rational numbers $a/b$ with these choices of $a$ and $b$, we see that the only possible rational roots of the given equation are

$$\pm 1, \pm 2, \pm 4, \pm \tfrac{1}{2}, \pm \tfrac{1}{3}, \pm \tfrac{2}{3}, \pm \tfrac{4}{3}, \pm \tfrac{1}{6}$$

Substituting these values into Equation (3) one at a time, we find that $-2$, $-\frac{2}{3}$, and $\frac{1}{2}$ are the only values from this list that are actually roots. (Check these.) As we will see shortly, a polynomial equation of degree 3 cannot have more than three roots, so these are not only all the *rational* roots of Equation (3) but also its *only* roots.

We can improve on the trial-and-error method of Example D.5 in various ways. For example, once we find one root $a$ of a given polynomial equation $f(x) = 0$, we know that $x - a$ is a factor of $f(x)$—say, $f(x) = (x - a)g(x)$. We can therefore divide $f(x)$ by $x - a$ (using long division) to find $g(x)$. Since $\deg g < \deg f$, the roots of $g(x) = 0$ [which are also roots of $f(x) = 0$] may be easier to find. In particular, if $g(x)$ is a quadratic polynomial, we have access to the ***quadratic formula.***

Suppose

$$ax^2 + bx + c = 0$$

(We may assume that $a$ is positive, since multiplying both sides by $-1$ would produce an equivalent equation otherwise.) Then, completing the square, we have

$$a\left(x^2 + \frac{b}{a}x + \frac{b^2}{4a^2}\right) = \frac{b^2}{4a} - c$$

(Verify this.) Equivalently,

$$a\left(x + \frac{b}{2a}\right)^2 = \frac{b^2}{4a} - c \quad \text{or} \quad \left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}$$

Therefore,

$$x + \frac{b}{2a} = \pm\sqrt{\frac{b^2 - 4ac}{4a^2}} = \frac{\pm\sqrt{b^2 - 4ac}}{2a}$$

or

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Let's revisit the equation from Example D.5 with the quadratic formula in mind.

## Example D.6

Find the roots of $6x^3 + 13x^2 - 4 = 0$.

**Solution** Let's suppose we use the Rational Roots Theorem to discover that $x = -2$ is a rational root of $6x^3 + 13x^2 - 4 = 0$. Then $x + 2$ is a factor of $6x^3 + 13x^2 - 4$, and long division gives

$$6x^3 + 13x^2 - 4 = (x + 2)(6x^2 + x - 2)$$

(Check this.) We can now apply the quadratic formula to the second factor to find that its zeros are

$$\begin{aligned} x &= \frac{-1 \pm \sqrt{1^2 - 4(6)(-2)}}{2 \cdot 6} \\ &= \frac{-1 \pm \sqrt{49}}{12} = \frac{-1 \pm 7}{12} \\ &= \tfrac{6}{12}, -\tfrac{8}{12} \end{aligned}$$

or, in lowest terms, $\frac{1}{2}$ and $-\frac{2}{3}$. Thus, the three roots of Equation (3) are $-2$, $\frac{1}{2}$, and $-\frac{2}{3}$, as we determined in Example D.5.

**Remark** The Factor Theorem establishes a connection between the zeros of a polynomial and its *linear* factors. However, a polynomial without linear factors may still have factors of higher degree. Furthermore, when asked to factor a polynomial, we need to know the number system to which the coefficients of the factors are supposed to belong.

For example, consider the polynomial

$$p(x) = x^4 + 1$$

Over the *rational numbers* $\mathbb{Q}$, the only possible zeros of $p$ are 1 and $-1$, by the Rational Roots Theorem. A quick check shows that neither of these actually works, so $p(x)$ has no *linear* factors with rational coefficients, by the Factor Theorem. However, $p(x)$ may still factor into a product of two *quadratics*. We will check for quadratic factors using the method of ***undetermined coefficients.***

Suppose that

$$x^4 + 1 = (x^2 + ax + b)(x^2 + cx + d)$$

Expanding the right-hand side and comparing coefficients, we obtain the equations

$$\begin{aligned} a + c &= 0 \\ b + ac + d &= 0 \\ bc + ad &= 0 \\ bd &= 1 \end{aligned}$$

If $a = 0$, then $c = 0$ and $d = -b$. This gives $-b^2 = 1$, which has no solutions in $\mathbb{Q}$. Hence, we may assume that $a \neq 0$. Then $c = -a$, and we obtain $d = b$. It now follows that $b^2 = 1$, so $b = 1$ or $b = -1$. This implies that $a^2 = 2$ or $a^2 = -2$, respectively, neither of which has solutions in $\mathbb{Q}$. It follows that $x^4 + 1$ cannot be factored over $\mathbb{Q}$. We say that it is ***irreducible*** over $\mathbb{Q}$.

However, over the *real numbers* $\mathbb{R}$, $x^4 + 1$ does factor. The calculations we have just done show that

$$x^4 + 1 = (x^2 + \sqrt{2}x + 1)(x^2 - \sqrt{2}x + 1)$$

(Why?) To see whether we can factor further, we apply the quadratic formula. We see that the first factor has zeros

$$x = \frac{-\sqrt{2} \pm \sqrt{(\sqrt{2})^2 - 4}}{2} = \frac{-\sqrt{2} \pm \sqrt{-2}}{2} = \frac{\sqrt{2}}{2}(-1 \pm i) = -\frac{1}{\sqrt{2}} \pm \frac{1}{\sqrt{2}}i$$

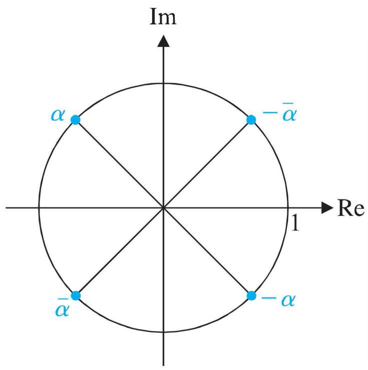



**Figure D.1**

which are in $\mathbb{C}$ but not in $\mathbb{R}$. Hence, $x^2 + \sqrt{2}x + 1$ cannot be factored into linear factors over $\mathbb{R}$. Similarly, $x^2 - \sqrt{2}x + 1$ cannot be factored into linear factors over $\mathbb{R}$.

Our calculations show that a complete factorization of $x^4 + 1$ is possible over the *complex numbers* $\mathbb{C}$. The four zeros of $x^4 + 1$ are

$$\alpha = -\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i, \quad \bar{\alpha} = -\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}i, \quad -\bar{\alpha} = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i,$$
$$-\alpha = \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}i$$

which, as Figure D.1 shows, all lie on the unit circle in the complex plane. Thus, the factorization of $x^4 + 1$ is

$$x^4 + 1 = (x - \alpha)(x - \bar{\alpha})(x + \bar{\alpha})(x + \alpha)$$

The preceding Remark illustrates several important properties of polynomials. Notice that the polynomial $p(x) = x^4 + 1$ satisfies $\deg p = 4$ and has exactly four zeros in $\mathbb{C}$. Furthermore, its complex zeros occur in *conjugate pairs*; that is, its complex zeros can be paired up as

$$\{\alpha, \bar{\alpha}\} \quad \text{and} \quad \{-\alpha, -\bar{\alpha}\}$$

These last two facts are true in general. The first is an instance of the ***Fundamental Theorem of Algebra (FTA)***, a result that was first proved by Gauss in 1797.

**Theorem D.4** **The Fundamental Theorem of Algebra**

Every polynomial of degree $n$ with real or complex coefficients has exactly $n$ zeros (counting multiplicities) in $\mathbb{C}$.

This important theorem is sometimes stated as

"Every polynomial with real or complex coefficients has a zero in $\mathbb{C}$."

Let's call this statement FTA′. Certainly, FTA implies FTA′. Conversely, if FTA′ is true, then if we have a polynomial $p$ of degree $n$, it has a zero $\alpha$ in $\mathbb{C}$. The Factor Theorem then tells us that $x - \alpha$ is a factor of $p(x)$, so

$$p(x) = (x - \alpha)q(x)$$

where $q$ is a polynomial of degree $n - 1$ (also with real or complex coefficients). We can now apply FTA′ to $q$ to get another zero, and so on, making FTA true. This argument can be made into a nice induction proof. (Try it.)

It is not possible to give a formula (along the lines of the quadratic formula) for the zeros of polynomials of degree 5 or more. (The work of Abel and Galois confirmed this; see page 311.) Consequently, other methods must be used to prove FTA. The proof that Gauss gave uses topological methods and can be found in more advanced mathematics courses.

Now suppose that

$$p(x) = a_0 + a_1x + \cdots + a_nx^n$$

is a polynomial with real coefficients. Let $\alpha$ be a complex zero of $p$ so that

$$a_0 + a_1\alpha + \cdots + a_n\alpha^n = p(\alpha) = 0$$

Then, using properties of conjugates, we have

$$\begin{aligned} p(\overline{\alpha}) &= a_0 + a_1\overline{\alpha} + \cdots + a_n\overline{\alpha}^n = \overline{a}_0 + \overline{a}_1\overline{\alpha} + \cdots + \overline{a}_n\overline{\alpha}^n \\ &= \overline{a_0 + a_1\alpha + \cdots + a_n\alpha^n} \\ &= \overline{p(\alpha)} = \overline{0} = 0 \end{aligned}$$

Thus, $\overline{\alpha}$ is also a zero of $p$. This proves the following result:

The complex zeros of a polynomial with real coefficients occur in conjugate pairs.

In some situations, we do not need to know *what* the zeros of a polynomial are—we only need to know *where* they are located. For example, we might only need to know whether the zeros are positive or negative (as in Theorem 4.35). One theorem that is useful in this regard is ***Descartes' Rule of Signs.*** It allows us to make certain predictions about the number of positive zeros of a polynomial with real coefficients based on the signs of these coefficients.

Descartes' stated this rule in his 1637 book *La Géometrie,* but did not give a proof. Several mathematicians later furnished a proof, and Gauss provided a somewhat sharper version of the theorem in 1828.

Given a polynomial $a_0 + a_1x + \cdots + a_nx^n$, write its nonzero coefficients in order. Replace each positive coefficient by a plus sign and each negative coefficient by a minus sign. We will say that the polynomial has $k$ ***sign changes*** if there are $k$ places where the coefficients change sign. For example, the polynomial $2 - 3x + 4x^3 + x^4 - 7x^5$ has the sign pattern

$$\underbrace{+\ -}\underbrace{\ +}\ \underbrace{+\ -}$$

so it has three sign changes, as indicated.

**Theorem D.5**

**Descartes' Rule of Signs**

Let $p$ be a polynomial with real coefficients that has $k$ sign changes. Then the number of positive zeros of $p$ (counting multiplicities) is at most $k$.

In words, Descartes' Rule of Signs says that a real polynomial cannot have more positive zeros than it has sign changes.

**Example D.7**

Show that the polynomial $p(x) = 4 + 2x^2 - 7x^4$ has exactly one positive zero.

**Solution** The coefficients of $p$ have the sign pattern $+ + -$, which has only one sign change. So, by Descartes' Rule of Signs, $p$ has at most one positive zero. But $p(0) = 4$ and $p(1) = -1$, so there is a zero somewhere in the interval $(0, 1)$. Hence, this is the only positive zero of $p$.

We can also use Descartes' Rule of Signs to give a bound on the number of *negative* zeros of a polynomial with real coefficients. Let

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

and let $b$ be a negative zero of $p$. Then $b = -c$ for $c > 0$, and we have

$$\begin{aligned} 0 = p(b) &= a_0 + a_1b + a_2b^2 + \cdots + a_nb^n \\ &= a_0 - a_1c + a_2c^2 - + \cdots + (-1)^n a_nc^n \end{aligned}$$

But $$p(-x) = a_0 - a_1x + a_2x^2 - + \cdots + (-1)^n a_nx^n$$

so $c$ is a positive zero of $p(-x)$. Therefore, $p(x)$ has exactly as many negative zeros as $p(-x)$ has positive zeros. Combined with Descartes' Rule of Signs, this observation yields the following:

Let $p$ be a polynomial with real coefficients. Then the number of negative zeros of $p$ is at most the number of sign changes of $p(-x)$.

## Example D.8

Show that the zeros of $p(x) = 1 + 3x + 2x^2 + x^5$ cannot all be real.

**Solution** The coefficients of $p(x)$ have no sign changes, so $p$ has no positive zeros. Since $p(-x) = 1 - 3x + 2x^2 - x^5$ has three sign changes among its coefficients, $p$ has at most three negative zeros. We note that 0 is not a zero of $p$ either, so $p$ has at most three real zeros. Therefore, it has at least two complex (nonreal) zeros.