Chapter 1

Integers and Permutations

God made the integers, and all the rest is the work of man.

—Leopold Kronecker

The use of arithmetic is a basic aspect of human culture. Anthropologists tell us that even the most primitive societies, because of their desire to count objects, have developed some sort of terminology for the numbers $1, 2,$ and 3, although many go no further. As a culture develops, it needs more sophisticated counting to deal with commerce, warfare, the calendar, and so on. This leads to methods of recording numbers often (but by no means always) based on groups of 10, presumably from counting on the fingers. Then the recording of numbers by making marks or notches becomes important (in bookkeeping, for example), and a variety of systems have been constructed for doing so. Many of these systems were not very useful for adding or multiplying (try multiplying with Roman numerals), and the development of our positional system, originating with the Babylonians using base 60 rather than 10, was a great advance.

In this chapter we assume the validity of the elementary arithmetic properties of the integers and use them to derive some more subtle facts related to divisibility and primes. Then two fundamental algebraic systems are described: the integers modulo n and the permutations of the set $\{1, 2, \ldots, n\}$. These are, respectively, excellent examples of *rings* and *groups,* two of the basic algebraic structures presented in detail in Chapters 2 and 3.

Introduction to Abstract Algebra, Fourth Edition. W. Keith Nicholson.

1.1 INDUCTION

> Great fleas have little fleas upon their backs to bite 'em, And little fleas have lesser fleas, and so ad infinitum.
>
> —Augustus De Morgan

Consider the sequence of equations:

$$\begin{aligned} 1 &= 1 \\ 1+3 &= 4 \\ 1+3+5 &= 9 \\ 1+3+5+7 &= 16 \\ &\vdots \end{aligned}$$

It is clear there is a pattern. The right sides are the squares $1^2, 2^2, 3^2, 4^2, \dots$, and, when the right side is n^2, the left side is the sum of the first n odd integers. As the nth odd integer is $2n-1$, the following expression is true for $n = 1, 2, 3$, and 4:

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

Now it is almost irresistible to ask whether the statement (p_n) is true for *every* $n \geq 1$. There is no hope of separately verifying all these statements, because there are infinitely many of them. A more subtle approach is required.

The idea is to prove that $p_k \Rightarrow p_{k+1}$ for every $k \geq 1$. Then the fact that p_1 is true implies that p_2 is true, which in turn implies that p_3 is true, then p_4, and so on. This is one of the most important axioms for the integers.

Principle of Mathematical Induction[6]. *Let p_n be a statement for each integer $n \geq 1$. Suppose that the following conditions are satisfied*:

(1) *p_1 is true.*

(2) *$p_k \Rightarrow p_{k+1}$ for every $k \geq 1$.*

Then p_n is true for every $n \geq 1$.

In the proof that $p_k \Rightarrow p_{k+1}$, we assume that p_k is true and use it to prove that p_{k+1} is also true. The assumption that p_k is true is called the **induction hypothesis**.

For a graphic illustration, consider an infinite row of dominoes labeled $1, 2, 3, \dots$ standing so that if one is knocked over, it will knock the next one over. If p_k is the statement that domino k falls over, this means that $p_k \Rightarrow p_{k+1}$ for each $k \geq 1$. The principle of induction asserts that knocking domino 1 over causes them all to fall.

As another illustration, let p_n be the statement $1+3+5+\cdots+(2n-1) = n^2$ mentioned above. Then p_1 has already been verified. To prove that $p_k \Rightarrow p_{k+1}$ for each $k \geq 1$, we assume that p_k is true (the induction hypothesis) and use it to simplify the left side of the sum p_{k+1}:

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

[6] One of the earliest uses of the principle is in the work of Francesco Maurolico in the 16th century. Augustus De Morgan coined the name *mathematical induction* in 1838.

This expression shows that p_{k+1} is true and hence, by the induction principle, that p_n is true for all $n \geq 1$.

Example 1. Prove **Gauss' Formula**[7]**:** $1 + 2 + \cdots + n = \frac{1}{2}n(n+1)$ for all $n \geq 1$.

Solution. Let p_n denote the statement $1 + 2 + \cdots + n = \frac{1}{2}n(n+1)$. Then p_1 is true because $1 = \frac{1}{2}(1+1)$. If we assume that p_k is true for some $k \geq 1$, we get

$$1 + 2 + 3 + \cdots + k + (k+1) = \tfrac{1}{2}k(k+1) + (k+1) = \tfrac{1}{2}(k+1)(k+2),$$

which shows that p_{k+1} is true. Hence, p_n is true for all $n \geq 1$ by the principle of mathematical induction. □

Example 2 gives an inductive proof of a useful formula for the sum of a geometric series $1 + x + \cdots + x^n$. We use the convention that $x^0 = 1$ for all numbers x.

Example 2. If x is any real number, show that

$$(1-x)(1 + x + \cdots + x^{n-1}) = 1 - x^n, \quad \text{for all } n \geq 1.$$

Solution. Let p_n be the given statement. Then p_1 is $(1-x)1 = 1 - x^1$, which is true. If we assume that p_k is true for some $k \geq 1$, then the left side of p_{k+1} becomes

$$\begin{aligned}(1-x)(1 + x + \cdots + x^{k-1} + x^k) &= (1-x)(1 + x + \cdots + x^{k-1}) + (1-x)x^k \\ &= (1 - x^k) + (1-x)x^k \\ &= 1 - x^{k+1}.\end{aligned}$$

This proves that p_{k+1} is true and so completes the induction. □

Example 3. Let w_n denote the number of n-letter words that can be formed using only the letters a and b. Show that $w_n = 2^n$ for all $n \geq 1$.

Solution. Clearly, a and b are the only such words with one letter, so $w_1 = 2 = 2^1$. If $k \geq 1$, we obtain each such word of $k+1$ letters by adjoining an a or a b to a word of k letters, and there are w_k of each type. Hence, $w_{k+1} = 2w_k$ for each $k \geq 1$ so, if we assume inductively that $w_k = 2^k$, we get $w_{k+1} = 2w_k = 2 \cdot 2^k = 2^{k+1}$, as required. □

The principle of induction starts at 1 in the sense that if p_1 is true and $p_k \Rightarrow p_{k+1}$ for all $k \geq 1$, then p_k is true for all $k \geq 1$. There is nothing special about 1.

Theorem 1. *If m is any integer, let $p_m, p_{m+1}, p_{m+2}, \ldots$ be statements such that*

(1) *p_m is true.*

(2) *$p_k \Rightarrow p_{k+1}$ for every $k \geq m$.*

Then p_n is true for each $n \geq m$.

[7] This formula was probably known to the ancient Greeks. However, the great mathematician Carl Friedrich Gauss is said to have derived a special case of the formula ($n = 100$) at age 7 by writing the sum $1 + 2 + \cdots + 100$ in two parts:

$$\begin{array}{ccccccccc} 1 & + & 2 & + \cdots + & 49 & + & 50 \\ 100 & + & 99 & + \cdots + & 52 & + & 51 \end{array}$$

and observing that each pair of terms, $1 + 100, 2 + 99, \ldots, 50 + 51$, adds to 101. As there are 50 such pairs, the sum is $50 \cdot 101 = 5050$.

Proof. Let $t_n = p_{m+n-1}$ for each $n \geq 1$. Then $t_1 = p_m$ is true, and $t_k \Rightarrow t_{k+1}$ because $p_{m+k-1} \Rightarrow p_{m+k}$. Hence, t_n is true for all $n \geq 1$ by induction; that is, p_n is true for all $n \geq m$. ■

Example 4. If $n \geq 8$, show that any postage of n cents can be made exactly using only 3- and 5 cent stamps.

Solution. The assertion clearly holds if $n = 8$. If it holds for some $k \geq 8$, we consider two cases:

Case 1. One or more 5 cent stamps are used to make up k cents postage.
Then replace one of them with two 3 cent stamps.

Case 2. Three or more 3 cent stamps are used to make up k cents postage.
Then replace three of them with two 5 cent stamps.

Because one of these cases must occur (as $k \geq 8$), the assertion holds for $k+1$ cents in both cases and the induction goes through. □

If $n \geq 1$ is an integer, the integer $n!$ (read n-**factorial**) is defined to be the product

$$n! = n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1$$

of all the integers from n to 1. Thus, $1! = 1$, $2! = 2$, $3! = 6$, and so on. Clearly,

$$(n+1)! = (n+1)n!, \quad \text{for each } n \geq 1,$$

which we extend to $n = 0$ by defining

$$0! = 1.$$

Example 5. Show that $2^n < n!$ for all $n \geq 4$.

Solution. If p_k is the statement $2^k < k!$, note that p_1, p_2, and p_3 are actually false, but p_4 is true because $2^4 = 16 < 24 = 4!$. If p_k is true where $k \geq 4$, then $2^k < k!$ so

$$2^{k+1} = 2 \cdot 2^k < 2 \cdot k! < (k+1)k! = (k+1)!$$

Hence, p_{k+1} is true and the induction is complete. □

Let n and r be integers with $0 < r \leq n$. The **binomial coefficient** $\binom{n}{r}$ is defined as follows:

$$\binom{n}{r} = \tfrac{n!}{r!(n-r)!}.$$

As $0! = 1$, we have $\binom{n}{0} = 1 = \binom{n}{n}$ and $\binom{n}{2} = \frac{n(n-1)}{2}$. It is easy to verify that

$$\binom{n}{r} = \binom{n}{n-r}, \quad \text{whenever } 0 \leq r \leq n.$$

We leave the proof of the following formula (the **Pascal identity**) as Exercise 13.

$$\binom{n}{r-1} + \binom{n}{r} = \binom{n+1}{r}, \quad \text{whenever } 1 \leq r \leq n.$$

The name honors Blaise Pascal. The identity leads to a way of displaying the binomial coefficients known as **Pascal's triangle**:

$$\begin{array}{ccccccccc} & & & & 1 & & & & \\ & & & 1 & & 1 & & & \\ & & 1 & & 2 & & 1 & & \\ & 1 & & 3 & & 3 & & 1 & \\ 1 & & 4 & & 6 & & 4 & & 1 \\ & & & & \vdots & & & & \end{array}$$

The n^{th} row of the triangle is $\binom{n}{0}\ \binom{n}{1}\ \binom{n}{2} \cdots \binom{n}{n-1}\ \binom{n}{n}$, starting at $n = 0$. The Pascal identity shows that each entry in a given row (except at the ends) can be found by adding the two entries adjacent to it in the row above. Hence, Pascal's triangle is easy to write down row by row.[8]

The entries in each row also arise in another way. The formulas

$$\begin{aligned} (1+x)^2 &= 1 + 2x + x^2, \\ (1+x)^3 &= 1 + 3x + 3x^2 + x^3, \\ (1+x)^4 &= 1 + 4x + 6x^2 + 4x^3 + x^4, \end{aligned}$$

are easily verified, and the coefficients on the right side in each case are the integers in rows $2, 3$, and 4 of Pascal's triangle. The general result follows by induction, and will be used several times in this book.

Example 6. Prove the **Binomial Theorem**:

$$(1+x)^n = \binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots + \binom{n}{n}x^n, \quad \text{for all } n \geq 0.$$

Solution. The theorem holds if $n = 0$ because $\binom{0}{0} = 1$ and $(1+x)^0 = 1$. If it holds for some $k \geq 0$ then, using the Pascal identity, we obtain

$$\begin{aligned} (1+x)^{k+1} &= (1+x)(1+x)^k \\ &= (1+x)\left[\binom{k}{0} + \binom{k}{1}x + \cdots + \binom{k}{k-1}x^{k-1} + \binom{k}{k}x^k\right] \\ &= \binom{k}{0} + \left[\binom{k}{0} + \binom{k}{1}\right]x + \cdots + \left[\binom{k}{k-1} + \binom{k}{k}\right]x^k + \binom{k}{k}x^{k+1} \\ &= \binom{k+1}{0} + \binom{k+1}{1}x + \cdots + \binom{k+1}{k}x^k + \binom{k+1}{k+1}x^{k+1}, \end{aligned}$$

which completes the induction. □

When proving inductively that statements $p_m, p_{m+1}, \ldots, p_k$ are true, the most difficult part is usually showing that $p_k \Rightarrow p_{k+1}$ for each $k \geq m$. Clearly, this task would be easier if we could assume the truth of $p_m, \ldots, p_{k-1}$ in addition to the truth of p_k when deducing p_{k+1}. This assumption leads to a useful variant of the principle of induction (in fact, it is equivalent to it).

[8] Note that this shows the binomial coefficients are all *integers*, a fact that is not clear from the definition.

Theorem 2. ***Principle of Strong Induction***. *Let m be an integer and, for each $n \geq m$, let p_n be a statement. Suppose the following conditions are satisfied.*

(1) *p_m is true.*

(2) *If $k \geq m$ and all of $p_m, p_{m+1}, \ldots, p_k$ are true, then p_{k+1} is also true.*

Then p_n is true for every $n \geq m$.

Proof. For each $n \geq m$, let t_n be the statement that $p_m, p_{m+1}, \ldots, p_n$ are all true. Then, t_m is true by (1). If t_k is true for some $k \geq m$, then (2) implies that p_{k+1} is true, so t_{k+1} is also true. Hence, t_n is true for all $n \geq m$ by Theorem 1, so certainly p_n is true for all $n \geq m$. ■

In the next example, we use strong induction to prove an important fact about primes that would be more difficult to deduce using (ordinary) induction. Recall that a *prime number* (or *prime*) is an integer $p \geq 2$ that cannot be factored as a product of two smaller positive integers.

Example 7. Show that every integer $n \geq 2$ is a product of (one or more) primes.

Solution. This assertion is true if $n = 2$ because 2 is a prime. If $k \geq 2$, we assume inductively that $2, 3, \ldots, k$ are all products of primes. To apply strong induction, we must show that $k + 1$ is a product of primes. This is clear if $k + 1$ is itself prime; otherwise, let $k + 1 = ab$, where $2 \leq a \leq k$ and $2 \leq b \leq k$. Then both a and b are products of primes by the (strong) induction hypothesis, so $k + 1 = ab$ is also a product of primes. □

We conclude with an intuitively clear property of $\mathbb{Z}$ that is equivalent to the principle of induction, and which is usually taken as an axiom.

Well-Ordering Principle. *Every nonempty set of nonnegative integers has a smallest member.*

Proof. If the principle is false, let $X \subseteq \{0, 1, 2, \ldots\}$ be a nonempty set that has no smallest member. For each $n \geq 0$, let p_n be the statement "$n \notin X$." It suffices to show that p_n is true for all $n \geq 0$—since then X is empty, contrary to our assumption. We prove this by strong induction. First, p_0 is true because if $0 \in X$, then it is the smallest member of X (because $X \subseteq \{0, 1, 2, \ldots\}$). Now assume inductively that p_0, $p_1, \ldots, p_k$ are all true, so that none of $0, 1, \ldots, k$ is in X. This implies that $k + 1 \notin X$ since otherwise it would be the smallest member of X. This means p_{k+1} is true, and so completes the induction. □

The way the well-ordering principle is used can be illustrated by the following frivolous example: Suppose that we want to show that every positive integer is interesting. If this assertion were false, the set of uninteresting positive integers would be nonempty and so would contain a smallest member by the axiom. But the smallest uninteresting integer would surely be interesting—a contradiction! This technique can also be applied to *serious* situations.

For example, the well-ordering principle implies the induction principle. Indeed, let p_1, p_2, $p_3, \ldots$ be statements such that p_1 is true and $p_k \Rightarrow p_{k+1}$ for every $k \geq 1$. If $X = \{n \geq 1 \mid p_n \text{ is false}\}$, we must show that X is empty. But if not, then X has a smallest member, which leads to a contradiction. The details are in Exercise 15.

We have proved the following implications (the first is Theorem 2):

$$\text{Induction} \quad \Rightarrow \quad \text{Strong Induction} \quad \Rightarrow \quad \text{Well Ordering.}$$

Moreover, well ordering implies induction (see above), so the three principles are logically equivalent. The validity of these principles is one of the basic **Peano axioms**[9] for the integers.

Inductive Definition

Many arguments in algebra (in fact, in mathematics generally) refer to **sequences** $a_0,\ a_1,\ a_2,\ a_3, \cdots, a_n, \cdots$ from a set A where each a_i is an element of A called the i^{th} *term* of the sequence. Hence $1,\ 2,\ 4,\ 8,\ 16, \dots$ are the first five terms of the sequence $a_n = 2^n$ from $\mathbb{Z}$. This sequence can be compactly described as follows:

$$a_0 = 1 \quad \text{and} \quad a_n = 2a_{n-1} \quad \text{for each } n \geq 1. \tag{*}$$

These conditions uniquely describe the sequence (the formula $a_n = 2^n$ for $n \geq 0$ can be proved by induction), and for this reason (*) is called an *inductive definition* of the sequence. More generally, a sequence is said to be **defined inductively** if the first term is specified and each later term is uniquely determined by the earlier terms (often by a formula). It is usually very difficult to give an explicit formula for the n^{th} term a_n in terms of the earlier terms; nevertheless, the following theorem shows that such a sequence always exists and is uniquely determined.

Theorem 3. ***Recursion Theorem***. *Given a set A and $a \in A$, there is exactly one sequence $a_0,\ a_1,\ a_2,\ a_3, \dots, a_n, \dots$ from A that satisfies the following requirements*:

(1) $a_0 = a$.

(2) *For each $n \geq 1$, the term a_n is uniquely determined by the preceding terms $a_0,\ a_1,\ a_2, \dots, a_{n-1}$.*

Proof. The existence of such a sequence is given in Appendix D; we prove uniqueness by strong induction on $n \geq 0$. Clearly, a_0 is uniquely determined by (1). If each of $a_0,\ a_1,\ a_2, \dots a_{n-1}$ has been uniquely specified, then a_n is uniquely determined by (2). Hence, the sequence is uniquely determined by (1) and (2). ■

Exercises 1.1

1. Prove each equation by induction on n.

(a) $1 + 5 + 9 + \cdots + (4n - 3) = n(2n - 1)$ for all $n \geq 1$.

(b) $1^2 + 2^2 + \cdots + n^2 = \frac{1}{6}n(n + 1)(2n + 1)$ for all $n \geq 1$.

(c) $1^3 + 2^3 + \cdots + n^3 = \frac{1}{4}n^2(n + 1)^2$ for all $n \geq 1$.

(d) $1 \cdot 2 + 2 \cdot 3 + \cdots + n \cdot (n + 1) = \frac{1}{3}n(n + 1)(n + 2)$ for all $n \geq 1$.

(e) $1 \cdot 2^2 + 2 \cdot 3^2 + \cdots + n \cdot (n + 1)^2 = \frac{1}{12}n(n + 1)(n + 2)(3n + 5)$ for all $n \geq 1$.

(f) $\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \cdots + \frac{1}{n\cdot(n+1)} = \frac{n}{n+1}$ for all $n \geq 1$.

(g) $1^2 + 3^2 + \cdots + (2n - 1)^2 = \frac{n}{3}(4n^2 - 1)$ for all $n \geq 1$.

[9] Named after Giuseppe Peano, an Italian mathematician and logician who, in 1889, reduced the theory of the natural numbers $\mathbb{N}$ to five simple axioms. For a discussion of this, see R.A. Beaumont and R.S. Pierce, *The Algebraic Foundations of Mathematics,* Addison-Wesley, 1963.

(h) $1^2 - 2^2 + 3^2 - \cdots + (-1)^{n+1}n^2 = \frac{1}{2}(-1)^{n+1}n(n+1)$ for all $n \geq 1$.

(i) $\frac{1}{2!} + \frac{2}{3!} + \frac{3}{4!} + \cdots + \frac{n}{(n+1)!} = 1 - \frac{1}{(n+1)!}$ for all $n \geq 1$.

2. Prove each inequality by induction on n.

(a) $n < 2^n$ for all $n \geq 0$.

(b) $n^2 \leq 2^n$ for all $n \geq 4$.

(c) $n! \leq 2^{n^2}$ for all $n \geq 4$ (compare with Example 5).

(d) $\frac{1}{1^2} + \frac{1}{2^2} + \cdots + \frac{1}{n^2} \leq 2 - \frac{1}{n}$ for all $n \geq 1$.

(e) $\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \cdots + \frac{1}{\sqrt{n}} \geq \sqrt{n}$ for all $n \geq 1$.

(f) $\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \cdots + \frac{1}{\sqrt{n}} \leq 2\sqrt{n} - 1$ for all $n \geq 1$.

3. Prove each statement by induction on n.

(a) $n^3 + (n+1)^3 + (n+2)^3$ is a multiple of 9 for all $n \geq 1$.

(b) $n^3 - n$ is a multiple of 3 for all $n \geq 1$.

(c) $3^{2n+1} + 2^{n+2}$ is a multiple of 7 for all $n \geq 0$.

4. Show that $\left(1 - \frac{1}{2^2}\right)\left(1 - \frac{1}{3^2}\right)\cdots\left(1 - \frac{1}{n^2}\right) = \frac{n+1}{2n}$ for all $n > 2$.

5. Show that $3^{3n} + 1$ is a multiple of 7 for all odd $n \geq 1$.

6. Suppose that n straight lines in the plane are positioned so that no two are parallel and no three pass through the same point. Show that they divide the plane into $\frac{1}{2}(n^2 + n + 2)$ distinct regions.

7. Show that there are 3^n positive integers with n digits, where each digit must be $4, 5$, or 6.

8. A polygon in the plane is called *convex* if every line joining two vertices is either an edge or lies entirely within the polygon. If $n \geq 3$, show that the sum of the interior angles of an n-sided convex polygon equals $(n-2) \cdot 180°$.

9. A straight line segment joining two distinct points on a circle is called a *secant*. For $n \geq 1$, draw n secants with no two identical. Show that the resulting regions can be unambiguously colored black and white (where *unambiguously* means that no two regions sharing a straight line boundary are of the same color).

10. (a) Show that any postage of $n \geq 2$ cents can be made of 2 and 3 cent stamps.

(b) Show that any postage of $n \geq 12$ cents can be made of 3 and 7 cent stamps.

(c) Show that any postage of $n \geq 18$ cents can be made of 4 and 7 cent stamps.

(d) Can you generalize from the results in (a)–(c)?

11. Let $a_n = 2^{3n} - 1$ for $n \geq 0$. Guess a common divisor of each a_n and prove your assertion.

12. (a) Try to prove the statement "$1^3 + 2^3 + \cdots + n^3$ is a perfect square" by induction. Now look at Exercise 1(c).

(b) Try to prove that $1 + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^n} < 2$ by induction. Now formulate a stronger equality for the sum on the left, prove it by induction, and use it to deduce the inequality.

13. Prove the **Pascal identity**: $\binom{n}{r-1} + \binom{n}{r} = \binom{n+1}{r}$ for $1 \leq r \leq n$.

14. (a) Show that $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n$ for all $n \geq 0$.

(b) Show that $\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \cdots \pm \binom{n}{n} = 0$ if $n > 0$.

15. Use the well-ordering principle to prove the principle of induction. [*Hint*: See the discussion following the well-ordering principle.]

16. Let X be a nonempty set of integers. Then X is said to be *bounded below* (*bounded above*) if an integer m exists such that $m \leq x$ for all $x \in X$ (respectively $m \geq x$ for all $x \in X$).

(a) If X is bounded below, show that it has a smallest member.

(b) If X is bounded above, show that it has a largest member.

17. Use strong induction to prove that every integer $n \geq 2$ has a prime factor.

18. In each case, conjecture a formula for a_n and prove it by induction.

(a) $a_0 = 2, a_{n+1} = -a_n$, $n \geq 0$.

(b) $a_0 = 1, a_1 = -2, a_{n+2} = 2a_n - a_{n+1}$, $n \geq 0$.

(c) $a_0 = 1, a_{n+1} = 1 - a_n$, $n \geq 0$.

(d) $a_0 = 3, a_{n+1} = (a_n)^2$, $n \geq 0$.

19. Let n lines in the plane be such that no two are parallel and no three are concurrent. Find the number a_n of regions into which the plane is divided by first showing that $a_{n+1} = a_n + (n+1)$.

20. Prove the following induction principle.

Let m be an integer and let p_n be a statement for all $n \geq m$. Assume that

(1) p_m and p_{m+1} are true.

(2) If $k \geq m$ and both p_k and p_{k+1} are true, then p_{k+2} is true.

Then p_n is true for all $n \geq m$.

21. Let a_n denote a number for each integer $n \geq 0$ and assume that $a_{n+2} = a_{n+1} + 2a_n$ holds for every $n \geq 0$. Use the principle in Exercise 20 to prove each assertion.

(a) If $a_0 = 1$ and $a_1 = -1$, then $a_n = (-1)^n$ for each $n \geq 0$.

(b) If $a_0 = 1$ and $a_1 = 2$, then $a_n = 2^n$ for each $n \geq 0$.

(c) If $a_0 = p$ and $a_1 = q$, then $a_n = \frac{1}{3}[(p+q)2^n + (2p-q)(-1)^n]$ for each $n \geq 0$.

22. Let p_n denote the statement: "$3n + 2$ is a multiple of 3." Show that $p_k \Rightarrow p_{k+1}$ for all $k \geq 1$. What does this say about Theorem 1?

23. Let p_n denote the statement: "In any class of n algebra students, every student obtains the same grade." Then p_1 is clearly true. If p_n is satisfied for $n > 1$, suppose that $x_1, x_2, \ldots, x_{n+1}$ denotes a class of $n+1$ students. Then $x_1, x_2, \ldots, x_n$ all have the same grade (by induction) as do $x_2, x_3, \ldots, x_{n+1}$. Thus $x_1, x_2, \ldots, x_{n+1}$ all have the same grade (the same as x_n), so p_{n+1} is true. Hence, p_n is true for all n. What is wrong with this argument?

24. Suppose that p_n is a statement about n for each $n \geq 1$. In each case what must be done to prove that p_n is true for all $n \geq 1$?

(a) $p_n \Rightarrow p_{n+2}$ for each $n \geq 1$.

(b) $p_n \Rightarrow p_{n+8}$ for each $n \geq 1$.

(c) $p_n \Rightarrow p_{n+1}$ for each $n \geq 10$.

25. If p_n is a statement about n for each $n \geq 1$, argue that p_n is true for all $n \geq 1$ if $p_n \Rightarrow p_{n-1}$ for each $n \geq 2$ and p_n is true for infinitely many values of n.

26. For a sequence $a_1, a_2, \ldots$, suppose that $a_1 + a_2 + \cdots + a_n$ is to be evaluated.

(a) If a sequence $b_1, b_2, \ldots$ can be found such that $a_n = b_{n+1} - b_n$ for all $n > 1$, prove by induction that $a_1 + a_2 + \cdots + a_n = b_{n+1} - b_1$.

(b) Use the technique in (a) to evaluate $1 \cdot 2 \cdot 3 + 2 \cdot 3 \cdot 4 + \cdots + n(n+1)(n+2)$. [*Hint:* Try $b_n = (n-1)n(n+1)(n+2)$.]

27. Suppose that a sequence a_0, $a_1, \ldots$ is given.

(a) Show that the sequence s_0, $s_1, \ldots$ exists where $s_0 = a_0$ and s_n is the sum of the first $n+1$ of a_i.

(b) Show that the sequence p_0, $p_1, \ldots$ exists where $p_0 = a_0$ and p_n is the product of the first $n+1$ of the a_i.

1.2 DIVISORS AND PRIME FACTORIZATION

> Mathematics is the queen of the sciences and number theory is the queen of mathematics.
>
> —Carl Friedrich Gauss

The set $\mathbb{Z}$ of integers will be used in several ways throughout this book: as a major source of examples of algebraic systems; to state definitions and prove theorems (often by induction); and as a prototype for results about more general systems. For the most part, the properties of $\mathbb{Z}$ that we need are familiar facts about addition, multiplication, and ordering of the integers, although we present a more detailed look at these properties in Section 3.2. However, we also utilize several less familiar properties of divisibility and primes in $\mathbb{Z}$ and so devote this section to them.

The Greatest Common Divisor

When we write 22/7 in the form $3\frac{1}{7}$ we are using the fact that $22 = 3 \cdot 7 + 1$; that is, 22 leaves a remainder of 1 when divided by 7. The general result is a consequence of the well-ordering axiom.

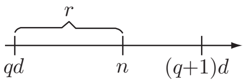
Theorem 1. ***Division Algorithm***. *Let n and $d \geq 1$ be integers. There exist uniquely determined integers q and r such that*

$$n = qd + r \quad \textit{and} \quad 0 \leq r < d.$$

Proof. Let $X = \{n - td \mid t \in \mathbb{Z},\ \ n - td \geq 0\}$. Then X is nonempty. In fact, if $n \geq 0$, then $n = n - 0d$ is in X; if $n < 0$, then $n - nd = n(1-d)$ is in X. Hence, by the well-ordering principle, let r be the smallest member of X. Then $r = n - qd$ for some q and $r \geq 0$, so it remains to show that $r < d$. But if $r \geq d$, then $0 \leq r - d = n - (q+1)d$. This means that $r - d$ is in X, contradicting the minimality of r. This result proves the existence of q and r.

To prove uniqueness, suppose also that $n = q'd + r'$ with $0 \leq r' < d$. Assume $r \leq r'$ (the case $r' \leq r$ is similar). Then $(q - q')d = r' - r$ is a nonnegative, integral multiple of d that is less than d (because $r' - r \leq r' < d$). This can occur only if $r = r'$, which implies that $q = q'$ and so proves uniqueness. ■

For n and $d \geq 1$, the integers q and r in Theorem 1 are called the **quotient** and **remainder**, respectively. Thus, for example, if we divide $n = -17$ by $d = 5$, the result is $-17 = (-4) \cdot 5 + 3$, so the quotient is -4 and the remainder is 3.



The division algorithm can also be seen geometrically. If the real line is marked off in multiples of d, n clearly falls either on a multiple qd of d or between qd and $(q+1)d$ (see the diagram). Hence, $qd \leq n < (q+1)d$, so $0 \leq n - qd < d$, and we take $r = n - qd$.

If both n and d are positive and a calculator is available, the quotient q and the remainder r can be easily found as follows: Calculate $\frac{n}{d}$ and let q denote the largest integer that is less than or equal to $\frac{n}{d}$. Hence,

$$0 \leq \tfrac{n}{d} - q < 1.$$

If we multiply through by d, we get $0 \leq n - qd < d$, so take $r = n - qd$.

Example 1. Find the quotient and remainder if $n = 4187$ and $d = 129$.

Solution. We have $\frac{n}{d} = 32.457$ approximately, so $q = 32$. Then $r = n - dq = 59$, and so $4187 = 32 \cdot 129 + 59$, as desired. □

If n and d are integers, d is called a **divisor** of n if $n = qd$ for some integer q. When this is the case, we write $d|n$. If $d|n$ is not true, we write $d \nmid n$. Thus, $7|84$ but $7 \nmid 85$. Note that $1|n$ and $n|0$ for all integers n. The following properties of divisors will be used frequently.

Theorem 2. *Let m, n and d denote integers.*

(1) *$n|n$ for all n.*

(2) *If $d|m$ and $m|n$, then $d|n$.*

(3) *If $d|n$ and $n|d$, then $d = \pm n$.*

(4) *If $d|n$ and $d|m$, then $d|(xn + ym)$ for all integers x and y.*

Proof. The proofs of (1) and (2) are left to the reader. In (3), let $n = qd$ and $d = pn$ for integers p and q. If $d = 0$, then $n = qd = 0 = d$. If $d \neq 0$, then $d = pn = pqd$, which implies that $1 = pq$. As p and q are integers, this means that $p = q = 1$ or $p = q = -1$, and so $d = n$ or $d = -n$, which proves (3). As to (4), if $n = ad$ and $m = bd$ in (4), then $xn + ym = (xa + yb)d$, so $d|(xn + ym)$, as required. ■

Expressions of the form $xn + ym$, where x and y are integers, are called **linear combinations** of n and m.

Example 2. If $d \geq 1$ is such that $d|(3k + 5)$ and $d \mid (7k + 2)$ for some k, show that $d = 1$ or $d = 29$.

Solution. The hypotheses and (4) of Theorem 2 imply that d divides the linear combination $7(3k + 5) - 3(7k + 2) = 35 - 6 = 29$. Hence, d is a positive divisor of 29, so $d = 1$ or $d = 29$. □

An integer d is called a **common divisor** of two integers m and n if $d|m$ and $d|n$. To motivate the next theorem, consider the positive divisors of 36 and 84:

• Positive divisors of 36:	$1, 2, 3, 4, 6, 9, 12, 18, 36$
• Positive divisors of 84:	$1, 2, 3, 4, 6, 7, 12, 14, 21, 28, 42, 84$
• Common divisors:	$1, 2, 3, 4, 6, 12$

We wish to focus attention on the fact that the largest common divisor 12 is actually a *multiple* of all the other positive common divisors. This idea is built into the following definition. Let m and n be integers.

An integer d is called a **greatest common divisor** of m and n if:

(1) $d \geq 1$

(2) *$d|m$ and $d|n$*

(3) *If $k|m$ and $k|n$, then $k|d$.*

When it exists we write $d = \gcd(m, n)$.

For example, $\gcd(18, 30) = 6$, $\gcd(6, 7) = 1$, and $\gcd(-9, 15) = 3$.

Conditions (2) and (3) can be stated as follows: $\gcd(m, n)$ is a common divisor of m and n by (2), which is a multiple of every common divisor by (3). If it exists,

$d = \gcd(m, n)$ is unique. In fact, if d' is another integer satisfying (1), (2), and (3), then $d'|d$ by (3). Similarly, $d|d'$ so $d = \pm d'$ by Theorem 2. But then $d' = d$ because we insist that greatest common divisors are positive.

The following fundamental theorem shows that, if m and n are not both zero, then $d = \gcd(m, n)$ does indeed exist and, surprisingly, that d is actually a linear combination of m and n.

Theorem 3. *Let m and n be integers, not both zero. Then $d = \gcd(m, n)$ exists and $d = xm + yn$ for some integers x and y.*

Proof. Let $X = \{xm + yn \mid x, y \in \mathbb{Z},\ xm + yn \geq 1\}$. Then X is not empty because $m^2 + n^2 \in X$, so let d be the smallest member of X (by the well-ordering principle). Since $d \in X$, we have $d \geq 1$ and $d = xm + yn$ for integers x and y. Also, if $k|m$ and $k|n$, then $k|(xm + yn) = d$ by Theorem 2. So it remains to show that $d|m$ and $d|n$.

To show that $d|m$, write $m = qd + r$ where $0 \leq r \leq d - 1$. Then,

$$r = m - qd = m - q(xm + yn) = (1 - qx)m + (-qy)n.$$

Hence, if $r \geq 1$, then $r \in X$ and $r < d$, contradicting the choice of d. So $r = 0$, that is, $m = qd$. Thus, $d|m$, and $d|n$ is proved similarly. ■

Note that $\gcd(m, n)$ does *not* exist if $m = 0 = n$ (verify), which explains the requirement in Theorem 3 that m and n are not both zero. Also, the greatest common divisor of m and n can be a linear combination of m and n in more than one way. For example, $\gcd(2, 3) = 1$ and we have $1 = 2 \cdot 1 - 3$ and $1 = 3 - 2$.

Example 3. If p and q are distinct primes, show that $\gcd(p, q) = 1$.

Solution. Write $d = \gcd(m, n)$. Then $d|p$, so $d = 1$ or p. Similarly, $d = 1$ or q, so $d = 1$ because, otherwise, $p = d = q$ is contrary to the assumption that $p \neq q$. □

The next example (which is needed later) illustrates how the definition of the greatest common divisor is used.

Example 4. If $m = qn + r$, show that $\gcd(m, n) = \gcd(n, r)$.

Solution. Write $d = \gcd(m, n)$ and $k = \gcd(n, r)$. Then k divides both n and r and so divides $m = qn + r$. Thus, k is a common divisor of m and n, so $k|d$ because $d = \gcd(m, n)$. A similar argument (using $r = -qn + m$) shows that $d|k$, so $d = \pm k$ by (3) of Theorem 2. Hence, $d = k$, because both d and k are positive. □

How do we compute $d = \gcd(m, n)$ in general? There is an efficient procedure for doing so, which also shows how to express d as a linear combination of m and n. To illustrate how it works, consider the numbers 78 and 30. The idea is to use the division algorithm repeatedly. First divide 78 by 30:

$$\begin{aligned} 78 &= 2 \cdot 30 + 18 \\ 30 &= 1 \cdot 18 + 12 \\ 18 &= 1 \cdot 12 + 6 \\ 12 &= 2 \cdot 6 + 0 \end{aligned}$$

At each stage (after the first) we divide the divisor at the previous stage by the remainder at that stage. The last nonzero remainder is 6, and this equals $\gcd(78, 30)$.

This is no coincidence as we shall see. To express 6 as a linear combination of 78 and 30, eliminate the remainders from the second last lineup:

$$\begin{aligned} 6 &= 18 - 1 \cdot 12 \\ &= 18 - (30 - 1 \cdot 18) \\ &= 2 \cdot 18 - 30 \\ &= 2(78 - 2 \cdot 30) - 30 \\ &= 2 \cdot 78 - 5 \cdot 30 \end{aligned}$$

This procedure is called the **euclidean algorithm**, and it works in general. For positive integers m and n, not both zero, we use the division algorithm repeatedly:

$$\begin{aligned} m &= q_1 n + r_1 & r_1 &< n \\ n &= q_2 r_1 + r_2 & r_2 &< r_1 \\ r_1 &= q_3 r_2 + r_3 & r_3 &< r_2 \\ &\vdots & &\vdots \end{aligned}$$

At each stage we divide the divisor at the previous stage by the remainder, so the remainders form a decreasing sequence of nonnegative integers:

$$n > r_1 > r_2 > r_3 > \cdots \geq 0.$$

Clearly, we must encounter a remainder of 0 (in at most n steps). If r_t denotes the last nonzero remainder, the last two equations are

$$r_{t-2} = q_t r_{t-1} + r_t \quad \text{and} \quad r_{t-1} = q_{t+1} r_t + 0.$$

Now, repeated application of the result in Example 4 gives

$$\gcd(m, n) = \gcd(n, r_1) = \gcd(r_1, r_2) = \cdots = \gcd(r_{t-1}, r_t) = r_t.$$

Hence, $\gcd(m, n)$ really is the last nonzero remainder.

Example 5. Find $\gcd(41, 12)$ and express it as a linear combination of 41 and 12.

Solution. The algorithm is not needed to find $\gcd(41, 12)$. In fact, 1 and 41 are the only positive divisors of 41, so $\gcd(41, 12) = 1$ because 41 does not divide 12. However, guessing a linear combination $1 = x \cdot 41 + y \cdot 12$ is not easy. The euclidean algorithm gives

$$\begin{aligned} 41 &= 3 \cdot 12 + 5 \\ 12 &= 2 \cdot 5 + 2 \\ 5 &= 2 \cdot 2 + 1 \\ 2 &= 2 \cdot 1 + 0 \end{aligned}$$

Hence, $\gcd(41, 12) = 1$ as expected. Elimination of remainders gives

$$\begin{aligned} 1 &= 5 - 2 \cdot 2 \\ &= 5 - 2(12 - 2 \cdot 5) \\ &= 5 \cdot 5 - 2 \cdot 12 \\ &= 5(41 - 3 \cdot 12) - 2 \cdot 12 \\ &= 5 \cdot 41 - 17 \cdot 12 \end{aligned}$$

which is the required linear combination. □

The following definition will be used frequently throughout this book.

Two integers m and n are called **relatively prime** if $\gcd(m, n) = 1$.

For example, 2 and 3 are relatively prime, as are 20 and 9. Note that 1 is relatively prime to every integer n. The condition in Theorem 4 is useful.

Theorem 4. *Let m and n be integers, not both zero. Then m and n are relatively prime if and only if $1 = xm + yn$ for some integers x and y.*

Proof. If $\gcd(m, n) = 1$, then $1 = xm + yn$ by Theorem 3. Conversely, if $1 = xm + yn$, then any common divisor of m and n must divide 1. In particular, $\gcd(m, n) = 1$. ■

For example, any two consecutive integers k and $k + 1$ are relatively prime because $(k + 1) - k = 1$. Similarly, $5(6k + 5) - 6(5k + 4) = 1$ shows that $6k + 5$ and $5k + 4$ are relatively prime for any integer k.

Corollary. *If $d = \gcd(m, n)$, $m, n \in \mathbb{Z}$, then $\frac{m}{d}$ and $\frac{n}{d}$ are relatively prime.*

Proof. If $d = xm + yn$, $x, y \in \mathbb{Z}$, dividing by d gives $1 = x\frac{m}{d} + y\frac{n}{d}$. □

The following theorem contains two very useful properties of relatively prime integers, and will be referred to several times below.

Theorem 5. *Let m and n be relatively prime integers.*

(1) *If $m|k$ and $n|k$ for some integer k, then $mn|k$.*

(2) *If $m|kn$ for some integer k, then $m|k$.*

Proof. We first prove (1). By Theorem 4, let $1 = xm + yn$, where x and y are integers. If $k = qm$ and $k = pn$ where p and q are integers, then

$$k = 1 \cdot k = xmk + ynk = xm(pn) + yn(qm) = (xp + yq)mn.$$

Hence, $mn|k$, proving (1). As to (2), let $nk = qm$ where q is an integer. Then,

$$k = 1 \cdot k = xmk + ynk = xmk + y(qm) = (xk + yq)m.$$

This shows that $m|k$, and so proves (2). ■

Prime Factorization

Clearly, every integer $n \geq 2$ has at least two positive divisors: 1 and n. The integers for which these are the *only* positive divisors are important. An integer p is called a **prime** if it satisfies the following conditions:

(1) $p \geq 2$.

(2) *If $d|p$ and $d > 0$, then either $d = 1$ or $d = p$.*

Thus, the first few primes are $2, 3, 5, 7, 11, 13, \ldots$. We know (Example 7 §1.1) that every integer greater than 1 is a product of primes; the reason for not regarding 1 as a prime is to ensure that this factorization is unique (see Theorem 7).

If the product of two integers is even, one of these integers must be even (because the product of two odd integers is odd). We can rephrase this statement as follows: If $2|mn$, where m and n are integers, then $2|m$ or $2|n$. This statement holds for any prime in place of 2.

Theorem 6. ***Euclid's Lemma***. *Let p denote a prime.*

(1) *If $p|mn$ where m and n are integers, then $p|m$ or $p|n$.*

(2) *If $p|m_1m_2\cdots m_r$ where each m_i is an integer, then $p|m_i$ for some i.*

Proof. (1) Write $d = \gcd(m, p)$. Then $d|p$, so $d = 1$ or $d = p$ because p is a prime. If $d = p$, then $p|m$ because $d|m$; if $d = 1$, then $p|n$ by (2) of Theorem 5.

(2) This assertion follows by induction on r. If $r = 1$, it is obvious. If (2) holds for some $r \geq 1$, let $p|m_1m_2\cdots m_rm_{r+1}$. Then (1) shows that either $p|m_1\cdots m_r$ or $p|m_{r+1}$. In the first case, $p|m_i$ for some $i = 1, 2, \dots, r$ by the induction hypothesis. Hence, in any case, $p|m_i$ for some $i = 1, 2, \dots, r + 1$, completing the induction. ■

Note that Euclid's lemma fails for nonprimes. For example, 6 is a divisor of $3 \cdot 4$, but 6 does not divide 3 or 4.

It is not too difficult to convince yourself that every integer $n \geq 2$ is either a prime itself or can be factored as a product of primes—just keep factoring as long as possible. For example, $12 = 2^2 \cdot 3$, $25 = 5^2$, and $360 = 2^3 \cdot 3^2 \cdot 5$. In fact, *every* integer greater than 1 is a product of primes, and this factorization is unique up to the order of the factors.

Theorem 7. ***Prime Factorization Theorem***.

(1) *Every integer $n \geq 2$ is a product of (one or more) primes.*

(2) *This factorization is unique up to the order of the factors. That is, if*

$$n = p_1p_2\cdots p_r \quad \textit{and} \quad n = q_1q_2\cdots q_s,$$

where p_i and q_j are primes, then $r = s$ and q_j can be relabeled so that $p_i = q_i$ for all $i = 1, 2, \dots, r$.

Proof. We proved (1) in Example 7 §1.1. If (2) fails, let (by the well-ordering principle) $m \geq 2$ be the smallest integer with two distinct factorizations into primes:

$$m = p_1p_2\cdots p_r = q_1q_2\cdots q_s.$$

Then m is not a prime (verify), so $r \geq 2$ and $s \geq 2$. We have $p_1|q_1q_2\cdots q_s$, so $p_1|q_j$ for some j by Euclid's lemma. By relabeling q_j, we may assume that $p_1|q_1$. Then $p_1 = q_1$ because both are primes, so

$$\tfrac{m}{p_1} = p_2\cdots p_r = q_2\cdots q_s$$

is an integer—smaller than m—that admits two distinct factorizations into primes. This result contradicts the choice of m, and so proves (2). ■

Corollary. *Two integers $m \geq 2$ and $n \geq 2$ are relatively prime if and only if no prime divides both m and n.*

Proof. Write $d = \gcd(m, n)$. If $d = 1$, then any common prime divisor would have to divide 1, so no such common divisor exists. Conversely, suppose no prime divides both m and n. If $d > 1$ and $p|d$ where p is a prime, then $p|m$ and $p|n$, contrary to our assumption. So $d = 1$, that is m and n are relatively prime. □

If $n \geq 2$ is an integer and $p_1, p_2, \dots, p_r$ are the distinct prime divisors of n, the prime factorization theorem asserts that n can be written uniquely in the form

$$n = p_1^{n_1}p_2^{n_2}\cdots p_r^{n_r},$$

where $n_i \geq 1$ for each i. This means that the primes p_i and the integers n_i are uniquely determined by n. For example, $60 = 2^2 \cdot 3 \cdot 5$ and $882 = 2 \cdot 3^2 \cdot 7^2$.

If n has only one prime divisor, we call it a **prime power**, examples being $7 = 7^1$, $9 = 3^2$, and $32 = 2^5$. At the other extreme, we say that n is **square free** if all the exponents $n_i = 1$. Hence, any prime is square free as are $6 = 2 \cdot 3$ and $70 = 2 \cdot 5 \cdot 7$.

If n is not prime, it must have a prime divisor $p \leq \sqrt{n}$ (it cannot have two prime divisors greater than $\sqrt{n}$). So to test whether n is prime, it suffices to verify that it has no prime divisor $p \leq \sqrt{n}$ (which is impractical if n is very large).

Example 6. Factor 1591 into primes.

Solution. We start dividing 1591 by the successive primes, $2, 3, 5, 7, \ldots$. Since $\sqrt{1591} < 40$ (because $40^2 = 1600$), we need go only as high as 37; in fact, the first prime that divides 1591 is 37. As $1591 = 37 \cdot 43$ and 43 is a prime, we have the required prime factorization. □

Obviously, the method in Example 6 requires that we have a list of the primes. Although large tables of primes are available, the method clearly fails for very large numbers. Finding the prime factorization of large integers is very difficult. Even so, on December 15, 2005 it was announced that $2^{30,402,457} - 1$ is a prime with 9,152,052 digits, the largest prime known to that date. Such a result requires a very large amount of computer time.[10]

The prime factorization theorem gives a systematic way of listing all the positive divisors of an integer n when the prime factorization of n is known. For example, if $n = 12 = 2^3 \cdot 3$, these divisors are $1, 2, 3, 4, 6$, and 12, and they can be written as

$$\begin{array}{lll} 1 = 2^0 3^0 & 2 = 2^1 3^0 & 4 = 2^2 3^0 \\ 3 = 2^0 3^1 & 6 = 2^1 3^1 & 12 = 2^2 3^1 \end{array}$$

Thus, they can all be expressed as $2^r 3^s$, where $0 \leq r \leq 2$ and $0 \leq s \leq 1$ (where $p^0 = 1$ for any prime p). The general situation is as follows:

Theorem 8. *Let n be an integer with prime factorization*

$$n = p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r},$$

where p_i are distinct primes and $n_i \geq 1$ for each i. Then the positive divisors of n are precisely the integers d of the form:

$$d = p_1^{d_1} p_2^{d_2} \cdots p_r^{d_r},$$

where $0 \leq d_i \leq n_i$ holds for each i.

Proof. The prime divisors of d are contained in $\{p_1, \ldots, p_r\}$ by Euclid's lemma, and d cannot contain a higher power of p_i than $p_i^{n_i}$ by Theorem 7. ■

In much the same way, the prime factorization theorem provides a simple way to compute the greatest common divisor of any finite set of positive integers (rather

[10] On the other hand, in 2002, Maninda Agrawal and two undergraduate students (Neeraj Kayal and Nitin Saxena) gave a simple algorithm that can decide whether a given integer n is prime or not. Moreover, the time taken is approximately a polynomial function of n. This is an important breakthrough in computer science.

than just two). It also provides the "dual" notion, the least common multiple. The definitions are as follows. Let $n_1, n_2, \dots, n_r$ be positive integers.

(1) *The* **greatest common divisor** $\gcd(n_1, n_2, \dots, n_r)$ *of these integers is the positive common divisor that is a multiple of every common divisor.*

(2) *The* **least common multiple** $\operatorname{lcm}(n_1, n_2, \dots, n_r)$ *of these integers is the positive common multiple that is a divisor of every common multiple.*

Thus, $\gcd(4, 6, 10) = 2$ and $\operatorname{lcm}(4, 6, 10) = 60$ by inspection. Theorem 9 below shows that the gcd and lcm always exist. They are uniquely determined in the same way as the gcd of two integers (see the discussion preceding Theorem 3). The next example illustrates a systematic method for finding the gcd and lcm.

Example 7. Find $d = \gcd(12, 20, 18)$ and $m = \operatorname{lcm}(12, 20, 18)$.

Solution. We might find $d = 2$ by experiment, but $m = 180$ is not clear. A systematic method involves writing the prime factorizations as follows:

$$\begin{aligned} 12 &= 2^2 \cdot 3^1 \cdot 5^0 \\ 20 &= 2^2 \cdot 3^0 \cdot 5^1 \\ 18 &= 2^1 \cdot 3^2 \cdot 5^0 \end{aligned}$$

We have $d = 2^a \cdot 3^b \cdot 5^c$ for some a, b, and c by Theorem 8. We have $a \leq 1$ because $d | 18$, and $b = c = 0$ because $d | 20$ and $d | 12$. Thus, $d = 2$ is the largest possibility. Similarly, write the prime factorization of m as $m = 2^p \cdot 3^q \cdot 5^r \cdot k$, where $k \geq 1$ is the factor involving primes (if any) other than $2, 3$, or 5. Then $p \geq 2$ because $12 | m$ (or because $20 | m$), $q \geq 2$ because $18 | m$, and $r \geq 1$ because $20 | m$. The smallest possibility is thus $m = 2^2 \cdot 3^2 \cdot 5^1 = 180$. □

In Example 7, the power of 2 in $d = \gcd(12, 20, 18)$ is the *smallest* of the powers of 2 occurring in $12, 20$, and 18; the same is true for the powers of 3 and 5 in d. Similarly, the power of 2 in $m = \operatorname{lcm}(12, 20, 18)$ is the *largest* of the powers of 2 in $12, 20$, and 18, with similar statements for the primes 3 and 5. This method works in general. For finitely many integers $a, b, c, \dots$, let

$$\max(a, b, c, \dots) \quad \text{and} \quad \min(a, b, c, \dots)$$

denote the largest and the smallest of these integers, respectively. For example, we have $\max(3, 1, -5, 3) = 3$ and $\min(1, 0, 5) = 0$.

Using Theorem 8, the solution to Example 7 extends to a proof of Theorem 9.

Theorem 9. *Let $\{a, b, c, \dots\}$ be a finite set of positive integers, and write*

$$\begin{aligned} a &= p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r} \\ b &= p_1^{b_1} p_2^{b_2} \cdots p_r^{b_r} \\ c &= p_1^{c_1} p_2^{c_2} \cdots p_r^{c_r} \\ &\vdots \end{aligned}$$

where p_i are primes dividing at least one of $a, b, c, \dots$, and where an exponent is zero if the prime in question does not occur in that number. Then,

$$\begin{aligned} \gcd(a, b, c, \dots) &= p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}, \\ \operatorname{lcm}(a, b, c, \dots) &= p_1^{m_1} p_2^{m_2} \cdots p_r^{m_r}, \end{aligned}$$

where $k_i = \min(a_i, b_i, c_i, \dots)$ and $m_i = \max(a_i, b_i, c_i, \dots)$ for each i.

Example 8. Find $\gcd(63, 60, 105)$ and $\operatorname{lcm}(63, 60, 105)$.

Solution. The prime factorizations are

$$63 = 2^0 3^2 5^0 7^1, \quad 60 = 2^2 3^1 5^1 7^0, \quad \text{and} \quad 105 = 2^0 3^1 5^1 7^1.$$

Hence, $\gcd(63, 60, 105) = 2^0 3^1 5^0 7^0 = 3$ and $\operatorname{lcm}(63, 60, 105) = 2^2 3^2 5^1 7^1 = 1260$. □

Of course we can use Theorem 9 to find $\operatorname{lcm}(a, b)$ and $\gcd(a, b)$ for two integers a and b. However, the euclidean algorithm is also available to compute $\gcd(a, b)$, so the next result is useful for finding $\operatorname{lcm}(a, b)$.

Corollary. *If a and b are positive integers, then* $\operatorname{lcm}(a, b) \cdot \gcd(a, b) = ab$.

Proof. The assertion follows from Theorem 9 and the fact that, for integers m and n, $\max(m, n) + \min(m, n) = m + n$. ■

Note that $\operatorname{lcm}(a, b, c) \cdot \gcd(a, b, c) \neq abc$ can occur (consider Example 8).

We conclude with one last application of the prime factorization theorem.

Theorem 10. ***Euclid's Theorem***. *There are infinitely many primes.*

Proof. Suppose, on the contrary, that there are only n primes, denoted $p_1, p_2, \ldots, p_n$. Then consider the integer $m = 1 + p_1 p_2 \cdots p_n$. Since $m \geq 2$, some prime divides m by Theorem 7. But if $p_i | m$, then p_i divides $m - p_1 p_2 \cdots p_m = 1$, a contradiction. Hence the assumption that there are only finitely many primes is untenable. ■

Euclid's theorem certainly implies that there are infinitely many odd primes, that is, primes of the form $2k + 1$, $k = 0, 1, \ldots$, and a natural question is whether there are infinitely many primes of the form $mk + n$ for any positive integers m and n. This clearly cannot happen unless m and n are relatively prime. However, in this case it is valid, a result first proved by P.G.L. Dirichlet. One instance of Dirichlet's theorem is treated in Exercise 39.

However, there are many unanswered questions about primes, among them the celebrated **Goldbach conjecture**, which asserts that every even integer greater than 2 is the sum of two primes. The conjecture dates from 1742 and originated in some correspondence between C. Goldbach and L. Euler. It is not known whether this assertion is true; the question appears to be extremely difficult to answer. The best result known is that every sufficiently large even number is the sum of a prime and a number that is the product of at most two primes.

Exercises 1.2

1. In each case find the quotient and remainder when n is divided by d.
(a) $n = 391$, $d = 17$ (b) $n = 401$, $d = 19$
(c) $n = -116$, $d = 13$ (d) $n = -162$, $d = 17$

2. In each case write $r = n - qd$, as in Example 1.
(a) $n = 51837$, $d = 386$ (b) $n = 39214$, $d = 871$

3. If n and $d \neq 0$ are integers, show that integers q and r exist such that $n = qd + r$ and $0 \leq r < |d|$.

4. Show that the negative divisors of an integer n are just the negatives of the positive divisors.

5. If m and n are odd integers, show that $m^2 - n^2$ is divisible by 8.

6. Given three consecutive integers, show that one must be a multiple of 3.

7. (a) If $d > 0$, $d|(11k+4)$, and $d|(10k+3)$ for some integer k, show that $d = 1$ or $d = 7$.

(b) If $d > 0$, $d|(35k+26)$, and $d|(7k+3)$ for some integer k, show that $d = 1$ or $d = 11$.

8. Explain why $\gcd(0,0)$ does not exist. If $n > 0$, what is $\gcd(0,n)$?

9. In each case, compute $\gcd(m,n)$ and express it as a linear combination of m and n.

(a) $m = 72$, $n = 42$ (b) $m = 41$, $n = 25$
(c) $m = 327$, $n = 54$ (d) $m = 198$, $n = 241$
(e) $m = 377$, $n = 29$ (f) $m = 527$, $n = 31$
(g) $m = 72$, $n = -175$ (h) $m = -231$, $n = 150$

10. If $m \geq 1$, show that $m|n$ if and only if $\gcd(m,n) = m$.

11. Let $d = \gcd(m,n)$. If $k|d$, $k \geq 1$, show that $\gcd(\frac{m}{k}, \frac{n}{k}) = \frac{d}{k}$.

12. If m and n are relatively prime and $k|m$, show that k and n are relatively prime.

13. Is $n^2 + n + 11$ prime for all $n \geq 1$? Support your answer.

14. Show that $\gcd(m+n, m) = \gcd(m,n)$.

15. If $m|m_1$ and $n|n_1$, show that $\gcd(m,n)|\gcd(m_1,n_1)$.

16. If $n|k(n+1)$, show that $n|k$.

17. If $\gcd(m,n) = 1$ and $\gcd(k,n) = 1$, show that $\gcd(mk,n) = 1$.

18. If $\gcd(m,n) = 1$, let $d = \gcd(m+n, m-n)$. Show that $d = 1$ or $d = 2$.

19. Show that $\gcd(km,kn) = k\gcd(m,n)$ if $k \geq 1$.

20. Show that m and n are relatively prime if and only if no prime divides both.

21. Suppose that $p \geq 2$ is an integer with the following property: If m and n are integers and $p|mn$, either $p|m$ or $p|n$. Show that p must be a prime.

22. If $d_1, \dots, d_r$ are all divisors of n and if $\gcd(d_i, d_j) = 1$ whenever $i \neq j$, show that $d_1 d_2 \cdots d_r$ divides n.

23. If $d = \gcd(a,n)$, must $\frac{a}{d}$ and n be relatively prime? Prove or disprove.

24. Show that any two consecutive odd integers are relatively prime.

25. Show that $3, 5$, and 7 is the only *prime triple* (that is, three consecutive odd integers, each of which is prime). It is not known if there are infinitely many *prime pairs.*

26. Let p be a prime. If n is any integer, show that either $p|n$ or $\gcd(p,n) = 1$.

27. If $\gcd(m,p) = 1$ and p is a prime, show that $\gcd(m,p^k) = 1$ for all $k \geq 1$.

28. Show that none of $n!+2, n!+3, \dots, n!+n$ are primes for any $n \geq 2$. Hence, show that there are arbitrarily long gaps in the primes.

29. Let $ab = a_1 b_1$, where a, b, a_1, and b_1 are positive integers. If $\gcd(a, b_1) = 1$ and $\gcd(a_1, b) = 1$, show that $a = a_1$ and $b = b_1$.

30. Find the prime factorizations of the following integers:

(a) 27783 (b) 1331 (c) 2431
(d) 18900 (e) 241 (f) 1457

31. Find the gcd and the lcm of the following pairs of numbers:

(a) $735, 110$ (b) $101, 113$ (c) $139, 278$ (d) $221, 187$

32. If $d = \gcd(a,b)$ and $m = ab/d$, show that $m = \operatorname{lcm}(a,b)$ using only Theorem 3.

33. Let n be a positive integer with prime factorization $n = p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}$ where the p_i are distinct primes and $n_i \geq 1$ for each i.

(a) Show that n has $(n_1+1)(n_2+1)\dots(n_r+1)$ distinct positive divisors.

(b) Write down all the positive divisors of 340, 108, p^n, p^2q, where p and q are distinct primes.

(c) How many positive divisors does n have if $n = 25200$; $n = 41472$?

34. If $m \geq 1$ and $n \geq 1$ are relatively prime integers and nm is the square of an integer, show that both m and n are squares. Is this result true if m and n are not relatively prime?

35. If $\gcd(m, n) = 1$, where $m \geq 1$ and $n \geq 1$, and if $d | mn$, show that $d = m_1 n_1$ for some $m_1 | m$ and $n_1 | n$. [*Hint:* Theorem 7.]

36. Do Exercise 35 without assuming that $\gcd(m, n) = 1$. [*Hint:* If $0 \leq e \leq f + g$, where $f \geq 0$ and $g \geq 0$ are integers, show that e can be written $e = f_1 + g_1$, where $0 \leq f_1 \leq f$ and $0 \leq g_1 \leq g$. Use Theorem 8.]

37. Let $a \geq 1$ and $b \geq 1$ be integers. Show that there exist integers $u \geq 1$ and $v \geq 1$ such that $u|a$, $v|b$, $\gcd(u, v) = 1$, and $\operatorname{lcm}(u, v) = ab$. [*Hint*: Theorem 9.]

38. If q is a rational number such that q^2 is an integer, show that q is an integer. [*Hint:* If $m^2 | n^2$, show that $m | n$ using Theorem 7.]

39. (a) Show that every prime $p > 2$ has the form $p = 4k + 1$ or $p = 4k + 3$.

(b) Modify the proof of Theorem 10 to show that there are infinitely many primes of the form $4k + 3$.

40. A school has n lockers in a row along one side of a hall. The n students run down the hall one after the other. The first student closes all the lockers; then the second opens doors $2, 4, 6, \ldots$; the third changes doors $3, 6, 9, \ldots$ (that is, opens a door if it is closed and closes it if it is open); the fourth student changes doors $4, 8, 12, \ldots$, and so on. When all n students have gone through, which locker doors remain closed? Prove your answer. [*Hint*: Exercise 33(a).]

41. Compute the following:

(a) $\gcd(28665, 22869)$ and $\operatorname{lcm}(28665, 22869)$

(b) $\gcd(231, 273, 429)$ and $\operatorname{lcm}(231, 273, 429)$

(c) $\gcd(1365, 1911, 1155, 1925)$ and $\operatorname{lcm}(1365, 1911, 1155, 1925)$

42. Show that $\gcd(a, b, c) = \gcd[a, \gcd(b, c)]$.

43. Let $d = \gcd(a_1, a_2, a_3, \ldots, a_k)$, where the a_i are positive integers. Show that integers $x_1, x_2, \ldots, x_k$ exist such that $d = x_1 a_1 + \cdots + x_k a_k$. [*Hint:* Let m be the smallest member of $X = \{x_1 a_1 + \cdots + x_k a_k |\, x_i \in \mathbb{Z},\ \ x_1 a_1 + \cdots + x_k a_k \geq 1\}$, and show that $m = d$. See the proof of Theorem 3.]

44. Let $b \geq 2$ be a fixed integer. If $n \geq 0$ is any integer, show that n can be written in the form $n = r_t b^t + r_{t-1} b^{t-1} + \cdots + r_1 b + r_0$, where $t \geq 0$ and $0 \leq r_i < b$ for all i. Show further that these integers r_i and t are uniquely determined by n. This expression is called the **base *b* representation** of n.

45. Let $m \geq 1$ and $n \geq 1$ be integers.

(a) If $m = qn + r$, $q, r \in \mathbb{Z}$, $0 \leq r < n$, show that $2^m - 1 = x(2^n - 1) + (2^r - 1)$ for some $x \in \mathbb{Z}$, where $0 \leq (2^r - 1) < 2^n - 1$.

(b) If $d = \gcd(m, n)$, show that $\gcd(2^m - 1, 2^n - 1) = 2^d - 1$. [*Hint:* Get d by the euclidean algorithm and use (a).]

1.3 INTEGERS MODULO n

Two integers a and b are said to have the same **parity** if both are even or both are odd, that is, if $2 | (a - b)$. The following definition extends this idea and introduces an important equivalence on the set $\mathbb{Z}$ of integers. Let $n \geq 2$ be an integer.

> *Then integers a and b are said to be* **congruent modulo** n *if* $n|(a-b)$. *In this case we write* $a \equiv b \pmod{n}$ *and refer to n as the* **modulus**.

Thus, we have $2 \equiv 5 \pmod 3$, $21 \equiv 16 \pmod 5$, and $-4 \equiv 2 \pmod 6$. The expression $21832 \equiv 32 \pmod{100}$ explains why we can test whether an integer is divisible by 100 by looking at the last two digits. Note that $a \equiv 0 \pmod n$ if and only if $n \mid a$. We assume that $n \geq 2$ because congruence modulo 0 or 1 is of no interest (verify).

As the notation $\equiv$ suggests, congruence modulo n is an equivalence relation on $\mathbb{Z}$.[11] The notation is justified in Theorem 1 and the proof is left as Exercise 6(a).

Theorem 1. *Congruence modulo n is an equivalence on $\mathbb{Z}$; that is*:

(1) $a \equiv a \pmod n$ *for every integer* a.

(2) *If* $a \equiv b \pmod n$, *then* $b \equiv a \pmod n$.

(3) *If* $a \equiv b \pmod n$ *and* $b \equiv c \pmod n$, *then* $a \equiv c \pmod n$.

If a is an integer, its equivalence class $[a]$ with respect to congruence modulo n is called its **residue class modulo** n, and we write $\bar{a} = [a]$ for convenience:

$$\bar{a} = [a] = \{x \in \mathbb{Z} \mid x \equiv a \pmod n\}.$$

The following result will be used frequently below.

Theorem 2. *Given* $n \geq 2$, $\bar{a} = \bar{b}$ *if and only if* $a \equiv b \pmod n$.

Proof. Suppose $\bar{a} = \bar{b}$. Since $a \in \bar{a}$, we have $a \in \bar{b}$, so $a \equiv b$. Conversely, let $a \equiv b$. Since $\bar{a}$ and $\bar{b}$ are sets, we must show that $\bar{a} \subseteq \bar{b}$ and $\bar{b} \subseteq \bar{a}$. If $x \in \bar{a}$, then $x \equiv a$; so, as $a \equiv b$, we have $x \equiv b$ by (3) of Theorem 1. This proves that $\bar{a} \subseteq \bar{b}$. Since $b \equiv a$ by (2) of Theorem 1, a similar proof shows that $\bar{b} \subseteq \bar{a}$. ■

Residue classes are easy to describe. For example, if $n = 2$,

$$\begin{aligned}\bar{0} &= \{x \in \mathbb{Z} \mid x \equiv 0 \pmod 2\} = \text{the set of even integers}\\ \bar{1} &= \{x \in \mathbb{Z} \mid x \equiv 1 \pmod 2\} = \text{the set of odd integers}\end{aligned}$$

In general, if a is an integer, the division algorithm gives $a = qn + r$, where $0 \leq r \leq n-1$, so $a \equiv r \pmod n$. Thus every residue class modulo n appears in the list $\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{n-1}$. In fact it appears exactly once.

Theorem 3. *Let $n \geq 2$ be an integer.*

(1) *If* $a \in \mathbb{Z}$, *then* $\bar{a} = \bar{r}$ *for some r where* $0 \leq r \leq n-1$.

(2) *The residue classes* $\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{n-1}$ *modulo n are distinct.*

Proof. It remains to verify (2). Suppose $\bar{r} = \bar{s}$, where $0 \leq r \leq n-1$ and $0 \leq s \leq n-1$. We may assume that $r \leq s$. Then $\bar{r} = \bar{s}$ means that $r \equiv s \pmod n$, so $s - r$ is an integral multiple of n such that $0 \leq s - r \leq n-1$. This implies that $r = s$. ■

The set of all residue classes modulo n is denoted

$$\mathbb{Z}_n = \{\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{n-1}\}$$

[11] See Section 0.4 for a discussion on equivalence relations.

and is called the set of **integers modulo** n. Thus, (2) of Theorem 3 is the assertion that $|\mathbb{Z}_n| = n$. In particular, $\mathbb{Z}_2 = \{\bar{0}, \bar{1}\}$, $\mathbb{Z}_3 = \{\bar{0}, \bar{1}, \bar{2}\}$, and so on.[12]

Example 1. Locate $\overline{48}$ and $\overline{-16}$ in $\mathbb{Z}_7 = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}, \bar{4}, \bar{5}, \bar{6}\}$.

Solution. It seems that $\overline{48}$ does not appear. However, $48 \equiv 6 \pmod 7$ means that $\overline{48} = \bar{6}$ does indeed occur. Similarly, $-16 \equiv 5 \pmod 7$, so $\overline{-16} = \bar{5}$ also appears. □

Example 2. If a is an odd integer, show that $\bar{a} = \bar{1}$ or $\bar{a} = \bar{3}$ in $\mathbb{Z}_4 = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}\}$.

Solution. We know that $\bar{a}$ is one of $\bar{0}, \bar{1}, \bar{2}$, or $\bar{3}$ in $\mathbb{Z}_4$. If $\bar{a} = \bar{2}$, then $a \equiv 2 \pmod 4$, so $a - 2 = 4q$ for some integer q. This means that a is even, contrary to assumption. So $\bar{a} \neq \bar{2}$ and, similarly, $\bar{a} \neq \bar{0}$. The only other possibilities are $\bar{a} = \bar{1}$ and $\bar{a} = \bar{3}$. □

Example 3. In $\mathbb{Z}_n$, show that $\bar{a} = \bar{0}$ if and only if $n|a$.

Solution. By Theorem 2, $\bar{a} = \bar{0}$ means that $a \equiv 0 \pmod n$, that is, $n|a$. □

Congruence modulo n is compatible with addition and multiplication of integers in the following sense. Let a, a_1, b, and b_1 denote integers.

$$\text{If } \begin{cases} a \equiv a_1 \pmod n \\ b \equiv b_1 \pmod n \end{cases} \quad \text{then} \quad \begin{matrix} a + b \equiv a_1 + b_1 \pmod n \\ ab \equiv a_1 b_1 \pmod n \end{matrix} \tag{*}$$

In fact, let $a - a_1 = pn$ and $b - b_1 = qn$, where p and q are integers. Adding these equations gives $(a + b) - (a_1 + b_1) = (p + q)n$, and this implies that $a + b \equiv a_1 + b_1 \pmod n$. Similarly, multiplying the equations $a = a_1 + pn$ and $b = b_1 + qn$ gives $ab \equiv a_1 b_1 \pmod n$.

Condition (*) means that the arithmetic of $\mathbb{Z}$ extends naturally to $\mathbb{Z}_n$ as follows: We define addition and multiplication of residue classes $\bar{a}$ and $\bar{b}$ in $\mathbb{Z}_n$ by

$$\bar{a} + \bar{b} = \overline{a + b} \quad \text{and} \quad \bar{a}\bar{b} = \overline{ab}. \tag{**}$$

Of course, we must verify that these operations are well defined, that is, we must check that they do not depend on which generators are used for the residue classes $\bar{a}$ and $\bar{b}$. More precisely, suppose that

$$\bar{a} = \bar{a}_1 \quad \text{and} \quad \bar{b} = \bar{b}_1,$$

where $a \neq a_1$ and $b \neq b_1$ are possible. If we add these classes as $\bar{a}$ and $\bar{b}$, (**) gives their sum as $\overline{a + b}$, but if we represent the classes as $\bar{a}_1$ and $\bar{b}_1$, their sum is $\overline{a_1 + b_1}$. Clearly, the definition of addition makes no sense unless $\overline{a + b} = \overline{a_1 + b_1}$. But $a \equiv a_1$ and $b \equiv b_1$ by Theorem 2, so $a + a_1 \equiv b + b_1$ by (*), so $\overline{a + b} = \overline{a_1 + b_1}$, as required. Similarly, (*) shows that $\overline{ab} = \overline{a_1 b_1}$, so the definition of multiplication also makes sense. In other words, addition and multiplication of residue classes are well defined by (**).

Example 4. In $\mathbb{Z}_6$ compute $\bar{3} + \bar{5}$ and $\bar{3} \cdot \bar{5}$.

Solution. The definition gives $\bar{3} + \bar{5} = \bar{8} = \bar{2}$, because $8 \equiv 2 \pmod 6$. Similarly, $\bar{3} \cdot \bar{5} = \overline{15} = \bar{3}$. □

[12] Note that $\bar{a}$ means different things in $\mathbb{Z}_2$, $\mathbb{Z}_3$, ..., so to avoid ambiguity, perhaps we should denote residue classes $\bar{a}$ in such a way that the modulus is apparent (say, ${}^2\bar{a}$ and ${}^3\bar{a}$). However, this is rarely done in practice as the modulus is usually clear from the context.

Theorem 4 collects several properties of these operations in $\mathbb{Z}_n$, each of which is the analogue of the corresponding property for $\mathbb{Z}$.

Theorem 4. *Let $n \geq 2$ be a fixed modulus and let a, b, and c denote arbitrary integers. Then the following hold in $\mathbb{Z}_n$.*

(1) $\bar{a} + \bar{b} = \bar{b} + \bar{a}$ *and* $\bar{a}\bar{b} = \bar{b}\bar{a}$.

(2) $\bar{a} + (\bar{b} + \bar{c}) = (\bar{a} + \bar{b}) + \bar{c}$ *and* $\bar{a}(\bar{b}\bar{c}) = (\bar{a}\bar{b})\bar{c}$.

(3) $\bar{a} + \bar{0} = \bar{a}$ *and* $\bar{a}\bar{1} = \bar{a}$.

(4) $\bar{a} + \overline{-a} = \bar{0}$.

(5) $\bar{a}(\bar{b} + \bar{c}) = \bar{a}\bar{b} + \bar{a}\bar{c}$.

Proof. We prove (5) and leave the rest as Exercise 6(b). Thus,

$$\begin{aligned}
\bar{a}(\bar{b} + \bar{c}) &= \bar{a}(\overline{b + c}) && \text{(definition of addition in } \mathbb{Z}_n) \\
&= \overline{a(b + c)} && \text{(definition of multiplication in } \mathbb{Z}_n) \\
&= \overline{ab + ac} && \text{(property of } \mathbb{Z}) \\
&= \overline{ab} + \overline{ac} && \text{(definition of addition in } \mathbb{Z}_n) \\
&= \bar{a}\bar{b} + \bar{a}\bar{c} && \text{(definition of multiplication in } \mathbb{Z}_n),
\end{aligned}$$

which proves (5). ■

These properties enable us to do arithmetic in $\mathbb{Z}_n$ in much the same way as in $\mathbb{Z}$. In particular, (3) shows that $\bar{0}$ and $\bar{1}$ play roles in $\mathbb{Z}_n$ analogous to those of 0 and 1 in $\mathbb{Z}$. For this reason, $\bar{0}$ and $\bar{1}$ are called the *zero* of $\mathbb{Z}_n$ and the *unity* of $\mathbb{Z}_n$, respectively. Similarly, because of (4), $\overline{-a}$ is called the *negative* of $\bar{a}$ in $\mathbb{Z}_n$, and is denoted $\overline{-a} = -\bar{a}$. Then *subtraction* in $\mathbb{Z}_n$ is defined by

$$\bar{a} - \bar{b} = \bar{a} + \overline{-b} = \overline{a - b},$$

an operation used much as it is in $\mathbb{Z}$.

Now consider the addition and multiplication tables for $\mathbb{Z}_6 = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}, \bar{4}, \bar{5}\}$:

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{4}$	$\bar{4}$	$\bar{5}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{5}$	$\bar{5}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$

$\times$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{4}$	$\bar{0}$	$\bar{2}$	$\bar{4}$
$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{0}$	$\bar{3}$
$\bar{4}$	$\bar{0}$	$\bar{4}$	$\bar{2}$	$\bar{0}$	$\bar{4}$	$\bar{2}$
$\bar{5}$	$\bar{0}$	$\bar{5}$	$\bar{4}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

These tables reveal many differences between the arithmetic of $\mathbb{Z}_6$ and that of $\mathbb{Z}$. For example, while 0 and 1 are the only integers k in $\mathbb{Z}$ with the property that $k^2 = k$, each of $\bar{0}$, $\bar{1}$, $\bar{3}$, and $\bar{4}$ enjoy this property in $\mathbb{Z}_6$. Another difference is that if $ab = ac$ in $\mathbb{Z}$ and $a \neq 0$, then $b = c$. But $\bar{4} \cdot \bar{2} = \bar{4} \cdot \bar{5}$ in $\mathbb{Z}_6$, and $\bar{4} \neq \bar{0}$, but $\bar{2} \neq \bar{5}$. Hence, we must be careful about "cancellation" in $\mathbb{Z}_n$. In fact, this concern is related to another difference between $\mathbb{Z}$ and $\mathbb{Z}_n$. If $ab = 0$ in $\mathbb{Z}$, then $a = 0$ or $b = 0$. However, this need not hold in $\mathbb{Z}_n$. For example, $\bar{2} \cdot \bar{3} = \bar{0}$ in $\mathbb{Z}_6$, but $\bar{2} \neq \bar{0}$ and $\bar{3} \neq \bar{0}$.

In Examples 5–7, we use the arithmetic of $\mathbb{Z}_n$ to deduce facts about $\mathbb{Z}$. The connection is the fact (in Theorem 2) that $\bar{a} = \bar{b}$ in $\mathbb{Z}_n$ means that $a \equiv b \pmod{n}$.

Example 5. Show that $a^5 \equiv a \pmod 5$ holds for all integers a.

Solution. For an integer a, it suffices by Theorem 2 to show that $\bar{a}^5 = \bar{a}$ in $\mathbb{Z}_5$. Because $\bar{a}$ equals $\bar{0}, \bar{1}, \bar{2}, \bar{3}$, or $\bar{4}$, we examine each case separately.

- If $\bar{a} = \bar{0}$, then $\bar{a}^5 = \bar{0}^5 = \bar{0} = \bar{a}$.
- If $\bar{a} = \bar{1}$, then $\bar{a}^5 = \bar{1}^5 = \bar{1} = \bar{a}$.
- If $\bar{a} = \bar{2}$, then $\bar{a}^5 = \bar{2}^5 = \bar{2}^3 \cdot \bar{2}^2 = \bar{3} \cdot \bar{4} = \bar{2} = \bar{a}$.
- If $\bar{a} = \bar{3}$, then $\bar{a}^5 = \bar{3}^5 = \bar{9} \cdot \overline{27} = \bar{4} \cdot \bar{2} = \bar{3} = \bar{a}$.
- If $\bar{a} = \bar{4}$, then $\bar{a}^5 = \bar{4}^5 = \overline{16} \cdot \overline{64} = \bar{1} \cdot \bar{4} = \bar{4} = \bar{a}$.

Hence, $\bar{a}^5 = \bar{a}$ in every case, so $a^5 \equiv a \pmod 5$ for all integers a. □

Example 5 is a special case of Fermat's theorem, which, for any prime p, asserts that $a^p \equiv a \pmod p$ for all integers a. We return to it later (Theorem 8).

Example 6. What is the remainder when 4^{119} is divided by 7?

Solution. If we can show that $4^{119} \equiv r \pmod 7$, where $0 \le r \le 6$, then r is the desired remainder. We do the computation in $\mathbb{Z}_7$. Note that, as $\bar{4}^2 = \bar{2}$ in $\mathbb{Z}_7$, we have $\bar{4}^3 = \bar{8} = \bar{1}$. With this in mind, divide the exponent 119 by 3 to get $119 = 3 \cdot 39 + 2$. Then,

$$\bar{4}^{119} = \bar{4}^{3\cdot 39+2} = (\bar{4}^3)^{39} \cdot \bar{4}^2 = \bar{1}^{39} \cdot \bar{2} = \bar{2}.$$

Hence, $4^{119} \equiv 2 \pmod 7$, so the required remainder is 2. □

If a is an integer in decimal notation, it is common knowledge that a is divisible by 2 or 5 if and only if the same is true of its unit digit. Example 7 gives a similar test for divisibility by 9.

Example 7. **Casting Out Nines**. Show that a positive integer is divisible by 9 if and only if the sum of its digits is divisible by 9.

Solution. If $a = d_r d_{r-1} \dots d_1 d_0$ in decimal notation, where $d_0, d_1, \cdots, d_r$ are the digits, then $a = d_0 + 10d_1 + 10^2 d_2 + \cdots + 10^r d_r$. Now $\overline{10} = \bar{1}$ in $\mathbb{Z}_9$, so $\overline{10}^k = \bar{1}^k = \bar{1}$ for each k. Hence, in $\mathbb{Z}_9$,

$$\bar{a} = \bar{d_0} + \bar{1} \cdot \bar{d_1} + \bar{1}^2 \cdot \bar{d_2} + \cdots + \bar{1}^r \cdot \bar{d_r} = \overline{d_0 + d_1 + \cdots + d_r}.$$

Thus, $a \equiv d_0 + d_1 + \cdots + d_r \pmod 9$, and the result follows from Example 3. □

These three examples show that the properties in Theorem 4 allow many of the operations of ordinary arithmetic to be carried out in $\mathbb{Z}_n$. However, these properties tell us nothing about how to solve an equation such as $\bar{a}x = \bar{b}$ in $\mathbb{Z}_n$. For example, consider

$$\bar{5}x = \bar{2}$$

in $\mathbb{Z}_{17}$. The desired solution (if there is one) is a residue class x in $\mathbb{Z}_{17}$, so x is one of $\bar{0}, \bar{1}, \bar{2}, \dots, \overline{16}$. Hence, one method is simply to try all these classes! If we do so, we find that $x = \overline{14}$ is the only solution. However, this method is impractical if the modulus is large.

A better approach is as follows. Suppose that a residue class $\bar{b}$ can be found such that $\bar{b} \cdot \bar{5} = \bar{1}$. Then if we multiply both sides of the equation $\bar{5}x = \bar{2}$ by $\bar{b}$, the result is $\bar{b} \cdot \bar{5}x = \bar{b} \cdot \bar{2}$, that is, $x = \bar{2}\bar{b}$. The class $\bar{b}$ (if it exists) can again be found by trial and error. In fact $\bar{b} = \bar{7}$ works, so $x = \overline{2b} = \overline{14}$, as before.

Fortunately, there is a systematic way of finding $\bar{b}$ in $\mathbb{Z}_{17}$ such that $\bar{b} \cdot \bar{5} = \bar{1}$. Note that 5 and 17 are relatively prime, so the euclidean algorithm can be used to express $\gcd(5, 17) = 1$ as a linear combination of 5 and 17. In fact, we have

$$17 = 3 \cdot 5 + 2 \quad \text{and then} \quad 5 = 2 \cdot 2 + 1;$$

so, eliminating remainders, $1 = 5 - 2(17 - 3 \cdot 5) = 7 \cdot 5 - 2 \cdot 17$. This implies that $7 \cdot 5 \equiv 1 \pmod{17}$, and so $\bar{7} \cdot \bar{5} = \bar{1}$ in $\mathbb{Z}_{17}$. This gives $\bar{b} = \bar{7}$.

This method clearly generalizes. For a modulus $n \geq 2$ and an integer a, a residue class $\bar{b}$ in $\mathbb{Z}_n$ is called an **inverse** of $\bar{a}$ if $\bar{b}\bar{a} = \bar{1}$ in $\mathbb{Z}_n$. If $\bar{a}$ has an inverse, that inverse is unique (Exercise 23) and we say $\bar{a}$ is **invertible**. Theorem 5 characterizes when an inverse exists, and the proof shows that (as above) the euclidean algorithm can be used to find it.

Theorem 5. *Let a and n be integers with $n \geq 2$. Then $\bar{a}$ has an inverse in $\mathbb{Z}_n$ if and only if a and n are relatively prime.*

Proof. If a and n are relatively prime, then $1 = \gcd(a, n)$ is a linear combination of a and n (by Theorem 4 §1.2), say $1 = ba + cn$, where b and c are integers. Hence, $ba \equiv 1 \pmod{n}$, so $\bar{b}\bar{a} = \bar{1}$ by Theorem 2. Conversely, if b exists such that $\bar{b}\bar{a} = \bar{1}$, then $ba \equiv 1 \pmod{n}$. Thus, $n|(1 - ba)$, say $1 - ba = qn$ for some integer q. But then $1 = ba + qn$, so a and n are relatively prime (again by Theorem 4 §1.2). ■

Example 8. Find the inverse of $\overline{16}$ in $\mathbb{Z}_{35}$ and use it to solve $\overline{16}x = \bar{9}$ in $\mathbb{Z}_{35}$.

Solution. The inverse exists as $\gcd(35, 16) = 1$. The euclidean algorithm gives

$$35 = 2 \cdot 16 + 3 \quad \text{and then} \quad 16 = 5 \cdot 3 + 1,$$

so $1 = 16 - 5(35 - 2 \cdot 16) = 11 \cdot 16 - 5 \cdot 35$. Thus, $11 \cdot 16 \equiv 1 \pmod{35}$, and so $\overline{11}$ is the inverse of $\overline{16}$ in $\mathbb{Z}_{35}$. Now multiply the equation $\overline{16}x = \bar{9}$ by $\overline{11}$ to obtain $\overline{11} \cdot \overline{16}x = \overline{11} \cdot \bar{9}$; that is, $x = \overline{99} = \overline{29}$. □

Example 9. Find the elements in $\mathbb{Z}_9$ that have inverses.

Solution. The members of $\mathbb{Z}_9$ are of the form $\bar{r}$, where $r = 0, 1, 2, \cdots, 8$. Since $9 = 3^2$, r is relatively prime to 9 if and only if r is not a multiple of 3. Hence, $\bar{1}, \bar{2}, \bar{4}, \bar{5}, \bar{7}$, and $\bar{8}$ will all have inverses. Indeed, $\bar{1}$ and $\bar{8}$ are both self-inverse, whereas $\bar{2}$ and $\bar{5}$ are inverses of each other as are $\bar{4}$ and $\bar{7}$. □

Example 10. Solve the system $\begin{cases} \bar{5}x + \bar{8}y = \bar{2} \\ \bar{3}x + \bar{2}y = \bar{1} \end{cases}$ of equations in $\mathbb{Z}_{11}$.

Solution. The usual techniques apply. Since $\bar{4} \cdot \bar{3} = \bar{1}$, we eliminate y by first multiplying the second equation by $\bar{4}$ to get $x + \bar{8}y = \bar{4}$. Subtract this from the first equation to get $\bar{4}x = -\bar{2} = \bar{9}$. Now $\bar{3}$ is the inverse of $\bar{4}$ in $\mathbb{Z}_{11}$, so multiplication by $\bar{3}$ gives $x = \bar{3} \cdot \bar{9} = \bar{5}$. Then the last equation gives $\bar{2}y = \bar{1} - \bar{3}x = \bar{8}$. Finally, $\bar{6}$ is the inverse of $\bar{2}$, so $y = \bar{6} \cdot \bar{8} = \bar{4}$. □

If a is a real number, an expression $x^2 + ax$ becomes a square if $(\frac{1}{2}a)^2$ is added: $x^2 + ax + \left(\frac{1}{2}a\right)^2 = \left(x + \frac{1}{2}a\right)^2$. This process is called **completing the square**, and it works in $\mathbb{Z}_n$ provided $\bar{2}$ has an inverse in $\mathbb{Z}_n$ (that is, if n is odd).

Example 11. Solve the quadratic $x^2 + \bar{3}x + \bar{9} = \bar{0}$ in $\mathbb{Z}_{13}$.

Solution. First subtract $\bar{9}$ from both sides to obtain $x^2 + \bar{3}x = -\bar{9} = \bar{4}$. The inverse of $\bar{2}$ in $\mathbb{Z}_{13}$ is $\bar{7}$, so we complete the square on the left by adding $(\bar{7} \cdot \bar{3})^2 = \bar{8}^2 = \overline{12}$ to both sides. The result is $x^2 + \bar{3}x + \overline{12} = \bar{4} + \overline{12}$, that is, $(x + \bar{8})^2 = \bar{3}$. Now $\mathbb{Z}_{13}$ has 13 elements and, by inspection, only 2 of them square to $\bar{3}$, namely, $\bar{4}$ and $-\bar{4} = \bar{9}$. Hence, $x + \bar{8} = \bar{4}$ or $x + \bar{8} = \bar{9}$, and so $x = \bar{9}$ and $x = \bar{1}$ are the solutions. □

Note that there are *two* solutions in Example 11. The reason is that $\bar{3}$ has two "square roots" in $\mathbb{Z}_{13}$: $\bar{4}$ and $-\bar{4} = \bar{9}$. However, other situations are possible: In $\mathbb{Z}_7$, $\bar{3}$ has no square root, whereas in $\mathbb{Z}_{27}$, $\bar{9}$ has six square roots, $\bar{3}$ and $-\bar{3} = \overline{24}$, $\bar{6}$ and $-\bar{6} = \overline{21}$, and finally $\overline{12}$ and $-\overline{12} = \overline{15}$.

The following fact about congruences is useful in number theory and computer science, and was known to the Chinese in the fourth century.

Theorem 6. ***Chinese Remainder Theorem***. *Let m and n be relatively prime integers. If s and t are arbitrary integers, there exists a solution $x \in \mathbb{Z}$ to the simultaneous congruences*

$$x \equiv s \pmod{m} \quad \textit{and} \quad x \equiv t \pmod{n}.$$

Proof. Since $\gcd(m, n) = 1$, the euclidean algorithm gives p and q in $\mathbb{Z}$ such that $1 = mp + nq$. Take

$$x = (mp)t + (nq)s.$$

Then $x - s = mpt + (nq - 1)s = mp(t - s)$, so $x \equiv s \pmod{m}$. A similar argument gives $x \equiv t \pmod{n}$. ■

The nice thing about Theorem 6 is that the proof gives an algorithm for finding the solution x: The euclidean algorithm gives p and q such that $1 = mp + nq$, and the solution is $x = mpt + nqs$. Furthermore, this method can be iterated to solve a system of more than two congruences, provided that only the moduli are relatively prime in pairs. To illustrate, let m_1, m_2, and m_3 be integers relatively prime in pairs. Given arbitrary integers s_1, s_2, and s_3, we want to find an integer x such that

$$x \equiv s_i \pmod{m_i} \quad \text{for each } i = 1, 2, 3.$$

The Chinese remainder theorem yields a such that $a \equiv s_i \pmod{m_i}$ for $i = 1, 2$. Since m_1m_2 and m_3 are relatively prime, apply the Chinese remainder theorem again to obtain x such that

$$x \equiv a \pmod{m_1m_2} \quad \text{and} \quad x \equiv s_3 \pmod{m_3}.$$

But then $x \equiv a \pmod{m_1}$, so since $a \equiv s_1 \pmod{m_1}$, we have $x \equiv s_1 \pmod{m_1}$. Similarly, $x \equiv s_2 \pmod{m_2}$.

In general, if $m_1, m_2, \ldots, m_k$ are relatively prime in pairs, and if $s_1, s_2, \ldots, s_k$ are arbitrary integers, then there exists $x \in \mathbb{Z}$ such that

$$x \equiv s_i \pmod{m_i} \quad \text{for each } i = 1, 2, \ldots, k.$$

These general systems of congruences are important in computer science because they provide a method for doing arithmetic with integers that exceed the *word size* of the computer (the largest integer that can be used in machine arithmetic).

The only elements of $\mathbb{Z}$ that have an inverse in $\mathbb{Z}$ are 1 and -1 (because $\frac{1}{k}$ does not lie in $\mathbb{Z}$ if $k \neq 1, -1$). Thus, $\mathbb{Z}$ resembles $\mathbb{Z}_6$ in this respect (see the table following Theorem 4). At the other extreme, *every* nonzero real number $x \neq 0$ has an inverse $\frac{1}{x}$ in $\mathbb{R}$. Theorem 7 characterizes when this happens in $\mathbb{Z}_n$.

Theorem 7. *The following are equivalent for an integer* $n \geq 2$.

(1) *Every element* $\bar{a} \neq \bar{0}$ *in* $\mathbb{Z}_n$ *has an inverse.*

(2) *If* $\bar{a}\bar{b} = \bar{0}$ *in* $\mathbb{Z}_n$, *then either* $\bar{a} = \bar{0}$ *or* $\bar{b} = \bar{0}$.

(3) n *is a prime.*

Proof. We prove that (1) ⇒ (2), (2) ⇒ (3), and (3) ⇒ (1).

(1) ⇒ (2). Assume (1) is true and let $\bar{a}\bar{b} = \bar{0}$ in $\mathbb{Z}_n$. If $\bar{a} = \bar{0}$, there is nothing to prove. Otherwise, $\bar{a}$ has an inverse by (1), say $\bar{c}\bar{a} = \bar{1}$. Then we multiply both sides of $\bar{a}\bar{b} = \bar{0}$ by $\bar{c}$ to get $\bar{c}\bar{a}\bar{b} = \bar{c}\bar{0}$; that is, $\bar{b} = \bar{0}$.

(2) ⇒ (3). If n is not prime, let $n = ab$, where $2 \leq a < n$ and $2 \leq b < n$. But then $\bar{a}\bar{b} = \bar{n} = \bar{0}$, where $\bar{a} \neq \bar{0}$ and $\bar{b} \neq \bar{0}$. This contradicts (2), so the assumption that n is not prime cannot be valid.

(3) ⇒ (1). If n is prime, let $\bar{a} \neq \bar{0}$ in $\mathbb{Z}_n$. Then $\gcd(a, n) = 1$ (because otherwise $\gcd(a, n) = n$, so $n|a$). But then $1 = ba + cn$ for integers b and c (by Theorem 4 §1.2), so $ba \equiv 1 \pmod{n}$. Thus, $\bar{b}\bar{a} = \bar{1}$ in $\mathbb{Z}_n$, proving (1). ■

Hence, if p is a prime, $\mathbb{Z}_p$ has the property that every nonzero element has an inverse. This is also true of the real numbers $\mathbb{R}$, and such systems are called **fields**.

The following consequence of Theorem 7 will be referred to later.

Corollary. ***Wilson's Theorem.*** *If* p *is a prime, then* $(p-1)! \equiv -1 \pmod{p}$.

Proof. We write $\bar{a} = a$ in $\mathbb{Z}_p$ for convenience. Since p is prime, each element $1, 2, 3, \dots, p-1$ in $\mathbb{Z}_p$ has an inverse by Theorem 7. Hence, pairs of inverses in the product $(p-1)! = 1\ 2\ 3 \cdots (p-1)$ will cancel leaving only the self-inverse elements 1 and -1 (Exercise 26). Thus, $(p-1)! = 1\,(-1) = -1$ in $\mathbb{Z}_p$, as required. ■

Example 12. Write down the multiplication table of $\mathbb{Z}_5$ and illustrate Theorem 7.

Solution. The first row and column of the table consist entirely of zeros (true for any modulus), but the fact that no other entry equals $\bar{0}$ verifies (2) of Theorem 7. Similarly, the fact that every row (or column) except the first contains $\bar{1}$ verifies (1) of Theorem 7.

$\times$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{4}$	$\bar{1}$	$\bar{3}$
$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{1}$	$\bar{4}$	$\bar{2}$
$\bar{4}$	$\bar{0}$	$\bar{4}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

The simplest situation in which Theorem 7 applies is when $n = 2$. In this case, $\mathbb{Z}_2 = \{\bar{0}, \bar{1}\}$ and the addition and multiplication tables are as follows:

$+$	$\bar{0}$	$\bar{1}$
$\bar{0}$	$\bar{0}$	$\bar{1}$
$\bar{1}$	$\bar{1}$	$\bar{0}$

$\times$	$\bar{0}$	$\bar{1}$
$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$

This is binary arithmetic, which is important in the design of computers.

We conclude with a famous theorem of Pierre de Fermat. In Example 5, we showed that $a^5 \equiv a \pmod 5$ holds for all integers a. In fact, it holds if we replace 5 by any prime.

Theorem 8. ***Fermat's Theorem.*** *If p is a prime, then*

$$a^p \equiv a \pmod p \quad \textit{for all integers } a.$$

In fact, $a^{p-1} \equiv 1 \pmod p$ for all integers a that are relatively prime to p.

Proof. We must show that $\bar{a}^p = \bar{a}$ in $\mathbb{Z}_p$. Because this equation is true if $\bar{a} = \bar{0}$, it suffices to show that $\bar{a}^{p-1} = \bar{1}$ in $\mathbb{Z}_p$ whenever $\bar{a} \neq \bar{0}$. But if $\bar{a} \neq \bar{0}$, then $\bar{a}$ has an inverse in $\mathbb{Z}_p$ by Theorem 7, say $\bar{b}\bar{a} = \bar{1}$. Now multiply all the nonzero elements in $\mathbb{Z}_p$ by $\bar{a}$ to obtain

$$\bar{a}\bar{1},\ \bar{a}\bar{2}, \dots, \bar{a}(\overline{p-1}).$$

These are all distinct (because $\bar{a}\bar{r} = \bar{a}\bar{s}$ yields $\bar{r} = \bar{s}$ after multiplication by $\bar{b}$) and none equals $\bar{0}$, so they must be the set of *all* nonzero elements $\bar{1},\ \bar{2}, \dots, \overline{p-1}$ in some order. In particular, the products are the same, and we obtain

$$\bar{a}^{p-1}(\bar{1}\ \bar{2}\ \cdots\ \overline{p-1}) = \bar{1}\ \bar{2}\ \cdots\ \overline{p-1}.$$

But the element $\bar{1}\,\bar{2}\ \cdots\ \overline{p-1}$ is invertible in $\mathbb{Z}_p$ (Exercise 24). Hence, multiplication by its inverse gives $\bar{a}^{p-1} = \bar{1}$, which is what we wanted. ■

Note that Fermat's theorem fails if p is not prime; for example, $2^4 \not\equiv 2 \pmod 4$.

Fermat's theorem is important in number theory, and the following result will be referred to several times. To state it, we use the following useful observation (Exercise 36): If prime $p > 2$ is a prime, then $p \equiv 1 \pmod 4$ or $p \equiv 3 \pmod 4$.

Corollary. *Let $p > 2$ be a prime.*

(1) *If $p \equiv 1 \pmod 4$, then $x^2 = -\bar{1}$ in $\mathbb{Z}_p$, where $x = \bar{1}\ \bar{2}\ \cdots\ \overline{\tfrac{1}{2}(p-1)}$.*

(2) *If $p \equiv 3 \pmod 4$, then the equation $x^2 = -\bar{1}$ has no solution in $\mathbb{Z}_p$.*

Proof. Write $\bar{a} = a$ in $\mathbb{Z}_p$ for convenience.

(1) We have $(p-1)! = -1$ by the Corollary to Theorem 7. Write

$$q = \tfrac{1}{2}(p+1)\ \cdots\ (p-2)\ (p-1).$$

Then,

$$xq = [1\ 2\ \cdots\ \tfrac{1}{2}(p-1)]\ \ [\tfrac{1}{2}(p+1)\ \cdots\ (p-2)\ (p-1)] = (p-1)! = -1.$$

Thus, it suffices to show that $q = x$. Now observe that we can write q as follows:

$$q = (-\tfrac{1}{2}(p-1)) \cdots (-2)\,(-1).$$

Since $p \equiv 1 \pmod 4$, the integer $\frac{1}{2}(p-1)$ is even. Hence, q has an even number of factors, and it follows that $q = x$ after all. This proves (1).

(2) Let $p = 4n+3$ in $\mathbb{Z}$. Suppose $a \in \mathbb{Z}_p$ satisfies $a^2 = -1$ in $\mathbb{Z}_p$; we look for a contradiction. Since $a^{p-1} = 1$ by Fermat's theorem, we have

$$1 = a^{p-1} = a^{4n+2} = (a^2)^{2n+1} = (-1)^{2n+1} = -1 \text{ in } \mathbb{Z}_p,$$

a contradiction because $p > 2$. So $x^2 = -1$ has no solution in $\mathbb{Z}_p$, proving (2). ■

Clearly, a residue class $\bar{a}$ is not the same thing as the integer a. However, because of the definitions $\bar{a}+\bar{b}=\overline{a+b}$ and $\bar{a}\bar{b}=\overline{ab}$ in $\mathbb{Z}_n$, the arithmetic of $\mathbb{Z}_n$ closely resembles that of $\mathbb{Z}$—so much so that in subsequent chapters we adopt the following convention (used above in the Corollaries to Theorems 7 and 8):

Notational Convention. *When working in $\mathbb{Z}_n$ we frequently write the residue class $\bar{a}$ simply as a.*

Then $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$, and equations such as $3 \cdot 4 = 2$ and $2 + 3 = 0$ appear. This notation is harmless, once everyone knows that we are using it, and it facilitates hand calculations (the reader as probably been using it already!). Of course, when the convention causes confusion, we revert to the more formal $\bar{a}$ notation.

Pierre De Fermat (1601–1685) Fermat was a lawyer by profession and served in the parliament in Toulouse, France. His mathematical work was a pastime, and he has been called "the prince of amateurs." This appellation should not be taken as diminishing his stature, because he did first-rate work in several areas. He invented analytic geometry prior to Descartes and made contributions to the development of calculus. Along with Pascal, he is credited with starting the theory of probability.

However, he is most remembered for his work in number theory. Theorem 8 first appeared in a letter in 1640, and a proof was first published much later by Euler. Fermat published virtually nothing, and his results became known through letters to his friends (many to Mersenne) and as notes jotted in the margin of his copy of Arithmetica by Diophantus, usually with no proof. The most famous of these notes is the assertion that, if $n \geq 3$, positive integers x, y, and z do not exist such that $x^n + y^n = z^n$. This assertion has become known as "Fermat's Last Theorem", and he wrote that "I have found a truly remarkable proof but the margin was too small to contain it." His intuition was so good that every other theorem that he claimed he could prove has been subsequently verified. However, despite the best efforts of the greatest mathematicians, the "Last Theorem" remained open for 300 years. But in 1997, in a spectacular display of mathematical virtuosity, Andrew Wiles of Princeton University finally proved the result. Wiles related Fermat's conjecture to a problem in geometry, which he solved.

Exercises 1.3

1. In each case determine whether the statement is true or false.
(a) $40 \equiv 13 \pmod 9$ (b) $-29 \equiv 1 \pmod 7$
(c) $-29 \equiv 6 \pmod 7$ (d) $132 \equiv 0 \pmod{11}$
(e) $8 \equiv 8 \pmod n$ (f) $3^4 \equiv 1 \pmod 5$
(g) $8^4 \equiv 2 \pmod{13}$

2. In each case find all integers k making the statement true.
(a) $4 \equiv 2k \pmod 7$ (b) $12 \equiv 3k \pmod{10}$
(c) $3k \equiv k \pmod 9$ (d) $5k \equiv k \pmod{15}$

3. Find all integers $k \geq 2$ such that
(a) $-3 \equiv 7 \pmod k$ (b) $7 \equiv -5 \pmod k$
(c) $3 \equiv k^2 \pmod k$ (d) $5 \equiv k \pmod{k^2}$

4. Find all integers $k \geq 2$ such that $k^2 \equiv 5k \pmod{15}$.

5. (a) Show that congruence modulo 0 is equality.
(b) What can you say about congruence modulo 1?

6. (a) Prove Theorem 1.
(b) Prove (1)–(4) of Theorem 4.

7. If $a \equiv b \pmod{n}$ and $m|n$, show that $a \equiv b \pmod{m}$.

8. Find the remainder when
(a) 10^{515} is divided by 7 (b) 8^{391} is divided by 5

9. Find the unit decimal digit of
(a) 3^{1027} (b) 27^{2113}

10. Show that the unit decimal digit of k^4 must be $0, 1, 5$, or 6 for all integers k.

11. If $p \neq 2, 3$ is prime, show that $\bar{p} = \bar{1}$ or $\bar{p} = \bar{5}$ in $\mathbb{Z}_6$.

12. (a) If a is an integer, show that $a^2 \equiv 0$ or $a^2 \equiv 1 \pmod{4}$.
(b) Show that none of $11, 111, 1111, 11111, \ldots,$ is a perfect square.

13. Show that a^5 is congruent to $0, 1$, or -1 mod 11 for every integer a.

14. Show that $\bar{a}^7 = \bar{a}$ in $\mathbb{Z}_7$ for every integer a using the method of Example 5.

15. Show that $\bar{a}(\bar{a} + \bar{1})(\bar{a} + \bar{2}) = \bar{0}$ in $\mathbb{Z}_6$ for every integer a.

16. Show that $a^3 + 2$ is not divisible by 7 for every integer a.

17. Show that $\bar{a}^3 = \bar{a}$ in $\mathbb{Z}_6$ for every integer a.

18. (a) Show that every integer a has a cube root in $\mathbb{Z}_5$ ($\bar{a} = \bar{b}^3$ for some integer b).
(b) If $n \geq 3$, show that some integer has no square root in $\mathbb{Z}_n$.

19. (a) Show that no integer of the form $k^2 + 1$ is a multiple of 7.
(b) Find all integers k such that $k^2 + 1$ is a multiple of 17.

20. If a space mission takes exactly 175 hours and the craft blasts off at 8 a.m., at what hour of the day will it land?

21. Let $n = d_k d_{k-1} \cdots d_2 d_1 d_0$ be the decimal representation of n.
(a) Show that $3|n$ if and only if 3 divides $(d_0 + d_1 + \cdots + d_k)$.
(b) Show that $11|n$ if and only if 11 divides $(d_0 - d_1 + d_2 - d_3 + \cdots \pm d_k)$.
(c) Show that $6|n$ if and only if 6 divides $[d_0 + 4(d_1 + d_2 + \cdots + d_k)]$.

22. (a) In $\mathbb{Z}_{35}$, find the inverse of $\overline{13}$ and use it to solve $\overline{13}x = \bar{9}$.
(b) In $\mathbb{Z}_{25}$, find the inverse of $\bar{7}$ and use it to solve $\bar{7}x = \overline{12}$.
(c) In $\mathbb{Z}_{20}$, find the inverse of $\overline{11}$ and use it to solve $\overline{11}x = \overline{16}$.
(d) In $\mathbb{Z}_{16}$, find the inverse of $\bar{9}$ and use it to solve $\bar{9}x = \overline{14}$.

23. (a) If $\bar{a}\bar{b} = \bar{a}\bar{c}$ in $\mathbb{Z}_n$ and if $\bar{a}$ has an inverse in $\mathbb{Z}_n$, show that $\bar{b} = \bar{c}$.
(b). If $\bar{a}$ has an inverse in $\mathbb{Z}_n$, show that the inverse is unique.

24. (a) If $\bar{a}$ and $\bar{b}$ both have inverses in $\mathbb{Z}_n$, show that the same is true for $\bar{a}\bar{b}$.
(b) If $\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_m$ all have inverses in $\mathbb{Z}_n$, show that the same is true of their product $\bar{a}_1 \bar{a}_2 \cdots \bar{a}_m$.

25. Find all solutions in $\mathbb{Z}_n$ (as indicated) for each of the given equations.

(a) $\begin{cases} \bar{3}x + \bar{2}y = \bar{1} \\ \bar{5}x + \ \ y = \bar{1} \end{cases}$ in $\mathbb{Z}_{11}$ (b) $\begin{cases} \bar{3}x + \bar{4}y = \bar{1} \\ \bar{2}x + \ \ y = \bar{1} \end{cases}$ in $\mathbb{Z}_7$

(c) $\begin{cases} \bar{3}x + \bar{2}y = \bar{1} \\ \bar{5}x + \ \ y = \bar{1} \end{cases}$ in $\mathbb{Z}_7$ (d) $\begin{cases} \bar{3}x + \bar{4}y = \bar{1} \\ \bar{2}x + \ \ y = \bar{1} \end{cases}$ in $\mathbb{Z}_5$

(e) $\begin{cases} \bar{3}x + \bar{2}y = \bar{1} \\ \bar{5}x + \ \ y = \bar{4} \end{cases}$ in $\mathbb{Z}_7$ (f) $\begin{cases} \bar{3}x + \bar{4}y = \bar{1} \\ \bar{2}x + \ \ y = \bar{4} \end{cases}$ in $\mathbb{Z}_5$

26. If p is a prime and $x^2 = \bar{a}^2$ in $\mathbb{Z}_p$, show that $x = \bar{a}$ or $x = -\bar{a}$.

27. (a) Find all x in $\mathbb{Z}_7$ such that $x^2 + \bar{5}x + \bar{4} = \bar{0}$.
(b) Find all x in $\mathbb{Z}_5$ such that $x^2 + x + \bar{3} = \bar{0}$.
(c) Find all x in $\mathbb{Z}_5$ such that $x^2 + x + \bar{2} = \bar{0}$.
(d) Find all x in $\mathbb{Z}_9$ such that $x^2 + x + \bar{7} = \bar{0}$.
(e) Let n be odd. Show that $\bar{2}$ has an inverse $\bar{r}$ in $\mathbb{Z}_n$. Show that $x^2 + \bar{a}x + \bar{b} = \bar{0}$ has a solution in $\mathbb{Z}_n$ if and only if $(\overline{r^2a^2 - b})$ is a square in $\mathbb{Z}_n$.

28. Find $x \in \mathbb{Z}$ such that $x \equiv 8 \pmod{10}$, $x \equiv 3 \pmod{9}$, and $x \equiv 2 \pmod{7}$.

29. (a) If $\bar{a}\bar{b} = \bar{0}$ in $\mathbb{Z}_n$ and $\gcd(a, n) = 1$, show that $\bar{b} = \bar{0}$.

(b) Show that $\bar{a}$ is invertible in $\mathbb{Z}_n$ if and only if $\bar{a}\bar{b} = \bar{0}$ implies that $\bar{b} = \bar{0}$.

30. Show that the following conditions on an integer $n \geq 2$ are equivalent.

(1) $\bar{a}^2 = \bar{0}$ in $\mathbb{Z}_n$ implies that $\bar{a} = \bar{0}$.

(2) n is square free (that is, a product of distinct primes).

[*Hint:* Theorem 5 §1.2.]

31. Show that the following conditions on an integer $n \geq 2$ are equivalent.

(1) If $\bar{a}$ is in $\mathbb{Z}_n$, then either $\bar{a}$ is invertible or $\bar{a}^k = \bar{0}$ for some $k \geq 1$.

(2) n is a power of a prime.

32. If $p \geq 3$ is a prime, show that every element of $\mathbb{Z}_p$ has a $(p-2)$th root. [*Hint:* Use Fermat's theorem to show that $f\colon \mathbb{Z}_p \to \mathbb{Z}_p$ is one-to-one, where $f(\bar{a}) = \bar{a}^{p-2}$. Apply Theorem 2 §0.3.]

33. Show that $2^{37} - 1$ is divisible by 223 and that $2^{32} + 1$ is divisible by 641. (Remarkably, $\frac{1}{223}(2^{37} - 1)$ is *also* prime.) *Note:* If p is a prime, numbers of the form $2^p - 1$ and $2^{2^n} + 1$ are called **Mersenne numbers** and **Fermat numbers**, respectively, and were once thought to be all primes.

34. Let a and n denote integers with $n \geq 2$, and write $d = \gcd(a, n)$.

(a) Show that $ax \equiv b \pmod{n}$ has a solution if and only if $d|b$.

(b) If $d = ra + sn$, r and s integers, show that $x_0 = r(b/d)$ is one solution.

(c) If x_0 is any solution, show that there are exactly d solutions that are distinct modulo n: $\left\{x_0, x_0 + \frac{n}{d}, x_0 + 2\frac{n}{d}, \ldots, x_0 + (d-1)\frac{n}{d}\right\}$. [*Hint:* If $ax \equiv b \pmod{n}$, show that $a(x - x_0) \equiv 0 \pmod{n}$, so $(a/d)(x - x_0) \equiv 0$ [mod(n/d)] by Exercise 11 §1.2. Conclude that $x - x_0 \equiv 0$ [mod(n/d)].]

(d) Find all solutions to $15x \equiv 25 \pmod{35}$.

(e) Find all solutions to $21x \equiv 14 \pmod{35}$.

(f) Find all solutions to $21x \equiv 8 \pmod{33}$.

35. Let p be a prime. If $x^2 = \bar{1}$ in $\mathbb{Z}_p$, show that $x = \bar{1}$ or $x = -\bar{1}$.

36. Let p be a prime, show that either $p \equiv 1 \pmod{4}$ or $p \equiv 3 \pmod{4}$.

37. (a) Show that if $a^n \equiv a \pmod{n}$ holds for all integers a, the modulus n must be square free, that is, a product of distinct primes.

(b) Show that $a^{561} \equiv a \pmod{561}$ for all integers a. [*Hint:* Use Theorem 5 §1.2 to reduce the problem to showing that $a^{561} \equiv a \pmod{p}$, where $p = 3, 11$, or 17. In each case, use Fermat's theorem in the form $a^{p-1} \equiv 1 \pmod{p}$ whenever p does not divide a.]

1.4 PERMUTATIONS

A permutation of the numbers 1, 2, and 3 is a rearrangement of these numbers in a definite order. Thus, the six possibilities are

$$1\ 2\ 3 \qquad 1\ 3\ 2 \qquad 2\ 1\ 3 \qquad 2\ 3\ 1 \qquad 3\ 1\ 2 \qquad 3\ 2\ 1$$

They can also be described as mappings $\{1, 2, 3\} \to \{1, 2, 3\}$:

$$\begin{array}{cccccc} 1 \to 1 & 1 \to 1 & 1 \to 2 & 1 \to 2 & 1 \to 3 & 1 \to 3 \\ 2 \to 2 & 2 \to 3 & 2 \to 1 & 2 \to 3 & 2 \to 1 & 2 \to 2 \\ 3 \to 3 & 3 \to 2 & 3 \to 3 & 3 \to 1 & 3 \to 2 & 3 \to 1 \end{array}$$

We use this terminology of mappings to describe permutations.

If X and Y are sets, recall that a mapping $\alpha : X \to Y$ is a rule that assigns to every element x of X exactly one element $\alpha(x)$ of Y, called the image of x under α. Hence, the diagram

$$\begin{aligned} 1 &\to 1 \\ 2 &\to 3 \\ 3 &\to 2 \end{aligned}$$

describes the mapping $\alpha : \{1,2,3\} \to \{1,2,3\}$ given by the rule $\alpha(1) = 1$, $\alpha(2) = 3$, $\alpha(3) = 2$.

Now consider a mapping $\alpha \colon \{1,2,\ldots,n\} \to \{1,2,\ldots,n\}$. Because such mappings occur frequently, we write $\alpha(k) = \alpha k$ for simplicity. Our interest is in when the images $\alpha 1,\ \alpha 2,\ \ldots\ , \alpha n$ are a *permutation* of the numbers $1,2,\ldots,n$; that is, each element of $\{1,2,\ldots,n\}$ occurs *exactly once* in the list $\alpha 1, \alpha 2, \ldots, \alpha n$. In other words, the function α is both one-to-one and onto (a **bijection**).[13]

Given an integer $n \geq 1$, write $X_n = \{1,2,\ldots,n\}$.

A **permutation** of X_n is a bijection $\sigma \colon X_n \to X_n$.

We call the set S_n of all permutations of X_n the **symmetric group of degree** n. Two permutations σ and τ in S_n are **equal** if they are equal as functions, that is, if $\sigma k = \tau k$ for all k in X_n.

To simplify the manipulation of these permutations, a matrix-type notation is useful. For example, if the permutation $\sigma \colon X_4 \to X_4$ is defined by $\sigma 1 = 3$, $\sigma 2 = 1$, $\sigma 3 = 4$, and $\sigma 4 = 2$, we write it as

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}.$$

Here the image of each element of $X_4 = \{1,2,3,4\}$ is written below that element. In general, a permutation $\sigma \in S_n$ is written in matrix form as

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma 1 & \sigma 2 & \cdots & \sigma n \end{pmatrix}.$$

Hence, a typical member of S_n takes this form, where $\sigma 1,\ \sigma 2,\ \ldots\ , \sigma n$ is the list of numbers $1,2,\ldots,n$ in a (possibly) different order.

Example 1. List the elements of S_3 in matrix notation.

Solution. There are six different permutations:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

In general, to construct a permutation

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma 1 & \sigma 2 & \cdots & \sigma n \end{pmatrix},$$

we must choose the numbers $\sigma 1,\ \sigma 2,\ \ldots,\ \sigma n$ from X_n so that they are all distinct. Hence, we have n choices for $\sigma 1$, then $n-1$ choices for $\sigma 2$, then $n-2$ choices for

[13] A review of one-to-one and onto mappings can be found in Section 0.3.

$\sigma 3$, and so on. Thus, σ can be chosen in $n(n-1)(n-2)\cdots 3\cdot 2\cdot 1 = n!$ ways, which proves the following theorem:

Theorem 1. *The set S_n of permutations of X_n has $|S_n| = n!$ elements.*

Let σ and τ be permutations in S_n. Both are mappings from X_n to X_n, and we write them as follows:

$$X_n \xrightarrow{\tau} X_n \xrightarrow{\sigma} X_n.$$

We then define the *composite* $\sigma\tau\colon X_n \to X_n$ by first applying τ and then σ:

$$(\sigma\tau)k = \sigma(\tau k), \quad \text{for all } k \in X_n.$$

Because both σ and τ are one-to-one and onto, these properties hold for the composite $\sigma\tau$ (see Theorem 3 §0.3). Hence, $\sigma\tau$ is again a permutation in S_n.

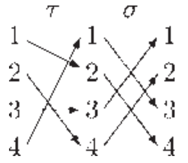
Example 2. Compute $\sigma\tau$ if

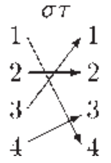
$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} \quad \text{and} \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}.$$

Solution. Consider the action of $\sigma\tau$ on 1: $(\sigma\tau)1 = \sigma 2 = 4$. We can compute it directly from the matrix forms:

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}.$$

It is important to remember that, in computing $\sigma\tau$, we apply τ first and then σ. Thus, we read $1 \xrightarrow{\tau} 2$ from the matrix for τ, then $2 \xrightarrow{\sigma} 4$ from the matrix for σ. The result is $1 \xrightarrow{\sigma\tau} 4$, as indicated. Similarly, $2 \xrightarrow{\tau} 4 \xrightarrow{\sigma} 2$ leads to $2 \xrightarrow{\sigma\tau} 2$. We can read the entire action of $\sigma\tau$ in this manner. The following diagrams illustrate what is happening:





The action of $\sigma\tau$ is read from the first diagram by following the arrows. □

Note that $\sigma\tau \neq \tau\sigma$ in general: If σ and τ are as in Example 2,

$$\tau\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}$$

is not the same as $\sigma\tau$ (computed in Example 2). If it happens that $\sigma\tau = \tau\sigma$, we say that σ and τ *commute*. Thus, two permutations need not commute (but see Theorem 3). On the other hand, if σ, τ, and μ are three permutations in S_n then we always have

$$(\sigma\tau)\mu = \sigma(\tau\mu),$$

which we can easily verify directly (see Theorem 3 §0.3).

The **identity permutation** ε in S_n is defined as

$$\varepsilon = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}.$$

In other words, $\varepsilon k = k$ holds for every $k \in X_n$. It is easy to verify that

$$\varepsilon\sigma = \sigma = \sigma\varepsilon$$

holds for all $\sigma \in S_n$, so ε plays the role in S_n that 1 plays for multiplication of numbers.

Consider the permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix}$$

in S_4. The action of σ is obtained by *reading down*: $\sigma 1 = 3$, $\sigma 2 = 4$, $\sigma 3 = 2$, and $\sigma 4 = 1$. There is clearly another permutation in S_4 obtained by *reading up* $3 \to 1$, $4 \to 2$, $2 \to 3$, and $1 \to 4$. This new permutation is determined uniquely by σ; In fact, it is the inverse of σ (denoted σ^{-1} as in Section 0.3). Thus,

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix}.$$

In general, if $\sigma \in S_n$, the fact that $\sigma \colon X_n \to X_n$ is one-to-one and onto implies (Theorem 6 §0.3) that a uniquely determined permutation $\sigma^{-1} \colon X_n \to X_n$ exists (called the **inverse** of σ), which satisfies

$$\sigma(\sigma^{-1}k) = k \quad \text{and} \quad \sigma^{-1}(\sigma k) = k, \quad \text{for all } k \in X_n. \qquad (*)$$

Equations (*) imply that each of σ and σ^{-1} reverses the action of the other and hence that we can indeed obtain the action of σ^{-1} from

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma 1 & \sigma 2 & \cdots & \sigma n \end{pmatrix}$$

by reading up.

Example 3. Find the inverse of $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 1 & 8 & 3 & 2 & 5 & 6 & 7 \end{pmatrix}$ in S_8.

Solution. Reversing the action of σ gives $\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 5 & 4 & 1 & 6 & 7 & 8 & 3 \end{pmatrix}$. □

If $\sigma \in S_n$, it is related to σ^{-1} by composition. Indeed, because the identity permutation ε in S_n satisfies $\varepsilon k = k$ for all $k \in X_n$, we can write equations (*) as

$$\sigma\sigma^{-1} = \varepsilon \quad \text{and} \quad \sigma^{-1}\sigma = \varepsilon.$$

This and other properties of composition discussed earlier are recorded in the following theorem for reference.

Theorem 2. *Let σ, τ, and μ denote permutations in S_n.*

(1) *$\sigma\tau$ is in S_n.*

(2) *$\sigma\varepsilon = \sigma = \varepsilon\sigma$.*

(3) *$\sigma(\tau\mu) = (\sigma\tau)\mu$.*

(4) *$\sigma\sigma^{-1} = \varepsilon = \sigma^{-1}\sigma$.*

By virtue of this, S_n is said to be a *group under composition* that explains the name "symmetric group." Groups in general are discussed in Chapter 2.

Example 4. Given

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 1 & 2 & 3 \end{pmatrix}$$

and

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 1 & 5 & 4 \end{pmatrix},$$

find χ in S_5 such that $\chi\sigma = \tau$.

Solution. Suppose that $\chi \in S_n$ exists such that $\tau = \chi\sigma$. Multiply on the right by σ^{-1} to get $\tau\sigma^{-1} = \chi\sigma\sigma^{-1} = \chi\varepsilon = \chi$. Thus,

$$\chi = \tau\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 1 & 5 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 5 & 4 & 3 & 2 \end{pmatrix}.$$

The reader should verify that χ actually works, that is, $\chi\sigma = \tau$. □

Let $\sigma \in S_n$ so that $\sigma\colon X_n \to X_n$ is a bijection. We say that an element $k \in X_n$ is **fixed** by σ if $\sigma k = k$. If $\sigma k \neq k$, we say that k is **moved** by σ, and we write $M_\sigma = \{k \in X_n \mid k \text{ is moved by } \sigma\}$. Two permutations σ and τ are called **disjoint** if no element of X_n is moved by both; that is, if $M_\sigma \cap M_\tau = \varnothing$.

Clearly, the identity permutation ε in S_n is the only permutation that fixes every element of X_n. By contrast,

$$\begin{pmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ 2 & 3 & 4 & \cdots & n & 1 \end{pmatrix}$$

moves every element of X_n, whereas

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 5 & 4 & 1 \end{pmatrix}$$

moves $1, 3$, and 5 and fixes 2 and 4. The following result is needed in the proof of Theorem 3.

Lemma 1[14]. *If $k \in M_\sigma$ then $\sigma k \in M_\sigma$.*

Proof. Otherwise, σk is fixed by σ; that is, $\sigma(\sigma k) = \sigma k$. But then the fact that σ is one-to-one gives $\sigma k = k$, which is contrary to the hypothesis. ■

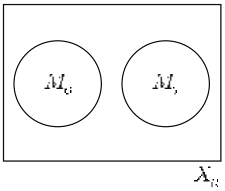
Theorem 3. *If σ and τ in S_n are disjoint, then $\sigma\tau = \tau\sigma$.*

Proof. For $k \in X_n$, we must show that $(\tau\sigma)k = (\sigma\tau)k$. Since $M_\sigma \cap M_\tau = \varnothing$ by hypothesis, there are three cases (see the diagram).

[14] The word "lemma" means a subsidiary proposition used in the proof of another proposition.

- *Case 1*: $k \in M_\sigma$. Then $\sigma k \in M_\sigma$ too (by Lemma 1), so neither lies in M_τ. Hence, both are fixed by τ, so $\tau k = k$ and $\tau(\sigma k) = \sigma k$. Hence,

$$(\tau\sigma)k = \tau(\sigma k) = \sigma k = \sigma(\tau k) = (\sigma\tau)k.$$



- *Case 2*: $k \in M_\tau$. This case is analogous to Case 1, and is left to the reader.

- *Case 3*: $k \notin M_\sigma$ and $k \notin M_\tau$. Then $\sigma k = k$ and $\tau k = k$, so

$$(\tau\sigma)k = \tau(\sigma k) = \tau k = k = \sigma k = \sigma(\tau k) = (\sigma\tau)k.$$

This completes the proof. ■

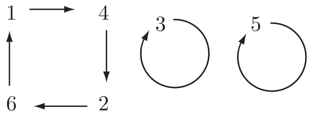
Note that the converse to Theorem 3 is not true. For example, $\sigma\sigma^{-1} = \sigma^{-1}\sigma$ for any σ in S_n, but σ and σ^{-1} are certainly not disjoint. Theorem 3 is important because it leads to a proof of the fact (Theorem 5 below) that every permutation in S_n can be written as a product of pairwise disjoint (and commuting) factors. We now turn our attention to this topic.

Cycles

Consider the permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 6 & 3 & 2 & 5 & 1 \end{pmatrix}$$

in S_6. The action of σ is described graphically as



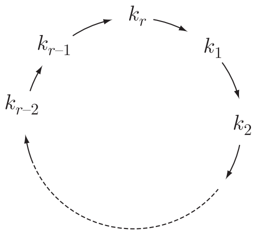
Thus, the elements σ moves are moved in a cycle, and σ is called a *cycle* for this reason. We write σ as $\sigma = (1 \;\; 4 \;\; 2 \;\; 6)$. This notation lists only elements moved by σ, and each is moved to its neighbor to the right, except the last element, which "cycles around" to the first. We generalize this type of permutation as follows.

Let $k_1, k_2, \ldots, k_r$ be distinct elements of X_n. Then, as shown in the diagram, the **cycle**

$$\sigma = (k_1 \;\; k_2 \;\; \cdots \;\; k_r)$$

is the permutation in S_n defined by

$$\begin{aligned} \sigma k_i &= k_{i+1}, \;\; \text{if } 1 \le i \le r-1. \\ \sigma k_r &= k_1 \\ \sigma k &= k, \qquad \text{if } k \notin \{k_1, k_2, \ldots, k_r\} \end{aligned}$$



We say that σ has **length** r and refer to σ as an ***r*-cycle**. Note that the only cycle of length 1 is ε, that is $(k) = \varepsilon$ for each $k \in X_n$.

Example 5. Write

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 7 & 1 & 6 & 5 & 2 & 3 \end{pmatrix}$$

in cycle notation.

Solution. $\tau = (1 \ \ 4 \ \ 6 \ \ 2 \ \ 7 \ \ 3)$. Note that τ fixes 5. □

Example 6. $S_3 = \{\varepsilon, (1 \ \ 2 \ \ 3), (1 \ \ 3 \ \ 2), (1 \ \ 2), (1 \ \ 3), (2 \ \ 3)\}$ from Example 1. Hence, S_3 consists of cycles; however, the same is not true of S_n in general, as we show later.

Example 7. The only cycle of length 1 is the identity permutation ε.

To reverse the action of a cycle, we simply go around the cycle in the opposite direction. Thus we obtain

Theorem 4. *If σ is an r-cycle, then σ^{-1} is also an r-cycle. More precisely, if $\sigma = (k_1 \ \ k_2 \ \ \cdots \ \ k_{r-1} \ \ k_r)$, then $\sigma^{-1} = (k_r \ \ k_{r-1} \ \ \cdots \ \ k_2 \ \ k_1)$.*

Cycle notation is much simpler than two-row matrix notation. However, we must briefly discuss two ambiguous aspects of cycle notation. First, the same permutation can be written in several ways in cycle notation. For example, $\sigma = (1 \ \ 4 \ \ 2 \ \ 3)$ in S_4 can be written as $\sigma = (4 \ \ 2 \ \ 3 \ \ 1) = (2 \ \ 3 \ \ 1 \ \ 4) = (3 \ \ 1 \ \ 4 \ \ 2)$. This is harmless once we are aware of it.

The second ambiguity can be illustrated as follows: Given $\sigma = (1 \ \ 2 \ \ 4)$, is it in S_4 (fixing 3) or in S_5 (fixing 3 and 5)? We introduce the following convention so that it does not matter.

Convention. *Every permutation in S_n is regarded as a permutation in S_{n+1} that fixes $n+1$. Thus,*

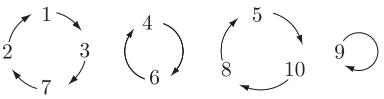
$$S_1 \subseteq S_2 \subseteq S_3 \subseteq \cdots .$$

We shall adhere to this convention throughout this book.

Of course, not every permutation is a cycle. For example, consider

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 3 & 1 & 7 & 6 & 10 & 4 & 2 & 5 & 9 & 8 \end{pmatrix}$$

in S_{10}. If we represent the action of σ geometrically, we obtain



The four cycles are $(1 \ \ 3 \ \ 7 \ \ 2)$, $(4 \ \ 6)$, $(5 \ \ 10 \ \ 8)$, and $(9) = \varepsilon$. These are pairwise disjoint, so each commutes with the others by Theorem 3. Even more remarkable is the fact that σ is the product of these cycles (where we omit $(9) = \varepsilon$):

$$\sigma = (1 \ \ 3 \ \ 7 \ \ 2)(4 \ \ 6)(5 \ \ 10 \ \ 8).$$

The reader should check this assertion. In fact, every permutation can be expressed as a product of disjoint cycles in this way. Here is another example.

Example 8. Factor

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\ 5 & 12 & 2 & 1 & 9 & 11 & 4 & 3 & 7 & 10 & 13 & 8 & 6 \end{pmatrix}$$

as a product of (pairwise) disjoint cycles.

Solution. Starting with 1, follow the action of σ: $1 \to 5 \to 9 \to 7 \to 4 \to 1$. Thus, it has cycled, and the first cycle is $(1\ \ 5\ \ 9\ \ 7\ \ 4)$. Now start with any member of X_{13} not already considered, say $2 \to 12 \to 8 \to 3 \to 2$; so the next cycle is $(2\ \ 12\ \ 8\ \ 3)$. However, 6 has still not been used. It provides the cycle $(6\ \ 11\ \ 13)$. The remaining member of X_{13} is 10 that is fixed by σ, so the corresponding cycle is $(10) = \varepsilon$. Hence,

$$\sigma = (1\ \ 5\ \ 9\ \ 7\ \ 4)(2\ \ 12\ \ 8\ \ 3)(6\ \ 11\ \ 13)$$

is the desired factorization (where we drop the 1-cycles as before). Of course, the action of σ can be sketched as shown previously. □

The method of Example 8 will express every permutation as a product of disjoint cycles because each cycle agrees with σ on the elements it moves, and these elements are fixed by the other cycles. In addition, the factorization is unique up to the order of the disjoint cycles, and we give a formal inductive proof of the following theorem at the end of this section.

Theorem 5. ***Cycle Decomposition Theorem***. *If $\sigma \neq \varepsilon$ is a permutation in S_n, then σ is a product of (one or more) disjoint cycles of length at least 2. This factorization is unique up to the order of the factors.*

Example 9. List all the elements of S_4, each factored into disjoint cycles.

Solution. The $4! = 24$ elements are as follows:

$$\begin{array}{ccccc}
\varepsilon & (1\ \ 2) & (1\ \ 2\ \ 3) & (1\ \ 2)(3\ \ 4) & (1\ \ 2\ \ 3\ \ 4) \\
 & (1\ \ 3) & (1\ \ 2\ \ 4) & (1\ \ 3)(2\ \ 4) & (1\ \ 2\ \ 4\ \ 3) \\
 & (1\ \ 4) & (1\ \ 3\ \ 4) & (1\ \ 4)(2\ \ 3) & (1\ \ 3\ \ 2\ \ 4) \\
 & (2\ \ 3) & (2\ \ 3\ \ 4) & & (1\ \ 3\ \ 4\ \ 2) \\
 & (2\ \ 4) & (1\ \ 3\ \ 2) & & (1\ \ 4\ \ 2\ \ 3) \\
 & (3\ \ 4) & (1\ \ 4\ \ 2) & & (1\ \ 4\ \ 3\ \ 2) \\
 & & (1\ \ 4\ \ 3) & & \\
 & & (2\ \ 4\ \ 3) & &
\end{array}$$

□

The permutations in Example 9 are classified according to the following notion: Two permutations in S_n have the same **cycle structure** if, when they are factored into disjoint cycles, they have the same number of cycles of each length. We refer to this notation again later.

The Alternating Group

A cycle of length 2 is called a **transposition**. Thus, each transposition δ has the form $\delta = (m\ \ n)$ where $m \neq n$. Hence,

$$\delta^2 = \varepsilon \quad \text{and} \quad \delta^{-1} = \delta, \quad \text{for every transposition } \delta.$$

Note, however, that $\sigma = (1\ \ 2)(3\ \ 4)$ also satisfies $\sigma^2 = \varepsilon$ and $\sigma^{-1} = \sigma$, so these properties do not characterize the transpositions.

One reason for studying transpositions is that every permutation is a product of transpositions. For example, the cycle $(1\ \ 2\ \ 3\ \ 4\ \ 5\ \ 6)$ factors as follows:

$$(1\ \ 2\ \ 3\ \ 4\ \ 5\ \ 6) = (1\ \ 2)(2\ \ 3)(3\ \ 4)(4\ \ 5)(5\ \ 6)$$

as is easily verified. This pattern works in general.

Theorem 6. *Every cycle of length $r > 1$ is a product of $r - 1$ transpositions*:

$$(k_1\ \ k_2\ \ \cdots\ \ k_r) = (k_1\ \ k_2)(k_2\ \ k_3)\cdots(k_{r-2}\ \ k_{r-1})(k_{r-1}\ \ k_r).$$

Hence, every permutation is a product of transpositions.

Proof. The verification of the cycle factorization is left to the reader. The rest follows because every permutation is a product of cycles by Theorem 5. ■

In contrast to the factorization into cycles, factorizations into transpositions are *not* unique. For example,

$$(2\ \ 3)(1\ \ 2)(2\ \ 5)(1\ \ 3)(2\ \ 4) = (1\ \ 2\ \ 4\ \ 5) = (1\ \ 5)(1\ \ 4)(1\ \ 2).$$

Indeed, any factorization into m transpositions gives rise to a factorization into $m + 2$ transpositions simply by inserting $\varepsilon = (1\ \ 2)(1\ \ 2)$ somewhere. This gives a glimpse (admittedly not convincing!) into why the next theorem is true. It asserts that if a permutation can be factored in one way as a product of an even (or odd) number of transpositions, then *any* factorization into transpositions must involve an even (respectively odd) number of factors.

Two integers m and n are said to have the **same parity** if they are both even or both odd; equivalently, if $m \equiv n \pmod 2$.

Theorem 7. ***Parity Theorem***. *If a permutation σ has two factorizations*

$$\sigma = \gamma_n \cdots \gamma_2\gamma_1 = \mu_m \cdots \mu_2\mu_1,$$

where each γ_i and μ_j is a transposition, then m and n have the same parity.

The proof of this astonishing fact is given at the end of this section.

A permutation σ is called **even** or **odd** accordingly as it can be written in some way as the product of an even or odd number of transpositions. The parity theorem ensures that this is unambiguous, that is no permutation is both even and odd.

The parity of a cycle γ is easy to determine: Theorem 6 shows that γ is even if its length is odd, and odd if its length is even. When combined with Theorem 5, this result provides a way to easily compute the parity of any permutation.

Example 10. Determine the parity of $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 4 & 6 & 1 & 7 & 8 & 2 & 9 & 3 \end{pmatrix}$.

Solution. The factorization of σ into disjoint cycles is $\sigma = (1\ \ 5\ \ 7\ \ 2\ \ 4)(3\ \ 6\ \ 8\ \ 9)$. Then, $(1\ \ 5\ \ 7\ \ 2\ \ 4)$ is even and $(3\ \ 6\ \ 8\ \ 9)$ is odd by Theorem 6, so σ is odd (because the sum of an even and an odd integer is odd). □

The set of all even permutations in S_n is denoted A_n. It is called the **alternating group of degree** n and plays an important role in the theory of groups (in Chapter 2). Theorem 8 collects several facts about A_n that will be needed later.

Theorem 8. *If $n \geq 2$, the set A_n has the following properties:*

(1) *ε is in A_n and, if σ and τ are in A_n, then both σ^{-1} and $\sigma\tau$ are in A_n.*

(2) *$|A_n| = \frac{1}{2}n!$*

Proof. (1) $\varepsilon = (1\ \ 2)(1\ \ 2)$, so it is even. If σ and τ are even, write $\sigma = \gamma_1\gamma_2\cdots\gamma_n$ and $\tau = \delta_1\delta_2\cdots\delta_m$, where n and m are even and γ_i and δ_j are transpositions. Then $\sigma\tau = \gamma_1\gamma_2\cdots\gamma_n\delta_1\delta_2\cdots\delta_m$ is a product of $n+m$ transpositions, and so is even. Finally, write $\mu = \gamma_n\cdots\gamma_2\gamma_1$. The fact that $\gamma_i^2 = \varepsilon$ for each i implies that $\sigma\mu = \varepsilon$ (verify). Hence, $\sigma^{-1} = \sigma^{-1}\varepsilon = \sigma^{-1}\sigma\mu = \varepsilon\mu = \mu$. But μ is even because n is even, so σ^{-1} is even.

(2) Let O_n denote the set of odd permutations in S_n. Then $S_n = A_n \cup O_n$ and the parity theorem guarantees that $A_n \cap O_n = \varnothing$. Since $|S_n| = n!$, it suffices to show that $|A_n| = |O_n|$. We do so by exhibiting a bijection $f\colon A_n \to O_n$. Let $\gamma = (1\ \ 2)$ and define f by $f(\sigma) = \gamma\sigma$ for all $\sigma \in A_n$. (Note that $\gamma\sigma$ is odd if σ is even.) The fact that $\gamma^2 = \varepsilon$ implies that f is a bijection. In fact, $\gamma\sigma = \gamma\sigma_1$ gives $\sigma = \gamma^2\sigma = \gamma^2\sigma_1 = \sigma_1$ (so f is one-to-one); if $\tau \in O_n$, then $\sigma = \gamma\tau \in A_n$ and $f(\sigma) = \gamma\sigma = \gamma^2\tau = \tau$ (so f is onto). Thus, $|A_n| = |O_n|$. ■

A set of permutations is called a *group* if it contains the identity permutation, the product of any two of its members, and the inverse of any member. Hence, S_n is a group, and the first part of Theorem 8 shows that A_n is a group. The general idea of a group is defined and discussed at length in Chapter 2.

Proof of the Cycle Decomposition Theorem

If $\sigma \neq \varepsilon$ is a permutation in S_n, we show it is a product of disjoint cycles by induction on $n \geq 2$. This is clear if $n = 2$. If $n > 2$, assume that the result is true for S_{n-1} and let $\sigma \in S_n$. If $\sigma n = n$, then $\sigma \in S_{n-1}$ and we are done. So assume $\sigma n \neq n$ and write $m = \sigma^{-1}n$. Then $\sigma m = \sigma(\sigma^{-1}n) = \varepsilon n = n$, and $m \neq n$ (because $\sigma n \neq n$). We write $\gamma = (m\ \ n)$ and consider $\tau = \sigma\gamma$. Because $\gamma^2 = \varepsilon$, we have $\tau\gamma = \sigma\gamma^2 = \sigma\varepsilon = \sigma$. Moreover, $\tau n = \sigma\gamma n = \sigma m = n$, so $\tau \in S_{n-1}$ and τ is a product of disjoint cycles by induction. There are two cases:

- *Case 1*: $\tau m = m$. In this case, γ and τ are disjoint (as $\tau n = n$) and we are done because $\sigma = \gamma\tau$.
- *Case 2*: $\tau m \neq m$. Then m is moved by (exactly one) cycle factor of τ. Hence we can write

$$\tau = \mu(m\ \ k_1\ \ k_2\ \ \cdots\ \ k_r),$$

where μ is a product of disjoint cycles fixing $m, k_1, k_2, \ldots, k_r$ (and also fixing n because $\tau n = n$). Finally, it is easy to verify that

$$\sigma = \tau\gamma = \mu(m\ \ k_1\ \ k_2\ \ \cdots\ \ k_r)(m\ \ n) = \mu(m\ \ n\ \ k_1\ \ \cdots\ \ k_r),$$

which gives σ as a product of disjoint cycles.

Turning to the uniqueness, suppose that $\sigma = \gamma_a \dots \gamma_2\gamma_1 = \delta_b \cdots \delta_2\delta_1$ are two factorizations into disjoint cycles. We proceed by induction on $\max(a, b)$. If this is 1, then $\sigma = \gamma_1 = \delta_1$. Otherwise, let σ move m. Then m occurs in exactly one γ_i and exactly one δ_j. By reordering the factors if necessary, assume that m occurs in γ_1 and in δ_1. Hence, we can write

$$\gamma_1 = (k_1 \quad k_2 \quad \cdots \quad k_r) \quad \text{and} \quad \delta_1 = (l_1 \quad l_2 \quad \cdots \quad l_s),$$

where $k_1 = m = l_1$. We may assume that $r \leq s$. Then, because $k_1 = l_1$,

$$\begin{aligned} k_2 &= \sigma k_1 = \sigma l_1 = l_2 \\ k_3 &= \sigma k_2 = \sigma l_2 = l_3 \\ &\ \,\vdots \qquad\quad \vdots \\ k_r &= \sigma k_{r-1} = \sigma l_{r-1} = l_r \end{aligned}$$

If $r < s$, the next step gives

$$l_1 = k_1 = \sigma k_r = \sigma l_r = l_{r+1},$$

a contradiction. Thus, $r = s$ and $\gamma_1 = \delta_1$. If we write $\lambda = \gamma_1 = \delta_1$, we obtain $\sigma = \gamma_a \dots \gamma_2 \lambda = \delta_b \cdots \delta_2 \lambda$. It follows that $\sigma\lambda^{-1} = \gamma_a \dots \gamma_2 = \delta_b \cdots \delta_2$ is a product of $a - 1$ (and $b - 1$) disjoint cycles. By induction, $a = b$ and (after possible reordering) $\gamma_i = \delta_i$ for $i = 2, 3, \cdots, a$, which completes the induction.

Proof of the Parity Theorem

The proof depends on two preliminary results about transpositions.

Lemma 2. *Let $\gamma_1 \neq \gamma_2$ be transpositions. If γ_1 moves k, transpositions δ_1 and λ_2 exist such that*

$$\gamma_2\gamma_1 = \lambda_2\delta_1, \qquad \textit{where } \delta_1 \textit{ fixes } k \textit{ and } \lambda_2 \textit{ moves } k.$$

Proof. Let $\gamma_1 = (k \quad a)$. Because $\gamma_1 \neq \gamma_2$, the transposition γ_2 has one of the forms $(k \quad b), (a \quad b)$, or $(b \quad c)$ where k, a, b, and c denote distinct integers. In these cases,

$$\begin{aligned} \gamma_2\gamma_1 &= (k \quad b)(k \quad a) = (k \quad a)(a \quad b) \\ \gamma_2\gamma_1 &= (a \quad b)(k \quad a) = (k \quad b)(a \quad b) \\ \gamma_2\gamma_1 &= (b \quad c)(k \quad a) = (k \quad a)(b \quad c) \end{aligned}$$

Hence the conclusion of Lemma 2 holds in every case. ■

Lemma 3. *If the identity permutation ε can be written as a product of $n \geq 3$ transpositions, then it can be written as a product of $n - 2$ transpositions.*

Proof. Let $\varepsilon = \gamma_n \cdots \gamma_4\gamma_3\gamma_2\gamma_1$, where $n \geq 3$ and γ_i are transpositions. Suppose that γ_1 moves k. If $\gamma_1 = \gamma_2$, then $\gamma_2\gamma_1 = \varepsilon$, so $\varepsilon = \gamma_n \cdots \gamma_4\gamma_3$ and we are done. Otherwise, Lemma 2 gives $\gamma_1\gamma_2 = \lambda_2\delta_1$, where δ_1 fixes k and λ_2 moves k. Thus,

$$\varepsilon = \gamma_n \cdots \gamma_4\gamma_3\lambda_2\delta_1.$$

Again, we are done if $\lambda_2 = \gamma_3$, so we let $\gamma_3\lambda_2 = \lambda_3\delta_2$, where δ_2 fixes k and λ_3 moves k. Hence,

$$\varepsilon = \gamma_n \cdots \gamma_5\gamma_4\lambda_3\delta_2\delta_1.$$

Continue in this way. Either we are done at some stage or we finally arrive at a factorization

$$\varepsilon = \lambda_n \delta_{n-1} \cdots \delta_2 \delta_1,$$

where each δ_i fixes k and λ_n moves k. But this cannot happen because, if it did,

$$k = \varepsilon k = \lambda_n \delta_{n-1} \cdots \delta_2 \delta_1 k = \lambda_n k \neq k,$$

a contradiction. This proves Lemma 3. ■

Proof of the parity theorem. Suppose a permutation σ has two factorizations into transpositions:

$$\sigma = \gamma_n \ldots \gamma_2 \gamma_1 = \mu_m \ldots \mu_2 \mu_1.$$

We must show that n and m are both even or both odd. The fact that $\mu_j^{-1} = \mu_j$ for all j gives $\varepsilon = \mu_1 \mu_2 \ldots \mu_m \gamma_n \ldots \gamma_2 \gamma_1$. Hence, it suffices to show that ε cannot be written as the product of an odd number of transpositions. But if ε is a product of p transpositions, where $p \geq 3$ is odd, then repeating Lemma 3 gives factorizations into $p-2, p-4, \ldots,$ transpositions. Ultimately we get a factorization of ε as *one* transposition, which is impossible. □

Exercises 1.4

1. Let

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 4 & 3 & 5 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 1 & 5 & 4 \end{pmatrix}, \quad \mu = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 1 & 2 \end{pmatrix}$$

be permutations. Compute:
(a) $\tau\sigma$ (b) $\sigma\tau$ (c) τ^{-1}
(d) μ^{-1} (e) $\mu\tau\sigma^{-1}$ (f) $\mu^{-1}\sigma\tau$

2. (a) Verify that any two of σ, τ, and μ commute:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}, \quad \mu = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}.$$

(b) Do (a) by first verifying that $\sigma = \tau^2$ and $\mu = \tau^3$.

3. Let

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}$$

and

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}.$$

In each case solve for χ in S_4.
(a) $\sigma\chi = \tau$ (b) $\chi\tau = \sigma$ (c) $\sigma^{-1}\chi = \tau$
(d) $\chi\tau\sigma = \varepsilon$ (e) $\tau\chi\sigma = \varepsilon$ (f) $\tau\chi\sigma^{-1} = \sigma$

4. Suppose that

$$\tau\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 1 & 4 & 2 \end{pmatrix}$$

and

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 3 & 5 & 1 \end{pmatrix}$$

in S_5. If $\sigma 1 = 2$, find σ and τ.

5. Show that

$$\tau\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$$

and

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

is impossible for σ and τ in S_4.

6. If σ and τ fix k, show that $\sigma\tau$ and σ^{-1} both fix k.

7. (a) How many permutations in S_5 fix 1?
(b) How many fix both 1 and 2?

8. (a) If $\sigma\tau = \varepsilon$ in S_n, show that $\sigma = \tau^{-1}$.
(b) If $\sigma^2 = \sigma$ in S_n, show that $\sigma = \varepsilon$.

9. In S_n, show that $\sigma = \tau$ if and only if $\sigma\tau^{-1} = \varepsilon$.

10. If σ and τ are disjoint in S_n and $\sigma\tau = \varepsilon$, what can you say about σ and τ? Support your answer.

11. Write the following in two-row matrix notation.
(a) $(1\ 8\ 7\ 4)(3\ 6\ 7\ 5\ 9)$ (b) $(1\ 3\ 5\ 7)(4\ 1\ 9)$

12. Let $\sigma = (1\ 2\ 3)$ and $\tau = (1\ 2)$ in S_3.
(a) Show that $S_3 = \{\varepsilon, \sigma, \sigma^2, \tau, \tau\sigma, \tau\sigma^2\}$ and that $\sigma^3 = \varepsilon = \tau^2$ and $\sigma\tau = \tau\sigma^2$.
(b) Use (a) to fill in the multiplication table for S_3.

13. Factor each of the following permutations into disjoint cycles, find its parity, and factor the inverse into disjoint cycles.

(a) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 7 & 9 & 8 & 2 & 1 & 6 & 3 & 5 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 8 & 9 & 5 & 2 & 1 & 6 & 4 & 7 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 8 & 6 & 9 & 4 & 7 & 3 & 1 & 5 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 6 & 4 & 8 & 9 & 3 & 1 & 7 & 5 & 2 \end{pmatrix}$

(e) $(1\ 3)(2\ 5\ 7)(3\ 8\ 5)$
(f) $(1\ 2\ 3\ 4\ 5)(6\ 7)(1\ 3\ 5\ 7)(1\ 6\ 3)$

14. If $\sigma\tau = \sigma\mu$ or $\tau\sigma = \mu\sigma$ in S_n, show that $\tau = \mu$. Does $\sigma\tau = \mu\sigma$ imply that $\tau = \mu$? Support your answer.

15. In each of (a) S_5, and (b) S_6, list one permutation of each possible cycle structure (see Example 9).

16. If $\sigma = (1\ 2\ 3\ \cdots\ n)$, show that $\sigma^n = \varepsilon$ and that n is the smallest positive integer with this property.

17. (a) If $\sigma = (1\ 2\ 3\ 4)(5\ 6\ 7)$, factor σ^{-1} into disjoint cycles.
(b) If $\sigma = \gamma_1\gamma_2\cdots\gamma_n$, where the γ_i are disjoint cycles, how is the factorization of σ^{-1} into disjoint cycles related to the γ_i? Support your answer.

18. Find the parity of

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ 5 & 11 & 6 & 1 & 15 & 13 & 2 & 9 & 4 & 10 & 14 & 3 & 12 & 7 & 8 \end{pmatrix}.$$

19. Find the parity of each permutation in Exercise 13.

20. Show that $(1 \ \ 2)$ is not a product of 3-cycles.

21. (a) If $\gamma_1, \gamma_2, \cdots, \gamma_m$ are transpositions, show that

$$(\gamma_1 \ \gamma_2 \ \cdots \ \gamma_m)^{-1} = \gamma_m \gamma_{m-1} \cdots \gamma_2 \gamma_1.$$

(b) Show that σ and σ^{-1} have the same parity for all σ in S_n.

(c) Show that σ and $\tau\sigma\tau^{-1}$ have the same parity for all σ and τ in S_n.

22. Show that $A_{n+1} \cap S_n = A_n$ for all $n \geq 3$ (regard $S_n \subseteq S_{n+1}$ in the usual way).

23. Let $\sigma \in S_n$, $\sigma \neq \varepsilon$. If $n \geq 3$, show that $\gamma \in S_n$ exists such that $\sigma\gamma \neq \gamma\sigma$. [*Hint:* If $\sigma k = l$ with $k \neq l$, choose $m \notin \{k, l\}$ and take $\gamma = (k \ \ m)$.]

24. If $\sigma \in S_n$, show that $\sigma^2 = \varepsilon$ if and only if σ is a product of disjoint transpositions.

25. If $n \geq 3$, show that every even permutation in S_n is a product of 3-cycles.

26. Let γ be any cycle of length r. If $\sigma \in S_n$, show that $\sigma\gamma\sigma^{-1}$ is also a cycle of length r. More precisely, if $\gamma = (k_1 \ \ k_2 \ \ \cdots \ \ k_r)$ show that $\sigma\gamma\sigma^{-1} = (\sigma k_1 \ \ \sigma k_2 \ \ \cdots \ \ \sigma k_r)$.

27. (a) Show that $(k_1 \ \ k_2 \ \ \cdots \ \ k_r) = (k_1 \ \ k_r)(k_1 \ \ k_{r-1}) \cdots (k_1 \ \ k_2)$.

(b) Show that each $\sigma \in S_n$ is a product of the transpositions $(1 \ \ 2), (1 \ \ 3), \ldots, (1 \ \ n)$. [*Hint:* Each transposition is such a product by (a) and Exercise 26.]

(c) Repeat (b) for the transpositions $(1 \ \ 2)$, $(2 \ \ 3), \ldots, (n-1 \ \ n)$. [*Hint:* Use (a) and Exercise 26.]

(d) If $\sigma = (1 \ \ 2 \ \ 3 \ \ \cdots \ \ n)$, show that each element of S_n is a product of the permutations $(1 \ \ 2), \sigma$, and σ^{-1}. [*Hint:* Use (b) and Exercise 26.]

28. Let $\sigma = (1 \ \ 2 \ \ 3 \ \ \cdots \ \ n)$ be a cycle of length $n \geq 2$.

(a) If $n = 2k$, find the factorization of σ^2 into disjoint cycles.

(b) If $n = mq$ with $m \geq 3$ and $q \geq 2$, show that σ^m is a product of m disjoint cycles, each of length q.

(c) If $1 \leq m \leq n$, show that $\sigma^m k \equiv k + m \pmod{n}$.

(d) If $n = p$ is a prime, show that σ^m is a cycle of length p for each $m = 1, 2, \ldots, p-1$.

29. Define the **sign** of a permutation σ to be

$$\operatorname{sgn}\sigma = \begin{cases} 1 & \text{if } \sigma \text{ is even} \\ -1 & \text{if } \sigma \text{ is odd} \end{cases}.$$

Prove that $\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}\sigma \operatorname{sgn}\tau$ for all σ and τ in S_n.

30. Consider a puzzle made up of five numbered squares in a 2×3 frame. Assume that the squares slide vertically and horizontally so that rearrangements are possible. For example, arrangement (2) can be obtained from (1) (in four moves). Call an arrangement "nice" if the lower right position is vacant. Then, the "nice" arrangements correspond to permutations in S_5. For example, arrangement (2) corresponds to $(2 \ \ 5 \ \ 3)$.

(1)

1	2	3
4	5	

(2)

1	5	2
4	3	

Show that every "nice" arrangement corresponds to an even permutation.[15]

[15] In fact, every even permutation arises in this way. (See Newman, J. R., *World of Mathematics*, New York: Simon & Schuster, 1956, p. 2431.)

1.5 AN APPLICATION TO CRYPTOGRAPHY

> How often have I said to you that when you have eliminated the impossible, whatever remains, however improbable, must be the truth.
>
> —Sir Arthur Conan Doyle

The ability to transmit messages in a way that cannot be recognized by adversaries has intrigued people for centuries. In this brief section, we outline a method that uses Fermat's theorem to encode information in a way that is very difficult to break. The idea is based on the following consequence of that theorem.

Theorem 1. *Let $n = pq$, where p and q are distinct primes, write $m = (p-1)(q-1)$, and let $e > 2$ be any integer such that $e \equiv 1 \pmod{m}$. Then*

$$x^e \equiv x \pmod{n} \qquad \textit{for all } x \textit{ such that} \qquad \gcd(x, n) = 1.$$

Proof. Because $e \equiv 1 \pmod{m}$, write $e - 1 = ym$, where y is an integer. Then $x^e = x \cdot (x^m)^y$, so it suffices to show that $x^m \equiv 1 \pmod{n}$ whenever $\gcd(x, n) = 1$. This condition certainly implies that p does not divide x. Hence, Fermat's theorem shows that $x^{p-1} \equiv 1 \pmod{p}$ and so $x^m = (x^{p-1})^{q-1} \equiv 1^{q-1} \equiv 1 \pmod{p}$. Similarly, $x^m \equiv 1 \pmod{q}$ and so, as p and q are relatively prime, Theorem 5 §1.2 shows that $x^m \equiv 1 \pmod{pq}$. This is what we wanted. ■

The coding process can be described as follows. Two distinct primes p and q are chosen, each very large in practice. Then the words available for transmission (and punctuation symbols) are paired with distinct integers $x \geq 2$. The integers x used may be assumed to be chosen relatively prime to p and q if these primes are large enough and, in practice, to be smaller than each of these primes. The idea is to use p and q to compute an integer r from x and then to transmit r rather than x. Clearly, r must be chosen in such a way that x (and hence the corresponding word) can be retrieved from r. The passage from x to r (called *encoding*) is carried out by the sender of a message, the integer r is transmitted, and the computation of x from r (*decoding*) is done by the receiver.

Here is how the process works. Given the distinct primes p and q, the cryptographer denotes

$$n = pq \quad \text{and} \quad m = (p-1)(q-1)$$

and then chooses any integer $k \geq 2$ such that $\gcd(k, m) = 1$. The sender is given only the numbers n and k. If the sender wants to transmit an integer x, he or she encodes it by reducing x^k modulo n, say,

$$x^k \equiv r \pmod{n}, \quad \text{where } 0 \leq r < n.$$

Then the sender transmits r to the receiver of the message who must use it to retrieve x. If the receiver knows the inverse k' of k in $\mathbb{Z}_m$, then $k'k \equiv 1 \pmod{m}$. Hence, Theorem 1 (with $e = k'k$) gives $x^{k'k} \equiv x \pmod{n}$ and

$$x \equiv x^{k'k} \equiv (x^k)^{k'} \equiv r^{k'}$$

modulo n. Knowing both r and k', the receiver can compute x (and hence the corresponding word in the message).

Note that all the sender really has to know are n and k. A third party intercepting the message r cannot retrieve x without k', and computing it requires p and q.

Even if the third party can extract the integers n and k from the sender, factoring $n = pq$ in practice is very time-consuming if the primes p and q are large, even with a computer. Hence, the code is extremely difficult to break. Example 1 illustrates how the process works, although the primes used are small.

Example 1. Let $p = 11$ and $q = 13$ so that $n = 143$ and $m = 120$. Then let $k = 7$, chosen so that $\gcd(k, m) = 1$. Encode the number $x = 9$ and then decode the result.

Solution. The sender reduces $x^k = 9^7$ modulo $n = 143$. Working modulo 143: $9^2 \equiv 81$, $9^3 \equiv 14$, $9^4 \equiv 126$, $9^7 \equiv 48$. Hence, $r = 48$ is transmitted. The receiver then finds k', the inverse of $k = 7$ modulo $m = 120$. In fact, the euclidean algorithm gives $1 = 120 - 17 \cdot 7$, so $k' \equiv -17 \equiv 103 \pmod{120}$ is the required inverse. Hence, x is retrieved (modulo n) by $x \equiv r^{k'} \equiv 48^{103} \pmod{143}$. One fairly efficient way to compute this is to note that $103 = 1100111$ in binary, so $103 = 1 + 2 + 2^2 + 2^5 + 2^6$. Then the receiver computes 48^t, where t is a power of 2 by successive squaring of 48 modulo 143:

$$48^2 \equiv 16,\ 48^{2^2} \equiv 113,\ 48^{2^3} \equiv 42,\ 48^{2^4} \equiv 48,\ 48^{2^5} \equiv 16,\ 48^{2^6} \equiv 113.$$

Again working modulo 143 gives

$$x \equiv 48^{103} \equiv 48^{1+2+2^2+2^5+2^6} \equiv 48 \cdot 16 \cdot 113 \cdot 16 \cdot 113 \equiv 9,$$

which retrieves the original 9. □

This system is called the RSA system after its inventors.[16] Other, more comprehensive coverage of cryptography is available,[17] including overviews of the subject, methods, and bibliographies.

The RSA system works by finding two large primes p and q and computing the number $n = pq$. The code is difficult to break because it is difficult to find p and q given n. However, in 2002, Maninda Agrawal and two undergraduate students (Neeraj Kayal and Nitin Saxena) gave a simple algorithm that can decide whether a given integer n is prime or not. Moreover, the time taken is approximately a polynomial function of n. This is an important breakthrough in computer science, and certainly affects algorthms like the RSA system.

Cryptography, in general, refers to the transmission of messages where the primary aim is to disguise the message to make its interpretation by an unauthorized interceptor very difficult. Coding theory, in contrast, aims at fast and correct transmission of messages; we briefly discuss this topic in Sections 2.11 and 6.7.

[16] Rivest, R. L., Shamir, A., and Adleman, L., A method for obtaining digital signatures and public-key cryptosystems, *Communication of the ACM*, 21 (1978), 120–126.

[17] For example, see the section on Algebraic Cryptography in Lidl, R. and Pilz, G., *Applied Abstract Algebra*, New York: Springer-Verlag, 1983.