## ELECTIVE V (3)

## ALGEBRAIC NUMBER THEORY

### Objectives

1. To expose the students to the charm, niceties and nuances in the world of numbers.
2. To highlight some of the Applications of the Theory of Numbers.

### UNIT I

Introduction – Divisibility – Primes – The Binomial Theorem – Congruences – Euler's totient – Fermat's, Euler's and Wilson's Theorems – Solutions of congruences – The Chinese Remainder theorem.

### UNIT II

Techniques of numerical calculations – Public key cryptography – Prime power Moduli – Primitive roots and Power Residues –Congruences of degree two.

### UNIT III

Number theory from an Algebraic Viewpoint – Groups, rings and fields – Quadratic Residues- The Legendre symbol (a/r) where r is an odd prime – Quadratic Reciprocity – The Jacobi Symbol (P/q) where q is an odd positive integer.

### UNIT IV

Binary Quadratic Forms – Equivalence and Reduction of Binary Quadratic Forms – Sums of three squares – Positive Definite Binary Quadratic forms – Greatest integer Function – Arithmetic Functions – The Mobius Inversion Formula – Recurrence Functions – Combinatorial number theory .

### UNIT V

Diophantine Equations – The equation $ax+by=c$ – Simultaneous Linear Diophantine Equations – Pythagorean Triangles – Assorted examples.

### TEXT BOOK

Ivan Niven, Herbert S. Zuckerman and Hugh L. Montgomery, An Introduction to the Theory of Numbers, Fifth edn., John Wiley & Sons Inc. 2004.

UNIT I Chapter 1 and Chapter 2 : Sections 2.1 to 2.3
UNIT II Chapter 2 : Sections 2.4 to 2.9
UNIT III Chapter 2 : Sections 2.10, 2.11 and Chapter 3: Sections 3.1 to 3.3
UNIT IV Chapter 3 : Sections 3.4 to 3.7 and Chapter 4
UNIT V Chapter 5: Sections 5.1 to 5.4.

### REFERENCES


1. Elementary Number Theory, David M. Burton W.M.C. Brown Publishers, Dubuque, Iowa, 1989.
2. Number Theory, George Andrews, Courier Dover Publications, 1994.
3. Fundamentals of Number Theory, William J. Leveque Addison-Wesley Publishing Company, Phillipines, 1977.


*****

/* HARIPRABHU.M. M.Sc., M.Phil., PGDCA.,
HEAD Dept. of MATHS.-Z %. MTC.
Algebraic Number Theory.

# CHAPTER 1

# Divisibility

## 1.1 INTRODUCTION

The theory of numbers is concerned with properties of the *natural numbers* $1, 2, 3, 4, \cdots$, also called the *positive integers*. These numbers, together with the negative integers and zero, form the set of integers. Properties of these numbers have been studied from earliest times. For example, an integer is divisible by 3 if and only if the sum of its digits is divisible by 3, as in the number 852 with sum of digits $8 + 5 + 2 = 15$. The equation $x^2 + y^2 = z^2$ has infinitely many solutions in positive integers, such as $3^2 + 4^2 = 5^2$, whereas $x^3 + y^3 = z^3$ and $x^4 + y^4 = z^4$ have none. There are infinitely many prime numbers, where a prime is a natural number such as 31 that cannot be factored into two smaller natural numbers. Thus, 33 is not a prime, because $33 = 3 \cdot 11$.

The fact that the sequence of primes, $2, 3, 5, 7, 11, 13, 17, \cdots$, is endless was known to Euclid, who lived about 350 B.C. Also known to Euclid was the result that $\sqrt{2}$ is an *irrational number*, that is, a number that cannot be expressed as the quotient $a/b$ of two integers. The numbers $2/7$, $13/5$, $-14/9$, and $99/100$ are examples of *rational numbers*. The integers are themselves rational numbers because, for example, 7 can be written in the form $7/1$. Another example of an irrational number is $\pi$, the ratio of the circumference to the diameter of any circle. The rational number $22/7$ is a good approximation to $\pi$, close but not precise. The fact that $\pi$ is irrational means that there is no fraction $a/b$ that is exactly equal to $\pi$, with $a$ and $b$ integers.

In addition to known results, number theory abounds with unsolved problems. Some background is needed just to state these problems in many cases. But there are a few unsolved problems that can be understood with essentially no prior knowledge. Perhaps the most famous of these is the conjecture known as *Fermat's last theorem*, which is not really a theorem at all because it has not yet been proved. Pierre de Fermat (1601–1665) stated that he had a truly wondrous proof that the equation $x^n + y^n = z^n$ has no solutions in positive integers $x, y, z$ for any exponent $n > 2$. Fermat added that the margin of the book was too small to hold the

proof. Whether Fermat really had a proof is not known, but it now seems unlikely, as the question has eluded mathematicians since his time.

Results in number theory often have their sources in empirical observations. We might notice, for example, that every natural number up to 1000 can be expressed as a sum of four squares of natural numbers, as illustrated by

$$1000 = 30^2 + 10^2 + 0^2 + 0^2, \qquad 999 = 30^2 + 9^2 + 3^2 + 3^2.$$

We might then feel confident enough to make the conjecture that every natural number is expressible as a sum of four squares. This turns out to be correct; it is presented as Theorem 6.25 in Chapter 6. The first proof of this result was given by J. L. Lagrange (1736–1813). We say that the four square theorem is *best possible*, because not every positive integer is expressible as a sum of three squares of integers, 7 for example.

Of course, a conjecture made on the basis of a few examples may turn out to be incorrect. For example, the expression $n^2 - n + 41$ is a prime number for $n = 1, 2, 3, \cdots, 40$ because it is easy to verify that $41, 43, 47, 53, \cdots, 1601$ are indeed prime numbers. But it would be hasty to conjecture that $n^2 - n + 41$ is a prime for every natural number $n$, because for $n = 41$ the value is $41^2$. We say that the case $n = 41$ is a *counterexample* to the conjecture.

Leonhard Euler (1707–1783) conjectured that no $n$th power is a sum of fewer than $n$ $n$th powers (the Swiss name Euler is pronounced "Oiler"). For $n = 3$, this would assert that no cube is the sum of two smaller cubes. This is true; it is proved in Theorem 9.35. However, a counterexample to Euler's conjecture was provided in 1968 by L. J. Lander and Thomas Parkin. As the result of a detailed computer search, they found that

$$144^5 = 27^5 + 84^5 + 110^5 + 133^5.$$

In 1987, N. J. Elkies used the arithmetic of elliptic curves to discover that

$$20615673^4 = 2682440^4 + 15365639^4 + 18796760^4,$$

and a subsequent computer search located the least counterexample to Euler's conjecture for fourth powers.

The *Goldbach conjecture* asserts that every even integer greater than 2 is the sum of two primes, as in the examples

$$4 = 2 + 2, \qquad 6 = 3 + 3, \qquad 20 = 7 + 13,$$
$$50 = 3 + 47, \qquad 100 = 29 + 71.$$

Stated by Christian Goldbach in 1742, verified up to 100,000 at least, this conjecture has evaded all attempts at proof.



Because it is relatively easy to make conjectures in number theory, the person whose name gets attached to a problem has often made a lesser contribution than the one who later solves it. For example, John Wilson (1741–1793) stated that every prime $p$ is a divisor of $(p-1)! + 1$, and this result has henceforth been known as Wilson's theorem, although the first proof was given by Lagrange.

However, empirical observations *are* important in the discovery of general results and in testing conjectures. They are also useful in understanding theorems. In studying a book on number theory, you are well advised to construct numerical examples of your own devising, especially if a concept or a theorem is not well understood at first.

Although our interest centers on integers and rational numbers, not all proofs are given within this framework. For example, the proof that $\pi$ is irrational makes use of the system of real numbers. The proof that $x^3 + y^3 = z^3$ has no solution in positive integers is carried out in the setting of complex numbers.

Number theory is not only a systematic mathematical study but also a popular diversion, especially in its elementary form. It is part of what is called *recreational mathematics*, including numerical curiosities and the solving of puzzles. This aspect of number theory is not emphasized in this book, unless the questions are related to general propositions. Nevertheless, a systematic study of the theory is certainly helpful to anyone looking at problems in recreational mathematics.

The theory of numbers is closely tied to the other areas of mathematics, most especially to abstract algebra, but also to linear algebra, combinatorics, analysis, geometry, and even topology. Consequently, proofs in the theory of numbers rely on many different ideas and methods. Of these, there are two basic principles to which we draw especial attention. The first is that any set of positive integers has a smallest element if it contains any members at all. In other words, if a set $\mathcal{S}$ of positive integers is not empty, then it contains an integer $s$ such that for any member $a$ of $\mathcal{S}$, the relation $s \leqslant a$ holds. The second principle, *mathematical induction*, is a logical consequence of the first.[1] It can be stated as follows: If a set $\mathcal{S}$ of positive integers contains the integer 1, and contains $n + 1$ whenever it contains $n$, then $\mathcal{S}$ consists of all the positive integers.

It also may be well to point out that a simple statement which asserts that there is an integer with some particular property may be easy to prove, by simply citing an example. For example, it is easy to demonstrate the proposition, "There is a positive number that is not the sum of three squares," by noting that 7 is such a number. On the other hand, a

[1] Compare G. Birkhoff and S. MacLane, *A Survey of Modern Algebra*, 4th ed., Macmillan (New York), 1977, 10–13.

statement which asserts that all numbers possess a certain property cannot be proved in this manner. The assertion, "Every prime number of the form $4n + 1$ is a sum of two squares," is substantially more difficult to establish (see Lemma 2.13 in Section 2.1).

Finally, it is presumed that you are familiar with the usual formulation of mathematical propositions. In particular, if $A$ and $B$ are two assertions, the following statements are logically equivalent—they are just different ways of saying the same thing.

$A$ implies $B$.
If $A$ is true, then $B$ is true.
In order that $A$ be true it is necessary that $B$ be true.
$B$ is a necessary condition for $A$.
$A$ is a sufficient condition of $B$.

If $A$ implies $B$ and $B$ implies $A$, then one can say that $B$ is a necessary and sufficient condition for $A$ to hold.

In general, we shall use letters of the roman alphabet, $a, b, c, \cdots, m, n, \cdots, x, y, z$ to designate integers unless otherwise specified. We let $\mathbb{Z}$ denote the set $\{\cdots, -2, -1, 0, 1, 2, \cdots\}$ of all integers, $\mathbb{Q}$ the set of all rational numbers, $\mathbb{R}$ the set of all real numbers, and $\mathbb{C}$ the set of all complex numbers.

## 1.2 DIVISIBILITY

Divisors, multiples, and prime and composite numbers are concepts that have been known and studied at least since the time of Euclid, about 350 B.C. The fundamental ideas are developed in this and the next section.

**Definition 1.1** *An integer $b$ is divisible by an integer $a$, not zero, if there is an integer $x$ such that $b = ax$, and we write $a|b$. In case $b$ is not divisible by $a$, we write $a \nmid b$.*

Other language for the divisibility property $a|b$ is that $a$ divides $b$, that $a$ is a divisor of $b$, and that $b$ is a multiple of $a$. If $a|b$ and $0 < a < b$, then $a$ is called a *proper divisor* of $b$. It is understood that we never use 0 as the left member of the pair of integers in $a|b$. On the other hand, not only may 0 occur as the right member of the pair, but also in such instances we always have divisibility. Thus $a|0$ for every integer $a$ not zero. The notation $a^K \| b$ is sometimes used to indicate that $a^K | b$ but $a^{K+1} \nmid b$.



**Theorem 1.1**

*(1) $a|b$ implies $a|bc$ for any integer $c$;*
*(2) $a|b$ and $b|c$ imply $a|c$;*
*(3) $a|b$ and $a|c$ imply $a|(bx + cy)$ for any integers $x$ and $y$;*
*(4) $a|b$ and $b|a$ imply $a = \pm b$;*
*(5) $a|b$, $a > 0$, $b > 0$, imply $a \leqslant b$;*
*(6) if $m \neq 0$, $a|b$ implies and is implied by $ma|mb$.*

*Proof* The proofs of these results follow at once from the definition of divisibility. Property 3 admits an obvious extension to any finite set, thus:

$$a|b_1, a|b_2, \cdots, a|b_n \text{ imply } a \left| \sum_{j=1}^{n} b_j x_j \right. \text{ for any integers } x_j.$$

Property 2 can be extended similarly.

To give a sample proof, consider item 3. Since $a|b$ and $a|c$ are given, this implies that there are integers $r$ and $s$ such that $b = ar$ and $c = as$. Hence, $bx + cy$ can be written as $a(rx + sy)$, and this proves that $a$ is a divisor of $bx + cy$.

The next result is a formal statement of the outcome when any integer $b$ is divided by any positive integer. For example, if 25 is divided by 7, the quotient is 3 and the remainder is 4. These numbers are related by the equality $25 = 7 \cdot 3 + 4$. Now we formulate this in the general case.

**Theorem 1.2** *The division algorithm. Given any integers $a$ and $b$, with $a > 0$, there exist unique integers $q$ and $r$ such that $b = qa + r$, $0 \leqslant r < a$. If $a \nmid b$, then $r$ satisfies the stronger inequalities $0 < r < a$.*

*Proof* Consider the arithmetic progression

$$\cdots, b - 3a, b - 2a, b - a, b, b + a, b + 2a, b + 3a, \cdots$$

extending indefinitely in both directions. In this sequence, select the smallest non-negative member and denote it by $r$. Thus by definition $r$ satisfies the inequalities of the theorem. But also $r$, being in the sequence, is of the form $b - qa$, and thus $q$ is defined in terms of $r$.

To prove the uniqueness of $q$ and $r$, suppose there is another pair $q_1$ and $r_1$ satisfying the same conditions. First we prove that $r_1 = r$. For if not, we may presume that $r < r_1$ so that $0 < r_1 - r < a$, and then we see

that $r_1 - r = a(q - q_1)$ and so $a|(r_1 - r)$, a contradiction to Theorem 1.1, part 5. Hence $r = r_1$, and also $q = q_1$.

We have stated the theorem with the assumption $a > 0$. However, this hypothesis is not necessary, and we may formulate the theorem without it: given any integers $a$ and $b$, with $a \neq 0$, there exist integers $q$ and $r$ such that $b = qa + r$, $0 \leqslant r < |a|$.

Theorem 1.2 is called the *division algorithm*. An *algorithm* is a mathematical procedure or method to obtain a result. We have stated Theorem 1.2 in the form "there exist integers $q$ and $r$," and this wording suggests that we have a so-called existence theorem rather than an algorithm. However, it may be observed that the proof does give a method for obtaining the integers $q$ and $r$, because the infinite arithmetic progression $\cdots, b - a, b, b + a, \cdots$ need be examined only in part to yield the smallest positive member $r$.

In actual practice the quotient $q$ and the remainder $r$ are obtained by the arithmetic division of $a$ into $b$.

***Remark on Calculation*** Given integers $a$ and $b$, the values of $q$ and $r$ can be obtained in two steps by use of a hand-held calculator. As a simple example, if $b = 963$ and $a = 428$, the calculator gives the answer 2.25 if 428 is divided into 963. From this we know that the quotient $q = 2$. To get the remainder, we multiply 428 by 2, and subtract the result from 963 to obtain $r = 107$. In case $b = 964$ and $a = 428$ the calculator gives 2.2523364 as the answer when 428 is divided into 964. This answer is approximate, not exact; the exact answer is an infinite decimal. Nevertheless, the value of $q$ is apparent, because $q$ is the largest integer not exceeding $964/428$; in this case $q = 2$. In symbols we write $q = [964/428]$. (In general, if $x$ is a real number then $[x]$ denotes the largest integer not exceeding $x$. That is, $[x]$ is the unique integer such that $[x] \leqslant x < [x] + 1$. Further properties of the function $[x]$ are discussed in Section 4.1.) The value of $r$ can then also be determined, as $r = b - qa = 964 - 2 \cdot 428 = 108$. Because the value of $q$ was obtained by rounding down a decimal that the calculator may not have determined to sufficient precision, there may be a question as to whether the calculated value of $q$ is correct. Assuming that the calculator performs integer arithmetic accurately, the proposed value of $q$ is confirmed by checking that the proposed remainder $b - qa = 108$ lies in the interval $0 \leqslant r < a = 428$. In case $r$ alone is of interest, it would be tempting to note that 428 times 0.2523364 is 107.99997, and then round to the nearest integer. The method we have described, though longer, is more reliable, as it depends only on integer arithmetic.



**Definition 1.2** *The integer $a$ is* a common divisor *of $b$ and $c$ in case $a|b$ and $a|c$. Since there is only a finite number of divisors of any nonzero integer, there is only a finite number of common divisors of $b$ and $c$, except in the case $b = c = 0$. If at least one of $b$ and $c$ is not 0, the greatest among their common divisors is called the* greatest common divisor *of $b$ and $c$ and is denoted by $(b, c)$. Similarly, we denote the greatest common divisor of the integers $b_1, b_2, \cdots, b_n$, not all zero, by $(b_1, b_2, \cdots, b_n)$.*

Thus the greatest common divisor $(b, c)$ is defined for every pair of integers $b, c$ except $b = 0$, $c = 0$, and we note that $(b, c) \geqslant 1$.

**Theorem 1.3** *If $g$ is the greatest common divisor of $b$ and $c$, then there exist integers $x_0$ and $y_0$ such that $g = (b, c) = bx_0 + cy_0$.*

Another way to state this very fundamental result is that the greatest common divisor (abbreviated g.c.d.) of two integers $b$ and $c$ is expressible as a linear combination of $b$ and $c$ with integral multipliers $x_0$ and $y_0$. This assertion holds not just for two integers but for any finite collection, as we shall see in Theorem 1.5.

*Proof* Consider the linear combinations $bx + cy$, where $x$ and $y$ range over all integers. This set of integers $\{bx + cy\}$ includes positive and negative values, and also 0 by the choice $x = y = 0$. Choose $x_0$ and $y_0$ so that $bx_0 + cy_0$ is the least positive integer $l$ in the set; thus $l = bx_0 + cy_0$.

Next we prove that $l|b$ and $l|c$. We establish the first of these, and the second follows by analogy. We give an indirect proof that $l|b$, that is, we assume $l \nmid b$ and obtain a contradiction. From $l \nmid b$ it follows that there exist integers $q$ and $r$, by Theorem 1.2, such that $b = lq + r$ with $0 < r < l$. Hence we have $r = b - lq = b - q(bx_0 + cy_0) = b(1 - qx_0) + c(-qy_0)$, and thus $r$ is in the set $\{bx + cy\}$. This contradicts the fact that $l$ is the least positive integer in the set $\{bx + cy\}$.

Now since $g$ is the greatest common divisor of $b$ and $c$, we may write $b = gB$, $c = gC$, and $l = bx_0 + cy_0 = g(Bx_0 + Cy_0)$. Thus $g|l$, and so by part 5 of Theorem 1.1, we conclude that $g \leqslant l$. Now $g < l$ is impossible, since $g$ is the *greatest* common divisor, so $g = l = bx_0 + cy_0$.

**Theorem 1.4** *The greatest common divisor $g$ of $b$ and $c$ can be characterized in the following two ways*: (1) *It is the least positive value of $bx + cy$ where $x$ and $y$ range over all integers*; (2) *it is the positive common divisor of $b$ and $c$ that is divisible by every common divisor.*

*Proof* Part 1 follows from the proof of Theorem 1.3. To prove part 2, we observe that if $d$ is any common divisor of $b$ and $c$, then $d|g$ by part 3 of Theorem 1.1. Moreover, there cannot be two distinct integers with property 2, because of Theorem 1.1, part 4.

If an integer $d$ is expressible in the form $d = bx + cy$, then $d$ is not necessarily the g.c.d. $(b, c)$. However, it does follow from such an equation that $(b, c)$ is a divisor of $d$. In particular, if $bx + cy = 1$ for some integers $x$ and $y$, then $(b, c) = 1$.

**Theorem 1.5** *Given any integers $b_1, b_2, \cdots, b_n$ not all zero, with greatest common divisor $g$, there exist integers $x_1, x_2, \cdots, x_n$ such that*

$$g = (b_1, b_2, \cdots, b_n) = \sum_{j=1}^{n} b_j x_j.$$

*Furthermore, $g$ is the least positive value of the linear form $\sum_{j=1}^{n} b_j y_j$ where the $y_j$ range over all integers; also $g$ is the positive common divisor of $b_1, b_2, \cdots, b_n$ that is divisible by every common divisor.*

*Proof* This result is a straightforward generalization of the preceding two theorems, and the proof is analogous without any complications arising in the passage from two integers to $n$ integers.

**Theorem 1.6** *For any positive integer $m$,*

$$(ma, mb) = m(a, b).$$

*Proof* By Theorem 1.4 we have

$$\begin{aligned}(ma, mb) &= \text{least positive value of } max + mby \\ &= m \cdot \{\text{least positive value of } ax + by\} \\ &= m(a, b).\end{aligned}$$

**Theorem 1.7** *If $d|a$ and $d|b$ and $d > 0$, then*

$$\left(\frac{a}{d}, \frac{b}{d}\right) = \frac{1}{d}(a, b).$$

*If $(a, b) = g$, then*

$$\left(\frac{a}{g}, \frac{b}{g}\right) = 1.$$



*Proof* The second assertion is the special case of the first obtained by using the greatest common divisor $g$ of $a$ and $b$ in the role of $d$. The first assertion in turn is a direct consequence of Theorem 1.6 obtained by replacing $m, a, b$ in that theorem by $d, a/d, b/d$ respectively.

**Theorem 1.8** *If $(a, m) = (b, m) = 1$, then $(ab, m) = 1$.*

*Proof* By Theorem 1.3 there exist integers $x_0, y_0, x_1, y_1$ such that $1 = ax_0 + my_0 = bx_1 + my_1$. Thus we may write $(ax_0)(bx_1) = (1 - my_0)(1 - my_1) = 1 - my_2$ where $y_2$ is defined by the equation $y_2 = y_0 + y_1 - my_0y_1$. From the equation $abx_0x_1 + my_2 = 1$ we note, by part 3 of Theorem 1.1, that any common divisor of $ab$ and $m$ is a divisor of 1, and hence $(ab, m) = 1$.

**Definition 1.3** *We say that $a$ and $b$ are* relatively prime *in case $(a, b) = 1$, and that $a_1, a_2, \cdots, a_n$ are* relatively prime *in case $(a_1, a_2, \cdots, a_n) = 1$. We say that $a_1, a_2, \cdots, a_n$ are* relatively prime in pairs *in case $(a_i, a_j) = 1$ for all $i = 1, 2, \cdots, n$ and $j = 1, 2, \cdots, n$ with $i \neq j$.*

The fact that $(a, b) = 1$ is sometimes expressed by saying that $a$ and $b$ are coprime, or by saying that $a$ is prime to $b$.

**Theorem 1.9** *For any integer $x$, $(a, b) = (b, a) = (a, -b) = (a, b + ax)$.*

*Proof* Denote $(a, b)$ by $d$ and $(a, b + ax)$ by $g$. It is clear that $(b, a) = (a, -b) = d$.

By Theorem 1.3, we know that there exist integers $x_0$ and $y_0$ such that $d = ax_0 + by_0$. Then we can write

$$d = a(x_0 - xy_0) + (b + ax)y_0.$$

It follows that the greatest common divisor of $a$ and $b + ax$ is a divisor of $d$, that is, $g|d$. Now we can also prove that $d|g$ by the following argument. Since $d|a$ and $d|b$, we see that $d|(b + ax)$ by Theorem 1.1, part 3. And from Theorem 1.4, part 2, we know that every common divisor of $a$ and $b + ax$ is a divisor of their g.c.d., that is, a divisor of $g$. Hence, $d|g$. From $d|g$ and $g|d$, we conclude that $d = \pm g$ by Theorem 1.1, part 4. However, $d$ and $g$ are both positive by definition, so $d = g$.

**Theorem 1.10** *If $c|ab$ and $(b, c) = 1$, then $c|a$.*

*Proof* By Theorem 1.6, $(ab, ac) = a(b, c) = a$. By hypothesis $c|ab$ and clearly $c|ac$, so $c|a$ by Theorem 1.4, part 2.

Given two integers $b$ and $c$, how can the greatest common divisor $g$ be found? Definition 1.2 gives no answer to this question. The investigation of the set of integers $\{bx + cy\}$ to find a smallest positive element is not practical for large values of $b$ and $c$. If $b$ and $c$ are small, values of $g$, $x_0$, and $y_0$ such that $g = bx_0 + cy_0$, can be found by inspection. For example, if $b = 10$ and $c = 6$, it is obvious that $g = 2$, and one pair of values for $x_0, y_0$ is $2, -3$. But if $b$ and $c$ are large, inspection is not adequate except in rather obvious cases such as $(963, 963) = 963$ and $(1000, 600) = 200$. However, Theorem 1.9 can be used to calculate $g$ effectively and also to get values of $x_0$ and $y_0$. (The reason we want values of $x_0$ and $y_0$ is to find integral solutions of linear equations. These turn up in many simple problems in number theory.) We now discuss an example to show how Theorem 1.9 can be used to calculate the greatest common divisor.

Consider the case $b = 963$, $c = 657$. If we divide $c$ into $b$, we get a quotient $q = 1$, and remainder $r = 306$. Thus $b = cq + r$, or $r = b - cq$, in particular $306 = 963 - 1 \cdot 657$. Now $(b, c) = (b - cq, c)$ by replacing $a$ and $x$ by $c$ and $-q$ in Theorem 1.9, so we see that

$$(963, 657) = (963 - 1 \cdot 657, 657) = (306, 657).$$

The integer 963 has been replaced by the smaller integer 306, and this suggests that the procedure be repeated. So we divide 306 into 657 to get a quotient 2 and a remainder 45, and

$$(306, 657) = (306, 657 - 2 \cdot 306) = (306, 45).$$

Next 45 is divided into 306 with quotient 6 and remainder 36, then 36 is divided into 45 with quotient 1 and remainder 9. We conclude that

$$(963, 657) = (306, 657) = (306, 45) = (36, 45) = (36, 9).$$

Thus $(963, 657) = 9$, and we can express 9 as a linear combination of 963 and 657 by sequentially writing each remainder as a linear combination of



the two original numbers:

$$\begin{aligned}
306 &= 963 - 657;\\
45 &= 657 - 2 \cdot 306 = 657 - 2 \cdot (963 - 657)\\
&= 3 \cdot 657 - 2 \cdot 963;\\
36 &= 306 - 6 \cdot 45 = (963 - 657) - 6 \cdot (3 \cdot 657 - 2 \cdot 963)\\
&= 13 \cdot 963 - 19 \cdot 657;\\
9 &= 45 - 36 = 3 \cdot 657 - 2 \cdot 963 - (13 \cdot 963 - 19 \cdot 657)\\
&= 22 \cdot 657 - 15 \cdot 963.
\end{aligned}$$

In terms of Theorem 1.3, where $g = (b, c) = bx_0 + cy_0$, beginning with $b = 963$ and $c = 657$ we have used a procedure called the *Euclidean algorithm* to find $g = 9$, $x_0 = -15$, $y_0 = 22$. Of course, these values for $x_0$ and $y_0$ are not unique: $-15 + 657k$ and $22 - 963k$ will do where $k$ is any integer.

To find the greatest common divisor $(b, c)$ of *any* two integers $b$ and $c$, we now generalize what is done in the special case above. The process will also give integers $x_0$ and $y_0$ satisfying the equation $bx_0 + cy_0 = (b, c)$. The case $c = 0$ is special: $(b, 0) = |b|$. For $c \neq 0$, we observe that $(b, c) = (b, -c)$ by Theorem 1.9, and hence, we may presume that $c$ is positive.

**Theorem 1.11** *The Euclidean algorithm. Given integers $b$ and $c > 0$, we make a repeated application of the division algorithm, Theorem 1.2, to obtain a series of equations*

$$\begin{aligned}
b &= cq_1 + r_1, & 0 < r_1 < c,\\
c &= r_1q_2 + r_2, & 0 < r_2 < r_1,\\
r_1 &= r_2q_3 + r_3, & 0 < r_3 < r_2,\\
&\cdots & \cdots\\
r_{j-2} &= r_{j-1}q_j + r_j, & 0 < r_j < r_{j-1},\\
r_{j-1} &= r_jq_{j+1}.
\end{aligned}$$

*The greatest common divisor $(b, c)$ of $b$ and $c$ is $r_j$, the last nonzero remainder in the division process. Values of $x_0$ and $y_0$ in $(b, c) = bx_0 + cy_0$ can be obtained by writing each $r_i$ as a linear combination of $b$ and $c$.*

*Proof* The chain of equations is obtained by dividing $c$ into $b$, $r_1$ into $c$, $r_2$ into $r_1, \cdots, r_j$ into $r_{j-1}$. The process stops when the division is exact, that is, when the remainder is zero. Thus in our application of Theorem 1.2 we have written the inequalities for the remainder without an equality sign. Thus, for example, $0 < r_1 < c$ in place of $0 \leq r_1 < c$, because if $r_1$

were equal to zero, the chain would stop at the first equation $b = cq_1$, in which case the greatest common divisor of $b$ and $c$ would be $c$.

We now prove that $r_j$ is the greatest common divisor $g$ of $b$ and $c$. By Theorem 1.9, we observe that

$$(b,c) = (b - cq_1, c) = (r_1, c) = (r_1, c - r_1q_2)$$

$$= (r_1, r_2) = (r_1 - r_2q_3, r_2) = (r_3, r_2).$$

Continuing by mathematical induction, we get $(b,c) = (r_{j-1}, r_j) = (r_j, 0) = r_j$.

To see that $r_j$ is a linear combination of $b$ and $c$, we argue by induction that each $r_i$ is a linear combination of $b$ and $c$. Clearly, $r_1$ is such a linear combination, and likewise $r_2$. In general, $r_i$ is a linear combination of $r_{i-1}$ and $r_{i-2}$. By the inductive hypothesis we may suppose that these latter two numbers are linear combinations of $b$ and $c$, and it follows that $r_i$ is also a linear combination of $b$ and $c$.

**Example 1** Find the greatest common divisor of 42823 and 6409.

*Solution* We apply the Euclidean algorithm, using a calculator. We divide $c$ into $b$, where $b = 42823$ and $c = 6409$, following the notation of Theorem 1.11. The quotient $q_1$ and remainder $r_1$ are $q_1 = 6$ and $r_1 = 4369$, with the details of this division as follows. Assuming the use of the simplest kind of hand-held calculator with only the four basic operations $+, -, \times, \div$, when 6409 is divided into 42823 the calculator gives 6.6816976, or some version of this with perhaps fewer decimal places. So we know that the quotient is 6. To get the remainder, we multiply 6 by 6409 to get 38454, and we subtract this from 42823 to get the remainder 4369.

Continuing, if we divide 4369 into 6409 we get a quotient $q_2 = 1$ and remainder $r_2 = 2040$. Dividing 2040 into 4369 gives $q_3 = 2$ and $r_3 = 289$. Dividing 289 into 2040 gives $q_4 = 7$ and $r_4 = 17$. Since 17 is an exact divisor of 289, the solution is that the g.c.d. is 17.

This can be put in tabular form as follows:

$$\begin{aligned} 42823 &= 6 \cdot 6409 + 4369 && (42823, 6409) \\ 6409 &= 1 \cdot 4369 + 2040 && = (6409, 4369) \\ 4369 &= 2 \cdot 2040 + 289 && = (4369, 2040) \\ 2040 &= 7 \cdot 289 + 17 && = (2040, 289) \\ 289 &= 17 \cdot 17 && = (289, 17) = 17 \end{aligned}$$



**Example 2** Find integers $x$ and $y$ to satisfy

$$42823x + 6409y = 17.$$

*Solution* We find integers $x_i$ and $y_i$ such that

$$42823x_i + 6409y_i = r_i.$$

Here it is natural to consider $i = 1, 2, \cdots$, but to initiate the process we also consider $i = 0$ and $i = -1$. We put $r_{-1} = 42823$, and write

$$42823 \cdot 1 + 6409 \cdot 0 = 42823.$$

Similarly, we put $r_0 = 6409$, and write

$$42823 \cdot 0 + 6409 \cdot 1 = 6409.$$

We multiply the second of these equations by $q_1 = 6$, and subtract the result from the first equation, to obtain

$$42823 \cdot 1 + 6409 \cdot (-6) = 4369.$$

We multiply this equation by $q_2 = 1$, and subtract it from the preceding equation to find that

$$42823 \cdot (-1) + 6409 \cdot 7 = 2040.$$

We multiply this by $q_3 = 2$, and subtract the result from the preceding equation to find that

$$42823 \cdot 3 + 6409 \cdot (-20) = 289.$$

Next we multiply this by $q_4 = 7$, and subtract the result from the preceding equation to find that

$$42823 \cdot (-22) + 6409 \cdot 147 = 17.$$

On dividing 17 into 289, we find that $q_5 = 17$ and that $289 = 17 \cdot 17$. Thus $r_4$ is the last positive remainder, so that $g = 17$, and we may take $x = -22$, $y = 147$. These values of $x$ and $y$ are not the only ones possible. In Section 5.1, an analysis of *all* solutions of a linear equation is given.

***Remark on Calculation.*** We note that $x_i$ is determined from $x_{i-1}$ and $x_{i-2}$ by the same formula that $r_i$ is determined from $r_{i-1}$ and $r_{i-2}$. That is,

$$r_i = r_{i-2} - q_i r_{i-1},$$

$$x_i = x_{i-2} - q_i x_{i-1},$$

and similarly

$$y_i = y_{i-2} - q_i y_{i-1}.$$

The only distinction between the three sequences $r_i$, $x_i$, and $y_i$ is that they start from different initial conditions:

$$r_{-1} = b, \quad r_0 = c,$$

$$x_{-1} = 1, \quad x_0 = 0,$$

and

$$y_{-1} = 0, \quad y_0 = 1.$$

Just as polynomial division may be effected symbolically, omitting the powers of the variable, we may generate the $q_i, r_i, x_i, y_i$ in a compact table. In the numerical example just considered, this would take the following form:

| $i$ | $q_{i+1}$ | $r_i$ | $x_i$ | $y_i$ |
|---|---|---|---|---|
| −1 | | 42823 | 1 | 0 |
| 0 | 6 | 6409 | 0 | 1 |
| 1 | 1 | 4369 | 1 | −6 |
| 2 | 2 | 2040 | −1 | 7 |
| 3 | 7 | 289 | 3 | −20 |
| 4 | 17 | 17 | −22 | 147 |
| 5 | | 0 | | |

When implemented on a computer, it is unnecessary to record the entire table. Each row is generated solely from the two preceding rows, so it suffices to keep only the two latest rows. In the numerical cases we have considered it has been the case that $b > c$. Although it is natural to start in this way, it is by no means necessary. If $b < c$, then $q_1 = 0$ and $r_1 = b$, which has the effect of interchanging $b$ and $c$.

**Example 3** Find $g = (b, c)$ where $b = 5033464705$ and $c = 3137640337$, and determine $x$ and $y$ such that $bx + cy = g$.



*Solution* We calculate:

| | | | |
|---|---|---|---|
| | 5033464705 | 1 | 0 |
| 1 | 3137640337 | 0 | 1 |
| 1 | 1895824368 | 1 | −1 |
| 1 | 1241815969 | −1 | 2 |
| 1 | 654008399 | 2 | −3 |
| 1 | 587807570 | −3 | 5 |
| 8 | 66200829 | 5 | −8 |
| 1 | 58200938 | −43 | 69 |
| 7 | 7999891 | 48 | −77 |
| 3 | 2201701 | −379 | 608 |
| 1 | 1394788 | 1185 | −1901 |
| 1 | 806913 | −1564 | 2509 |
| 1 | 587875 | 2749 | −4410 |
| 2 | 219038 | −4313 | 6919 |
| 1 | 149799 | 11375 | −18248 |
| 2 | 69239 | −15688 | 25167 |
| 6 | 11321 | 42751 | −68582 |
| 8 | 1313 | −272194 | 436659 |
| 1 | 817 | 2220303 | −3561854 |
| 1 | 496 | −2492497 | 3998513 |
| 1 | 321 | 4712800 | −7560367 |
| 1 | 175 | −7205297 | 11558880 |
| 1 | 146 | 11918097 | −19119247 |
| 5 | 29 | −19123394 | 30678127 |
| 29 | 1 | 107535067 | −172509882 |

Thus $g = 1$, and we may take $x = 107535067$, $y = -172509882$.

The exact number of iterations $j$ of the Euclidean algorithm required to calculate $(b, c)$ depends in an intricate manner on $b$ and $c$, but it is easy to establish a rough bound for $j$ as follows: If $r_i$ is small compared with $r_{i-1}$, say $r_i \leqslant r_{i-1}/2$, then substantial progress has been made at this step. Otherwise $r_{i-1}/2 < r_i < r_{i-1}$, in which case $q_{i+1} = 1$, and $r_{i+1} = r_{i-1} - r_i < r_{i-1}/2$. Thus we see that $r_{i+1} < r_{i-1}/2$ in either case. From this it can be deduced that $j < 3 \log c$. (Here, and throughout this book, we employ the natural logarithm, to the base $e$. Some writers denote this function $\ln x$.) With more care we could improve on the constant 3 (see Problem 10 in Section 4.4), but it is nevertheless the case that $j$ is comparable to $\log c$

for most pairs $b, c$. Since the logarithm increases very slowly, the practical consequence is that one can calculate the g.c.d. quickly, even when $b$ and $c$ are very large.

**Definition 1.4** *The integers $a_1, a_2, \cdots, a_n$, all different from zero, have a common multiple $b$ if $a_i|b$ for $i = 1, 2, \cdots, n$. (Note that common multiples do exist; for example the product $a_1 a_2 \cdots a_n$ is one.) The least of the positive common multiples is called the* least common multiple, *and it is denoted by $[a_1, a_2, \cdots, a_n]$.*

**Theorem 1.12** *If $b$ is any common multiple of $a_1, a_2, \cdots, a_n$, then $[a_1, a_2, \cdots, a_n]|b$. This is the same as saying that if $h$ denotes $[a_1, a_2, \cdots, a_n]$, then $0, \pm h, \pm 2h, \pm 3h, \cdots$ comprise all the common multiples of $a_1, a_2, \cdots, a_n$.*

*Proof* Let $m$ be any common multiple and divide $m$ by $h$. By Theorem 1.2 there is a quotient $q$ and a remainder $r$ such that $m = qh + r$, $0 \leqslant r < h$. We must prove that $r = 0$. If $r \neq 0$ we argue as follows. For each $i = 1, 2, \cdots, n$ we know that $a_i|h$ and $a_i|m$, so that $a_i|r$. Thus $r$ is a positive common multiple of $a_1, a_2, \cdots, a_n$ contrary to the fact that $h$ is the least of all the positive common multiples.

**Theorem 1.13** *If $m > 0$, $[ma, mb] = m[a, b]$. Also $[a, b] \cdot (a, b) = |ab|$.*

*Proof* Let $H = [ma, mb]$, and $h = [a, b]$. Then $mh$ is a multiple of $ma$ and $mb$, so that $mh \geqslant H$. Also, $H$ is a multiple of both $ma$ and $mb$, so $H/m$ is a multiple of $a$ and $b$. Thus, $H/m \geqslant h$, from which it follows that $mh = H$, and this establishes the first part of the theorem.

It will suffice to prove the second part for positive integers $a$ and $b$, since $[a, -b] = [a, b]$. We begin with the special case where $(a, b) = 1$. Now $[a, b]$ is a multiple of $a$, say $ma$. Then $b|ma$ and $(a, b) = 1$, so by Theorem 1.10 we conclude that $b|m$. Hence $b \leqslant m$, $ba \leqslant ma$. But $ba$, being a positive common multiple of $b$ and $a$, cannot be less than the least common multiple, so $ba = ma = [a, b]$.

Turning to the general case where $(a, b) = g > 1$, we have $(a/g, b/g) = 1$ by Theorem 1.7. Applying the result of the preceding paragraph, we obtain

$$\left[\frac{a}{g}, \frac{b}{g}\right]\left(\frac{a}{g}, \frac{b}{g}\right) = \frac{a}{g}\frac{b}{g}.$$

Multiplying by $g^2$ and using Theorem 1.6 as well as the first part of the present theorem, we get $[a, b](a, b) = ab$.



## PROBLEMS

1. By using the Euclidean algorithm, find the greatest common divisor (g.c.d.) of
   (*a*) 7469 and 2464; (*b*) 2689 and 4001;
   (*c*) 2947 and 3997; (*d*) 1109 and 4999.
2. Find the greatest common divisor $g$ of the numbers 1819 and 3587, and then find integers $x$ and $y$ to satisfy
   $$1819x + 3587y = g.$$
3. Find values of $x$ and $y$ to satisfy
   (*a*) $423x + 198y = 9$;
   (*b*) $71x - 50y = 1$;
   (*c*) $43x + 64y = 1$;
   (*d*) $93x - 81y = 3$;
   (*e*) $6x + 10y + 15z = 1$.
4. Find the least common multiple (l.c.m.) of (*a*) 482 and 1687, (*b*) 60 and 61.
5. How many integers between 100 and 1000 are divisible by 7?
6. Prove that the product of three consecutive integers is divisible by 6; of four consecutive integers by 24.
7. Exhibit three integers that are relatively prime but not relatively prime in pairs.
8. Two integers are said to be of the same *parity* if they are both even or both odd; if one is even and the other odd, they are said to be of opposite parity, or of different parity. Given any two integers, prove that their sum and their difference are of the same parity.
9. Show that if $ac|bc$ then $a|b$.
10. Given $a|b$ and $c|d$, prove that $ac|bd$.
11. Prove that $4 \nmid (n^2 + 2)$ for any integer $n$.
12. Given that $(a, 4) = 2$ and $(b, 4) = 2$, prove that $(a + b, 4) = 4$.
13. Prove that $n^2 - n$ is divisible by 2 for every integer $n$; that $n^3 - n$ is divisible by 6; that $n^5 - n$ is divisible by 30.
14. Prove that if $n$ is odd, $n^2 - 1$ is divisible by 8.
15. Prove that if $x$ and $y$ are odd, then $x^2 + y^2$ is even but not divisible by 4.
16. Prove that if $a$ and $b$ are positive integers satisfying $(a, b) = [a, b]$ then $a = b$.
17. Evaluate $(n, n + 1)$ and $[n, n + 1]$ where $n$ is a positive integer.

plane. Similarly, draw $n-2$ additional regular pentagons on the base sides $P_1P_3, P_1P_4, \cdots, P_1P_n$, all pentagons lying on the same side of the line $P_1P_n$. Mark dots at each vertex and at unit intervals along the sides of these pentagons. Prove that the total number of dots in the array is $(3n^2-n)/2$. In general, if regular $k$-gons are constructed on the sides $P_1P_2, P_1P_3, \cdots, P_1P_n$, with dots marked again at unit intervals, prove that the total number of dots is $1 + kn(n-1)/2 - (n-1)^2$. This is the $n$th $k$-gonal number.

*49. Prove that if $m > n$ then $a^{2^n}+1$ is a divisor of $a^{2^m}-1$. Show that if $a, m, n$ are positive with $m \neq n$, then

$$(a^{2^m}+1, a^{2^n}+1) = \begin{cases} 1 \text{ if } a \text{ is even} \\ 2 \text{ if } a \text{ is odd.} \end{cases}$$

*50. Show that if $(a,b)=1$ then $(a+b, a^2-ab+b^2) = 1$ or 3.

*51. Show that if $(a,b)=1$ and $p$ is an odd prime, then

$$\left(a+b, \frac{a^p+b^p}{a+b}\right) = 1 \text{ or } p.$$

*52. Suppose that $2^n+1 = xy$, where $x$ and $y$ are integers $>1$ and $n>0$. Show that $2^a | (x-1)$ if and only if $2^a | (y-1)$.

*53. Show that $(n!+1, (n+1)!+1) = 1$.

**54. Let $a$ and $b$ be positive integers such that $(1+ab) | (a^2+b^2)$. Show that the integer $(a^2+b^2)/(1+ab)$ must be a perfect square.

## 1.3 PRIMES

**Definition 1.5** *An integer $p>1$ is called a* prime number, *or a* prime, *in case there is no divisor $d$ of $p$ satisfying $1<d<p$. If an integer $a>1$ is not a prime, it is called a* composite number.

Thus, for example, 2, 3, 5, and 7 are primes, whereas 4, 6, 8, and 9 are composite.

**Theorem 1.14** *Every integer $n$ greater than 1 can be expressed as a product of primes (with perhaps only one factor).*

*Proof* If the integer $n$ is a prime, then the integer itself stands as a "product" with a single factor. Otherwise $n$ can be factored into, say,

**Problems marked with a double asterisk are much more difficult.



$n_1n_2$, where $1<n_1<n$ and $1<n_2<n$. If $n_1$ is a prime, let it stand; otherwise it will factor into, say, $n_3n_4$ where $1<n_3<n_1$ and $1<n_4<n_1$; similarly for $n_2$. This process of writing each composite number that arises as a product of factors must terminate because the factors are smaller than the composite number itself, and yet each factor is an integer greater than 1. Thus we can write $n$ as a product of primes, and since the prime factors are not necessarily distinct, the result can be written in the form

$$n = p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_r^{\alpha_r},$$

where $p_1, p_2, \cdots, p_r$ are distinct primes and $\alpha_1, \alpha_2, \cdots, \alpha_r$ are positive.

This representation of $n$ as a product of primes is called the *canonical factoring of $n$ into prime powers*. It turns out that the representation is unique in the sense that, for fixed $n$, any other representation is merely a reordering or permutation of the factors. Although it may appear obvious that the factoring of an integer into a product of primes is unique, nevertheless, it requires proof. Historically, mathematicians took the unique factorization theorem for granted, but the great mathematician Gauss stated the result and proved it in a systematic way. It is proved later in the chapter as Theorem 1.16. The importance of this result is suggested by one of the names given to it, *the fundamental theorem of arithmetic*. This unique factorization property is needed to establish much of what comes later in the book. There are mathematical systems, notably in algebraic number theory, which is discussed in Chapter 9, where unique factorization fails to hold, and the absence of this property causes considerable difficulty in a systematic analysis of the subject. To demonstrate that unique factorization need not hold in a mathematical system, we digress from the main theme for a moment to present two examples in which factorization is not unique. The first example is easy; the second is much harder to follow, so it might well be omitted on a first reading of this book.

First consider the class $\mathscr{E}$ of positive even integers, so that the elements of $\mathscr{E}$ are $2, 4, 6, 8, 10, \cdots$. Note that $\mathscr{E}$ is a multiplicative system, the product of any two elements in $\mathscr{E}$ being again in $\mathscr{E}$. Now let us confine our attention to $\mathscr{E}$ in the sense that the only "numbers" we know are members of $\mathscr{E}$. Then $8 = 2 \cdot 4$ is "composite," whereas 10 is a "prime" since 10 is not the product of two or more "numbers." The "primes" are $2, 6, 10, 14, \cdots$, the "composite numbers" are $4, 8, 12, \cdots$. Now the "number" 60 has two factorings into "primes," namely $60 = 2 \cdot 30 = 6 \cdot 10$, and so factorization is not unique.

A somewhat less artificial, but also rather more complicated, example is obtained by considering the class $\mathscr{E}$ of numbers $a + b\sqrt{-6}$ where $a$ and $b$ range over all integers. We say that this system $\mathscr{E}$ is *closed* under

addition and multiplication, meaning that the sum and product of two elements in $\mathscr{E}$ are elements of $\mathscr{E}$. By taking $b = 0$ we note that the integers form a subset of the class $\mathscr{E}$.

First we establish that there are primes in $\mathscr{E}$, and that every number in $\mathscr{E}$ can be factored into primes. For any number $a + b\sqrt{-6}$ in $\mathscr{E}$ it will be convenient to have a norm, $N(a + b\sqrt{-6})$, defined as

$$N(a + b\sqrt{-6}) = (a + b\sqrt{-6})(a - b\sqrt{-6}) = a^2 + 6b^2.$$

Thus the norm of a number in $\mathscr{E}$ is the product of the complex number $a + b\sqrt{-6}$ and its conjugate $a - b\sqrt{-6}$. Another way of saying this, perhaps in more familiar language, is that the norm is the square of the absolute value. Now the norm of every number in $\mathscr{E}$ is a positive integer greater than 1, except for the numbers $0, 1, -1$ for which we have $N(0) = 0$, $N(1) = 1$, $N(-1) = 1$. We say that we have a factoring of $a + b\sqrt{-6}$ if we can write

$$a + b\sqrt{-6} = (x_1 + y_1\sqrt{-6})(x_2 + y_2\sqrt{-6}) \tag{1.1}$$

where $N(x_1 + y_1\sqrt{-6}) > 1$ and $N(x_2 + y_2\sqrt{-6}) > 1$. This restriction on the norms of the factors is needed to rule out such trivial factorings as $a + b\sqrt{-6} = (1)(a + b\sqrt{-6}) = (-1)(-a - b\sqrt{-6})$. The norm of a product can be readily calculated to be the product of the norms of the factors, so that in the factoring (1.1) we have $N(a + b\sqrt{-6}) = N(x_1 + y_1\sqrt{-6})N(x_2 + y_2\sqrt{-6})$. It follows that

$$1 < N(x_1 + y_1\sqrt{-6}) < N(a + b\sqrt{-6}),$$

$$1 < N(x_2 + y_2\sqrt{-6}) < N(a + b\sqrt{-6})$$

so any number $a + b\sqrt{-6}$ will break up into only a finite number of factors since the norm of each factor is an integer.

We remarked above that the norm of any number in $\mathscr{E}$, apart from 0 and $\pm 1$, is greater than 1. More can be said. Since $N(a + b\sqrt{-6})$ has the value $a^2 + 6b^2$, we observe that

$$N(a + b\sqrt{-6}) \geqslant 6 \qquad \text{if } b \neq 0, \tag{1.2}$$

that is, the norm of any nonreal number in $\mathscr{E}$ is not less than 6.

A number of $\mathscr{E}$ having norm $> 1$, but that cannot be factored in the sense of (1.1), is called a *prime in* $\mathscr{E}$. For example, 5 is a prime in $\mathscr{E}$, for in the first place, 5 cannot be factored into real numbers in $\mathscr{E}$. In the second



place, if we had a factoring $5 = (x_1 + y_1\sqrt{-6})(x_2 + y_2\sqrt{-6})$ into complex numbers, we could take norms to get

$$25 = N(x_1 + y_1\sqrt{-6})N(x_2 + y_2\sqrt{-6}),$$

which contradicts (1.2). Thus, 5 is a prime in $\mathscr{E}$, and a similar argument establishes that 2 is a prime.

We are now in a position to show that not all numbers of $\mathscr{E}$ factor uniquely into primes. Consider the number 10 and its two factorings:

$$10 = 2 \cdot 5 = (2 + \sqrt{-6})(2 - \sqrt{-6}).$$

The first product $2 \cdot 5$ has factors that are prime in $\mathscr{E}$, as we have seen. Thus we can conclude that there is not unique factorization of the number 10 in $\mathscr{E}$. Note that this conclusion does not depend on our knowing that $2 + \sqrt{-6}$ and $2 - \sqrt{-6}$ are primes; they actually are, but it is unimportant in our discussion.

This example may also seem artificial, but it is, in fact, taken from an important topic, algebraic number theory, discussed in Chapter 9.

We now return to the discussion of unique factorization in the ordinary integers $0, \pm 1, \pm 2, \cdots$. It will be convenient to have the following result.

**Theorem 1.15** *If $p|ab$, $p$ being a prime, then $p|a$ or $p|b$. More generally, if $p|a_1a_2 \cdots a_n$, then $p$ divides at least one factor $a_i$ of the product.*

*Proof* If $p \nmid a$, then $(a, p) = 1$ and so by Theorem 1.10, $p|b$. We may regard this as the first step of a proof of the general statement by mathematical induction. So we assume that the proposition holds whenever $p$ divides a product with fewer than $n$ factors. Now if $p|a_1a_2 \cdots a_n$, that is, $p|a_1c$ where $c = a_2a_3 \cdots a_n$, then $p|a_1$ or $p|c$. If $p|c$ we apply the induction hypothesis to conclude that $p|a_i$ for some subscript $i$ from 2 to $n$.

**Theorem 1.16** *The fundamental theorem of arithmetic, or the unique factorization theorem. The factoring of any integer $n > 1$ into primes is unique apart from the order of the prime factors.*

*First Proof* Suppose that there is an integer $n$ with two different factorings. Dividing out any primes common to the two representations, we would have an equality of the form

$$p_1p_2 \cdots p_r = q_1q_2 \cdots q_s, \tag{1.3}$$

where the factors $p_i$ and $q_j$ are primes, not necessarily all distinct, but where no prime on the left side occurs on the right side. But this is impossible because $p_1 | q_1 q_2 \cdots q_s$, so by Theorem 1.15, $p_1$ is a divisor of at least one of the $q_j$. That is, $p_1$ must be identical with at least one of the $q_j$.

*Second Proof* Suppose that the theorem is false and let $n$ be the smallest positive integer having more than one representation as the product of primes, say

$$n = p_1 p_2 \cdots p_r = q_1 q_2 \cdots q_s. \tag{1.4}$$

It is clear that $r$ and $s$ are greater than 1. Now the primes $p_1, p_2, \cdots, p_r$ have no members in common with $q_1, q_2, \cdots, q_s$ because if, for example, $p_1$ were a common prime, then we could divide it out of both sides of (1.4) to get two distinct factorings of $n/p_1$. But this would contradict our assumption that all integers smaller than $n$ are uniquely factorable.

Next, there is no loss of generality in presuming that $p_1 < q_1$, and we define the positive integer $N$ as

$$N = (q_1 - p_1) q_2 q_3 \cdots q_s = p_1(p_2 p_3 \cdots p_r - q_2 q_3 \cdots q_s). \tag{1.5}$$

It is clear that $N < n$, so that $N$ is uniquely factorable into primes. But $p_1 \nmid (q_1 - p_1)$, so (1.5) gives us two factorings of $N$, one involving $p_1$ and the other not, and thus we have a contradiction.

In the application of the fundamental theorem we frequently write any integer $a \geqslant 1$ in the form

$$a = \prod_p p^{\alpha(p)}$$

where $\alpha(p)$ is a non-negative integer, and it is understood that $\alpha(p) = 0$ for all sufficiently large primes $p$. If $a = 1$ then $\alpha(p) = 0$ for all primes $p$, and the product may be considered to be empty. For brevity we sometimes write $a = \prod p^{\alpha}$, with the tacit understanding that the exponents $\alpha$ depend on $p$ and, of course on $a$. If

$$a = \prod_p p^{\alpha(p)}, \qquad b = \prod_p p^{\beta(p)}, \qquad c = \prod_p p^{\gamma(p)}, \tag{1.6}$$

and $ab = c$, then $\alpha(p) + \beta(p) = \gamma(p)$ for all $p$, by the fundamental theorem. Here $a | c$, and we note that $\alpha(p) \leqslant \gamma(p)$ for all $p$. If, conversely, $\alpha(p) \leqslant \gamma(p)$ for all $p$, then we may define an integer $b = \prod p^{\beta(p)}$ with



$\beta(p) = \gamma(p) - \alpha(p)$. Then $ab = c$, which is to say that $a | c$. Thus we see that the divisibility relation $a | c$ is equivalent to the family of inequalities $\alpha(p) \leqslant \gamma(p)$. As a consequence, the greatest common divisor and the least common multiple can be written as

$$(a, b) = \prod_p p^{\min(\alpha(p), \beta(p))}, \qquad [a, b] = \prod_p p^{\max(\alpha(p), \beta(p))}. \tag{1.7}$$

For example, if $a = 108$ and $b = 225$, then

$$a = 2^2 3^3 5^0, \qquad b = 2^0 3^2 5^2,$$

$$(a, b) = 2^0 3^2 5^0 = 9, \qquad [a, b] = 2^2 3^3 5^2 = 2700.$$

The first part of Theorem 1.13, like many similar identities, follows easily from the fundamental theorem in conjunction with (1.7). Since $\min(\alpha, \beta) + \max(\alpha, \beta) = \alpha + \beta$ for any real numbers $\alpha, \beta$, the relations (1.7) also provide a means of establishing the second part of Theorem 1.13. On the other hand, for calculational purposes the identifies (1.7) should only be used when the factorizations of $a$ and $b$ are already known, as in general the task of factoring $a$ and $b$ will involve much more computation than is required if one determines $(a, b)$ by the Euclidean algorithm.

We call $a$ a *square* (or alternatively a *perfect square*) if it can be written in the form $n^2$. By the fundamental theorem we see that $a$ is a square if and only if all the exponents $\alpha(p)$ in (1.6) are even. We say that $a$ is *square-free* if 1 is the largest square dividing $a$. Thus $a$ is square-free if and only if the exponents $\alpha(p)$ take only the values 0 and 1. Finally, we observe that if $p$ is prime, then the assertion $p^k \| a$ is equivalent to $k = \alpha(p)$.

**Theorem 1.17** *Euclid. The number of primes is infinite. That is, there is no end to the sequence of primes*

$$2, 3, 5, 7, 11, 13, \cdots.$$

*Proof* Suppose that $p_1, p_2, \cdots, p_r$ are the first $r$ primes. Then form the number

$$n = 1 + p_1 p_2 \cdots p_r.$$

Note that $n$ is not divisible by $p_1$ or $p_2$ or $\cdots$ or $p_r$. Hence any prime divisor $p$ of $n$ is a prime distinct from $p_1, p_2, \cdots, p_r$. Since $n$ is either a prime or has a prime factor $p$, this implies that there is a prime distinct from $p_1, p_2, \cdots, p_r$. Thus we see that for any finite $r$, the number of primes is not exactly $r$. Hence the number of primes is infinite.

Students often note that the first few of the numbers $n$ here are primes. However, $1 + 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 = 59 \cdot 509$.

**Theorem 1.18** *There are arbitrarily large gaps in the series of primes. Stated otherwise, given any positive integer $k$, there exist $k$ consecutive composite integers.*

*Proof* Consider the integers

$$(k+1)!+2, (k+1)!+3, \cdots, (k+1)!+k, (k+1)!+k+1.$$

Every one of these is composite because $j$ divides $(k+1)!+j$ if $2 \leqslant j \leqslant k+1$.

The primes are spaced rather irregularly, as the last theorem suggests. If we denote the number of primes that do not exceed $x$ by $\pi(x)$, we may ask about the nature of this function. Because of the irregular occurrence of the primes, we cannot expect a simple formula for $\pi(x)$, but we may seek to estimate its rate of growth. The proof of Theorem 1.17 can be used to derive a lower bound for $\pi(x)$, but the estimate obtained, $\pi(x) > c \log\log x$, is very weak. We now derive an inequality that is more suggestive of the true state of affairs.

**Theorem 1.19** *For every real number $y \geqslant 2$,*

$$\sum_{p \leqslant y} \frac{1}{p} > \log\log y - 1.$$

Here it is understood that the sum is over all primes $p \leqslant y$. From this it follows that the infinite series $\Sigma 1/p$ diverges, which provides a second proof of Theorem 1.17.

*Proof* Let $y$ be given, $y \geqslant 2$, and let $\mathcal{N}$ denote the set of all those positive integers $n$ that are composed entirely of primes $p$ not exceeding $y$. Since there are only finitely many primes $p \leqslant y$, and since the terms of an absolutely convergent infinite series may be arbitrarily rearranged, we see that

$$\prod_{p \leqslant y} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \cdots\right) = \sum_{n \in \mathcal{N}} \frac{1}{n}. \tag{1.8}$$



If $n$ is a positive integer $\leqslant y$ then $n \in \mathcal{N}$, and thus the sum above includes the sum $\Sigma_{n \leqslant y} 1/n$. Let $N$ denote the largest integer not exceeding $y$. By the integral test,

$$\sum_{n=1}^{N} \frac{1}{n} \geqslant \int_1^{N+1} \frac{dx}{x} = \log(N+1) > \log y.$$

Thus the right side of (1.8) is $> \log y$. On the other hand, the sum on the left side of (1.8) is a geometric series, whose value is $(1 - 1/p)^{-1}$, so we see that

$$\prod_{p \leqslant y} \left(1 - \frac{1}{p}\right)^{-1} > \log y.$$

We assume for the moment that the inequality

$$e^{v+v^2} \geqslant (1-v)^{-1} \tag{1.9}$$

holds for all real numbers $v$ in the interval $0 \leqslant v \leqslant 1/2$. Taking $v = 1/p$, we deduce that

$$\prod_{p \leqslant y} \exp\left(\frac{1}{p} + \frac{1}{p^2}\right) > \log y.$$

Since $\prod \exp(a_i) = \exp(\Sigma a_i)$, and since the logarithm function is monotonically increasing, we may take logarithms of both sides and deduce that

$$\sum_{p \leqslant y} \frac{1}{p} + \sum_{p \leqslant y} \frac{1}{p^2} > \log\log y.$$

By the comparison test we see that the second sum is

$$< \sum_{n=2}^{\infty} \frac{1}{n^2},$$

and by the integral test this is

$$< \int_1^{\infty} \frac{dx}{x^2} = 1.$$

This gives the stated inequality, but it remains to prove (1.9). We need

show that $f(v) \geqslant 1$ for $0 \leqslant v \leqslant 1/2$, where $f(v) = (1 - v)\exp(v + v^2)$. Since $f(0) = 1$, it suffices to show that $f(v)$ is increasing for $0 \leqslant v \leqslant 1/2$. To this end it is enough to observe that

$$f'(v) = v(1 - 2v)\exp(v + v^2) \geqslant 0.$$

Thus we have (1.9), and the proof is complete.

With more work it can be shown that the difference

$$\sum_{p<y} \frac{1}{p} - \log\log y$$

is a bounded function of $y$, for $y \geqslant 2$. Deeper still lies the *Prime Number Theorem*, which asserts that

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log x} = 1.$$

We say that $f(x)$ is asymptotic to $g(x)$, or write $f(x) \sim g(x)$, if $\lim_{x\to\infty} f(x)/g(x) = 1$. Thus the prime number theorem may be expressed by writing $\pi(x) \sim x/\log x$. This is one of the most important results of analytic number theory. We do not prove it in this book, but in Section 8.1 we establish a weaker estimate in this direction.

**PROBLEMS**

1. With $a$ and $b$ as in (1.6) what conditions on the exponents must be satisfied if $(a, b) = 1$?
2. What is the largest number of consecutive square-free positive integers? What is the largest number of consecutive cube-free positive integers, where $a$ is cube-free if it is divisible by the cube of no integer greater than 1?
3. In any positive integer, such as 8347, the last digit is called the *units* digit, the next the *tens* digit, the next the *hundreds* digit, and so forth. In the example 8347, the units digit is 7, the tens digit is 4, the hundreds digit is 3, and the thousands digit is 8. Prove that a number is divisible by 2 if and only if its units digit is divisible by 2; that a number is divisible by 4 if and only if the integer formed by its tens digit and its units digit is divisible by 4; that a number is divisible by 8 if and only if the integer formed by its last three digits is divisible by 8.



4. Prove that an integer is divisible by 3 if and only if the sum of its digits is divisible by 3. Prove that an integer is divisible by 9 if and only if the sum of its digits is divisible by 9.
5. Prove that an integer is divisible by 11 if and only if the difference between the sum of the digits in the odd places and the sum of the digits in the even places is divisible by 11.
6. Show that every positive integer $n$ has a unique expression of the form $n = 2^r m$, $r \geqslant 0$, $m$ a positive odd integer.
7. Show that every positive integer $n$ can be written uniquely in the form $n = ab$, where $a$ is square-free and $b$ is a square. Show that $b$ is then the largest square dividing $n$.
8. A test for divisibility by 7. Starting with any positive integer $n$, subtract double the units digit from the integer obtained from $n$ by removing the units digit, giving a smaller integer $r$. For example, if $n = 41283$ with units digit 3, we subtract 6 from 4128 to get $r = 4122$. The problem is to prove that if either $n$ or $r$ is divisible by 7, so is the other. This gives a test for divisibility by 7 by repeating the process. From 41283 we pass to 4122, then to 408 by subtracting 4 from 412, and then to 24 by subtracting 16 from 40. Since 24 is not divisible by 7, neither is 41283. (H)
9. Prove that any prime of the form $3k + 1$ is of the form $6k + 1$.
10. Prove that any positive integer of the form $3k + 2$ has a prime factor of the same form; similarly for each of the forms $4k + 3$ and $6k + 5$.
11. If $x$ and $y$ are odd, prove that $x^2 + y^2$ cannot be a perfect square.
12. If $x$ and $y$ are prime to 3, prove that $x^2 + y^2$ cannot be a perfect square.
13. If $(a, b) = p$, a prime, what are the possible values of $(a^2, b)$? Of $(a^3, b)$? Of $(a^2, b^3)$?
14. Evaluate $(ab, p^4)$ and $(a + b, p^4)$ given that $(a, p^2) = p$ and $(b, p^3) = p^2$ where $p$ is a prime.
15. If $a$ and $b$ are represented by (1.6), what conditions must be satisfied by the exponents if $a$ is to be a cube? For $a^2|b^2$?
16. Find a positive integer $n$ such that $n/2$ is a square, $n/3$ is a cube, and $n/5$ is a fifth power.
17. Twin primes are those differing by 2. Show that 5 is the only prime belonging to two such pairs. Show also that there is a one-to-one correspondence between twin primes and numbers $n$ such that $n^2 - 1$ has just four positive divisors.
18. Prove that $(a^2, b^2) = c^2$ if $(a, b) = c$.

*48. Prove that there are infinitely many primes by considering the sequence $2^{2^1}+1, 2^{2^2}+1, 2^{2^3}+1, 2^{2^4}+1, \cdots$. (H)

*49. If $g$ is a divisor of each of $ab$, $cd$, and $ac+bd$, prove that it is also a divisor of $ac$ and $bd$, where $a, b, c, d$ are integers.

*50. Show that

$$(ab, cd) = (a,c)(b,d)\left(\frac{a}{(a,c)}, \frac{d}{(b,d)}\right)\left(\frac{c}{(a,c)}, \frac{b}{(b,d)}\right).$$

*51. Show that 24 is the largest integer divisible by all integers less than its square root. (H)

*52. (For readers familiar with the rudiments of point-set topology.) We topologize the integers as follows: a set $\mathcal{N}$ of integers is open if for every $n \in \mathcal{N}$ there is an arithmetic progression $\mathcal{A}$ such that $n \in \mathcal{A} \subseteq \mathcal{N}$. (An arithmetic progression is a set of the form $\{dk + r: k \in \mathbb{Z}\}$ with $d \neq 0$.) Prove that arbitrary unions of open sets are open, and that finite intersections of open sets are open, so that these open sets define a topology in the usual sense. (From a more advanced perspective, this is known as a *profinite* topology.) As is usual in topology, we call a set $\mathcal{N}$ *closed* if its complement $\mathbb{Z} \setminus \mathcal{N}$ is open. Let $\mathcal{A}$ be an arithmetic progression. Prove that the complement of $\mathcal{A}$ is a union of arithmetic progressions. Deduce that $\mathcal{A}$ is both open and closed. Let $\mathcal{U}$ denote the union over all prime numbers $p$ of the arithmetic progressions $\{np: n \in \mathbb{Z}\}$, and let $\mathcal{V}$ denote the complement of $\mathcal{U}$. In symbols, $\mathcal{U} = \bigcup_p p\mathbb{Z}$ and $\mathcal{V} = \mathbb{Z} \setminus \mathcal{U}$. Show that $\mathcal{V} = \{-1, 1\}$. Show that if there were only finitely many prime numbers then the set $\mathcal{U}$ would be closed. From the observation that $\mathcal{V}$ is not an open set, conclude that there exist infinitely many prime numbers.

*53. Let $\pi(x)$ denote the number of primes not exceeding $x$. Show that

$$\sum_{p \leq x} 1/p = \frac{\pi(x)}{x} + \int_2^x \pi(u)/u^2\, du.$$

Using Theorem 1.19, deduce that

$$\limsup_{x \to \infty} \frac{\pi(x)}{x/\log x} \geqslant 1.$$



## 1.4 THE BINOMIAL THEOREM

We first define the *binomial coefficients* and describe them combinatorially.

**Definition 1.6** *Let $\alpha$ be any real number, and let $k$ be a non-negative integer. Then the* binomial coefficient $\binom{\alpha}{k}$ *is given by the formula*

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}.$$

Suppose that $n$ and $k$ are both integers. From the formula we see that if $0 \leqslant k \leqslant n$ then $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$, whereas if $0 \leqslant n < k$, then $\binom{n}{k} = 0$. Here we employ the convention $0! = 1$.

**Theorem 1.20** *Let $\mathcal{S}$ be a set containing exactly $n$ elements. For any non-negative integer $k$, the number of subsets of $\mathcal{S}$ containing precisely $k$ elements is $\binom{n}{k}$.*

By the definition, $\binom{4}{2} = \dfrac{4 \cdot 3}{2!} = 6$, whereas if $\mathcal{S} = \{1,2,3,4\}$ then the subsets containing two elements are $\{1,2\}, \{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}, \{3,4\}$. Because of this combinatorial interpretation, the binomial coefficient $\binom{n}{k}$ is read "$n$ choose $k$."

*Proof* Suppose that $\mathcal{S} = \{1, 2, \cdots, n\}$. These numbers may be listed in various orders, called *permutations*, here denoted by $\pi$. There are $n!$ of these permutations $\pi$, because the first term may be any one of the $n$ numbers, the second term any one of the $n-1$ remaining numbers, and the third term any one of the still remaining $n-2$ numbers, and so on. We count the permutations in a way that involves the number $X$ of subsets containing precisely $k$ elements. Let $\mathcal{A}$ be a specific subset of $\mathcal{S}$ with $k$ elements. There are $k!$ permutations of the elements of $\mathcal{A}$, each permutation having $k$ terms. Similarly there are $(n-k)!$ permutations of the $n-k$ elements not in $\mathcal{A}$. If we attach any one of these $(n-k)!$ permutations to the right end of any one of the $k!$ previous permutations, the ordered sequence of $n$ elements thus obtained is one of the permutations $\pi$ of $\mathcal{S}$. Thus we can generate $k!(n-k)!$ of the permutations $\pi$ in this way. To get all the permutations $\pi$ of $\mathcal{S}$, we repeat this procedure with $\mathcal{A}$ replaced by each of the subsets in question. Let $X$ denote the

number of these subsets. Then there are $k!(n-k)!X$ permutations $\pi$, and equating this to $n!$ we find that $X = \binom{n}{k}$.

We now see that the quotient $\dfrac{n!}{k!(n-k)!}$ is an integer, because it represents the number of ways of doing something. In this way, combinatorial interpretations can be useful in number theory. We now use Theorem 1.20 to derive the following result, which we shall need in Section 2.6.

**Theorem 1.21** *The product of any $k$ consecutive integers is divisible by $k!$.*

*Proof* Write the product as $n(n-1)\cdots(n-k+1)$. If $n \geqslant k$, then we write this in the form $\binom{n}{k}k!$, and note that $\binom{n}{k}$ is an integer, by Theorem 1.20. If $0 \leqslant n < k$, then one of the factors of our product is 0, so the product vanishes, and is therefore a multiple of $k!$ in this case also. Finally, if $n < 0$, we note that the product may be written

$$(-1)^k(-n)(-n+1)\cdots(-n+k-1) = (-1)^k\binom{-n+k-1}{k}k!.$$

Note that in this case the upper member $-n+k-1$ is at least $k$, so that by Theorem 1.20 the binomial coefficient is an integer.

In the formula for the binomial coefficients we note a symmetry:

$$\binom{n}{k} = \binom{n}{n-k}. \tag{1.10}$$

This is also evident from the combinatorial interpretation, since the subsets of $\mathscr{A}$ containing $k$ elements are in one-to-one correspondence with the complementary subsets $\mathscr{S}\setminus\mathscr{A} = \{i \in \mathscr{S}: i \notin \mathscr{A}\}$ containing $n-k$ elements.

**Theorem 1.22** *The binomial theorem. For any integer $n \geqslant 1$ and any real numbers $x$ and $y$,*

$$(x+y)^n = \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}. \tag{1.11}$$

*Proof* We consider first the product

$$\prod_{i=1}^{n}(x_i + y_i).$$



On multiplying this out, we obtain $2^n$ monomial terms of the form

$$\prod_{i\in\mathscr{A}} x_i \prod_{i\notin\mathscr{A}} y_i$$

where $\mathscr{A}$ is any subset of $\{1,2,\cdots,n\}$. For each fixed $k$, $0 \leqslant k \leqslant n$, we consider the monomial terms obtained from those subsets $\mathscr{A}$ of $\{1,2,\cdots,n\}$ having exactly $k$ elements. We set $x_i = x$ and $y_i = y$ for all $i$ and note that such a monomial has value $x^k y^{n-k}$ for the subsets in question. Since there are $\binom{n}{k}$ such subsets, we see that the contribution of such subsets is $\binom{n}{k}x^k y^{n-k}$, which gives (1.11).

The binomial theorem can also be proved analytically by appealing to the following simple result.

**Lemma 1.23** *Let $P(z) = \sum_{k=0}^{n} a_k z^k$ be a polynomial with real coefficients. Then $a_r = P^{(r)}(0)/r!$ for $0 \leqslant r \leqslant n$, where $P^{(r)}(0)$ is the $r$th derivative of $P(z)$ at $z = 0$.*

*Proof* By differentiating repeatedly, we see that

$$P^{(r)}(z) = \sum_{k=r}^{n} k(k-1)\cdots(k-r+1)a_k z^{k-r}.$$

On setting $z = 0$ we see that $P^{(r)}(0) = r!a_r$, as desired.

If we take $P(z) = (1+z)^n$, then

$$P^{(r)}(z) = n(n-1)\cdots(n-r+1)(1+z)^{n-r},$$

so that $P^{(r)}(0) = n(n-1)\cdots(n-r+1)$, and hence by the Lemma, $a_r = n(n-1)\cdots(n-r+1)/r! = \binom{n}{r}$. That is,

$$(1+z)^n = \sum_{k=0}^{n}\binom{n}{k}z^k. \tag{1.12}$$

This is a form of the binomial theorem. We can recover (1.11) by taking $z = x/y$, and then multiplying both sides by $y^n$. This gives the identity when $y \neq 0$. The case $y = 0$ of (1.11) is obvious. In our first (combinatorial) proof of this theorem, the binomial coefficients arose in the context of Theorem 1.20, but in our second (analytic) proof, they occurred in the

form described in Definition 1.6. Thus the two proofs of Theorem 1.22 may be combined to provide a second proof of Theorem 1.20.

As a matter of logic, we require only one proof of each theorem, but additional proofs often provide new insights, and the various proofs may generalize in different directions. In the present case, the first proof can be used whenever $x$ and $y$ are members of a commutative ring, whereas the second proof can be used to derive a more general form of the binomial theorem, which asserts that

$$(1+z)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} z^k \tag{1.13}$$

for $|z| < 1$. Here $\alpha$ is an arbitrary real or complex number. This is consistent with (1.12) if $\alpha$ is a non-negative integer. As a function of $\alpha$, the quantity $\binom{\alpha}{k}$ is a polynomial of degree $k$ with rational coefficients. By Theorem 1.21 we see that this polynomial takes integral values whenever $\alpha$ is an integer. A polynomial with this property is called *integer-valued*.

The series (1.13) is the Taylor series of the function on the left. To demonstrate that it converges to the desired value, one may use the integral form of the remainder, which states that if $f(z)$ is a function for which $f^{(K+1)}(z)$ is continuous, then

$$f(z) = \sum_{k=0}^{K} \frac{f^{(k)}(0)}{k!} z^k + R_K(z)$$

where

$$R_K(z) = \frac{z^{K+1}}{K!} \int_0^1 (1-t)^K f^{(K+1)}(tz)\, dt.$$

We take $f(z) = (1+z)^\alpha$, so that

$$f^{(k)}(z) = \alpha(\alpha-1)\cdots(\alpha-k+1)(1+z)^{\alpha-k}.$$

Hence

$$R_K(z) = \alpha\binom{\alpha-1}{K} z^{K+1} \int_0^1 (1-t)^K (1+tz)^{\alpha-K-1}\, dt.$$

From the hypothesis $|z| < 1$ it follows that $|1+tz| \geqslant 1 - |tz| \geqslant 1 - t$.



Hence $|1+tz|^{-K} \leqslant (1-t)^{-K}$, and we see that

$$|R_K(z)| \leqslant \left|\alpha\binom{\alpha-1}{K} z^{K+1}\right| \int_0^1 |(1+tz)^{\alpha-1}|\, dt = T_K,$$

say. Here the integral is independent of $K$, and

$$\frac{T_{K+1}}{T_K} = \left|\frac{(\alpha-K-1)z}{K+1}\right| \to |z|$$

as $K \to \infty$. Taking $r$ so that $|z| < r < 1$, we deduce that $T_{K+1} \leqslant rT_K$ for all large $K$, say $K \geqslant L$. By induction it follows that $T_K \leqslant Cr^K$ for $K \geqslant L$, where $C = T_L/r^L$. Thus $T_K \to 0$ as $K \to \infty$, and we conclude that $R_K(z) \to 0$ as $K \to \infty$. Thus (1.13) holds when $|z| < 1$.

The binomial coefficients arise in many identities, both in analysis and in combinatorics. One of the simplest of these is the recursion

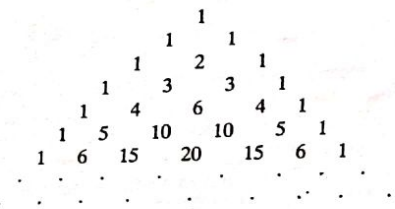
$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}, \tag{1.14}$$

used in many ways, for example, to construct *Pascal's triangle*. We define this triangle below, but first we give three short proofs of identity (1.14). Since all members vanish if $k > n$, and since the identity is clear when $k = -1$, we may assume that $0 \leqslant k \leqslant n$. First, we may simply use the formula of Definition 1.6, and then simplify the expressions. Second, we can interpret the identity combinatorially. To this end, observe that if $\mathscr{A}$ contains $k+1$ elements of $\mathscr{S} = \{1, 2, \cdots, n+1\}$, then one can consider two cases: either $n+1 \in \mathscr{A}$, or $n+1 \notin \mathscr{A}$. In the first case, $\mathscr{A}$ is determined by choosing $k$ of the numbers $1, 2, \cdots, n$; there are $\binom{n}{k}$ ways of doing this. In the second case, $\mathscr{A}$ is determined by choosing $k+1$ numbers from among $1, 2, \cdots, n$, which gives $\binom{n}{k+1}$ subsets of this type. This again gives the identity, by Theorem 1.20. Third, we note that the right side is the coefficient of $z^{k+1}$ in $(1+z)^{n+1}$. But this polynomial may be written

$$(1+z)(1+z)^n = (1+z)^n + z(1+z)^n = \sum_{k=0}^{n} \binom{n}{k} z^k + \sum_{k=0}^{n} \binom{n}{k} z^{k+1}.$$

In this last expression, the coefficient of $z^{k+1}$ is $\binom{n}{k+1} + \binom{n}{k}$. From Lemma 1.23 we see that the coefficient of $z^{k+1}$ is uniquely defined. Thus we again have (1.14).

Pascal's triangle is the infinite array of numbers



where, for example, the last row exhibited gives the binomial coefficients in the expansion of $(x+y)^6$. The identity (1.14) can be used to generate as many further rows as we please. Apart from the 1's at the ends of each row, the numbers can be obtained by adding the two integers on the preceding row, one just to the left and one just to the right. For example the next row is $1, 1+6, 6+15, 15+20$, and so on, or $1, 7, 21, 35, 35, 21, 7, 1$. The $n$th row has $n$ entries, namely the coefficients in the binomial expansion of $(x+y)^{n-1}$.

PROBLEMS

1. Use the binomial theorem to show that

$$\sum_{k=0}^{n}\binom{n}{k} = 2^n.$$

Can you give a combinatorial proof of this?

2. Show that if $n \geqslant 1$ then $\sum_{k=0}^{n}(-1)^k\binom{n}{k} = 0$.

3. (*a*) By comparing the coefficient of $z^k$ in the polynomial identity

$$\sum_{k=0}^{m+n}\binom{m+n}{k}z^k = (1+z)^{m+n} = (1+z)^m(1+z)^n$$

$$= \left(\sum_{k=0}^{m}\binom{m}{k}z^k\right)\left(\sum_{k=0}^{n}\binom{n}{k}z^k\right)$$

show that

$$\sum_{i=0}^{k}\binom{m}{i}\binom{n}{k-i} = \binom{m+n}{k}.$$



(*b*) Let $\mathscr{U}$ and $\mathscr{V}$ be disjoint sets containing $m$ and $n$ elements, respectively, and put $\mathscr{S} = \mathscr{U} \cup \mathscr{V}$. Show that the number of subsets $\mathscr{A}$ of $\mathscr{S}$ that contain $k$ elements and that also have the property that $\mathscr{A} \cap \mathscr{U}$ contains $i$ elements is $\binom{m}{i}\binom{n}{k-i}$. Interpret this identity combinatorially.

(*c*) Show that for $n \geqslant 0$,

$$\sum_{k=0}^{n}\binom{n}{k}^2 = \binom{2n}{n}.$$

4. (*a*) Suppose that $\mathscr{S}$ contains $2n$ elements, and that $\mathscr{S}$ is partitioned into $n$ disjoint subsets each one containing exactly two elements of $\mathscr{S}$. Show that this can be done in precisely

$$(2n-1)(2n-3)\cdots 5\cdot 3\cdot 1 = \frac{(2n)!}{2^n n!}$$

ways.

(*b*) Show that $(n+1)(n+2)\cdots(2n)$ is divisible by $2^n$, but not by $2^{n+1}$.

5. Show that if $a$ and $b$ are positive integers, then $a!^b b!\,|\,(ab)!$. (H)

6. Let $f(x)$ and $g(x)$ be $n$-times differentiable functions. Show that the $n$th derivative of $f(x)g(x)$ is

$$\sum_{k=0}^{n}\binom{n}{k}f^{(k)}(x)g^{(n-k)}(x).$$

7. Show that $\binom{-\alpha-1}{k} = (-1)^k\binom{\alpha+k}{k}$ for $k \geqslant 0$. Deduce that if $|z| < 1$ then

$$\frac{1}{(1-z)^{\alpha+1}} = \sum_{k=0}^{\infty}\binom{\alpha+k}{k}z^k. \qquad (1.15)$$

8. Give three proofs that

$$\sum_{m=0}^{M}\binom{m+k}{k} = \binom{k+M+1}{k+1}.$$

(*a*) With $k$ fixed, induct on $M$, using Theorem 1.20.

(*b*) Let $\mathscr{S} = \{1, 2, \cdots, k+M+1\}$. Count the number of subsets $\mathscr{A}$ of $\mathscr{S}$ containing $k+1$ elements, with the maximum one being $k+m+1$.

(*c*) Compute the coefficient of $z^M$ in the identity

$$(1+z+z^2+\cdots)\cdot\frac{1}{(1-z)^{k+1}} = \frac{1}{(1-z)^{k+2}}.$$

# CHAPTER 2

# Congruences

## 2.1 CONGRUENCES

It is apparent from Chapter 1 that divisibility is a fundamental concept of number theory, one that sets it apart from many other branches of mathematics. In this chapter we continue the study of divisibility, but from a slightly different point of view. A *congruence* is nothing more than a statement about divisibility. However, it is more than just a convenient notation. It often makes it easier to discover proofs, and we shall see that congruences can suggest new problems that will lead us to new and interesting topics.

The theory of congruences was introduced by Carl Friedrich Gauss (1777–1855), one of the greatest mathematicians of all time. Gauss contributed to the theory of numbers in many outstanding ways, including the basic ideas of this chapter and the next. Although Pierre de Fermat (1601–1665) had earlier studied number theory in a somewhat systematic way, Gauss was the first to develop the subject as a branch of mathematics rather than just a scattered collection of interesting problems. In his book *Disquisitiones Arithmeticae*, written at age 24, Gauss introduced the theory of congruences, which gained ready acceptance as a fundamental tool for the study of number theory.

Some fundamental ideas of congruences are included in this first section. The theorems of Fermat and Euler are especially noteworthy, providing powerful techniques for analyzing the multiplicative aspects of congruences. These two pioneers in number theory worked in widely contrasting ways. Mathematics was an avocation for Fermat, who was a lawyer by profession. He communicated his mathematical ideas by correspondence with other mathematicians, giving very few details of the proofs of his assertions. (One of his claims is known as Fermat's "last theorem," although it is not a theorem at all as yet, having never been proved. This situation is discussed in Section 5.4.) Leonard Euler (1707–1783), on the other hand, wrote prolifically in almost all the known branches of mathematics of his time. For example, although Fermat undoubtedly was able to

prove the result attributed to him as Theorem 2.7 below, Euler in 1736 was the first to publish a proof. Years later, in 1760, Euler stated and proved his generalization of Fermat's result, which is given as Theorem 2.8 here.

**Definition 2.1** *If an integer $m$, not zero, divides the difference $a - b$, we say that $a$ is* congruent *to $b$ modulo $m$ and write $a \equiv b \pmod{m}$. If $a - b$ is not divisible by $m$, we say that $a$ is* not congruent *to $b$ modulo $m$, and in this case we write $a \not\equiv b \pmod{m}$.*

Since $a - b$ is divisible by $m$ if and only if $a - b$ is divisible by $-m$, we can generally confine our attention to a positive modulus. Indeed, we shall assume throughout the present chapter that the modulus $m$ is a positive integer.

Congruences have many properties in common with equalities. Some properties that follow easily from the definition are listed in the following theorem.

**Theorem 2.1** *Let $a, b, c, d$ denote integers. Then:*

*(1) $a \equiv b \pmod{m}$, $b \equiv a \pmod{m}$, and $a - b \equiv 0 \pmod{m}$ are equivalent statements.*
*(2) If $a \equiv b \pmod{m}$ and $b \equiv c \pmod{m}$, then $a \equiv c \pmod{m}$.*
*(3) If $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, then $a + c \equiv b + d \pmod{m}$.*
*(4) If $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, then $ac \equiv bd \pmod{m}$.*
*(5) If $a \equiv b \pmod{m}$ and $d \mid m$, $d > 0$, then $a \equiv b \pmod{d}$.*
*(6) If $a \equiv b \pmod{m}$ then $ac \equiv bc \pmod{mc}$ for $c > 0$.*

**Theorem 2.2** *Let $f$ denote a polynomial with integral coefficients. If $a \equiv b \pmod{m}$ then $f(a) \equiv f(b) \pmod{m}$.*

*Proof* We can suppose $f(x) = c_n x^n + c_{n-1}x^{n-1} + \cdots + c_0$ where the $c_i$ are integers. Since $a \equiv b \pmod{m}$ we can apply Theorem 2.1, part 4, repeatedly to find $a^2 \equiv b^2$, $a^3 \equiv b^3, \cdots, a^n \equiv b^n \pmod{m}$, and then $c_i a^i \equiv c_i b^i \pmod{m}$, and finally $c_n a^n + c_{n-1}a^{n-1} + \cdots + c_0 \equiv c_n b^n + c_{n-1}b^{n-1} + \cdots + c_0 \pmod{m}$, by Theorem 2.1 part 3.

You are, of course, well aware of the property of real numbers that if $ax = ay$ and $a \neq 0$ then $x = y$. More care must be used in dividing a congruence through by $a$.



**Theorem 2.3**

*(1) $ax \equiv ay \pmod{m}$ if and only if $x \equiv y \left(\mathrm{mod}\ \dfrac{m}{(a,m)}\right)$.*
*(2) If $ax \equiv ay \pmod{m}$ and $(a, m) = 1$, then $x \equiv y \pmod{m}$.*
*(3) $x \equiv y \pmod{m_i}$ for $i = 1, 2, \cdots, r$ if and only if $x \equiv y \pmod{[m_1, m_2, \cdots, m_r]}$.*

*Proof* (1) If $ax \equiv ay \pmod{m}$ then $ay - ax = mz$ for some integer $z$. Hence we have

$$\frac{a}{(a,m)}(y - x) = \frac{m}{(a,m)}z,$$

and thus

$$\frac{m}{(a,m)} \Big| \frac{a}{(a,m)}(y - x).$$

But $(a/(a, m), m/(a, m)) = 1$ by Theorem 1.7 and therefore $(m/(a, m)) | (y - x)$ by Theorem 1.10. That is,

$$x \equiv y \left(\mathrm{mod}\ \frac{m}{(a,m)}\right).$$

Conversely, if $x \equiv y \pmod{m/(a, m)}$, we multiply by $a$ to get $ax \equiv ay \pmod{am/(a, m)}$ by use of Theorem 2.1, part 6. But $(a, m)$ is a divisor of $a$, so we can write $ax \equiv ay \pmod{m}$ by Theorem 2.1, part 5.

For example, $15x \equiv 15y \pmod{10}$ is equivalent to $x \equiv y \pmod{2}$, which amounts to saying that $x$ and $y$ have the same parity.

(2) This is a special case of part 1. It is listed separately because we shall use it very often.

(3) If $x \equiv y \pmod{m_i}$ for $i = 1, 2, \cdots, r$, then $m_i | (y - x)$ for $i = 1, 2, \cdots, r$. That is, $y - x$ is a common multiple of $m_1, m_2, \cdots, m_r$, and therefore (see Theorem 1.12) $[m_1, m_2, \cdots, m_r] | (y - x)$. This implies $x \equiv y \pmod{[m_1, m_2, \cdots, m_r]}$.

If $x \equiv y \pmod{[m_1, m_2, \cdots, m_r]}$ then $x \equiv y \pmod{m_i}$ by Theorem 2.1 part 5, since $m_i | [m_1, m_2, \cdots, m_r]$.

In dealing with integers modulo $m$, we are essentially performing the ordinary operations of arithmetic but are disregarding multiples of $m$. In a sense we are not distinguishing between $a$ and $a + mx$, where $x$ is any integer. Given any integer $a$, let $q$ and $r$ be the quotient and remainder on

division by $m$; thus $a = qm + r$ by Theorem 1.2. Now $a \equiv r \pmod{m}$ and, since $r$ satisfies the inequalities $0 \leqslant r < m$, we see that every integer is congruent modulo $m$ to one of the values $0, 1, 2, \cdots, m-1$. Also it is clear that no two of these $m$ integers are congruent modulo $m$. These $m$ values constitute a complete residue system modulo $m$, and we now give a general definition of this term.

**Definition 2.2** *If $x \equiv y \pmod{m}$ then $y$ is called a* residue *of $x$ modulo $m$. A set $x_1, x_2, \cdots, x_m$ is called a* complete residue system *modulo $m$ if for every integer $y$ there is one and only one $x_j$ such that $y \equiv x_j \pmod{m}$.*

It is obvious that there are infinitely many complete residue systems modulo $m$, the set $1, 2, \cdots, m-1, m$ being another example.

A set of $m$ integers forms a complete residue system modulo $m$ if and only if no two integers in the set are congruent modulo $m$.

For fixed integers $a$ and $m > 0$, the set of all integers $x$ satisfying $x \equiv a \pmod{m}$ is the arithmetic progression

$$\cdots, a - 3m, a - 2m, a - m, a, a + m, a + 2m, a + 3m, \cdots.$$

This set is called a *residue class*, or *congruence class*, modulo $m$. There are $m$ distinct residue classes modulo $m$, obtained for example by taking successively $a = 1, 2, 3, \cdots, m$.

**Theorem 2.4** *If $b \equiv c \pmod{m}$, then $(b, m) = (c, m)$.*

*Proof* We have $c = b + mx$ for some integer $x$. To see that $(b, m) = (b + mx, m)$, take $a = m$ in Theorem 1.9.

**Definition 2.3** *A* reduced residue system *modulo $m$ is a set of integers $r_i$ such that $(r_i, m) = 1$, $r_i \not\equiv r_j \pmod{m}$ if $i \neq j$, and such that every $x$ prime to $m$ is congruent modulo $m$ to some member $r_i$ of the set.*

In view of Theorem 2.4 it is clear that a reduced residue system modulo $m$ can be obtained by deleting from a complete residue system modulo $m$ those members that are not relatively prime to $m$. Furthermore, all reduced residue systems modulo $m$ will contain the same number of members, a number that is denoted by $\phi(m)$. This function is called *Euler's $\phi$-function*, sometimes the *totient*. By applying this definition of $\phi(m)$ to the complete residue system $1, 2, \cdots, m$ mentioned in the paragraph following Definition 2.2, we can get what amounts to an alternative definition of $\phi(m)$, as given in the following theorem.



**Theorem 2.5** *The number $\phi(m)$ is the number of positive integers less than or equal to $m$ that are relatively prime to $m$.*

Euler's function $\phi(m)$ is of considerable interest. We shall consider it further in Sections 2.3, 4.2, 8.2, and 8.3.

**Theorem 2.6** *Let $(a, m) = 1$. Let $r_1, r_2, \cdots, r_n$ be a complete, or a reduced, residue system modulo $m$. Then $ar_1, ar_2, \cdots, ar_n$ is a complete, or a reduced, residue system, respectively, modulo $m$.*

For example, since $1, 2, 3, 4$ is a reduced residue system modulo 5, so also is $2, 4, 6, 8$. Since $1, 3, 7, 9$ is a reduced residue system modulo 10, so is $3, 9, 21, 27$.

*Proof* If $(r_i, m) = 1$, then $(ar_i, m) = 1$ by Theorem 1.8.

There are the same number of $ar_1, ar_2, \cdots, ar_n$ as of $r_1, r_2, \cdots, r_n$. Therefore we need only show that $ar_i \not\equiv ar_j \pmod{m}$ if $i \neq j$. But Theorem 2.3, part 2, shows that $ar_i \equiv ar_j \pmod{m}$ implies $r_i \equiv r_j \pmod{m}$ and hence $i = j$.

**Theorem 2.7** *Fermat's theorem. Let $p$ denote a prime. If $p \nmid a$ then $a^{p-1} \equiv 1 \pmod{p}$. For every integer $a$, $a^p \equiv a \pmod{p}$.*

We shall postpone the proof of this theorem and shall obtain it as a corollary to Theorem 2.8.

**Theorem 2.8** *Euler's generalization of Fermat's theorem. If $(a, m) = 1$, then*

$$a^{\phi(m)} \equiv 1 \pmod{m}.$$

*Proof* Let $r_1, r_2, \cdots, r_{\phi(m)}$ be a reduced residue system modulo $m$. Then by Theorem 2.6, $ar_1, ar_2, \cdots, ar_{\phi(m)}$ is also a reduced residue system modulo $m$. Hence, corresponding to each $r_i$ there is one and only one $ar_j$ such that $r_i \equiv ar_j \pmod{m}$. Furthermore, different $r_i$ will have different corresponding $ar_j$. This means that the numbers $ar_1, ar_2, \cdots, ar_{\phi(m)}$ are just the residues modulo $m$ of $r_1, r_2, \cdots, r_{\phi(m)}$, but not necessarily in the same order. Multiplying and using Theorem 2.1, part 4, we obtain

$$\prod_{j=1}^{\phi(m)} (ar_j) \equiv \prod_{i=1}^{\phi(m)} r_i \pmod{m},$$

and hence

$$a^{\phi(m)} \prod_{j=1}^{\phi(m)} r_j \equiv \prod_{j=1}^{\phi(m)} r_j \pmod{m}.$$

Now $(r_j, m) = 1$, so we can use Theorem 2.3, part 2, to cancel the $r_j$ and we obtain $a^{\phi(m)} \equiv 1 \pmod{m}$.

*Proof of Theorem 2.7* If $p \nmid a$, then $(a, p) = 1$ and $a^{\phi(p)} \equiv 1 \pmod{p}$. To find $\phi(p)$, we refer to Theorem 2.5. All the integers $1, 2, \cdots, p-1, p$ with the exception of $p$ are relatively prime to $p$. Thus we have $\phi(p) = p - 1$, and the first part of Fermat's theorem follows. The second part is now obvious.

**Theorem 2.9** *If $(a, m) = 1$ then there is an $x$ such that $ax \equiv 1 \pmod{m}$. Any two such $x$ are congruent (mod $m$). If $(a, m) > 1$ then there is no such $x$.*

*Proof* If $(a, m) = 1$, then there exist $x$ and $y$ such that $ax + my = 1$. That is, $ax \equiv 1 \pmod{m}$. Conversely, if $ax \equiv 1 \pmod{m}$, then there is a $y$ such that $ax + my = 1$, so that $(a, m) = 1$. Thus if $ax_1 \equiv ax_2 \equiv 1 \pmod{m}$, then $(a, m) = 1$, and it follows from part 2 of Theorem 2.3 that $x_1 \equiv x_2 \pmod{m}$.

The relation $ax \equiv 1 \pmod{m}$ is equivalent to the assertion that the residue class $x \pmod{m}$ is the multiplicative inverse of the residue class $a \pmod{m}$. To avoid confusion with the rational number $a^{-1} = 1/a$, we denote this residue class by $\bar{a} \pmod{m}$. The value of $\bar{a}$ is quickly found by employing the Euclidean algorithm, as described in Section 1.2. The existence of $\bar{a}$ is also evident from Theorem 2.6, for if $(a, m) = 1$, then the numbers $a, 2a, \ldots, ma$ form a complete system of residues, which is to say that one of them is $\equiv 1 \pmod{m}$. In addition, the existence of $\bar{a}$ can also be inferred from Theorem 2.8, by taking $\bar{a} = a^{\phi(m)-1}$.

**Lemma 2.10** *Let $p$ be a prime number. Then $x^2 \equiv 1 \pmod{p}$ if and only if $x \equiv \pm 1 \pmod{p}$.*

In Section 2.7 we establish a more general result (Theorem 2.26) from which the foregoing is easily derived, but we give a direct proof now, since this observation has many useful applications.

*Proof* This quadratic congruence may be expressed as $x^2 - 1 \equiv 0 \pmod{p}$. That is, $(x - 1)(x + 1) \equiv 0 \pmod{p}$, which is to say that



$p \mid (x - 1)(x + 1)$. By Theorem 1.15 it follows that $p \mid (x - 1)$ or $p \mid (x + 1)$. Equivalently, $x \equiv 1 \pmod{p}$ or $x \equiv -1 \pmod{p}$. Conversely, if either one of these latter congruences holds, then $x^2 \equiv 1 \pmod{p}$.

**Theorem 2.11** *Wilson's theorem. If $p$ is a prime, then $(p - 1)! \equiv -1 \pmod{p}$.*

*Proof* If $p = 2$ or $p = 3$, the congruence is easily verified. Thus we may assume that $p \geq 5$. Suppose that $1 < a < p - 1$. Then $(a, p) = 1$, so that by Theorem 2.9 there is a unique integer $\bar{a}$ such that $1 < \bar{a} < p - 1$ and $a\bar{a} \equiv 1 \pmod{p}$. By a second application of Theorem 2.9 we find that if $\bar{a}$ is given then there is exactly one $a$, $1 < a < p - 1$, such that $a\bar{a} \equiv 1 \pmod{p}$. Thus $a$ and $\bar{a}$ form a pair whose combined contribution to $(p - 1)!$ is $\equiv 1 \pmod{p}$. However, a little care is called for because it may happen that $a = \bar{a}$. This is equivalent to the assertion that $a^2 \equiv 1 \pmod{p}$, and by Lemma 2.10 we see that this is in turn equivalent to $a = 1$ or $a = p - 1$. That is, $\bar{1} = 1$ and $\overline{p-1} = p - 1$, but if $2 < a < p - 2$ then $\bar{a} \neq a$. By pairing these latter residues in this manner we find that $\prod_{a=2}^{p-2} a \equiv 1 \pmod{p}$, so that $(p - 1)! \equiv 1 \cdot (\prod_{a=2}^{p-2} a) \cdot (p - 1) \equiv -1 \pmod{p}$.

We give a second proof of Wilson's theorem in our remarks following Corollary 2.30 in Section 2.7, and a third proof is outlined in Problem 22 of Section 2.8.

**Theorem 2.12** *Let $p$ denote a prime. Then $x^2 \equiv -1 \pmod{p}$ has solutions if and only if $p = 2$ or $p \equiv 1 \pmod{4}$.*

*Proof* If $p = 2$ we have the solution $x = 1$.

For any odd prime $p$, we can write Wilson's theorem in the form

$$\left(1 \cdot 2 \cdots j \cdots \frac{p-1}{2}\right)\left(\frac{p+1}{2} \cdots (p-j) \cdots (p-2)(p-1)\right) \equiv -1 \pmod{p}.$$

The product on the left has been divided into two parts, each with the same number of factors. Pairing off $j$ in the first half with $p - j$ in the second half, we can rewrite the congruence in the form

$$\prod_{j=1}^{(p-1)/2} j(p-j) \equiv -1 \pmod{p}.$$

But $j(p-j) \equiv -j^2 \pmod{p}$, and so the above is

$$\equiv \prod_{j=1}^{(p-1)/2} (-j^2) \equiv (-1)^{(p-1)/2} \left( \prod_{j=1}^{(p-1)/2} j \right)^2 \pmod{p}.$$

If $p \equiv 1 \pmod 4$ then the first factor on the right is 1, and we see that $x = \left(\frac{p-1}{2}\right)!$ is a solution of $x^2 \equiv -1 \pmod{p}$.

Suppose, conversely, that there is an $x$ such that $x^2 \equiv -1 \pmod{p}$. We note that for such an $x$, $p \nmid x$. We suppose that $p > 2$, and raise both sides of the congruence to the power $(p-1)/2$ to see that

$$(-1)^{(p-1)/2} \equiv (x^2)^{(p-1)/2} = x^{p-1} \pmod{p}.$$

By Fermat's congruence, the right side here is $\equiv 1 \pmod{p}$. The left side is $\pm 1$, and since $-1 \not\equiv 1 \pmod{p}$, we deduce that

$$(-1)^{(p-1)/2} = 1.$$

Thus $(p-1)/2$ is even; that is, $p \equiv 1 \pmod 4$.

In case $p \equiv 1 \pmod 4$, we have explicitly constructed a solution of the congruence $x^2 \equiv -1 \pmod{p}$. However, the amount of calculation required to evaluate $\left(\frac{p-1}{2}\right)! \pmod{p}$ is no smaller than would be required by exhaustively testing $x = 2, x = 3, \cdots, x = (p-1)/2$. In Section 2.9 we develop a method by which the desired $x$ can be quickly determined.

Theorem 2.12 provides the key piece of information needed to determine which integers can be written as the sum of the squares of two integers. We begin by showing that a certain class of prime numbers can be represented in this manner.

**Lemma 2.13** *If $p$ is a prime number and $p \equiv 1 \pmod 4$, then there exist positive integers $a$ and $b$ such that $a^2 + b^2 = p$.*

This was first stated in 1632 by Albert Girard, on the basis of numerical evidence. The first proof was given by Fermat in 1654.

*Proof* Let $p$ be a prime number, $p \equiv 1 \pmod 4$. By Theorem 2.12 we know that there exists an integer $x$ such that $x^2 \equiv -1 \pmod{p}$. Define $f(u, v) = u + xv$, and $K = [\sqrt{p}]$. Since $\sqrt{p}$ is not an integer, it follows that



$K < \sqrt{p} < K + 1$. We consider pairs $(u, v)$ of integers for which $0 \leqslant u \leqslant$ and $0 \leqslant v \leqslant K$. Since $u$ and $v$ each take on $K+1$ values, we have $(K+1)^2$ pairs. Since $K + 1 > \sqrt{p}$, the number of pairs is $> p$. If we consider $f(u, v) \pmod{p}$, we have more numbers under consideration than we have residue classes to put them in, so there must be some residue class that contains the number $f(u, v)$ for two different pairs $(u, v)$. (This is known as the *pigeonhole principle*, which we discuss in greater detail in Section 4.5.) Suppose, for example, that $(u_1, v_1)$ and $(u_2, v_2)$ are distinct pairs with coordinates in the interval $[0, K]$, for which $f(u_1, v_1) \equiv f(u_2, v_2) \pmod{p}$. That is, $u_1 + xv_1 \equiv u_2 + xv_2 \pmod{p}$, which gives $(u_1 - u_2) \equiv -x(v_1 - v_2) \pmod{p}$. Take $a = u_1 - u_2$ and $b = v_1 - v_2$. Then $a \equiv -xb \pmod{p}$, and on squaring both sides we see that $a^2 \equiv (-xb)^2 \equiv x^2 b^2 \equiv -b^2 \pmod{p}$ since $x^2 \equiv -1 \pmod{p}$. That is, $a^2 + b^2 \equiv 0 \pmod{p}$, which is to say that $p \mid (a^2 + b^2)$. Since the ordered pair $(u_1, v_1)$ is distinct from the pair $(u_2, v_2)$, it follows that not both $a$ and $b$ vanish, so that $a^2 + b^2 > 0$. On the other hand, $u_1 \leqslant K$ and $u_2 \geqslant 0$, so that $a = u_1 - u_2 \leqslant K$. Similarly, we may show that $a \geqslant -K$, and in the same manner that $-K \leqslant b \leqslant K$. But $K < \sqrt{p}$, so this gives $|a| < \sqrt{p}$ and $|b| < \sqrt{p}$. On squaring these inequalities we find that $a^2 < p$ and $b^2 < p$, which gives $a^2 + b^2 < 2p$. Thus altogether we have shown that $0 < a^2 + b^2 < 2p$ and that $p \mid (a^2 + b^2)$. But the only multiple of $p$ in the interval $(0, 2p)$ is $p$, so we conclude that $a^2 + b^2 = p$.

We now establish a similar result in the converse direction.

**Lemma 2.14** *Let $q$ be a prime factor of $a^2 + b^2$. If $q \equiv 3 \pmod 4$ then $q \mid a$ and $q \mid b$.*

*Proof* We prove the contrapositive, that is, that if $q$ does not divide both $a$ and $b$ then $q \not\equiv 3 \pmod 4$. By interchanging $a$ and $b$, if necessary, we may suppose that $(a, q) = 1$. Let $\bar{a}$ be chosen so that $a\bar{a} \equiv 1 \pmod{q}$. We multiply both sides of the congruence $a^2 \equiv -b^2 \pmod{q}$ by $\bar{a}^2$ to see that $1 \equiv (a\bar{a})^2 \equiv -(b\bar{a})^2 \pmod{q}$. Thus if $x = b\bar{a}$ then $x$ is a solution of the congruence $x^2 \equiv -1 \pmod{q}$, and by Theorem 2.12 it follows that $q = 2$ or $q \equiv 1 \pmod 4$.

**Theorem 2.15** *Fermat. Write the canonical factorization of $n$ in the form*

$$n = 2^\alpha \prod_{p \equiv 1(4)} p^\beta \prod_{q \equiv 3(4)} q^\gamma.$$

*Then $n$ can be expressed as a sum of two squares of integers if and only if all the exponents $\gamma$ are even.*

*Proof* We note that the identity

$$(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad - bc)^2$$

holds for any real numbers. In particular, it follows that if $m$ and $n$ are both sums of two squares then $mn$ is also a sum of two squares. The prime number $2 = 1^2 + 1^2$ is a sum of two squares, and every prime number $p \equiv 1 \pmod 4$ is a sum of two squares. If $q$ is a prime number, $q \equiv 3 \pmod 4$, then $q^2 = q^2 + 0^2$ is a sum of two squares. Hence any number that may be expressed as a product of 2's, $p$'s, and $q^2$'s is a sum of two squares. Conversely, suppose that $n$ is a sum of two squares, say $n = a^2 + b^2$. If $q$ is a prime number, $q \equiv 3 \pmod 4$, for which $\gamma > 0$, then $q|n$, and by Lemma 2.14 it follows that $q|a$ and $q|b$, which implies that $q^2|n$. That is, $\gamma \geq 2$, and we may write $n/q^2 = (a/q)^2 + (b/q)^2$. By applying this same argument to $n/q^2$ we discover that if $\gamma > 2$ then $\gamma \geq 4$ and that $q^2|a$ and $q^2|b$. Since this process must terminate, we conclude that $\gamma$ must be even, and additionally that $q^{\gamma/2}|a$ and $q^{\gamma/2}|b$.

This theorem of Fermat is the first of many similar such theorems. The object of constructing a coherent theory of quadratic forms was the primary influence on research in number theory for several centuries. The first step in the theory is to generalize Theorem 2.12. This is accomplished in the law of quadratic reciprocity, which we study in the initial sections of Chapter 3. With this tool in hand, we develop some of the fundamentals concerning quadratic forms in the latter part of Chapter 3. In Section 3.6 we apply the general theory to sums of two squares, to give not only a second proof of Theorem 2.15, but also some further results.

## PROBLEMS

1. List all integers $x$ in the range $1 \leq x \leq 100$ that satisfy $x \equiv 7 \pmod{17}$.
2. Exhibit a complete residue system modulo 17 composed entirely of multiples of 3.
3. Exhibit a reduced residue system for the modulus 12; for 30.
4. If an integer $x$ is even, observe that it must satisfy the congruence $x \equiv 0 \pmod 2$. If an integer $y$ is odd, what congruence does it satisfy? What congruence does an integer $z$ of the form $6k + 1$ satisfy?
5. Write a single congruence that is equivalent to the pair of congruences $x \equiv 1 \pmod 4$, $x \equiv 2 \pmod 3$.
6. Prove that if $p$ is a prime and $a^2 \equiv b^2 \pmod p$, then $p|(a + b)$ or $p|(a - b)$.



7. Show that if $f(x)$ is a polynomial with integral coefficients and if $f(a) \equiv k \pmod m$, then $f(a + tm) \equiv k \pmod m$ for every integer $t$.
8. Prove that any number that is a square must have one of the following for its units digit: 0, 1, 4, 5, 6, 9.
9. Prove that any fourth power must have one of 0, 1, 5, 6 for its units digit.
10. Evaluate $\phi(m)$ for $m = 1, 2, 3, \cdots, 12$.
11. Find the least positive integer $x$ such that $13|(x^2 + 1)$.
12. Prove that 19 is not a divisor of $4n^2 + 4$ for any integer $n$.
13. Exhibit a reduced residue system modulo 7 composed entirely of powers of 3.
14. Show that $7|(3^{2n+1} + 2^{n+2})$ for all $n$.
15. Find integers $a_1, \cdots, a_5$ such that every integer $x$ satisfies at least one of the congruences $x \equiv a_1 \pmod 2$, $x \equiv a_2 \pmod 3$, $x \equiv a_3 \pmod 4$, $x \equiv a_4 \pmod 6$, $x \equiv a_5 \pmod{12}$.
16. Illustrate the proof of Theorem 2.11 for $p = 11$ and $p = 13$ by actually determining the pairs of associated integers.
17. Show that $61! + 1 \equiv 63! + 1 \equiv 0 \pmod{71}$.
18. Show that if $p \equiv 3 \pmod 4$, then $\left(\frac{p-1}{2}\right)! \equiv \pm 1 \pmod p$.
19. Prove that $n^6 - 1$ is divisible by 7 if $(n, 7) = 1$.
20. Prove that $n^7 - n$ is divisible by 42, for any integer $n$.
21. Prove that $n^{12} - 1$ is divisible by 7 if $(n, 7) = 1$.
22. Prove that $n^{6k} - 1$ is divisible by 7 if $(n, 7) = 1$, $k$ being any positive integer.
23. Prove that $n^{13} - n$ is divisible by 2, 3, 5, 7 and 13 for any integer $n$.
24. Prove that $n^{12} - a^{12}$ is divisible by 13 if $n$ and $a$ are prime to 13.
25. Prove that $n^{12} - a^{12}$ is divisible by 91 if $n$ and $a$ are prime to 91.
26. Show that the product of three consecutive integers is divisible by 504 if the middle one is a cube.
27. Prove that $\frac{1}{5}n^5 + \frac{1}{3}n^3 + \frac{7}{15}n$ is an integer for every integer $n$.
28. What is the last digit in the ordinary decimal representation of $3^{400}$? (H)
29. What is the last digit in the ordinary decimal representation of $2^{400}$?
30. What are the last two digits in the ordinary decimal representation of $3^{400}$? (H)
31. Show that $-(m-1)/2, -(m-3)/2, \cdots, (m-3)/2, (m-1)/2$ is a complete residue system modulo $m$ if $m$ is odd, and that $-(m-2)/2, -(m-4)/2, \cdots, (m-2)/2, m/2$ is a complete residue system modulo $m$ if $m$ is even.

pairs $(u, v)$, show that at least one of the equations $a^2 + 2b^2 = p$, $a^2 + 2b^2 = 2p$ has a solution.

57. Show that $(a + b\sqrt{-2})(c + d\sqrt{-2}) = (ac - 2bd) + (bc + ad)\sqrt{-2}$. Thus or otherwise show that $(a^2 + 2b^2)(c^2 + 2d^2) = (ac - 2bd)^2 + 2(bc + ad)^2$.

58. Show that if $p$ is an odd prime and $a^2 + 2b^2 = 2p$, then $a$ is even and $b$ is odd. Deduce that $(2b)^2 + 2a^2 = 4p$, and hence that $b^2 + 2(a/2)^2 = p$.

59. Let $p$ be a prime factor of $a^2 + 2b^2$. Show that if $p$ does not divide both $a$ and $b$ then the congruence $x^2 \equiv -2 \pmod{p}$ has a solution.

60. Combine the results of the foregoing problems to show that a prime number $p$ can be expressed in the form $a^2 + 2b^2$ if and only if the congruence $x^2 \equiv -2 \pmod{p}$ is solvable. (In Chapter 3 we show that this congruence is solvable if and only if $p = 2$ or $p \equiv 1$ or $3 \pmod 8$.)

## 2.2 SOLUTIONS OF CONGRUENCES

In analogy with the solution of algebraic equations it is natural to consider the problem of solving a congruence. In the rest of this chapter we shall let $f(x)$ denote a polynomial with integral coefficients, and we shall write $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$. If $u$ is an integer such that $f(u) \equiv 0 \pmod m$, then we say that $u$ is a solution of the congruence $f(x) \equiv 0 \pmod m$. Whether or not an integer is a solution of a congruence depends on the modulus $m$ as well as on the polynomial $f(x)$. If the integer $u$ is a solution of $f(x) \equiv 0 \pmod m$, and if $v \equiv u \pmod m$, Theorem 2.2 shows that $v$ is also a solution. Because of this we shall say that $x \equiv u \pmod m$ is a solution of $f(x) \equiv 0 \pmod m$; meaning that every integer congruent to $u$ modulo $m$ satisfies $f(x) \equiv 0 \pmod m$.

For example, the congruence $x^2 - x + 4 \equiv 0 \pmod{10}$ has the solution $x = 3$ and also the solution $x = 8$. It also has the solutions $x = 13$, $x = 18$, and all other numbers obtained from 3 and 8 by adding and subtracting 10 as often as we wish. In counting the number of solutions of a congruence, we restrict attention to a complete residue system belonging to the modulus. In the example $x^2 - x + 4 \equiv 0 \pmod{10}$, we say that there are two solutions because $x = 3$ and $x = 8$ are the only numbers among $0, 1, 2, \cdots, 9$ that are solutions. The two solutions can be written in equation form, $x = 3$ and $x = 8$, or in congruence form, $x \equiv 3 \pmod{10}$ and $x \equiv 8 \pmod{10}$. As a second example, the congruence $x^2 - 7x + 2 \equiv 0 \pmod{10}$ has exactly four solutions $x = 3, 4, 8, 9$. The reason for counting the number of solutions in this way is that if $f(x) \equiv 0 \pmod m$ has a solution $x = a$, then it follows that all integers $x$ satisfying $x \equiv a \pmod m$ are automatically solutions, so this entire congruence class is counted as a single solution.



**Definition 2.4** *Let $r_1, r_2, \cdots, r_m$ denote a complete residue system modulo $m$. The number of solutions of $f(x) \equiv 0 \pmod m$ is the number of the $r_i$ such that $f(r_i) \equiv 0 \pmod m$.*

It is clear from Theorem 2.2 that the number of solutions is independent of the choice of the complete residue system. Furthermore, the number of solutions cannot exceed the modulus $m$. If $m$ is small it is a simple matter to just compute $f(r_i)$ for each of the $r_i$ and thus to determine the number of solutions. In the foregoing example the congruence has just two solutions. Some other examples are

$$x^2 + 1 \equiv 0 \pmod 7 \text{ has no solution,}$$

$$x^2 + 1 \equiv 0 \pmod 5 \text{ has two solutions,}$$

$$x^2 - 1 \equiv 0 \pmod 8 \text{ has four solutions.}$$

**Definition 2.5** *Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$. If $a_n \not\equiv 0 \pmod m$ the degree of the congruence $f(x) \equiv 0 \pmod m$ is $n$. If $a_n \equiv 0 \pmod m$, let $j$ be the largest integer such that $a_j \not\equiv 0 \pmod m$; then the degree of the congruence is $j$. If there is no such integer $j$, that is, if all the coefficients of $f(x)$ are multiples of $m$, no degree is assigned to the congruence.*

It should be noted that the degree of the congruence $f(x) \equiv 0 \pmod m$ is not the same thing as the degree of the polynomial $f(x)$. The degree of the congruence depends on the modulus; the degree of the polynomial does not. Thus if $g(x) = 6x^3 + 3x^2 + 1$, then $g(x) \equiv 0 \pmod 5$ is of degree 3, and $g(x) \equiv 0 \pmod 2$ is of degree 2, whereas $g(x)$ is of degree 3.

**Theorem 2.16** *If $d \mid m$, $d > 0$, and if $u$ is a solution of $f(x) \equiv 0 \pmod m$, then $u$ is a solution of $f(x) \equiv 0 \pmod d$.*

*Proof* This follows directly from Theorem 2.1, part 5.

There is a distinction made in the theory of algebraic equations that has an analogue for congruences. A conditional equation, such as $x^2 - 5x + 6 = 0$, is true for only certain values of $x$, namely $x = 2$ and $x = 3$. An identity or identical equation, such as $(x - 2)^2 = x^2 - 4x + 4$, holds for all real numbers $x$, or for all complex numbers for that matter.

Similarly, we say that $f(x) \equiv 0 \pmod m$ is an *identical congruence* if it holds for all integers $x$. If $f(x)$ is a polynomial all of whose coefficients are divisible by $m$, then $f(x) \equiv 0 \pmod m$ is an identical congruence. A different type of identical congruence is illustrated by $x^p \equiv x \pmod p$, true for all integers $x$ by Fermat's theorem.

Before considering congruences of higher degree, we first describe the solutions in the linear case.

**Theorem 2.17** *Let $a$, $b$, and $m > 0$ be given integers, and put $g = (a, m)$. The congruence $ax \equiv b \pmod m$ has a solution if and only if $g|b$. If this condition is met, then the solutions form an arithmetic progression with common difference $m/g$, giving $g$ solutions* (mod $m$).

*Proof* The question is whether there exist integers $x$ and $y$ such that $ax + my = b$. Since $g$ divides the left side, for such integers to exist we must have $g|b$. Suppose that this condition is met, and write $a = g\alpha$, $b = g\beta$, $m = g\mu$. Then by the first part of Theorem 2.3, the desired congruence holds if and only if $\alpha x \equiv \beta \pmod \mu$. Here $(\alpha, \mu) = 1$ by Theorem 1.7, so by Theorem 2.9 there is a unique number $\bar\alpha \pmod \mu$ such that $\alpha\bar\alpha \equiv 1 \pmod \mu$. On multiplying through by $\bar\alpha$, we find that $x \equiv \bar\alpha\alpha x \equiv \bar\alpha\beta \pmod \mu$. Thus the set of integers $x$ for which $ax \equiv b \pmod m$ is precisely the arithmetic progression of numbers of the form $\bar\alpha\beta + k\mu$. If we allow $k$ to take on the values $0, 1, \ldots, g - 1$, we obtain $g$ values of $x$ that are distinct (mod $m$). All other values of $x$ are congruent (mod $m$) to one of these, so we have precisely $g$ solutions.

Since $\bar\alpha$ can be located by an application of the Euclidean algorithm, the solutions are easily found.

## PROBLEMS

1. If $f(x) \equiv 0 \pmod p$ has exactly $j$ solutions with $p$ a prime, and $g(x) \equiv 0 \pmod p$ has no solution, prove that $f(x)g(x) \equiv 0 \pmod p$ has exactly $j$ solutions.
2. Denoting the number of solutions of $f(x) \equiv k \pmod m$ by $N(k)$, prove that $\sum_{k=1}^{m} N(k) = m$.
3. If a polynomial congruence $f(x) \equiv 0 \pmod m$ has $m$ solutions, prove that any integer whatsoever is a solution.
4. The fact that the product of any three consecutive integers is divisible by 3 leads to the identical congruence $x(x + 1)(x + 2) \equiv 0 \pmod 3$. Generalize this, and write an identical congruence modulo $m$.



5. Find all solutions of the congruences
   (a) $20x \equiv 4 \pmod{30}$;
   (b) $20x \equiv 30 \pmod 4$;
   (c) $353x \equiv 254 \pmod{400}$;
   (d) $57x \equiv 87 \pmod{105}$;
   (e) $64x \equiv 83 \pmod{105}$;
   (f) $589x \equiv 209 \pmod{817}$;
   (g) $49x \equiv 5000 \pmod{999}$.
6. How many solutions are there to each of the following congruences:
   (a) $15x \equiv 25 \pmod{35}$;
   (b) $15x \equiv 24 \pmod{35}$;
   (c) $15x \equiv 0 \pmod{35}$?
7. If $a$ is selected at random from $1, 2, 3, \cdots, 14$, and $b$ is selected at random from $1, 2, 3, \cdots, 15$, what is the probability that $ax \equiv b \pmod{15}$ has at least one solution? Exactly one solution?
8. Show that if $p$ is an odd prime then the congruence $x^2 \equiv 1 \pmod{p^\alpha}$ has only the two solutions $x \equiv 1$, $x \equiv -1 \pmod{p^\alpha}$.
9. Show that the congruence $x^2 \equiv 1 \pmod{2^\alpha}$ has one solution when $\alpha = 1$, two solutions when $\alpha = 2$, and precisely the four solutions $1, 2^{\alpha-1} - 1, 2^{\alpha-1} + 1, -1$ when $\alpha \geqslant 3$.
10. Show that if $p$ is an odd prime then the number of solutions (ordered pairs) of the congruence $x^2 - y^2 \equiv a \pmod p$ is $p - 1$ unless $a \equiv 0 \pmod p$, in which case the number of solutions is $2p - 1$. (H)
11. Suppose $(a, m) = 1$, and let $x_1$ denote a solution of $ax \equiv 1 \pmod m$. For $s = 1, 2, \cdots$, let $x_s = 1/a - (1/a)(1 - ax_1)^s$. Prove that $x_s$ is an integer and that it is a solution of $ax \equiv 1 \pmod{m^s}$.

*12. Suppose that $(a, m) = 1$. If $a = \pm 1$, the solution of $ax \equiv 1 \pmod{m^s}$ is obviously $x \equiv a \pmod{m^s}$. If $a = \pm 2$, then $m$ is odd and $x \equiv \frac{1}{2}(1 - m^s)\frac{1}{2}a \pmod{m^s}$ is the solution of $ax \equiv 1 \pmod{m^s}$. For all other $a$ use Problem 11 to show that the solution of $ax \equiv 1 \pmod{m^s}$ is $x \equiv k \pmod{m^s}$ where $k$ is the nearest integer to $-(1/a)(1 - ax_1)^s$.

13. Solve $3x \equiv 1 \pmod{125}$ by Problem 12, taking $x_1 = 2$.

*14. Show that $\binom{p^\alpha}{k} \equiv 0 \pmod p$ for $0 < k < p^\alpha$. (H)

*15. Show that $\binom{p^\alpha - 1}{k} \equiv (-1)^k \pmod p$ for $0 \leqslant k \leqslant p^\alpha - 1$. (H)

*16. Show that if $r$ is a non-negative integer then all coefficients of the polynomial $(1 + x)^{2^r} - (1 + x^{2^r})$ are even. Write a positive integer $n$ in binary, $n = \sum_{r \in \mathscr{R}} 2^r$. Show that all coefficients of the polynomial $(1 + x)^n - \prod_{r \in \mathscr{R}} (1 + x^{2^r})$ are even. Write $k = \sum_{s \in \mathscr{S}} 2^s$ in binary. Show that $\binom{n}{k}$ is odd if and only if $\mathscr{S} \subseteq \mathscr{R}$. Conclude that if $n$ is given,

then $\binom{n}{k}$ is odd for precisely $2^{w(n)}$ values of $k$, where $w(n)$, called the *binary weight* of $n$, is the number of 1's in the binary expansion of $n$. In symbols, $w(n) = \text{card}(\mathscr{R})$.

*Note* This is a special case of a result of E. Lucas, proved in 1891. See N. J. Fine, "Binomial coefficients modulo a prime," *Amer. Math. Monthly*, 54 (1947), 589–592.

*17. Let the numbers $c_i$ be defined by the power series identity

$$(1 + x + \cdots + x^{p-1})/(1 - x)^{p-1} = 1 + c_1 x + c_2 x^2 + \cdots .$$

Show that $c_i \equiv 0 \pmod{p}$ for all $i \geq 1$.

## 2.3 THE CHINESE REMAINDER THEOREM

We now consider the important problem of solving simultaneous congruences. The simplest case of this is to find those $x$ (if there are any) that satisfy the simultaneous congruences

$$\begin{aligned} x &\equiv a_1 \pmod{m_1}, \\ x &\equiv a_2 \pmod{m_2}, \\ &\vdots \\ x &\equiv a_r \pmod{m_r}. \end{aligned} \tag{2.1}$$

This is the subject of the next result, called the *Chinese Remainder Theorem* because the method was known in China in the first century A.D.

**Theorem 2.18** *The Chinese Remainder Theorem. Let $m_1, m_2, \cdots, m_r$ denote $r$ positive integers that are relatively prime in pairs, and let $a_1, a_1, \cdots, a_r$ denote any $r$ integers. Then the congruences* (2.1) *have common solutions. If $x_0$ is one such solution, then an integer $x$ satisfies the congruences* (2.1) *if and only if $x$ is of the form $x = x_0 + km$ for some integer $k$. Here $m = m_1 m_2 \cdots m_r$.*

Using the terminology introduced in the previous section, the last assertion of the Theorem would be expressed by saying that the solution $x$ is unique modulo $m = m_1 m_2 \cdots m_r$.

*Proof* Writing $m = m_1 m_2 \cdots m_r$, we see that $m/m_j$ is an integer and that $(m/m_j, m_j) = 1$. Hence by Theorem 2.9 for each $j$ there is an integer



$b_j$ such that $(m/m_j)b_j \equiv 1 \pmod{m_j}$. Clearly $(m/m_i)b_j \equiv 0 \pmod{m_i}$ if $i \neq j$. Put

$$x_0 = \sum_{j=1}^{r} \frac{m}{m_j} b_j a_j. \tag{2.2}$$

We consider this number modulo $m_i$, and find that

$$x_0 \equiv \frac{m}{m_i} b_i a_i \equiv a_i \pmod{m_i}.$$

Thus $x_0$ is a solution of the system (2.1).

If $x_0$ and $x_1$ are two solutions of the system (2.1), then $x_0 \equiv x_1 \pmod{m_i}$ for $i = 1, 2, \cdots, r$, and hence $x_0 \equiv x_1 \pmod{m}$ by part 3 of Theorem 2.3. This completes the proof.

**Example 1** Find the least positive integer $x$ such that $x \equiv 5 \pmod 7$, $x \equiv 7 \pmod{11}$, and $x \equiv 3 \pmod{13}$.

*Solution* We follow the proof of the theorem, taking $a_1 = 5$, $a_2 = 7$, $a_3 = 3$, $m_1 = 7$, $m_2 = 11$, $m_3 = 13$, and $m = 7 \cdot 11 \cdot 13 = 1001$. Now $(m_2 m_3, m_1) = 1$, and indeed by the Euclidean algorithm we find that $(-2) \cdot m_2 m_3 + 21 \cdot m_1 = 1$, so we may take $b_1 = -2$. Similarly, we find that $4 \cdot m_1 m_3 + (-33) \cdot m_2 = 1$, so we take $b_2 = 4$. By the Euclidean algorithm a third time we find that $(-1) \cdot m_1 m_2 + 6 \cdot m_3 = 1$, so we may take $b_3 = -1$. Then by (2.2) we see that $11 \cdot 13 \cdot (-2) \cdot 5 + 7 \cdot 13 \cdot 4 \cdot 7 + 7 \cdot 11 \cdot (-1) \cdot 3 = 887$ is a solution. Since this solution is unique modulo $m$, this is the only solution among the numbers $1, 2, \cdots, 1001$. Thus 887 is the least positive solution.

In the Chinese Remainder Theorem, the hypothesis that the moduli $m_j$ should be pairwise relatively prime is absolutely essential. When this hypothesis fails, the existence of a solution $x$ of the simultaneous system (2.1) is no longer guaranteed, and when such an $x$ does exist, we see from Part 3 of Theorem 2.3 that it is unique modulo $[m_1, m_2, \cdots, m_r]$, not modulo $m$. In case there is no solution of (2.1), we call the system *inconsistent*. In the following two examples we explore some of the possibilities that arise when the $m_j$ are allowed to have common factors. An extension of the Chinese Remainder Theorem to the case of unrestricted $m_j$ is laid out in Problems 19–23.

**Example 2** Show that there is no $x$ for which both $x \equiv 29 \pmod{52}$ and $x \equiv 19 \pmod{72}$.

*Solution* Since $52 = 4 \cdot 13$, we see by Part 3 of Theorem 2.3 that the first congruence is equivalent to the simultaneous congruences $x \equiv 29 \pmod 4$ and $x \equiv 29 \pmod{13}$, which reduces to $x \equiv 1 \pmod 4$ and $x \equiv 3 \pmod{13}$. Similarly, $72 = 8 \cdot 9$, and the second congruence given is equivalent to the simultaneous congruences $x \equiv 19 \pmod 8$ and $x \equiv 19 \pmod 9$. These reduce to $x \equiv 3 \pmod 8$ and $x \equiv 1 \pmod 9$. By the Chinese Remainder Theorem we know that the constraints (mod 13) and (mod 9) are independent of those (mod 8). The given congruences are inconsistent because there is no $x$ for which both $x \equiv 1 \pmod 4$ and $x \equiv 3 \pmod 8$.

Once an inconsistency has been identified, a brief proof can be constructed: The first congruence implies that $x \equiv 1 \pmod 4$ while the second congruence implies that $x \equiv 3 \pmod 4$.

**Example 3** Determine whether the system $x \equiv 3 \pmod{10}$, $x \equiv 8 \pmod{15}$, $x \equiv 5 \pmod{84}$ has a solution, and find them all, if any exist.

*First Solution* We factor each modulus into prime powers. By Part 3 of Theorem 2.3, we see that the first congruence of the system is equivalent to the two simultaneous congruences $x \equiv 3 \pmod 2$, $x \equiv 3 \pmod 5$. Similarly, the second congruence of the system is equivalent to the two conditions $x \equiv 8 \pmod 3$, $x \equiv 8 \pmod 5$, while the third congruence is equivalent to the three congruences $x \equiv 5 \pmod 4$, $x \equiv 5 \pmod 3$, $x \equiv 5 \pmod 7$. The new system of seven simultaneous congruences is equivalent to the ones given, but now all moduli are prime powers. We consider the powers of 2 first. The two conditions are $x \equiv 3 \pmod 2$ and $x \equiv 1 \pmod 4$. These two are consistent, but the second one implies the first, so that the first one may be dropped. The conditions modulo 3 are $x \equiv 8 \pmod 3$ and $x \equiv 5 \pmod 3$. These are equivalent, and may be expressed as $x \equiv 2 \pmod 3$. Third, the conditions modulo 5 are $x \equiv 3 \pmod 5$, $x \equiv 8 \pmod 5$. These are equivalent, so we drop the second of them. Finally, we have the condition $x \equiv 5 \pmod 7$. Hence our system of seven congruences is equivalent to the four conditions $x \equiv 1 \pmod 4$, $x \equiv 2 \pmod 3$, $x \equiv 3 \pmod 5$, and $x \equiv 5 \pmod 7$. Here the moduli are relatively prime in pairs, so we may apply the formula (2.2) used in the proof of the Chinese Remainder Theorem. Proceeding as in the solution of Example 1, we find that $x$ satisfies the given congruences if and only if $x \equiv 173 \pmod{420}$.

The procedure we employed here provides useful insights concerning the way that conditions modulo powers of the same prime must mesh, but when the numbers involved are large, it requires a large amount of computation (because the moduli must be factored). A superior method is provided by the iterative use of Theorem 2.17. This avoids the need to factor the moduli, and requires only $r - 1$ applications of the Euclidean algorithm.



*Second Solution* The $x$ that satisfy the third of the given congruences are precisely those $x$ of the form $5 + 84u$ where $u$ is an integer. On substituting this into the second congruence, we see that the requirement is that $5 + 84u \equiv 8 \pmod{15}$. That is, $84u \equiv 3 \pmod{15}$. By the Euclidean algorithm we find that $(84, 15) = 3$, and indeed we find that $2 \cdot 84 + (-11) \cdot 15 = 3$. By Theorem 2.17 we deduce that $u$ is a solution of the congruence if and only if $u \equiv 2 \pmod 5$. That is, $u$ is of the form $u = 2 + 5v$, and hence $x$ satisfies both the second and the third of the given congruences if and only if $x$ is of the form $5 + 84(2 + 5v) = 173 + 420v$. The first congruence now requires that $173 + 420v \equiv 3 \pmod{10}$. That is, $420v \equiv -170 \pmod{10}$. By the Euclidean algorithm we find that $(420, 10) = 10$. Since $10 \mid 170$, we deduce that this congruence holds for all $v$. That is, in this example, any $x$ that satisfies the second and third of the given congruences also satisfies the first. The set of solutions consists of those $x$ of the form $173 + 420v$. That is, $x \equiv 173 \pmod{420}$.

This procedure can be applied to general systems of the sort (2.1). In case the system is inconsistent, the inconsistency is revealed by a failure of the condition $g \mid b$ in Theorem 2.17. Alternatively, if it happens that the moduli are pairwise relatively prime, then $g = 1$ in each application of Theorem 2.17, and we obtain a second (less symmetric) proof of the Chinese Remainder Theorem.

Returning to Theorem 2.18, we take a fixed set of positive integers $m_1, m_2, \cdots, m_r$, relatively prime in pairs, with product $m$. But instead of considering just one set of equations (2.1), we consider all possible systems of this type. Thus $a_1$ may be any integer in a complete residue system modulo $m_1$, $a_2$ any integer in a complete residue system modulo $m_2$, and so on. To be specific, let us consider $a_1$ to be any integer among $1, 2, \cdots, m_1$, and $a_2$ any integer among $1, 2, \cdots, m_2, \cdots$, and $a_r$ any integer among $1, 2, \cdots, m_r$. The number of such $r$-tuples $(a_1, a_2, \cdots, a_r)$ is $m_1 m_2 \cdots m_r = m$. By the Chinese Remainder Theorem, each $r$-tuple determines precisely one residue class $x$ modulo $m$. Moreover, distinct $r$-tuples determine different residue classes. To see this, suppose that $(a_1, a_2, \cdots, a_r) \neq (a_1', a_2', \cdots, a_r')$. Then $a_i \neq a_i'$ for some $i$, and we see that no integer $x$ satisfies both the congruences $x \equiv a_i \pmod{m_i}$ and $x \equiv a_i' \pmod{m_i}$.

Thus we have a one-to-one correspondence between the $r$-tuples $(a_1, a_2, \cdots, a_r)$ and a complete residue system modulo $m$, such as the integers $1, 2, \cdots, m$. It is perhaps not surprising that two sets, each having

elements, can be put into one-to-one correspondence. However, this correspondence is particularly natural, and we shall draw some important consequences from it.

For any positive integer $n$ let $\mathscr{C}(n)$ denote the complete residue system $\mathscr{C}(n) = \{1, 2, \cdots, n\}$. The $r$-tuples we have considered are precisely the members of the *Cartesian product* (or *direct product*) of the sets $\mathscr{C}(m_1), \mathscr{C}(m_2), \cdots, \mathscr{C}(m_r)$. In symbols, this Cartesian product is denoted $\mathscr{C}(m_1) \times \mathscr{C}(m_2) \times \cdots \times \mathscr{C}(m_r)$. For example, if $\mathbb{R}$ denotes the set of real numbers, then $\mathbb{R} \times \mathbb{R}$, abbreviated $\mathbb{R}^2$, describes the ordinary Euclidean plane with the usual rectangular coordinates belonging to any point $(x, y)$. In this notation, we may express the one-to-one correspondence in question by writing

$$\mathscr{C}(m_1) \times \mathscr{C}(m_2) \times \cdots \times \mathscr{C}(m_r) \leftrightarrow \mathscr{C}(m).$$

Example 4 Exhibit the foregoing one-to-one correspondence explicitly, when $m_1 = 7$, $m_2 = 9$, $m = 63$.

*Solution* Consider the following matrix with 7 rows and 9 columns. At the intersection of the $i$th row and $j$th column we place the element $c_{ij}$, where $c_{ij} \equiv i \pmod 7$ and $c_{ij} \equiv j \pmod 9$. According to Theorem 2.18 we can select the element $c_{ij}$ from the complete residue system $\mathscr{C}(63) = \{1, 2, \cdots, 63\}$. Thus the element 40, for example, is at the intersection of the fifth row and the fourth column, because $40 \equiv 5 \pmod 7$ and $40 \equiv 4 \pmod 9$. Note that the element 41 is at the intersection of the sixth row and fifth column, since $41 \equiv 6 \pmod 7$ and $41 \equiv 5 \pmod 9$. Thus the element $c + 1$ in the matrix is just southeast from the element $c$, allowing for periodicity when $c$ is in the last row or column. For example, 42 is in the last row, so 43 turns up in the first row, one column later. Similarly, 45 is in the last column, so 46 turns up in the first column, one row lower. This gives us an easy way to construct the matrix: just write 1 in the $c_{11}$ position and proceed downward and to the right with 2, 3, and so on.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **1** | **29** | 57 | 22 | 50 | 15 | 43 | **8** | 36 |
| **37** | **2** | 30 | 58 | 23 | 51 | 16 | **44** | 9 |
| **10** | **38** | 3 | 31 | 59 | 24 | 52 | **17** | 45 |
| **46** | **11** | 39 | 4 | 32 | 60 | 25 | **53** | 18 |
| **19** | **47** | 12 | 40 | 5 | 33 | 61 | **26** | 54 |
| **55** | **20** | 48 | 13 | 41 | 6 | 34 | **62** | 27 |
| **28** | **56** | 21 | 49 | 14 | 42 | 7 | **35** | 63 |

Here the correspondence between the pair $(i, j)$ and the entry $c_{ij}$ provides a solution to the problem.



In the matrix, the entry $c_{ij}$ is entered in boldface if $(c_{ij}, 63) = 1$. We note that these entries are precisely those for which $i$ is one of the numbers $\{1, 2, \cdots, 6\}$, and $j$ is one of the numbers $\{1, 2, 4, 5, 7, 8\}$. That is, $(c_{ij}, 63) = 1$ if and only if $(i, 7) = 1$ and $(j, 9) = 1$. Since there are exactly 6 such $i$, and for each such $i$ there are precisely 6 such $j$, we deduce that $\phi(63) = 36 = \phi(7)\phi(9)$. We now show that this holds in general, and we derive a formula for $\phi(m)$ in terms of the prime factorization of $m$.

**Theorem 2.19** *If $m_1$ and $m_2$ denote two positive, relatively prime integers, then $\phi(m_1 m_2) = \phi(m_1)\phi(m_2)$. Moreover, if $m$ has the canonical factorization $m = \prod p^a$, then $\phi(m) = \prod_{p \mid m} (p^a - p^{a-1}) = m \prod_{p \mid m} (1 - 1/p)$.*

If $m = 1$, then the products are empty, and by convention an empty product has value 1. Thus the formula gives $\phi(1) = 1$ in this case, which is correct.

*Proof* Put $m = m_1 m_2$, and suppose that $(x, m) = 1$. By reducing $x$ modulo $m_1$ we see that there is a unique $a_1 \in \mathscr{C}(m_1)$ for which $x \equiv a_1 \pmod{m_1}$. Here, as before, $\mathscr{C}(m_1)$ is the complete system of residues $\mathscr{C}(m_1) = \{1, 2, \cdots, m_1\}$. Similarly, there is a unique $a_2 \in \mathscr{C}(m_2)$ for which $x \equiv a_2 \pmod{m_2}$. Since $(x, m_1) = 1$, it follows by Theorem 2.4 that $(a_1, m_1) = 1$. Similarly $(a_2, m_2) = 1$. For any positive integer $n$, let $\mathscr{R}(n)$ be the system of reduced residues formed of those numbers $a \in \mathscr{C}(n)$ for which $(a, n) = 1$. That is, $\mathscr{R}(n) = \{a \in \mathscr{C}(n): (a, n) = 1\}$. Thus we see that any $x \in \mathscr{R}(m)$ gives rise to a pair $(a_1, a_2)$ with $a_i \in \mathscr{R}(m_i)$ for $i = 1, 2$. Suppose, conversely, that we start with such a pair. By the Chinese Remainder Theorem (Theorem 2.18) there exists a unique $x \in \mathscr{C}(m)$ such that $x \equiv a_i \pmod{m_i}$ for $i = 1, 2$. Since $(a_1, m_1) = 1$ and $x \equiv a_1 \pmod{m_1}$, it follows by Theorem 2.4 that $(x, m_1) = 1$. Similarly we find that $(x, m_2) = 1$, and hence $(x, m) = 1$. That is, $x \in \mathscr{R}(m)$. In this way we see that the Chinese Remainder Theorem enables us to establish a one-to-one correspondence between the reduced residue classes modulo $m$ and pairs of reduced residue classes modulo $m_1$ and $m_2$, provided that $(m_1, m_2) = 1$. Since $a_1 \in \mathscr{R}(m_1)$ can take any one of $\phi(m_1)$ values, and $a_2 \in \mathscr{R}(m_2)$ can take any one of $\phi(m_2)$ values, there are $\phi(m_1)\phi(m_2)$ pairs, so that $\phi(m) = \phi(m_1)\phi(m_2)$.

We have now established the first identity of the theorem. If $m = \prod p^a$ is the canonical factorization of $m$, then by repeated use of this identity we see that $\phi(m) = \prod \phi(p^a)$. To complete the proof it remains to determine the value of $\phi(p^a)$. If $a$ is one of the $p^a$ numbers $1, 2, \cdots, p^a$, then $(a, p^a) = 1$ unless $a$ is one of the $p^{a-1}$ numbers $p, 2p, \cdots, p^{a-1} \cdot p$. On subtracting, we deduce that the number of reduced residue classes modulo $p^a$ is $p^a - p^{a-1} = p^a(1 - 1/p)$. This gives the stated formulae.

We shall derive further properties of Euler's $\phi$-function in Sections 4.2, 4.3, and an additional proof of the formula for $\phi(n)$ will be given in Section 4.5, by means of the inclusion–exclusion principle of combinatorial mathematics.

Let $f(x)$ denote a polynomial with integral coefficients, and let $N(m)$ denote the number of solutions of the congruence $f(x) \equiv 0 \pmod{m}$ as counted in Definition 2.4. We suppose that $m = m_1 m_2$, where $(m_1, m_2) = 1$. By employing the same line of reasoning as in the foregoing proof, we show that the roots of the congruence $f(x) \equiv 0 \pmod{m}$ are in one-to-one correspondence with pairs $(a_1, a_2)$ in which $a_1$ runs over all roots of the congruence $f(x) \equiv 0 \pmod{m_1}$ and $a_2$ runs over all roots of the congruence $f(x) \equiv 0 \pmod{m_2}$. In this way we are able to relate $N(m)$ to $N(m_1)$ and $N(m_2)$.

**Theorem 2.20** *Let $f(x)$ be a fixed polynomial with integral coefficients, and for any positive integer $m$ let $N(m)$ denote the number of solutions of the congruence $f(x) \equiv 0 \pmod{m}$. If $m = m_1 m_2$ where $(m_1, m_2) = 1$, then $N(m) = N(m_1)N(m_2)$. If $m = \prod p^\alpha$ is the canonical factorization of $m$, then $N(m) = \prod N(p^\alpha)$.*

The possibility that one or more of the $N(p^\alpha)$ may be 0 is not excluded in this formula. Indeed, from Theorem 2.16 we see that if $d|m$ and $N(d) = 0$, then $N(m) = 0$. One immediate consequence of this is that the congruence $f(x) \equiv 0 \pmod{m}$ has solutions if and only if it has solutions (mod $p^\alpha$) for each prime-power $p^\alpha$ exactly dividing $m$.

*Proof* Suppose that $x \in \mathscr{C}(m)$, where $\mathscr{C}(m)$ is the complete residue system $\mathscr{C}(m) = \{1, 2, \cdots, m\}$. If $f(x) \equiv 0 \pmod{m}$ and $m = m_1 m_2$, then by Theorem 2.16 it follows that $f(x) \equiv 0 \pmod{m_1}$. Let $a_1$ be the unique member of $\mathscr{C}(m_1) = \{1, 2, \ldots, m_1\}$ for which $x \equiv a_1 \pmod{m_1}$. By Theorem 2.2 it follows that $f(a_1) \equiv 0 \pmod{m_1}$. Similarly, there is a unique $a_2 \in \mathscr{C}(m_2)$ such that $x \equiv a_2 \pmod{m_2}$, and $f(a_2) \equiv 0 \pmod{m_2}$. Thus for each solution of the congruence modulo $m$ we construct a pair $(a_1, a_2)$ in which $a_i$ is a solution of the congruence modulo $m_i$, for $i = 1, 2$. Thus far we have not used the hypothesis that $m_1$ and $m_2$ are relatively prime. It is in the converse direction that this latter hypothesis becomes vital.

Suppose now that $m = m_1 m_2$, where $(m_1, m_2) = 1$, and that for $i = 1$ and 2, numbers $a_i \in \mathscr{C}(m_i)$ are chosen so that $f(a_i) \equiv 0 \pmod{m_i}$. By the Chinese Remainder Theorem (Theorem 2.18), there is a unique $x \in \mathscr{C}(m)$ such that $x \equiv a_i \pmod{m_i}$ for $i = 1, 2$. By Theorem 2.2 we see that this $x$ is a solution of the congruence $f(x) \equiv 0 \pmod{m_i}$, for $i = 1, 2$. Then by Part 3 of Theorem 2.3 we conclude that $f(x) \equiv 0 \pmod{m}$. We have now



established a one-to-one correspondence between the solutions $x$ of the congruence modulo $m$ and pairs $(a_1, a_2)$ of solutions modulo $m_1$ and $m_2$ respectively. Since $a_1$ runs over $N(m_1)$ values, and $a_2$ runs over $N(m_2)$ values, there are $N(m_1)N(m_2)$ such pairs, and we have the first assertion of the theorem. The second assertion follows by repeated application of the first part.

**Example 5** Let $f(x) = x^2 + x + 7$. Find all roots of the congruence $f(x) \equiv 0 \pmod{15}$.

*Solution* Trying the values $x = 0, \pm 1, \pm 2$, we find that $f(x) \equiv 0 \pmod{5}$ has no solution. Since $5 | 15$, it follows that there is no solution (mod 15).

**Example 6** Let $f(x)$ be as in Example 5. Find all roots of $f(x) \equiv 0 \pmod{189}$, given that $189 = 3^3 \cdot 7$, that the roots (mod 27) are 4, 13, and 22, and that the roots (mod 7) are 0 and 6.

*Solution* In a situation of this kind it is more efficient to proceed as we did in the solution of Example 1, rather than employ the method adopted in the second solution of Example 3. By the Euclidean algorithm and (2.2) we find that $x \equiv a_1 \pmod{27}$ and $x \equiv a_2 \pmod{7}$ if and only if $x \equiv 28a_1 - 27a_2 \pmod{189}$. We let $a_1$ take on the three values 4, 13, and 22, while $a_2$ takes on the values 0 and 6. Thus we obtain the six solutions $x \equiv 13, 49, 76, 112, 139, 175 \pmod{189}$.

We have now reduced the problem of locating the roots of a polynomial congruence modulo $m$ to the case in which the modulus is a prime power. In Section 2.6 we reduce this further, to the case of a prime modulus, and finally in Section 2.7 we consider some of the special properties of congruences modulo a prime number $p$.

## PROBLEMS

1. Find the smallest positive integer (except $x = 1$) that satisfies the following congruences simultaneously: $x \equiv 1 \pmod{3}$, $x \equiv 1 \pmod{5}$, $x \equiv 1 \pmod{7}$.
2. Find all integers that satisfy simultaneously: $x \equiv 2 \pmod{3}$, $x \equiv 3 \pmod{5}$, $x \equiv 5 \pmod{2}$.
3. Solve the set of congruences: $x \equiv 1 \pmod{4}$, $x \equiv 0 \pmod{3}$, $x \equiv 5 \pmod{7}$.

40. Prove that for $n \geq 2$ the sum of all positive integers less than $n$ and prime to $n$ is $n\phi(n)/2$.

*41. Define $f(n)$ as the sum of the positive integers less than $n$ and prime to $n$. Prove that $f(m) = f(n)$ implies that $m = n$.

*42. Find all positive integers $n$ such that $\phi(n) | n$.

*43. If $d|n$ and $0 < d < n$, prove that $n - \phi(n) > d - \phi(d)$.

*44. Prove the following generalization of Euler's theorem:

$$a^m \equiv a^{m-\phi(m)} \pmod{m}$$

for any integer $a$.

*45. Find the number of solutions of $x^2 \equiv x \pmod{m}$ for any positive integer $m$.

*46. Let $\psi(n)$ denote the number of integers $a$, $1 \leq a \leq n$, for which both $(a, n) = 1$ and $(a + 1, n) = 1$. Show that $\psi(n) = n \prod_{p|n} (1 - 2/p)$. For what values of $n$ is $\psi(n) = 0$?

*47. Let $f(x)$ be a polynomial with integral coefficients, let $N(m)$ denote the number of solutions of the congruence $f(x) \equiv 0 \pmod{m}$, and let $\phi_f(m)$ denote the number of integers $a$, $1 \leq a \leq m$, such that $(f(a), m) = 1$. Show that if $(m, n) = 1$ then $\phi_f(mn) = \phi_f(m)\phi_f(n)$. Show that if $\alpha > 1$ then $\phi_f(p^\alpha) = p^{\alpha-1}\phi_f(p)$. Show that $\phi_f(p) = p - N(p)$. Conclude that for any positive integer $n$, $\phi_f(n) = n \prod_{p|n} (1 - N(p)/p)$. Show that for an appropriate choice of $f(x)$, this reduces to Theorem 2.19.

## 2.4 TECHNIQUES OF NUMERICAL CALCULATION

When investigating properties of integers, it is often instructive to examine a few examples. The underlying patterns may be more evident if one extends the numerical data by the use of a programmable calculator or electronic computer. For example, after considering a long list of those odd primes $p$ for which the congruence $x^2 \equiv 2 \pmod{p}$ has a solution, one might arrive at the conjecture that it is precisely those primes that are congruent to $\pm 1$ modulo 8. (This is true, and forms an important part of quadratic reciprocity, proved in Section 3.2.) By extending the range of the calculation, one may provide further evidence in favor of a conjecture. Computers are also useful in constructing proofs. For example, one might formulate an argument to show that there is a particular number $n_0$ such that if $n > n_0$, then $n$ is not divisible by all numbers less than $\sqrt{n}$ (recall Problem 50 in Section 1.3). Then by direct calculation one might show that this is also true if $n$ lies in the interval $24 < n \leq n_0$, in order to conclude



that 24 is the largest number divisible by all numbers less than its square root. In this example, it is not hard to show that one may take $n_0 = 210$, and hence one might check the intermediate range by hand, but in other cases of this kind the $n_0$ may be very large, making a computer essential.

We assume that our calculators and computers perform integer arithmetic accurately, as long as the integers involved have at most $d$ digits. We refer to $d$ as the *word length*. This assumption applies not only to addition, subtraction, and multiplication, but also to division, provided that the resulting quotient is also an integer. That is, if $a|b$, the computer will accurately find $b/a$, with no round-off error. We also assume that our computer has a facility for determining the integral part $[x]$ of a real number. Thus in the division algorithm, $b = qa + r$, the computer will accurately find $q = [b/a]$. Use of the fractional part $\{x\} = x - [x]$ should be avoided, since in general the decimal (or binary) expansion of $\{x\}$ will not terminate, with the result that the computer will provide only an approximation to this function. In particular, as we indicated earlier, the remainder in the division algorithm should be calculated as $r = b - a[b/a]$, not as $r = a\{b/a\}$.

We have noted that the Euclidean algorithm does not require many steps. Indeed, when it is applied to very large numbers, the main constraint is the time involved in performing accurate multiple-precision arithmetic. The Euclidean algorithm provides a very efficient means of locating the solutions of linear congruences, and also of finding the root in the Chinese remainder theorem. Since the Euclidean algorithm has so many applications, it is worth spending some effort to optimize it. One way of improving the Euclidean algorithm is to form $q_{i+1}$ by rounding to the nearest integer, rather than rounding down. The resulting $r_i$ is generally smaller, although it may be negative. This modified form of the Euclidean algorithm requires fewer iterations to determine $(b, c)$, but the order of magnitude is still usually $\log c$ when $b > c$. Example 3 of Section 1.2 required 24 iterations, but with the modified algorithm only 15 would be needed. (*Warning:* The integral part function conveniently provided on most machines rounds toward 0. That is, when asked for the integer part of a decimal (or binary) number $\pm a_k a_{k-1} \cdots a_0.b_1 b_2 \cdots b_r$, the machine will return $\pm a_k a_{k-1} \cdots a_0$. This is $[x]$ when $x$ is non-negative, but it is $-[-x]$ when $x$ is negative. For example, $[-3.14159] = -4$, but the machine will round toward 0, giving an answer $-3$. To avoid this trap, ensure that a number is non-negative before asking a machine to give you the integer part. Alternatively, one could employ a conditional instruction, "Put $y = \text{int}(x)$. If $y > x$, then replace $y$ by $y - 1$." This has the effect of setting $y = [x]$.)

In performing congruence arithmetic, we observe that if $0 \leq a < m$ and $0 \leq b < m$ then either $a + b$ is already reduced or else $m \leq a + b <$

...in which case $a + b - m$ is reduced. To calculate $ab \pmod m$, we may set $c = ab$, and then reduce $c \pmod m$. However, $c$ may be as large as $(m-1)^2$, which means that if we are limited to integers $< 10^d$ then we can calculate $ab \pmod m$ in this way only for $m < 10^{d/2}$, that is, half the word length. The sensible solution to this problem is to employ multiple-precision arithmetic, but in the short term one may instead use an algorithm such as that described in Problem 21 at the end of this section.

Another situation in which we may introduce a modest saving is in the evaluation of a polynomial $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$. The naive approach would involve constructing the sequence of powers $x^k$, and as we do so, forming the partial sums $a_0, a_0 + a_1 x, \cdots$, until one arrives at $f(x)$. This requires $n$ additions and $2n - 1$ multiplications. A more efficient process is suggested by observing that

$$f(x) = (\cdots((a_n x + a_{n-1})x + a_{n-2})x + \cdots)x + a_0.$$

Here we still have $n$ additions, but now only $n$ multiplications. This procedure is known as *Horner's method*.

A much greater saving can be introduced when computing a power $a^k$, when $k$ is large. The naive approach would involve $k - 1$ multiplications. This is fine if $k$ is small, but for large $k$ one should repeatedly square to form the sequence of numbers $d_j = a^{2^j}$. Writing the binary expansion of $k$ in the form $k = \sum_{j \in \mathcal{J}} 2^j$, we see that $a^k = \prod_{j \in \mathcal{J}} d_j$. Here the number of multiplications required is of the order of magnitude $\log k$, a great savings if $k$ is large. This procedure can be made still more efficient if the machine in use automatically converts numbers to binary, for then the binary digits of $k$ can be accessed, rather than computed. It might seem at first that this device is of limited utility. After all, if $a^k$ is encountered in the context of real arithmetic, one would simply compute $\exp(k \log a)$. Even if $a$ and $k$ are integers, one is unlikely to examine $a^k$ when $k$ is large, unless one is willing to perform multiple-precision arithmetic. However, this device is extremely useful when computing $a^k \pmod m$.

**Example 7** Determine the value of $999^{179} \pmod{1763}$.

*Solution* We find that $179 = 1 + 2 + 2^4 + 2^5 + 2^7$, that $999^2 \equiv 143 \pmod{1763}$, $999^4 \equiv 143^2 \equiv 1056 \pmod{1763}$, $999^8 \equiv 1056^2 \equiv 920 \pmod{1763}$, $999^{16} \equiv 920^2 \equiv 160 \pmod{1763}$, $999^{32} \equiv 160^2 \equiv 918 \pmod{1763}$, $999^{64} \equiv 918^2 \equiv 10 \pmod{1763}$, so that $999^{128} \equiv 10^2 \equiv 100 \pmod{1763}$. Hence $999^{179} \equiv 999 \cdot 143 \cdot 160 \cdot 918 \cdot 100 \equiv 54 \cdot 160 \cdot 918 \cdot 100 \equiv 1588 \cdot 918 \cdot 100 \equiv 1546 \cdot 100 \equiv 1219 \pmod{1763}$.



When implemented, it would be a mistake to first list the binary digits of $k$, then form a list of the numbers $d_j$, and finally multiply the appropriate $d_j$ together, as we have done above. Instead, one should perform these three tasks concurrently, as follows:

1. Set $x = 1$. (Here $x$ is the product being formed.)
2. While $k > 0$, repeat the following steps:
   (*a*) Set $e = k - 2\lfloor k/2 \rfloor$. (Thus $e = 0$ or 1, according as $k$ is even or odd.)
   (*b*) If $e = 1$ then replace $x$ by $ax$, and reduce this (mod $m$). (If $e = 0$ then $x$ is not altered.)
   (*c*) Replace $a$ by $a^2$, and reduce this (mod $m$).
   (*d*) Replace $k$ by $(k - e)/2$. (i.e., drop the unit digit in the binary expansion, and shift the remaining digits one place to the right.)

When this is completed, we see that $x \equiv a^k \pmod m$.

Our ability to evaluate $a^k \pmod m$ quickly can be applied to provide an easy means of establishing that a given number is composite.

**Example 8** Show that 1763 is composite.

*Solution* By Fermat's congruence, if $p$ is an odd prime number then $2^{p-1} \equiv 1 \pmod p$. In other words, if $n$ is an odd number for which $2^{n-1} \not\equiv 1 \pmod n$, then $n$ is composite. We calculate that $2^{1762} \equiv 742 \pmod{1763}$, and deduce that 1763 is composite. Alternatively, we might search for a divisor of 1763, but the use we have made here of Fermat's congruence provides a quicker means of establishing compositeness when $n$ is large, provided, of course that the test succeeds. Since the empirical evidence is that the test detects most composite numbers, if $2^{n-1} \equiv 1 \pmod n$ then we call $n$ a *probable prime to the base* 2. A composite probable prime is called a *pseudoprime*. That such numbers exist is seen in the following example.

**Example 9** Show that 1387 is composite.

*Solution* We may calculate that $2^{1386} \equiv 1 \pmod{1387}$. Thus 1387 is a probable prime to the base 2. To demonstrate that it is composite, we may try a different base, but a more efficient procedure is provided by applying Lemma 2.10. We have a number $x = 2^{693}$ with the property that $x^2 \equiv 1 \pmod{1387}$. Since $2^{693} \equiv 512 \not\equiv \pm 1 \pmod{1387}$, we conclude that 1387 is composite.

When used systematically, this technique yields the *strong pseudoprime test*. If we wish to show that an odd number $m$ is composite, we divide $m-1$ by 2 repeatedly, in order to write $m-1=2^j d$, with $d$ odd. We form $a^d \pmod m$, and by repeatedly squaring and reducing, we construct the numbers

$$a^d, a^{2d}, a^{4d}, \cdots, a^{2^j d} \pmod m.$$

If the last number here is $\not\equiv 1 \pmod m$, then $m$ is composite. If this last member is $\equiv 1 \pmod m$, then $m$ is a probable prime to the base $a$, but if the entry immediately preceding the first 1 is $\not\equiv -1 \pmod m$, then we may still conclude (by Lemma 2.10) that $m$ is composite. When this test is inconclusive, we call $m$ a *strong probable prime*. An odd, composite, strong probable prime is called a *strong pseudoprime to the base a*, abbreviated spsp$(a)$. Such numbers exist, but numerical evidence suggests that they are much rarer than pseudoprimes. In our remarks following Problem 54 in Section 2.1, we noted the existence of numbers $m$, called Carmichael numbers, which are pseudoprime to every base $a$ that is relatively prime to $m$. Such a phenomenon does not persist with strong pseudoprimes, as it can be shown that if $m$ is odd and composite then $m$ is a spsp$(a)$ for at most $m/4$ values of $a \pmod m$. For most $m$, the number of such $a$ is much smaller. Expressed as an algorithm, the strong pseudoprime test for $m$ takes the following shape:

1. Find $j$ and $d$ with $d$ odd, so that $m-1=2^j d$.
2. Compute $a^d \pmod m$. If $a^d \equiv \pm 1 \pmod m$, then $m$ is a strong probable prime; stop.
3. Square $a^d$ to compute $a^{2d} \pmod m$. If $a^{2d} \equiv 1 \pmod m$, then $m$ is composite; stop. If $a^{2d} \equiv -1$, then $m$ is a strong probable prime; stop.
4. Repeat step 3 with $a^{2d}$ replaced by $a^{4d}, a^{8d}, \cdots, a^{2^{j-1}d}$.
5. If the procedure has not already terminated, then $m$ is composite.

Let $X = 25 \cdot 10^9$. Integers in the interval $[1, X]$ have been examined in detail, and it has been found that the number of prime numbers in this interval is $\pi(X) = 1{,}091{,}987{,}405$, that the number of odd pseudoprimes in this interval is 21,853, and that the number of Carmichael numbers in this interval is 2163. On the other hand, in this interval there are 4842 numbers of the class spsp(2), 184 that are both spsp(2) and spsp(3), 13 that are spsp$(a)$ for $a = 2, 3, 5$, only 1 that is spsp$(a)$ for $a = 2, 3, 5, 7$, and none that is spsp$(a)$ for $a = 2, 3, 5, 7, 11$.

The strong pseudoprime test provides a very efficient means for proving that an odd integer $m$ is composite. With further information one



can sometimes use it to demonstrate that a number is prime. If $m$ is a strong probable prime base 2, and if $m < 2047$, then $m$ is prime. If $m$ is 2047 is the least spsp(2). If $m$ is larger, apply the test to the base 3. If $m$ is again found to be a strong probable prime, then $m$ is prime provided that $m < 1{,}373{,}653$. This latter number is the least integer that is both spsp(2) and spsp(3). If $m$ is larger, then apply the test to the base 5. If $m$ is again found to be a strong probable prime, then $m$ is prime provided that $m < 25{,}326{,}001$. This is the least number that is simultaneously spsp for $a = 2$, 3, and 5. If $m$ is still larger, then apply the test to the base 7. If $m$ is once more found to be a probable prime, then $m$ is prime provided that $m < X = 25 \cdot 10^9$ and that $m \neq 3{,}215{,}031{,}751$. This last number is the only number $< X$ that is spsp$(a)$ for $a = 2$, 3, 5, and 7. It is not known in general how many applications of the strong test suffice to ensure that a number $m$ is prime. but it is conjectured that if $m$ is a strong probable prime for all bases $a$ in the range $1 < a \leqslant 2(\log m)^2$ then $m$ is prime.

Suppose that $m$ is a large composite number. By the strong pseudoprime test we may establish that $m$ is composite without exhibiting a proper divisor of $m$. In general, finding the factorization of $m$ involves much more calculation. If $p$ denotes the least prime factor of $m$, then we locate the proper divisor $p$ after $p$ trial divisions. Since $p$ may be nearly as large as $\sqrt{m}$, this may require up to $\sqrt{m}$ operations. We now describe a method which usually locates the smallest prime factor $p$ in just a little more than $\sqrt{p}$ steps. As in many such factoring algorithms, our estimate for the running time is not proved, but is instead based on heuristic probabilistic models, and experience. For our present purposes, the relevant probabilistic result is expressed in the following lemma.

**Lemma 2.21** *Suppose that $1 \leqslant k \leqslant n$, and that the numbers $u_1, u_2, \cdots, u_k$ are independently chosen from the set $\{1, 2, \cdots, n\}$. Then the probability that the numbers $u_k$ are distinct is*

$$\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right).$$

*Proof* Consider a sequence $u_1, \cdots, u_k$ in which each $u_i$ is one of the numbers $1, 2, \cdots, n$. Since each $u_i$ is one of $n$ numbers, there are $n^k$ such sequences. From among these, we count those for which the $u_i$ are distinct. We see that $u_1$ can be any one of $n$ numbers. If $u_2$ is to be distinct from $u_1$, then $u_2$ is one of $n-1$ numbers. If $u_3$ is to be distinct from both $u_1$ and $u_2$, then $u_3$ is one of $n-2$ numbers, and so on. Hence the total number of such sequences is $n(n-1)\cdots(n-k+1)$. We divide this by $n^k$ to obtain the stated probability.

As an application, we note that if $n = 365$ and $k = 23$, then the probability in question is less than $1/2$. That is, if 23 people are chosen at random, then the probability of two of them having the same birthday is greater than $1/2$. It may seem counterintuitive that such a small number of people suffices, but it can be shown that the product is approximately $\exp(-k^2/(2n))$. (A derivation of a precise estimate of this sort is outlined in Problem 22 at the end of this section.) Hence the $u_i$ are likely to be distinct if $k$ is small compared with $\sqrt{n}$, but unlikely to be distinct if $k$ is large compared with $\sqrt{n}$.

Suppose that $m$ is a large composite number whose smallest prime divisor is $p$. If we choose $k$ integers $u_1, u_2, \cdots, u_k$ "at random," with $k$ large compared to $\sqrt{p}$ but small compared to $\sqrt{m}$, then it is likely that the $u_i$ will be distinct (mod $m$), but not distinct (mod $p$). That is, there probably are integers $i, j$, with $1 \leqslant i < j \leqslant k$ such that $1 < (u_i - u_j, m) < m$. Each pair $(i, j)$ is easily tested by the Euclidean algorithm, but the task of inspecting all $\binom{k}{2}$ pairs is painfully long. To shorten our work, we adopt the following scheme: We generate the $u_i$ by a recursion of the form $u_{i+1} = f(u_i)$ where $f(u)$ is a polynomial with integral coefficients. The precise choice of $f(u)$ is unimportant, except that it should be easy to compute, and it should give rise to a sequence of numbers that "looks random." Here some experimentation is called for, but it has been found that $f(u) = u^2 + 1$ works well. (In general, polynomials of first degree do not.)

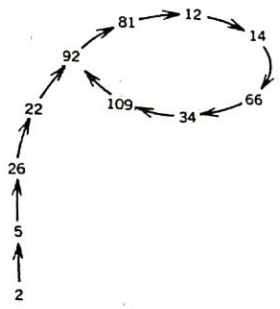
The advantage of generating the $u_i$ in this way is that if $u_i \equiv u_j \pmod{d}$, then $u_{i+1} = f(u_i) \equiv f(u_j) = u_{j+1} \pmod{d}$, so the sequence $u_i$ becomes periodic (mod $d$) with period $j - i$. In other words, if we put $r = j - i$, then $u_s \equiv u_t \pmod{d}$ whenever $s \equiv t \pmod{r}$, $s \geqslant i$, and $t \geqslant i$. In particular, if we let $s$ be the least multiple of $r$ that is $\geqslant i$, and we take $t = 2s$, then $u_s \equiv u_{2s} \pmod{d}$. That is, among the numbers $u_{2s} - u_s$ we expect to find one for which $1 < (u_{2s} - u_s, m) < m$, with $s$ of size roughly comparable to $\sqrt{p}$.

**Example 10** Use this method to locate a proper divisor of the number $m = 36{,}287$.

*Solution* We take $u_0 = 1$, $u_{i+1} \equiv u_i^2 + 1 \pmod{m}$, $0 \leqslant u_{i+1} < m$. Then the numbers $u_i$, $i = 1, 2, \ldots, 14$ are 2, 5, 26, 677, 22886, 2439, 33941, 24380, 3341, 22173, 25652, 26685, 29425, 22806. We find that $(u_{2s} - u_s, m) = 1$ for $s = 1, 2, \cdots, 6$, but that $(u_{14} - u_7, m) = 131$. That is, 131 is a divisor of $m$. In this example, it turns out that 131 is the smallest prime divisor of $m$, because the division of 36,287 by 131 gives the other prime factor, 277.



If we reduce the $u_i \pmod{131}$, we obtain the numbers 2, 5, 26, 22, 92, 81, 12, 14, 66, 34, 109, 92, 81, 12. Hence $u_{12} \equiv u_5 \pmod{131}$, and the sequence has period 7 from $u_5$ on. We might diagram this as follows:




This method was proposed by J. M. Pollard in 1975. Since the pattern above resembles the Greek letter $\rho$ ("rho"), this approach is known as the *Pollard rho method*. It should be applied only to numbers $m$ that are already known to be composite (e.g., by the strong pseudoprime test), for if $m$ is prime then the method will run for roughly $\sqrt{m}$ cycles, without proving anything. Since the method may be expected to disclose the smallest prime factor $p$ of $m$ in roughly $\sqrt{p}$ cycles, this method is faster than trial division for large composite $m$. Note that there is no guarantee that the divisor found will be the smallest prime factor of $m$. The divisor located may be some other prime factor, it may be composite, and it may even be $m$ itself. In the latter eventuality, one may start over with a new value of $u_0$, or with a new function $f(u)$, say $f(u) = u^2 + c$ with some new value for $c$. (The two values $c = 0$, $c = -2$ should be avoided.)

As of this writing, the most efficient factoring strategies are expected to locate a proper divisor of a composite number $m$ in no more than $\exp(c(\log m)^{1/2}(\log\log m)^{1/2})$ bit operations. (Here $c$ is some positive constant.) In Section 5.8 we use elliptic curves to find proper divisors this quickly. If $\varepsilon$ is a given positive number, then the function of $m$ above is $< m^\varepsilon$ for all sufficiently large $m$. Nevertheless, it remains the case that we can perform congruence arithmetic, compositeness tests, and so forth for much larger $m$ than we can factor.

(c) Replace $c$ by $c + s$.
(d) If $c \geq m$, replace $c$ by $c - m$.
(e) Replace $k$ by $gk - m[gk/m]$.
(f) Replace $a$ by $(a - r)/g$.

•22. Show that the product in Lemma 2.21 is smaller than $\exp\left(-\frac{k^2}{2n} + \frac{k}{2n}\right)$, but larger than $\exp\left(-\frac{k^2}{2n} - \frac{k^3}{3n^2}\right)$. (H)

## 2.5 PUBLIC-KEY CRYPTOGRAPHY

We now apply our knowledge of congruence arithmetic to construct a method of encrypting messages. The mathematical principle we use is formulated in the following lemma.

**Lemma 2.22** *Suppose that $m$ is a positive integer and that $(a, m) = 1$. If $k$ and $\bar{k}$ are positive integers such that $k\bar{k} \equiv 1 \pmod{\phi(m)}$, then $a^{k\bar{k}} \equiv a \pmod{m}$.*

*Proof* Write $k\bar{k} = 1 + r\phi(m)$, where $r$ is a non-negative integer. Then by Euler's congruence

$$a^{k\bar{k}} = a \cdot a^{r\phi(m)} = a\left(a^{\phi(m)}\right)^r \equiv a \cdot 1^r = a \pmod{m}.$$

If $(a, m) = 1$ and $k$ is a positive integer, then $(a^k, m) = 1$. Thus if $n = \phi(m)$ and $r_1, r_2, \cdots, r_n$ is a system of reduced residues (mod $m$), then the numbers $r_1^k, r_2^k, \cdots, r_n^k$ are also relatively prime to $m$. These $k$th powers may not all be distinct (mod $m$), as we see by considering the special case $k = \phi(m)$. On the other hand, from Lemma 2.22 we can deduce that these $k$th powers are distinct (mod $m$) provided that $(k, \phi(m)) = 1$. For, suppose that $r_i^k \equiv r_j^k \pmod{m}$ and $(k, \phi(m)) = 1$. By Theorem 2.9 we may determine a positive integer $\bar{k}$ such that $k\bar{k} \equiv 1 \pmod{\phi(m)}$, and then it follow from the lemma that

$$r_i \equiv r_i^{k\bar{k}} = \left(r_i^k\right)^{\bar{k}} \equiv \left(r_j^k\right)^{\bar{k}} = r_j^{k\bar{k}} \equiv r_j \pmod{m}.$$

This implies that $i = j$. (From our further analysis in Section 2.8 it will become apparent that the converse also holds: the numbers $r_1^k, r_2^k, \cdots, r_n^k$ are distinct (mod $m$) only if $(k, \phi(m)) = 1$.) Suppose that $(k, \phi(m)) = 1$. Since the numbers $r_1^k, r_2^k, \cdots, r_n^k$ are distinct (mod $m$), they form a system of reduced residues (mod $m$). That is, the map $a \mapsto a^k$ permutes the reduced residues (mod $m$) if $(k, \phi(m)) = 1$. The significance of the lemma is that the further map $b \mapsto b^{\bar{k}}$ is the inverse permutation.



To apply these observations to cryptography, we take two distinct large primes, $p_1, p_2$, say each one with about 100 digits, and multiply them to form a composite modulus $m = p_1 p_2$ of about 200 digits. Since we know the prime factorization of $m$, from Theorem 2.19 we see that $\phi(m) = (p_1 - 1)(p_2 - 1)$. Here $\phi(m)$ is somewhat smaller than $m$. We choose a big number, $k$, from the interval $0 < k < \phi(m)$, and check by the Euclidean algorithm that $(k, \phi(m)) = 1$. If a proposed $k$ does not have this property, we try another, until we obtain one for which this holds. We make the numbers $m$ and $k$ publicly available, but keep $p_1$, $p_2$, and $\phi(m)$ secret. Suppose now that some associate of ours wants to send us a message, say "Gauss was a genius!" The associate first converts the characters of the message to numbers in some standard way, say by employing the three digit American Standard Code for Information Interchange (ASCII) used on many computers. Then "G" becomes 071, "a" becomes 097, $\cdots$, and "!" becomes 033. Concatenate these codes to form a number

$$a = 071097117115115126119097115126097126103101110105117115033.$$

Since $a$ has only 56 digits, we see that $0 < a < m$. If the message were longer, it could be divided into a number of blocks. Our associate could send us the number $a$, and then we could reconstruct the original characters, but suppose that the message contains some sensitive material that would make it desirable to ensure the privacy of the transmission. In that case, our associate would use the numbers $k$ and $m$ that we have provided. Being acquainted with the ideas discussed in the preceding section, our associate quickly finds the unique number $b$, $0 \leq b < m$, such that $b \equiv a^k \pmod{m}$, and sends this $b$ to us. We use the Euclidean algorithm to find a positive number $\bar{k}$ such that $k\bar{k} \equiv 1 \pmod{\phi(m)}$, and then we find the unique number $c$ such that $0 \leq c < m$, $c \equiv b^{\bar{k}} \pmod{m}$. From Lemma 2.22 we deduce that $a = c$. In theory it might happen that $(a, m) > 1$, in which case the lemma does not apply, but the chances of this are remote ($\approx 1/p_i \approx 10^{-100}$). (In this unlikely event, one could still appeal to Problem 4 at the end of this section.) Suppose that some inquisitive third party gains access to the numbers $m$, $k$, and $b$, and seeks to recover the number $a$. In principle, all that need be done is to factor $m$, which yields $\phi(m)$, and hence $\bar{k}$, just as we have done. In practice, however, the task of locating the factors of $m$ is prohibitively long. Using the best algorithms known and fastest computers, it would take centuries to factor our 200 digit modulus $m$. Of course, we hope that faster factoring algorithms may yet be discovered, but here one can only speculate.

**PROBLEMS**

1. Suppose that $b \equiv a^{67} \pmod{91}$, and that $(a, 91) = 1$. Find a positive number $\bar{k}$ such that $b^{\bar{k}} \equiv a \pmod{91}$. If $b = 53$, what is $a \pmod{91}$?
2. Suppose that $m = pq$, and $\phi = (p-1)(q-1)$ where $p$ and $q$ are real numbers. Find a formula for $p$ and $q$, in terms of $m$ and $\phi$. Supposing that $m = 39{,}247{,}771$ is the product of two distinct primes, deduce the factors of $m$ from the information that $\phi(m) = 39{,}233{,}944$.
3. Show that if $d|m$, then $\phi(d)|\phi(m)$.
4. Suppose that $m$ is square-free, and that $k$ and $\bar{k}$ are positive integers such that $k\bar{k} \equiv 1 \pmod{\phi(m)}$. Show that $a^{k\bar{k}} \equiv a \pmod{m}$ for all integers $a$. (H)
5. Suppose that $m$ is a positive integer that is not square-free. Show that there exist integers $a_1$ and $a_2$ such that $a_1 \not\equiv a_2 \pmod{m}$, but $a_1^k \equiv a_2^k \pmod{m}$ for all integers $k > 1$.

## 2.6 PRIME POWER MODULI

The problem of solving a congruence was reduced in Section 2.3 to the case of a prime-power modulus. To solve a polynomial congruence $f(x) \equiv 0 \pmod{p^k}$, we start with a solution modulo $p$, then move on to modulo $p^2$, then to $p^3$, and by iteration to $p^k$. Suppose that $x = a$ is a solution of $f(x) \equiv 0 \pmod{p^j}$ and we want to use it to get a solution modulo $p^{j+1}$. The idea is to try to get a solution $x = a + tp^j$, where $t$ is to be determined, by use of Taylor's expansion

$$f(a + tp^j) = f(a) + tp^j f'(a) + t^2 p^{2j} f''(a)/2! + \cdots + t^n p^{nj} f^{(n)}(a)/n! \tag{2.3}$$

where $n$ is the presumed degree of the polynomial $f(x)$. All derivatives beyond the $n$th are identically zero.

Now with respect to the modulus $p^{j+1}$, equation (2.3) gives

$$f(a + tp^j) \equiv f(a) + tp^j f'(a) \pmod{p^{j+1}} \tag{2.4}$$

as the following argument shows. What we want to establish is that the coefficients of $t^2, t^3, \cdots t^n$ in equation (2.3) are divisible by $p^{j+1}$ and so can be omitted in (2.4). This is almost obvious because the powers of $p$ in those terms are $p^{2j}, p^{3j}, \cdots, p^{nj}$. But this is not quite immediate because of the denominators $2!, 3!, \cdots, n!$ in these terms. The explanation is that



$f^{(k)}(a)/k!$ is an integer for each value of $k$, $2 \leq k \leq n$. To see this, let $cx^r$ be a representative term from $f(x)$. The corresponding term in $f^{(k)}(a)$ is

$$cr(r-1)(r-2)\cdots(r-k+1)a^{r-k}$$

According to Theorem 1.21, the product of $k$ consecutive integers is divisible by $k!$, and the argument is complete. Thus, we have proved that the coefficients of $t^2, t^3, \cdots$ in (2.3) are divisible by $p^{j+1}$.

The congruence (2.4) reveals how $t$ should be chosen if $x = a + tp^j$ is to be a solution of $f(x) \equiv 0 \pmod{p^{j+1}}$. We want $t$ to be a solution of

$$f(a) + tp^j f'(a) \equiv 0 \pmod{p^{j+1}}.$$

Since $f(x) \equiv 0 \pmod{p^j}$ is presumed to have the solution $x = a$, we see that $p^j$ can be removed as a factor to give

$$tf'(a) \equiv -\frac{f(a)}{p^j} \pmod{p} \tag{2.5}$$

which is a linear congruence in $t$. This congruence may have no solution, one solution, or $p$ solutions. If $f'(a) \not\equiv 0 \pmod{p}$, then this congruence has exactly one solution, and we obtain

**Theorem 2.23** *Hensel's lemma. Suppose that $f(x)$ is a polynomial with integral coefficients. If $f(a) \equiv 0 \pmod{p^j}$ and $f'(a) \not\equiv 0 \pmod{p}$, then there is a unique $t \pmod{p}$ such that $f(a + tp^j) \equiv 0 \pmod{p^{j+1}}$.*

If $f(a) \equiv 0 \pmod{p^j}$, $f(b) \equiv 0 \pmod{p^k}$, $j < k$, and $a \equiv b \pmod{p^j}$, then we say that $b$ *lies above* $a$, or $a$ *lifts to* $b$. If $f(a) \equiv 0 \pmod{p^j}$, then the root $a$ is called *nonsingular* if $f'(a) \not\equiv 0 \pmod{p}$; otherwise it is *singular*. By Hensel's lemma we see that a nonsingular root $a \pmod{p}$ lifts to a unique root $a_2 \pmod{p^2}$. Since $a_2 \equiv a \pmod{p}$, it follows (by Theorem 2.2) that $f'(a_2) \equiv f'(a) \not\equiv 0 \pmod{p}$. By a second application of Hensel's lemma we may lift $a_2$ to form a root $a_3$ of $f(x)$ modulo $p^3$, and so on. In general we find that a nonsingular root $a$ modulo $p$ lifts to a unique root $a_j$ modulo $p^j$ for $j = 2, 3, \cdots$. By (2.5) we see that this sequence is generated by means of the recursion

$$a_{j+1} = a_j - f(a_j)\overline{f'(a)} \tag{2.6}$$

where $\overline{f'(a)}$ is an integer chosen so that $f'(a)\overline{f'(a)} \equiv 1 \pmod{p}$. This is

entirely analogous to Newton's method for locating the root of a differentiable function.

**Example 11** Solve $x^2 + x + 47 \equiv 0 \pmod{7^3}$.

*Solution* First we note that $x \equiv 1 \pmod 7$ and $x \equiv 5 \pmod 7$ are the only solutions of $x^2 + x + 47 \equiv 0 \pmod 7$. Since $f'(x) = 2x + 1$, we see that $f'(1) = 3 \not\equiv 0 \pmod 7$ and $f'(5) = 1 \not\equiv 0 \pmod 7$, so these roots are nonsingular. Taking $\overline{f'(1)} = 5$, we see by (2.6) that the root $a \equiv 1 \pmod 7$ lifts to $a_2 = 1 - 49 \cdot 5$. Since $a_2$ is considered $(\bmod 7^2)$, we may take instead $a_2 = 1$. Then $a_3 = 1 - 49 \cdot 5 \equiv 99 \pmod{7^3}$. Similarly, we take $\overline{f'(5)} = 2$, and see by (2.6) that the root $5 \pmod 7$ lifts to $5 - 77 \cdot 2 = -149 \equiv 47 \pmod{7^2}$, and that $47 \pmod{7^2}$ lifts to $47 - f(47) \cdot 2 = 47 - 2303 \cdot 2 = -4559 \equiv 243 \pmod{7^3}$. Thus we conclude that 99 and 243 are the desired roots and that there are no others.

We now turn to the more difficult problem of lifting singular roots. Suppose that $f(a) \equiv 0 \pmod{p^j}$ and that $f'(a) \equiv 0 \pmod p$. From the Taylor expansion (2.3) we see that $f(a + tp^j) \equiv f(a) \pmod{p^{j+1}}$ for all integers $t$. Thus if $f(a) \equiv 0 \pmod{p^{j+1}}$ then $f(a + tp^j) \equiv 0 \pmod{p^{j+1}}$, so that the single root $a \pmod{p^j}$ lifts to $p$ roots $\pmod{p^{j+1}}$. But if $f(a) \not\equiv 0 \pmod{p^{j+1}}$, then none of the $p$ residue classes $a + tp^j$ is a solution $(\bmod p^{j+1})$, and then there are no roots $(\bmod p^{j+1})$ lying above $a (\bmod p^j)$.

**Example 12** Solve $x^2 + x + 7 \equiv 0 \pmod{81}$.

*Solution* Starting with $x^2 + x + 7 \equiv 0 \pmod 3$, we note that $x = 1$ is the only solution. Here $f'(1) = 3 \equiv 0 \pmod 3$, and $f(1) \equiv 0 \pmod 9$, so that we have the roots $x = 1$, $x = 4$, and $x = 7 \pmod 9$. Now $f(1) \not\equiv 0 \pmod{27}$, and hence there is no root $x \pmod{27}$ for which $x \equiv 1 \pmod 9$. As $f(4) \equiv 0 \pmod{27}$, we obtain three roots, 4, 13, and 22 $(\bmod 27)$, which are $\equiv 4 \pmod 9$. On the other hand, $f(7) \not\equiv 0 \pmod{27}$, so there is no root $(\bmod 27)$ that is $\equiv 7 \pmod 9$. We are now in a position to determine which, if any, of the roots 4, 13, 22 $(\bmod 27)$ can be lifted to roots $(\bmod 81)$. We find that $f(4) = 27 \not\equiv 0 \pmod{81}$, $f(13) = 189 \equiv 27 \not\equiv 0 \pmod{81}$, and that $f(22) = 513 \equiv 27 \not\equiv 0 \pmod{81}$, from which we deduce that the congruence has no solution $(\bmod 81)$.

In this example, we see that a singular solution $a \pmod p$ may lift to some higher powers of $p$, but not necessarily to arbitrarily high powers of $p$. We now show that if the power of $p$ dividing $f(a)$ is sufficiently large compared with the power of $p$ in $f'(a)$, then the solution can be lifted without limit.



**Theorem 2.24** *Let $f(x)$ be a polynomial with integral coefficients. Suppose that $f(a) \equiv 0 \pmod{p^j}$, that $p^\tau \| f'(a)$, and that $j \geqslant 2\tau + 1$. If $b \equiv a \pmod{p^{j-\tau}}$ then $f(b) \equiv f(a) \pmod{p^j}$ and $p^\tau \| f'(b)$. Moreover, there is a unique $t \pmod p$ such that $f(a + tp^{j-\tau}) \equiv 0 \pmod{p^{j+1}}$.*

In this situation, a collection of $p^\tau$ solutions $(\bmod p^j)$ give rise to $p^\tau$ solutions $(\bmod p^{j+1})$, while the power of $p$ dividing $f'$ remains constant. Since the hypotheses of the theorem apply with $a$ replaced by $a + tp^{j-\tau}$ and $(\bmod p^j)$ replaced by $(\bmod p^{j+1})$ but with $\tau$ unchanged, the lifting may be repeated and continues indefinitely.

*Proof* By Taylor's expansion (2.3), we see that

$$f(b) = f(a + tp^{j-\tau}) \equiv f(a) + tp^{j-\tau} f'(a) \pmod{p^{2j-2\tau}}.$$

Here the modulus is divisible by $p^{j+1}$, since $2j - 2\tau = j + (j - 2\tau) \geqslant j + 1$. Hence

$$f(a + tp^{j-\tau}) \equiv f(a) + tp^{j-\tau} f'(a) \pmod{p^{j+1}}.$$

Since both terms on the right side are divisible by $p^j$, the left side is also. Moreover, on dividing through by $p^j$ we find that

$$\frac{f(a + tp^{j-\tau})}{p^j} \equiv \frac{f(a)}{p^j} + t\frac{f'(a)}{p^\tau} \pmod p,$$

and the coefficient of $t$ is relatively prime to $p$, so that there is a unique $t \pmod p$ for which the right side is divisible by $p$. This establishes the final assertion of the theorem. To complete the proof, we note that $f'(x)$ is a polynomial with integral coefficients, so that

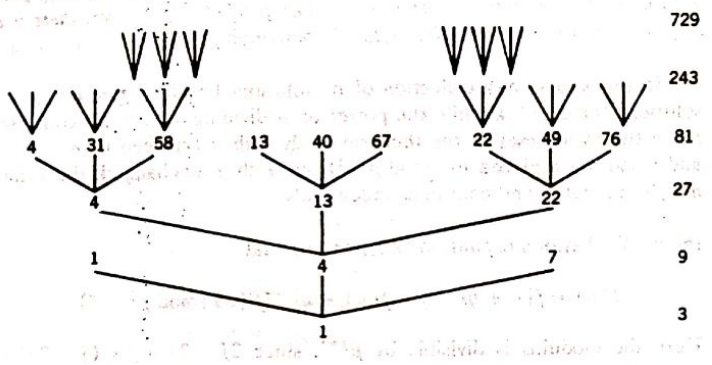
$$f'(a + tp^{j-\tau}) \equiv f'(a) \pmod{p^{j-\tau}}$$

for any integer $t$. But $j - \tau \geqslant \tau + 1$, so this congruence holds $(\bmod p^{\tau+1})$. Since $p^\tau$ exactly divides $f'(a)$ (in symbols, $p^\tau \| f'(a)$), we conclude that $p^\tau \| f'(a + tp^{j-\tau})$.

**Example 13** Discuss the solutions of $x^2 + x + 223 \equiv 0 \pmod{3^j}$.

*Solution* Since $223 \equiv 7 \pmod{27}$, the solutions $(\bmod 27)$ are the same as in Example 12. For this new polynomial, we find that $f(4) \equiv 0 \pmod{81}$, and thus we have three solutions $4, 31, 58 \pmod{81}$. Similarly $f(13) \equiv 0 \pmod{81}$, giving three solutions $13, 40, 67 \pmod{81}$. Moreover, $f(22) \equiv$

**Table 1** Solutions of $x^2 + x + 223 \equiv 0 \pmod{3^j}$.




$0 \pmod{81}$, yielding the solutions $22, 49, 76 \pmod{81}$. Thus we find that the congruence has exactly nine solutions $\pmod{81}$. In fact we note that $f(4) \equiv 0 \pmod{3^5}$, $3^2 \| f'(4)$, so by Theorem 2.20 the solution $4 \pmod{243}$ is one of nine solutions of the form $4 + 27t \pmod{243}$. We may further verify that there is precisely one value of $t \pmod 3$, namely $t = 2$, for which $f(4 + 27t) \equiv 0 \pmod{3^6}$. This gives nine solutions of the form $58 + 81t \pmod{3^6}$. Similarly, $f(22) \equiv 0 \pmod{3^5}$, $3^2 \| f'(22)$, so that $22 \pmod{243}$ is one of nine solutions of the form $22 + 27t \pmod{243}$. Moreover, we can verify that there is precisely one value of $t \pmod 3$, namely $t = 0$, for which $22 + 27t$ is a solution $\pmod{3^6}$. That is, we have nine solutions $\pmod{3^6}$ of the form $22 + 81t$. On the other hand, $f'(13) \equiv 0 \pmod{27}$, so that $f(13 + 27t) \equiv f(13) \pmod{3^6}$. As $3^4 \| f(13)$, we find that none of the three solutions $13 + 27t \pmod{81}$ lifts to a solution $\pmod{243}$. In conclusion, we have found that for each $j \geqslant 5$ there are precisely 18 solutions $\pmod{3^j}$, of which 12 do not lift to $3^{j+1}$, while each of the remaining six lifts to three solutions $\pmod{3^{j+1}}$. These results are depicted in Table 1.

Suppose that $f(a) \equiv 0 \pmod p$, and that $f'(a) \equiv 0 \pmod p$. We wish to know whether $a$ can be lifted to solutions modulo arbitrarily high powers of $p$. The situation is resolved if we can reach a point at which Theorem 2.24 applies, that is, $j \geqslant 2\tau + 1$. However, there is nothing in our discussion thus far to preclude the possibility that the power of $p$ in $f'$ might steadily increase with that in $f$, so that Theorem 2.24 might never take effect. In Appendix A.2 we define the *discriminant* $D(f)$ of the polynomial, and show that the critical inequality $j \geqslant 2\tau + 1$ holds whenever $j$ is larger than the power of $p$ in $D(f)$.



**PROBLEMS**

1. Solve the congruence $x^2 + x + 7 \equiv 0 \pmod{27}$ by using the method of completing the square from elementary algebra, thus $4x^2 + 4x + 28 = (2x + 1)^2 + 27$. Solve this congruence $\pmod{81}$ by the same method.
2. Solve $x^5 + x^4 + 1 \equiv 0 \pmod{3^4}$.
3. Solve $x^3 + x + 57 \equiv 0 \pmod{5^3}$.
4. Solve $x^2 + 5x + 24 \equiv 0 \pmod{36}$.
5. Solve $x^3 + 10x^2 + x + 3 \equiv 0 \pmod{3^3}$.
6. Solve $x^3 + x^2 - 4 \equiv 0 \pmod{7^3}$.
7. Solve $x^3 + x^2 - 5 \equiv 0 \pmod{7^3}$.
8. Apply the theory of this section to solve $1000x \equiv 1 \pmod{101^3}$, using a calculator.
9. Suppose that $f(a) \equiv 0 \pmod{p^j}$ and that $f'(a) \not\equiv 0 \pmod p$. Let $\overline{f'(a)}$ be an integer chosen so that $f'(a)\overline{f'(a)} \equiv 1 \pmod{p^j}$, and put $b = a - f(a)\overline{f'(a)}$. Show that $f(b) \equiv 0 \pmod{p^{2j}}$.
10. Let $p$ be an odd prime, and suppose that $a \not\equiv 0 \pmod p$. Show that if the congruence $x^2 \equiv a \pmod{p^j}$ has a solution when $j = 1$, then it has a solution for all $j$.

*11. Let $f(\mathbf{x})$ be a polynomial with integral coefficients in the $n$ variables $x_1, x_2, \cdots, x_n$. Suppose that $f(\mathbf{a}) \equiv 0 \pmod p$ where $\mathbf{a} = (a_1, a_2, \cdots, a_n)$, and that $\dfrac{\partial}{\partial x_i} f(\mathbf{a}) \not\equiv 0 \pmod p$ for at least one $i$. Show that the congruence $f(\mathbf{x}) \equiv 0 \pmod{p^j}$ has a solution for every $j$.

## 2.7 PRIME MODULUS

We have now reduced the problem of solving $f(x) \equiv 0 \pmod m$ to its last stage, congruences with prime moduli. Although we have no general method for solving such congruences, there are some interesting facts concerning the solutions. A natural question about polynomial congruences of the type $f(x) \equiv 0 \pmod m$ is whether there is any analogue to the well-known theorem in algebra that a polynomial equation of degree $n$ whose coefficients are complex numbers has exactly $n$ roots or solutions, allowing for multiple roots. For congruences the situation is more complicated. In the first place, for any modulus $m > 1$, there are polynomial congruences having no solutions. An example of this is given by $x^p - x + 1 \equiv 0 \pmod m$, where $p$ is any prime factor of $m$. This congruence has no solutions because $x^p - x + 1 \equiv 0 \pmod p$ has none, by Fermat's theorem.

Moreover, we have already seen that a congruence can have more solutions than its degree, for example, $x^2 - 7x + 2 = 0 \pmod{10}$ with four solutions $x = 3, 4, 8, 9$, and also $x^2 + x + 7 \equiv 0 \pmod{27}$ with three solutions $x = 4, 13, 22$. But if the modulus is a *prime*, a congruence cannot have more solutions than its degree. This is proved in Theorem 2.26 later in the section. It is important here to note carefully the meaning of "degree of congruence," given in Definition 2.5 in Section 2.2. Such a polynomial as $5x^3 + x^2 - x$ has degree 3, but the congruence $5x^3 + x^2 - x \equiv 0 \pmod{5}$ has degree 2.

Consider the congruence $5x^2 + 10x + 15 \equiv 0 \pmod{5}$, having five solutions $x = 0, 1, 2, 3$, and 4. At first glance, this might appear to be a counterexample to Theorem 2.26. However, by Definition 2.5, this congruence is assigned no degree, so that Theorem 2.26 does not apply.

With this background, we proceed to prove some fundamental results. As before, we write $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$, and we assume that $p$ is a prime not dividing $a_n$, so that the congruence $f(x) \equiv 0 \pmod{p}$ has degree $n$. In Theorem 2.25, we divide such a polynomial $f(x)$ of degree $n \geqslant p$ by $x^p - x$ to get a quotient and a remainder, both polynomials. This is a limited use of the *division algorithm for polynomials*, which is discussed more fully in Theorem 9.1. By "limited use," we mean that the only idea involved is the division of one polynomial into another, as in elementary algebra. The uniqueness of the quotient and the remainder are not needed.

**Theorem 2.25** *If the degree $n$ of $f(x) \equiv 0 \pmod{p}$ is greater than or equal to $p$, then either every integer is a solution of $f(x) \equiv 0 \pmod{p}$ or there is a polynomial $g(x)$ having integral coefficients, with leading coefficient 1, such that $g(x) \equiv 0 \pmod{p}$ is of degree less than $p$ and the solutions of $g(x) \equiv 0 \pmod{p}$ are precisely those of $f(x) \equiv 0 \pmod{p}$.*

*Proof* Dividing $f(x)$ by $x^p - x$, we get a quotient $q(x)$ and a remainder $r(x)$ such that $f(x) = (x^p - x)q(x) + r(x)$. Here $q(x)$ and $r(x)$ are polynomials with integral coefficients, and $r(x)$ is either zero or a polynomial of degree less than $p$. Since every integer is a solution of $x^p \equiv x \pmod{p}$ by Fermat's theorem, we see that the solutions of $f(x) \equiv 0 \pmod{p}$ are the same as those of $r(x) \equiv 0 \pmod{p}$. If $r(x) = 0$ or if every coefficient of $r(x)$ is divisible by $p$, then every integer is a solution of $f(x) \equiv 0 \pmod{p}$.

On the other hand, if at least one coefficient of $r(x)$ is not divisible by $p$, then the congruence $r(x) \equiv 0 \pmod{p}$ has a degree, and that degree is less than $p$. The polynomial $g(x)$ in the theorem can be obtained from $r(x)$ by getting leading coefficient 1, as follows. We may discard all terms



in $r(x)$ whose coefficients are divisible by $p$, since the congruence properties modulo $p$ are unaltered. Then let $bx^m$ be the term of highest degree in $r(x)$, with $(b, p) = 1$. Choose $\bar{b}$ so that $b\bar{b} \equiv 1 \pmod{p}$, and note that $(\bar{b}, p) = 1$ also. Then the congruence $\bar{b}r(x) \equiv 0 \pmod{p}$ has the same solutions as $r(x) \equiv 0 \pmod{p}$, and so has the same solutions as $f(x) \equiv 0 \pmod{p}$. Define $g(x)$ to be $\bar{b}r(x)$ with its leading coefficient $b\bar{b}$ replaced by 1, that is,

$$g(x) = \bar{b}r(x) - (b\bar{b} - 1)x^m.$$

**Theorem 2.26** *The congruence $f(x) \equiv 0 \pmod{p}$ of degree $n$ has at most $n$ solutions.*

*Proof* The proof is by induction on the degree of $f(x) \equiv 0 \pmod{p}$. If $n = 0$, the polynomial $f(x)$ is just $a_0$ with $a_0 \not\equiv 0 \pmod{p}$, and hence the congruence has no solution. If $n = 1$, the congruence has exactly one solution by Theorem 2.17. Assuming the truth of the theorem for all congruences of degree $< n$, suppose that there are more than $n$ solutions of the congruence $f(x) \equiv 0 \pmod{p}$ of degree $n$. Let the leading term of $f(x)$ be $a_n x^n$ and let $u_1, u_2, \cdots, u_n, u_{n+1}$ be solutions of the congruence, with $u_i \not\equiv u_j \pmod{p}$ for $i \neq j$. We define $g(x)$ by the equation

$$g(x) = f(x) - a_n(x - u_1)(x - u_2) \cdots (x - u_n),$$

noting the cancellation of $a_n x^n$ on the right.

Note that $g(x) \equiv 0 \pmod{p}$ has at least $n$ solutions, namely $u_1, u_2, \cdots, u_n$. We consider two cases, first where every coefficient of $g(x)$ is divisible by $p$, and second where at least one coefficient is not divisible by $p$. (The first case includes the situation where $g(x)$ is identically zero.) We show that both cases lead to a contradiction. In the first case, every integer is a solution of $g(x) \equiv 0 \pmod{p}$, and since $f(u_{n+1}) \equiv 0 \pmod{p}$ by assumption, it follows that $x = u_{n+1}$ is a solution of

$$a_n(x - u_1)(x - u_2) \cdots (x - u_n) \equiv 0 \pmod{p}.$$

This contradicts Theorem 1.15.

In the second case, we note that the congruence $g(x) \equiv 0 \pmod{p}$ has a degree, and that degree is less than $n$. By the induction hypothesis, this congruence has fewer than $n$ solutions. This contradicts the earlier observation that this congruence has at least $n$ solutions. Thus the proof is complete.

We have already noted, using the example $5x^2 + 10x + 15 \equiv 0 \pmod{5}$, that the conclusion of Theorem 2.26 need not hold if the

assumption is just that the polynomial $f(x)$ has degree $n$. The following corollary describes the situation.

**Corollary 2.27** *If* $b_n x^n + b_{n-1}x^{n-1} + \cdots + b_0 \equiv 0 \pmod p$ *has more than* $n$ *solutions, then all the coefficients* $b_j$ *are divisible by* $p$.

The reason for this is that if some coefficient is not divisible by $p$, then the polynomial congruence has a degree, and that degree is at most $n$. Theorem 2.26 implies that the congruence has at most $n$ solutions, and this is a contradiction.

**Theorem 2.28** *If* $F(x)$ *is a function that maps residue classes* (mod $p$) *to residue classes* (mod $p$), *then there is a polynomial* $f(x)$ *with integral coefficients and degree at most* $p-1$ *such that* $F(x) \equiv f(x) \pmod p$ *for all residue classes* $x$ (mod $p$).

*Proof* By Fermat's congruence we see that

$$1 - (x-a)^{p-1} \equiv \begin{cases} 1 \pmod p & \text{if } x \equiv a \pmod p, \\ 0 \pmod p & \text{otherwise.} \end{cases}$$

Hence the polynomial $f(x) = \sum_{i=1}^{p} F(i)(1 - (x-i)^{p-1})$ has the desired properties.

**Theorem 2.29** *The congruence* $f(x) \equiv 0 \pmod p$ *of degree* $n$, *with leading coefficient* $a_n = 1$, *has* $n$ *solutions if and only if* $f(x)$ *is a factor of* $x^p - x$ *modulo* $p$, *that is, if and only if* $x^p - x = f(x)q(x) + ps(x)$, *where* $q(x)$ *and* $s(x)$ *have integral coefficients,* $q(x)$ *has degree* $p-n$ *and leading coefficient* 1, *and where either* $s(x)$ *is a polynomial of degree less than* $n$ *or* $s(x)$ *is zero.*

*Proof* First assume that $f(x) \equiv 0 \pmod p$ has $n$ solutions. Then $n \leqslant p$, by Definition 2.4 of Section 2.2. Dividing $x^p - x$ by $f(x)$, we get a quotient $q(x)$ and a remainder $r(x)$ satisfying $x^p - x = f(x)q(x) + r(x)$, where $r(x)$ is either identically zero or a polynomial of degree less than $n$. This equation implies, by application of Fermat's theorem to $x^p - x$, that every solution of $f(x) \equiv 0 \pmod p$ is a solution of $r(x) \equiv 0 \pmod p$. Thus, $r(x) \equiv 0 \pmod p$ has at least $n$ solutions, and by Corollary 2.27, it follows that every coefficient in $r(x)$ is divisible by $p$, so $r(x) = ps(x)$ as in the theorem.

Conversely, assume that $x^p - x = f(x)q(x) + ps(x)$, as in the statement of the theorem. By Fermat's theorem, the congruence $f(x)q(x) \equiv 0 \pmod p$ has $p$ solutions. This congruence has leading term $x^p$. The leading term of $f(x)$ is $x^n$ by hypothesis, and hence the leading term of $q(x)$ is $x^{p-n}$. By Theorem 2.26, the congruences $f(x) \equiv 0 \pmod p$ and $q(x) \equiv 0 \pmod p$ have at most $n$ solutions and $p-n$ solutions, respectively. But every one of the $p$ solutions of $f(x)q(x) \equiv 0 \pmod p$ is a solution of at least one of the congruences $f(x) \equiv 0 \pmod p$ and $q(x) \equiv 0 \pmod p$. It follows that these two congruences have exactly $n$ solutions and $p-n$ solutions, respectively.



The restriction $a_n = 1$ in this theorem is needed so that we may divide $x^p - x$ by $f(x)$ and obtain a polynomial $q(x)$ with integral coefficients. However, it is not much of a restriction. We can always find an integer $\overline{a_n}$ such that $a_n\overline{a_n} \equiv 1 \pmod p$. Put $g(x) = \overline{a_n}f(x) - (a_n\overline{a_n} - 1)x^n$. Then $g(x) \equiv 0 \pmod p$ has the same solutions as $f(x) \equiv 0 \pmod p$, and $g(x)$ has leading coefficient 1.

As an example, we see that $x^5 - 5x^3 + 4x \equiv 0 \pmod 5$ has five solutions, and $x^5 - x = (x^5 - 5x^3 + 4x) + (5x^3 - 5x)$. As a second example, we cite $x^3 - x \equiv 0 \pmod 5$ with three solutions, and $x^5 - x = (x^3 - x)(x^2 + 1)$. Theorem 2.29 has many important applications. We now consider one that will be crucial to our discussion of primitive roots in Section 2.8.

**Corollary 2.30** *If* $d \mid (p-1)$, *then* $x^d \equiv 1 \pmod p$ *has* $d$ *solutions.*

*Proof* Choose $e$ so that $de = p - 1$. Since $(y-1)(1 + y + \cdots + y^{e-1}) = y^e - 1$, on taking $y = x^d$ we see that $x(x^d - 1)(1 + x^d + \cdots + x^{d(e-1)}) = x^p - x$.

A further application of Theorem 2.29 arises by considering the polynomial

$$f(x) = (x-1)(x-2)\cdots(x-p+1).$$

For convenience we assume that $p > 2$. On expanding, we find that

$$f(x) = x^{p-1} - \sigma_1 x^{p-2} + \sigma_2 x^{p-3} - \cdots + \sigma_{p-1} \tag{2.7}$$

where $\sigma_j$ is the sum of all products of $j$ distinct members of the set $\{1, 2, \cdots, p-1\}$. In the two extreme cases we have $\sigma_1 = 1 + 2 + \cdots + (p-1) = p(p-1)/2$, and $\sigma_{p-1} = 1 \cdot 2 \cdot \cdots \cdot (p-1) = (p-1)!$. The polynomial $f(x)$ has degree $p-1$ and has the $p-1$ roots $1, 2, \cdots, p-1 \pmod p$. Consequently the polynomial $xf(x)$ has degree $p$ and has $p$ roots. By applying Theorem 2.29 to this latter polynomial, we see that

there are polynomials $q(x)$ and $s(x)$ such that $x^p - x = xf(x)q(x) + ps(x)$. Since $q(x)$ has degree $p - p = 0$ and leading coefficient 1, we see that $q(x) = 1$. That is, $x^p - x = xf(x) + ps(x)$, which is to say that the coefficients of $x^p - x$ are congruent (mod $p$) to those of $xf(x)$. On comparing the coefficients of $x$, we deduce that $\sigma_{p-1} = (p-1)! \equiv -1 \pmod{p}$, which provides a second proof of Wilson's congruence. On comparing the remaining coefficients, we deduce that $\sigma_j \equiv 0 \pmod{p}$ for $1 \leqslant j \leqslant p-2$. To these useful observations we may add one further remark: if $p \geqslant 5$ then

$$\sigma_{p-2} \equiv 0 \pmod{p^2}. \tag{2.8}$$

This is *Wolstenholme's congruence*. To prove it, we note that $f(p) = (p-1)(p-2)\cdots(p-p+1) = (p-1)!$. On taking $x = p$ in (2.7), we have

$$(p-1)! = p^{p-1} - \sigma_1 p^{p-2} + \cdots + \sigma_{p-3}p^2 - \sigma_{p-2}p + \sigma_{p-1}.$$

We have already observed that $\sigma_{p-1} = (p-1)!$. On subtracting this amount from both sides and dividing through by $p$, we deduce that

$$p^{p-2} - \sigma_1 p^{p-3} + \cdots + \sigma_{p-3}p - \sigma_{p-2} = 0.$$

All terms except the last two contain visible factors of $p^2$. Thus $\sigma_{p-3}p \equiv \sigma_{p-2} \pmod{p^2}$. This gives the desired result, since $\sigma_{p-3} \equiv 0 \pmod{p}$.

**PROBLEMS**

1. Reduce the following congruences to equivalent congruences of degree $\leqslant 6$:
   (*a*) $x^{11} + x^8 + 5 \equiv 0 \pmod 7$;
   (*b*) $x^{20} + x^{13} + x^7 + x \equiv 2 \pmod 7$;
   (*c*) $x^{15} - x^{10} + 4x - 3 \equiv 0 \pmod 7$.
2. Prove that $2x^3 + 5x^2 + 6x + 1 \equiv 0 \pmod 7$ has three solutions by use of Theorem 2.29.
3. Prove that $x^{14} + 12x^2 \equiv 0 \pmod{13}$ has 13 solutions and so it is an identical congruence.
4. Prove that if $f(x) \equiv 0 \pmod p$ has $j$ solutions $x \equiv a_1, x \equiv a_2, \cdots, x \equiv a_j \pmod p$, there is a polynomial $q(x)$ such that $f(x) \equiv (x-a_1)(x-a_2)\cdots(x-a_j)q(x) \pmod p$. (H)



5. With the assumptions and notation of the preceding problem, prove that if the degree of $f(x)$ is $j$, then $q(x)$ is a constant and can be taken as the leading coefficient of $f(x)$.
6. Let $m$ be composite. Prove that Theorem 2.26 is false if "mod $p$" is replaced by "mod $m$."
7. Show that if the prime number $p$ in Theorem 2.28 is replaced by a composite number $m$ then the statement becomes false.
8. Explain why the proof of Wolstenholme's congruence fails when $p = 3$.
9. For $p = 5$, compute the values of the numbers $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ in (2.7).
10. Write $1/1 + 1/2 + \cdots + 1/(p-1) = a/b$ with $(a, b) = 1$. Show that $p^2 | a$ if $p \geqslant 5$.

*11. Let $p$ be a prime, $p \geqslant 5$, and suppose that the numbers $\sigma_j$ are as in (2.7). Show that $\sigma_{p-2} \equiv p\sigma_{p-3} \pmod{p^3}$.

*12. Show that if $p \geqslant 5$ and $m$ is a positive integer then $\binom{mp-1}{p-1} \equiv 1 \pmod{p^3}$.

*13. Show that if $p \geqslant 5$ then $(mp)! \equiv m!p!^m \pmod{p^{m+3}}$.

*14. Suppose that $p$ is an odd prime, and write $1/1 - 1/2 + 1/3 - \cdots - 1/(p-1) = a/(p-1)!$. Show that $a \equiv (2-2^p)/p \pmod p$.

## 2.8 PRIMITIVE ROOTS AND POWER RESIDUES

**Definition 2.6** *Let $m$ denote a positive integer and $a$ any integer such that $(a, m) = 1$. Let $h$ be the smallest positive integer such that $a^h \equiv 1 \pmod m$. We say that the order of $a$ modulo $m$ is $h$, or that $a$ belongs to the exponent $h$ modulo $m$.*

The terminology "$a$ belongs to the exponent $h$" is the classical language of number theory. This language is being replaced more and more in the current literature by "the order of $a$ is $h$," a usage that is standard in group theory. (In Sections 2.10 and 2.11 we shall explore the relationships between the ideas of number theory and those of group theory.)

Suppose that $a$ has order $h \pmod m$. If $k$ is a positive multiple of $h$, say $k = qh$, then $a^k = a^{qh} = (a^h)^q \equiv 1^q \equiv 1 \pmod m$. Conversely, if $k$ is a positive integer such that $a^k \equiv 1 \pmod m$, then we apply the division algorithm to obtain integers $q$ and $r$ such that $k = qh + r$, $q \geqslant 0$, and $0 \leqslant r < h$. Thus $1 \equiv a^k = a^{qh+r} = (a^h)^q a^r \equiv 1^q a^r \equiv a^r \pmod m$. But $0 \leqslant$

$r < h$ and $h$ is the least positive power of $a$ that is congruent to 1 modulo $m$, so it follows that $r = 0$. Thus $h$ divides $k$, and we have proved the following lemma.

**Lemma 2.31** *If $a$ has order $h$ (mod $m$), then the positive integers $k$ such that $a^k \equiv 1 \pmod m$ are precisely those for which $h|k$.*

**Corollary 2.32** *If $(a, m) = 1$, then the order of $a$ modulo $m$ divides $\phi(m)$.*

*Proof* Each reduced residue class $a$ modulo $m$ has finite order, for by Euler's congruence $a^{\phi(m)} \equiv 1 \pmod m$. Moreover, if $a$ has order $h$ then by taking $k = \phi(m)$ in the lemma we deduce that $h|\phi(m)$.

**Lemma 2.33** *If $a$ has order $h$ modulo $m$, then $a^k$ has order $h/(h, k)$ modulo $m$.*

Since $h/(h, k) = 1$ if and only if $h|k$, we see that Lemma 2.33 contains Lemma 2.31 as a special case.

*Proof* According to Lemma 2.31, $(a^k)^j \equiv 1 \pmod m$ if and only if $h|kj$. But $h|kj$ if and only if $\{h/(h, k)\}|\{k/(h, k)\}j$. As the divisor is relatively prime to the first factor of the dividend, this relation holds if and only if $\{h/(h, k)\}|j$. Therefore the least positive integer $j$ such that $(a^k)^j \equiv 1 \pmod m$ is $j = h/(h, k)$.

If $a$ has order $h$ and $b$ has order $k$, both modulo $m$, then $(ab)^{hk} = (a^h)^k(b^k)^h \equiv 1 \pmod m$, and from Lemma 2.31 we deduce that the order of $ab$ is a divisor of $hk$. If $h$ and $k$ are relatively prime, then we can say more.

**Lemma 2.34** *If $a$ has order $h$ (mod $m$), $b$ has order $k$ (mod $m$), and if $(h, k) = 1$, then $ab$ has order $hk$ (mod $m$).*

*Proof* Let $r$ denote the order of $ab \pmod m$. We have shown that $r|hk$. To complete the proof it suffices to show that $hk|r$. We note that $b^{rh} \equiv (a^h)^r b^{rh} = (ab)^{rh} \equiv 1 \pmod m$. Thus $k|rh$ by Lemma 2.31. As $(h, k) = 1$, it follows that $k|r$. By a similar argument we see that $h|r$. Using again the hypothesis $(h, k) = 1$, we conclude that $hk|r$.

We have already seen that the order of $a$ modulo $m$ is a divisor of $\phi(m)$. For certain values of $m$, there are integers $a$ such that the order of $a$ is equal to $\phi(m)$. These cases are of considerable importance, so a special label is used.



**Definition 2.7** *If $g$ belongs to the exponent $\phi(m)$ modulo $m$, then $g$ is called a* primitive root modulo $m$.

(In algebraic language, this definition can be stated: If the order of $a$ modulo $m$ is $\phi(m)$, then the multiplicative group of reduced residues modulo $m$ is a cyclic group generated by the element $g$. Readers not too familiar with group theory can find a more detailed explanation of this in Section 2.10.)

In view of Lemma 2.31, the number $a$ is a solution of the congruence $x^k \equiv 1 \pmod m$ if and only if the order of $a \pmod m$ divides $k$. In one special case, namely the situation of Corollary 2.30, we have determined the number of solutions of this congruence. That is, if $p$ is prime and $k|(p - 1)$, then there are precisely $k$ residue classes $a \pmod p$ such that the order of $a$ modulo $p$ is a divisor of $k$. If $k$ happens to be a prime power, we can then determine the exact number of residues $a \pmod p$ of order $k$.

**Lemma 2.35** *Let $p$ and $q$ be primes, and suppose that $q^\alpha|(p - 1)$, where $\alpha \geqslant 1$. Then there are precisely $q^\alpha - q^{\alpha-1}$ residue classes $a \pmod p$ of order $q^\alpha$.*

*Proof* The divisors of $q^\alpha$ are the numbers $q^\beta$ with $\beta = 0, 1, \cdots, \alpha$. Of these, $q^\alpha$ is the only one that is not a divisor of $q^{\alpha-1}$. There are $q^\alpha$ residues (mod $p$) of order dividing $q^\alpha$, and among these there are $q^{\alpha-1}$ residues of order dividing $q^{\alpha-1}$. On subtracting we see that there are precisely $q^\alpha - q^{\alpha-1}$ residues $a$ of order $q^\alpha \pmod p$.

**Theorem 2.36** *If $p$ is a prime then there exist $\phi(p - 1)$ primitive roots modulo $p$.*

*Proof* We first establish the existence of at least one primitive root. Let $p - 1 = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_j^{\alpha_j}$ be the canonical factorization of $p - 1$. By Lemma 2.35 we may choose numbers $a_i \pmod p$ so that $a_i$ has order $p_i^{\alpha_i}$, $i = 1, 2, \cdots, j$. The numbers $p_i^{\alpha_i}$ are pairwise relatively prime, so by repeated use of Lemma 2.34 we see that $g = a_1 a_2 \cdots a_j$ has order $p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_j^{\alpha_j} = p - 1$. That is, $g$ is a primitive root (mod $p$).

To complete the proof, we determine the exact number of primitive roots (mod $p$). Let $g$ be a primitive root (mod $p$). Then the numbers $g, g^2, g^3, \cdots, g^{p-1}$ form a system of reduced residues (mod $p$). By Lemma 2.33 we see that $g^k$ has order $(p - 1)/(k, p - 1)$. Thus $g^k$ is a primitive root if and only if $(k, p - 1) = 1$. By definition of Euler's phi function, there are exactly $\phi(p - 1)$ such values of $k$ in the interval $1 \leqslant k \leqslant p - 1$.

***Remark on Calculation*** Suppose that we wish to show that $a$ has order $h \pmod m$, where $a$, $h$, and $m$ are given. By using the repeated squaring device discussed in Section 2.4, we may quickly verify that $a^h \equiv 1 \pmod m$. If $h$ is small, then we simply examine $a, a^2, \cdots, a^{h-1} \pmod m$, but if $h$ is large (e.g., $h = \phi(m)$), then the amount of calculation here would be prohibitively long. Instead, we note by Lemma 2.31 that the order of $a$ must be a divisor of $h$. If the order of $a$ is a proper divisor of $h$ then the order of $a$ divides $h/p$ for some prime factor $p$ of $h$. That is, the order of $a \pmod m$ is $h$ if and only if the following two conditions are satisfied: (i) $a^h \equiv 1 \pmod m$, and (ii) for each prime factor $p$ of $h$, $a^{h/p} \not\equiv 1 \pmod m$. In case $m$ is prime, we may take $h = m - 1$ in this criterion to determine whether $a$ is a primitive root. To locate a primitive root we simply try $a = 2$, $a = 3, \cdots$, and in general a primitive root is quickly found. For example, to show that 2 is a primitive root (mod 101), we note that 2 and 5 are the primes dividing 100. Then we calculate that $2^{50} \equiv -1 \not\equiv 1 \pmod{101}$, and that $2^{20} \equiv 95 \not\equiv 1 \pmod{101}$.

The techniques discussed in Section 2.4 allow us to prove very quickly that a given number $m$ is composite, but they are not so useful in establishing primality. Suppose that a given number $p$ is a strong pseudoprime to several bases, and is therefore expected to be prime. To show that $p$ is prime it suffices to exhibit a number $a$ of order $p - 1 \pmod p$, for then $\phi(p) \geqslant p - 1$, and hence $p$ must be prime. Here the hard part is to factor $p - 1$. (If the desired primitive root is elusive, then $p$ is probably composite.) This approach is developed further in Problems 38 and 39 at the end of this Section.

Up to $10^9$ or so one may construct primes by sieving. Larger primes (such as those used in public-key cryptography) can be constructed as follows: Multiply several small primes together, add 1 to this product, and call the result $p$. This number has no greater chance of being prime than a randomly chosen number of the same size, and indeed it is likely that a pseudoprime test will reveal that $p$ is composite (in which case we try again with a new product of small primes). However, if $p$ passes several such tests, then one may proceed as above to show that $p$ is prime, since the factorization of $p - 1$ is known in advance.

**Definition 2.8** *If $(a, p) = 1$ and $x^n \equiv a \pmod p$ has a solution, then $a$ is called an $n$th power residue modulo $p$.*

If $(g, m) = 1$ then the sequence $g, g^2, \cdots \pmod m$ is periodic. If $g$ is a primitive root $\pmod m$ then the least period of this sequence is $\phi(m)$, and we see that $g, g^2, \cdots, g^{\phi(m)}$ form a system of reduced residues $\pmod m$. Thus $g^i \equiv g^j \pmod m$ if and only if $i \equiv j \pmod{\phi(m)}$. By expressing numbers as powers of $g$, we may convert a multiplicative congruence $\pmod m$ to an additive congruence $\pmod{\phi(m)}$, just as we apply logarithms to real numbers. In this way we determine whether $a$ is an $n$th power residue $\pmod p$.



**Theorem 2.37** *If $p$ is a prime and $(a, p) = 1$, then the congruence $x^n \equiv a \pmod p$ has $(n, p - 1)$ solutions or no solution according as*

$$a^{(p-1)/(n,p-1)} \equiv 1 \pmod p$$

*or not.*

*Proof* Let $g$ be a primitive root $\pmod p$, and choose $i$ so that $g^i \equiv a \pmod p$. If there is an $x$ such that $x^n \equiv a \pmod p$ then $(x, p) = 1$, so that $x \equiv g^u \pmod p$ for some $u$. Thus the proposed congruence is $g^{nu} \equiv g^i \pmod p$, which is equivalent to $nu \equiv i \pmod{p-1}$. Put $k = (n, p - 1)$. By Theorem 2.17, this has $k$ solutions if $k|i$, and no solution if $k \nmid i$. If $k|i$, then $i(p-1)/k \equiv 0 \pmod{p-1}$, so that $a^{(p-1)/k} \equiv g^{i(p-1)/k} = (g^{p-1})^{i/k} \equiv 1 \pmod p$. On the other hand, if $k \nmid i$ then $i(p-1)/k \not\equiv 0 \pmod{p-1}$, and hence $a^{(p-1)/k} \equiv g^{i(p-1)/k} \not\equiv 1 \pmod p$.

**Example 14** Show that the congruence $x^5 \equiv 6 \pmod{101}$ has 5 solutions.

*Solution* It suffices to verify that $6^{20} \equiv 1 \pmod{101}$. This is easily accomplished using the technique discussed in Section 2.4. Note that we do not need to find a primitive root $g$, or to find $i$ such that $g^i \equiv 6 \pmod{101}$. The mere fact that $6^{20} \equiv 1 \pmod{101}$ assures us that $5|i$. (With more work one may prove that $g = 2$ is a primitive root $\pmod{101}$, and that $2^{70} \equiv 6 \pmod{101}$. Hence the five solutions are $x \equiv 2^{14+20j} \pmod{101}$ where $j = 0, 1, 2, 3, 4$. That is, $x \equiv 22, 70, 85, 96, 30 \pmod{101}$.)

**Corollary 2.38** *Euler's criterion. If $p$ is an odd prime and $(a, p) = 1$, then $x^2 \equiv a \pmod p$ has two solutions or no solution according as $a^{(p-1)/2} \equiv 1$ or $\equiv -1 \pmod p$.*

*Proof* Put $b = a^{(p-1)/2}$. Thus $b^2 = a^{p-1} \equiv 1 \pmod p$ by Fermat's congruence. From Lemma 2.10 it follows that $b \equiv \pm 1 \pmod p$. If $b \equiv -1 \pmod p$ then the congruence $x^2 \equiv a \pmod p$ has no solution, by Theorem 2.37. If $b \equiv 1 \pmod p$ then the congruence has exactly two solutions, by Theorem 2.37.

By taking $a = -1$ in Euler's criterion we obtain a second proof of Theorem 2.12. In the next section we give an algorithm for solving the congruence $x^2 \equiv a \pmod p$. In Sections 3.1 and 3.2 a quite different approach of Gauss is developed, which offers an alternative to Euler's

criterion for determining whether a given number $a$ is a quadratic residue (mod $p$).

We have seen that primitive roots provide a valuable tool for analyzing certain congruences (mod $p$). We now investigate the extent to which this can be generalized to other moduli.

**Theorem 2.39** *If $p$ is a prime then there exist $\phi(\phi(p^2)) = (p-1)\phi(p-1)$ primitive roots modulo $p^2$.*

*Proof* We show that if $g$ is a primitive root (mod $p$) then $g + tp$ is a primitive root (mod $p^2$) for exactly $p - 1$ values of $t$ (mod $p$). Let $h$ denote the order of $g + tp$ (mod $p^2$). (Thus $h$ may depend on $t$.) Since $(g + tp)^h \equiv 1 \pmod{p^2}$, it follows that $(g + tp)^h \equiv 1 \pmod{p}$, which in turn implies that $g^h \equiv 1 \pmod{p}$, and hence that $(p-1) | h$. On the other hand, by Corollary 2.32 we know that $h | \phi(p^2) = p(p-1)$. Thus $h = p - 1$ or $h = p(p-1)$. In the latter case $g + tp$ is a primitive root (mod $p^2$), and in the former case it is not. We prove that the former case arises for only one of the $p$ possible values of $t$. Let $f(x) = x^{p-1} - 1$. In the former case, $g + tp$ is a solution of the congruence $f(x) \equiv 0 \pmod{p^2}$ lying above $g$ (mod $p$). Since $f'(g) = (p-1)g^{p-2} \not\equiv 0 \pmod{p}$, we know from Hensel's lemma (Theorem 2.23) that $g$ (mod $p$) lifts to a unique solution $g + tp \pmod{p^2}$. For all other values of $t$ (mod $p$), the number $g + tp$ is a primitive root (mod $p^2$).

Since each of the $\phi(p-1)$ primitive roots (mod $p$) give rise to exactly $p - 1$ primitive roots (mod $p^2$), we have now shown that there exist at least $(p-1)\phi(p-1)$ primitive roots (mod $p^2$). To show that there are no other primitive roots (mod $p^2$), it suffices to argue as in the preceding proof. Let $g$ denote a primitive root (mod $p^2$), so that the numbers $g, g^2, \cdots, g^{p(p-1)}$ form a system of reduced residues (mod $p^2$). By Lemma 2.33, we know that $g^k$ is a primitive root if and only if $(k, p(p-1)) = 1$. By the definition of Euler's phi function, there are precisely $\phi(p(p-1))$ such values of $k$ among the numbers $1, 2, \cdots, p(p-1)$. Since $(p, p-1) = 1$, we deduce from Theorem 2.19 that $\phi(p(p-1)) = \phi(p)\phi(p-1) = (p-1)\phi(p-1)$.

**Theorem 2.40** *If $p$ is an odd prime and $g$ is a primitive root modulo $p^2$, then $g$ is a primitive root modulo $p^\alpha$ for $\alpha = 3, 4, 5, \cdots$.*

*Proof* Suppose that $g$ is a primitive root (mod $p^2$), and that $h$ is the order of $g$ (mod $p^\alpha$) where $\alpha > 2$. From the congruence $g^h \equiv 1 \pmod{p^\alpha}$ we deduce that $g^h \equiv 1 \pmod{p^2}$, and hence that $\phi(p^2) | h$. By Corollary 2.32 we also know that $h | \phi(p^\alpha)$. Thus $h = p^\beta(p-1)$ for some $\beta$ among



$\beta = 1, 2, \cdots$, or $\alpha - 1$. To prove that $\beta = \alpha - 1$, it suffices to show that

$$g^{p^{\alpha-2}(p-1)} \not\equiv 1 \pmod{p^\alpha}. \tag{2.9}$$

We use induction to show that this holds for all $\alpha \geq 2$. By hypothesis, the order of $g$ (mod $p^2$) is $\phi(p^2) = p(p-1)$. Hence $g^{p-1} \not\equiv 1 \pmod{p^2}$, and we have (2.9) when $\alpha = 2$. By Fermat's congruence $g^{p-1} \equiv 1 \pmod{p}$, so we may write $g^{p-1} = 1 + b_1 p$ with $p \nmid b_1$. By the binomial theorem,

$$g^{p(p-1)} = (1 + b_1 p)^p = 1 + \binom{p}{1} b_1 p + \binom{p}{2} b_1^2 p^2 + \cdots.$$

Since $p > 2$ by hypothesis, $\binom{p}{2} = p(p-1)/2 \equiv 0 \pmod{p}$, and hence the above is $\equiv 1 + b_1 p^2 \pmod{p^3}$. This gives (2.9) when $\alpha = 3$. Thus we may write $g^{p(p-1)} = 1 + b_2 p^2$ with $p \nmid b_2$. We raise both sides of this to the $p$th power and repeat this procedure to find that $g^{p^2(p-1)} \equiv 1 + b_2 p^3 \pmod{p^4}$, which gives (2.9) for $\alpha = 4$. Continuing in this way, we conclude that (2.9) holds for all $\alpha \geq 2$, and the proof is complete.

The prime $p = 2$ must be excluded, for $g = 3$ is a primitive root (mod 4), but not (mod 8). Indeed it is easy to verify that $a^2 \equiv 1 \pmod{8}$ for any odd number $a$. As $\phi(8) = 4$, it follows that there is no primitive root (mod 8). Suppose that $a$ is odd. Since $8 | (a^2 - 1)$ and $2 | (a^2 + 1)$, it follows that $16 | (a^2 - 1)(a^2 + 1) = a^4 - 1$. That is, $a^4 \equiv 1 \pmod{16}$. On repeating this argument we see that $a^8 \equiv 1 \pmod{32}$, and in general that $a^{2^{\alpha-2}} \equiv 1 \pmod{2^\alpha}$ for $\alpha \geq 3$. Since $\phi(2^\alpha) = 2^{\alpha-1}$, we conclude that if $\alpha \geq 3$ then

$$a^{\phi(2^\alpha)/2} \equiv 1 \pmod{2^\alpha} \tag{2.10}$$

for all odd $a$, and hence that there is no primitive root (mod $2^\alpha$) for $\alpha = 3, 4, 5, \cdots$.

Suppose that $p$ is an odd prime and that $g$ is a primitive root (mod $p^\alpha$). We may suppose that $g$ is odd, for if $g$ is even then we have only to replace $g$ by $g + p^\alpha$, which is odd. The numbers $g, g^2, \cdots, g^{\phi(p^\alpha)}$ form a reduced residue system (mod $p^\alpha$). Since these numbers are odd, they also form a reduced residue system (mod $2p^\alpha$). Thus $g$ is a primitive root (mod $2p^\alpha$).

We have established that a primitive root exists modulo $m$ when $m = 1, 2, 4, p^\alpha$, or $2p^\alpha$, ($p$ an odd prime), but that there is no primitive root (mod $2^\alpha$) for $\alpha \geq 3$. Suppose now that $m$ is not a prime power or twice a prime power. Then $m$ can be expressed as a product, $m = m_1 m_2$

with $(m_1, m_2) = 1$, $m_1 > 2$, $m_2 > 2$. Let $e = \text{l.c.m.}(\phi(m_1), \phi(m_2))$. If $(a, m) = 1$ then $(a, m_1) = 1$, so that $a^{\phi(m_1)} \equiv 1 \pmod{m_1}$, and hence $a^e \equiv 1 \pmod{m_1}$. Similarly $a^e \equiv 1 \pmod{m_2}$, and hence $a^e \equiv 1 \pmod{m}$. Since $2|\phi(n)$ for all $n > 2$, we see that $2|(\phi(m_1), \phi(m_2))$, so that by Theorem 1.13,

$$e = \frac{\phi(m_1)\phi(m_2)}{(\phi(m_1), \phi(m_2))} < \phi(m_1)\phi(m_2) = \phi(m).$$

Thus there is no primitive root in this case. We have now determined precisely which $m$ possess primitive roots.

**Theorem 2.41** *There exists a primitive root modulo $m$ if and only if $m = 1, 2, 4, p^\alpha$, or $2p^\alpha$, where $p$ is an odd prime.*

Theorem 2.37 (and its proof) generalizes to any modulus $m$ possessing a primitive root.

**Corollary 2.42** *Suppose that $m = 1, 2, 4, p^\alpha$, or $2p^\alpha$, where $p$ is an odd prime. If $(a, m) = 1$ then the congruence $x^n \equiv a \pmod{m}$ has $(n, \phi(m))$ solutions or no solution, according as*

$$a^{\phi(m)/(n,\phi(m))} \equiv 1 \pmod{m} \tag{2.11}$$

*or not.*

For the general composite $m$ possessing no primitive root, we factor $m$ and apply the above to the prime powers dividing $m$.

**Example 15** Determine the number of solutions of the congruence $x^4 \equiv 61 \pmod{117}$.

*Solution* We note that $117 = 3^2 \cdot 13$. As $\phi(9)/(4, \phi(9)) = 6/(4, 6) = 3$ and $61^3 \equiv (-2)^3 \equiv 1 \pmod{9}$, we deduce that the congruence $x^4 \equiv 61 \pmod{9}$ has $(4, \phi(9)) = 2$ solutions. Similarly $\phi(13)/(4, \phi(13)) = 3$ and $61^3 \equiv (-4)^3 \equiv 1 \pmod{13}$, so the congruence $x^4 \equiv 61 \pmod{13}$ has $(4, \phi(13)) = 4$ solutions. Thus by Theorem 2.20, the number of solutions modulo 117 is $2 \cdot 4 = 8$.

This method fails in case the modulus is divisible by 8, as Corollary 2.42 does not apply to the higher powers of 2. In order to establish an analogue of Corollary 2.42 for the higher powers of 2, we first show that 5 is nearly a primitive root $(\text{mod } 2^\alpha)$.



**Theorem 2.43** *Suppose that $\alpha \geqslant 3$. The order of $5 \pmod{2^\alpha}$ is $2^{\alpha-2}$. The numbers $\pm 5, \pm 5^2, \pm 5^3, \cdots, \pm 5^{2^{\alpha-2}}$ form a system of reduced residues $(\text{mod } 2^\alpha)$. If $a$ is odd, then there exist $i$ and $j$ such that $a \equiv (-1)^i 5^j \pmod{2^\alpha}$. The values of $i$ and $j$ are uniquely determined $(\text{mod } 2)$ and $(\text{mod } 2^{\alpha-2})$, respectively.*

*Proof* We first show that $2^\alpha \| (5^{2^{\alpha-2}} - 1)$ for $\alpha \geqslant 2$. This is clear for $\alpha = 2$. If $a \equiv 1 \pmod 4$ then $2 \| (a + 1)$, and hence the power of 2 dividing $a^2 - 1 = (a - 1)(a + 1)$ is exactly one more than the power of 2 dividing $a - 1$. Taking $a = 5$, we deduce that $2^3 \| (5^2 - 1)$. Taking $a = 5^2$, we then deduce that $2^4 \| (5^4 - 1)$, and so on. Now let $h$ denote the order of $5 \pmod{2^\alpha}$. Since $h | \phi(2^\alpha)$ and $\phi(2^\alpha) = 2^{\alpha-1}$, we know that $h = 2^\beta$ for some $\beta$. But the least $\beta$ for which $5^{2^\beta} \equiv 1 \pmod{2^\alpha}$ is $\beta = \alpha - 2$. Thus 5 has order $2^{\alpha-2} \pmod{2^\alpha}$, so that the numbers $5, 5^2, 5^3, \cdots, 5^{2^{\alpha-2}}$ are mutually incongruent $(\text{mod } 2^\alpha)$. Of the $2^{\alpha-1}$ integers in a reduced residue system $(\text{mod } 2^\alpha)$, half are $\equiv 1 \pmod 4$, and half are $\equiv 3 \pmod 4$. The numbers $5^j$ are all $\equiv 1 \pmod 4$. Since the powers of 5 lie in $2^{\alpha-2}$ distinct residue classes $(\text{mod } 2^\alpha)$, and since $2^{\alpha-2}$ of the integers $(\text{mod } 2^\alpha)$ are $\equiv 1 \pmod 4$, for any $a \equiv 1 \pmod 4$ there is a $j$ such that $a \equiv 5^j \pmod{2^\alpha}$. For any integer $a \equiv 3 \pmod 4$, we observe that $-a \equiv 1 \pmod 4$, and hence that $-a \equiv 5^j \pmod{2^\alpha}$ for some $j$.

**Corollary 2.44** *Suppose that $\alpha \geqslant 3$ and that $a$ is odd. If $n$ is odd, then the congruence $x^n \equiv a \pmod{2^\alpha}$ has exactly one solution. If $n$ is even, then choose $\beta$ so that $(n, 2^{\alpha-2}) = 2^\beta$. The congruence $x^n \equiv a \pmod{2^\alpha}$ has $2^{\beta+1}$ solutions or no solution according as $a \equiv 1 \pmod{2^{\beta+2}}$ or not.*

*Proof* Since $a$ is odd, we may choose $i$ and $j$ so that $a \equiv (-1)^i 5^j \pmod{2^\alpha}$. As any $x$ for which $x^n \equiv a \pmod{2^\alpha}$ is necessarily odd, we may suppose that $x \equiv (-1)^u 5^v \pmod{2^\alpha}$. The desired congruence then takes the form $(-1)^{nu} 5^{nv} \equiv (-1)^i 5^j \pmod{2^\alpha}$. By Theorem 2.43, this is equivalent to the pair of congruences $nu \equiv i \pmod 2$, $nv \equiv j \pmod{2^{\alpha-2}}$. If $n$ is odd, then by Theorem 2.17 there exists exactly one $u \pmod 2$ for which the first congruence holds, and exactly one $v \pmod{2^{\alpha-2}}$ for which the second congruence holds, and hence there exists precisely one solution $x$ in this case.

Suppose now that $n$ is even. We apply Theorem 2.17 two more times. If $i \equiv 0 \pmod 2$ then the congruence $nu \equiv i \pmod 2$ has two solutions. Otherwise it has none. If $j \equiv 0 \pmod{2^\beta}$ then the congruence $nv \equiv j \pmod{2^{\alpha-2}}$ has exactly $2^\beta$ solutions. Otherwise it has none. Thus the congruence $x^n \equiv a \pmod{2^\alpha}$ has $2^{\beta+1}$ solutions or no solution, according as $a \equiv 5^j \pmod{2^\alpha}$, $j \equiv 0 \pmod{2^\beta}$, or not. From Theorem 2.43 we know

that 5 has order $2^\beta \pmod{2^{\beta+2}}$. Thus by Lemma 2.31, $5^j \equiv 1 \pmod{2^{\beta+2}}$ if and only if $2^\beta | j$. Since $2^{\beta+2} | 2^\alpha$, the condition on $a$ is precisely that $a \equiv 1 \pmod{2^{\beta+2}}$.

**PROBLEMS**

1. Find a primitive root of the prime 3; the prime 5; the prime 7; the prime 11; the prime 13.
2. Find a primitive root of 23.
3. How many primitive roots does the prime 13 have?
4. To what exponents do each of $1, 2, 3, 4, 5, 6$ belong modulo 7? To what exponents do they belong modulo 11?
5. Let $p$ be an odd prime. Prove that $a$ belongs to the exponent 2 modulo $p$ if and only if $a \equiv -1 \pmod{p}$.
6. If $a$ belongs to the exponent $h$ modulo $m$, prove that no two of $a, a^2, a^3, \cdots, a^h$ are congruent modulo $m$.
7. If $p$ is an odd prime, how many solutions are there to $x^{p-1} \equiv 1 \pmod{p}$; to $x^{p-1} \equiv 2 \pmod{p}$?
8. Use Theorem 2.37 to determine how many solutions each of the following congruences has:
   (a) $x^{12} \equiv 16 \pmod{17}$ (b) $x^{48} \equiv 9 \pmod{17}$
   (c) $x^{20} \equiv 13 \pmod{17}$ (d) $x^{11} \equiv 9 \pmod{17}$.
9. Show that $3^8 \equiv -1 \pmod{17}$. Explain why this implies that 3 is a primitive root of 17.
10. Show that the powers of 3 (mod 17) are 3, 9, 10, 13, 5, 15, 11, 16, 14, 8, 7, 4, 12, 2, 6, 1. Use this information to find the solutions of the congruences in Problem 8.
11. Using the data in the preceding problem, decide which of the congruences $x^2 \equiv 1, x^2 \equiv 2, x^2 \equiv 3, \cdots, x^2 \equiv 16 \pmod{17}$, have solutions.
12. Prove that if $p$ is a prime, $(a, p) = 1$ and $(n, p-1) = 1$, then $x^n \equiv a \pmod{p}$ has exactly one solution.
13. Show that the numbers $1^k, 2^k, \cdots, (p-1)^k$ form a reduced residue system (mod $p$) if and only if $(k, p-1) = 1$.
14. Suppose that $a$ has order $h \pmod{p}$, and that $a\bar{a} \equiv 1 \pmod{p}$. Show that $\bar{a}$ also has order $h$. Suppose that $g$ is a primitive root (mod $p$), and that $a \equiv g^i \pmod{p}$, $0 \leqslant i < p-1$. Show that $\bar{a} \equiv g^{p-1-i} \pmod{p}$.
15. Prove that if $a$ belongs to the exponent $h$ modulo a prime $p$, and if $h$ is even, then $a^{h/2} \equiv -1 \pmod{p}$.



16. Let $m$ and $n$ be positive integers. Show that $(2^m - 1, 2^n + 1) = 1$ if $m$ is odd.
17. Show that if $a^k + 1$ is prime, $k > 0$, and $a > 1$ then $k$ is a power of 2. Show that if $p | (a^{2n} + 1)$ then $p = 2$ or $p \equiv 1 \pmod{2^{n+1}}$. (H)
18. Show that if $g$ and $g'$ are primitive roots modulo an odd prime $p$, then $gg'$ is not a primitive root of $p$.
19. Show that if $a^h \equiv 1 \pmod{p}$ then $a^{ph} \equiv 1 \pmod{p^2}$. Show that if $g$ is a primitive root (mod $p^2$) then it is a primitive root (mod $p$).
20. Of the 101 integers in a complete residue system (mod $101^2$) that are $\equiv 2 \pmod{101}$, which one is not a primitive root (mod $101^2$)?
21. Let $g$ be a primitive root of the odd prime $p$. Show that $-g$ is a primitive root, or not, according as $p \equiv 1 \pmod 4$ or $p \equiv 3 \pmod 4$.
22. Let $g$ be a primitive root (mod $p$). Show that $(p-1)! \equiv g \cdot g^2 \cdot \cdots \cdot g^{p-1} \equiv g^{p(p-1)/2} \pmod{p}$. Use this to give another proof of Wilson's congruence (Theorem 2.11).

*23. Prove that if $a$ belongs to the exponent 3 modulo a prime $p$, then $1 + a + a^2 \equiv 0 \pmod{p}$, and $1 + a$ belongs to the exponent 6.

*24. Let $a$ and $n > 1$ be any integers such that $a^{n-1} \equiv 1 \pmod{n}$ but $a^d \not\equiv 1 \pmod{n}$ for every proper divisor $d$ of $n - 1$. Prove that $n$ is a prime.

*25. Show that the number of reduced residues $a \pmod{m}$ such that $a^{m-1} \equiv 1 \pmod{m}$ is exactly $\prod_{p|m} (p - 1, m - 1)$.

*26. (Recall that $m$ is a Carmichael number if $a^{m-1} \equiv 1 \pmod{m}$ for all reduced residues $a \pmod{m}$.) Show that $m$ is a Carmichael number if and only if $m$ is square-free and $(p-1) | (m-1)$ for all primes $p$ dividing $m$. Deduce that $2821 = 7 \cdot 13 \cdot 31$ is a Carmichael number.

*27. Show that $m$ is a Carmichael number if and only if $m$ is composite and $a^m \equiv a \pmod{m}$ for all integers $a$.

*28. Show that the following are equivalent statements concerning the positive integer $n$:
   (i) $n$ is square-free and $(p-1) | n$ for all primes $p$ dividing $n$;
   (ii) If $j$ and $k$ are positive integers such that $j \equiv k \pmod{n}$, then

$$a^j \equiv a^k \pmod{n} \text{ for all integers } a.$$

(The numbers $1, 2, 6, 42, 1806$ have this property, but there are no others. See J. Dyer-Bennet, "A theorem on partitions of the set of positive integers," *Amer. Math. Monthly*, 47 (1940), 152–154.)

*29. Show that the sequence $1^1, 2^2, 3^3, \cdots$, considered (mod $p$) is periodic with least period $p(p-1)$.

*39. Let $m$ be given, and let $s$ be a product of prime powers $q^a$ each having the property described in the preceding problem. Show that if $s > m^{1/2}$ then $m$ is prime.

## 2.9 CONGRUENCES OF DEGREE TWO, PRIME MODULUS

If $f(x) \equiv 0 \pmod p$ is of degree 2, then $f(x) = ax^2 + bx + c$, and $a$ is relatively prime to $p$. We shall suppose $p > 2$ since the case $p = 2$ offers no difficulties. Then $p$ is odd, and $4af(x) = (2ax + b)^2 + 4ac - b^2$. Hence $u$ is a solution of $f(x) \equiv 0 \pmod p$ if and only if $2au + b \equiv v \pmod p$, where $v$ is a solution of $v^2 \equiv b^2 - 4ac \pmod p$. Furthermore, since $(2a, p) = 1$, for each solution $v$ there is one, and only one, $u$ modulo $p$ such that $2au + b \equiv v \pmod p$. Clearly different $v$ modulo $p$ yield different $u$ modulo $p$. Thus the problem of solving the congruence of degree 2 is reduced to that of solving a congruence of the form $v^2 \equiv k \pmod p$. Following some preliminary observations on this congruence, we turn to an algorithm, called RESSOL, for finding its solutions.

If $a \equiv 0 \pmod p$, then this has the sole solution $x \equiv 0 \pmod p$. If $a \not\equiv 0 \pmod p$, then the congruence $x^2 \equiv a \pmod p$ may have no solution, but if $x$ is a solution then $-x$ is also a solution. Since $p$ is odd, $x \not\equiv -x \pmod p$, and thus the congruence has two distinct solutions in this case. It cannot have more than two, by Corollary 2.27.

If $p$ is a small prime then the solutions of the congruence $x^2 \equiv a \pmod p$ may be found by simply trying $x = 0, x = 1, \cdots, x = (p-1)/2$ until one is found. Since this involves $\approx p$ multiplications, for large $p$ it is desirable to have a more efficient procedure. If $p = 2$ then it suffices to take $x = a$. Thus we may suppose that $p > 2$. By Euler's criterion we may suppose that $a^{(p-1)/2} \equiv 1 \pmod p$, for otherwise the congruence has no solution.

Suppose first that $p \equiv 3 \pmod 4$. In this case we can verify that $x \equiv \pm a^{(p+1)/4}$ are the solutions, for

$$(\pm a^{(p+1)/4})^2 = a^{(p+1)/2} = a \cdot a^{(p-1)/2} \equiv a \pmod p.$$

Note that it is not necessary to verify in advance that $a^{(p-1)/2} \equiv 1 \pmod p$. It suffices to calculate $x \equiv a^{(p+1)/4} \pmod p$. If $x^2 \equiv a \pmod p$, then the solutions are $\pm x$. Otherwise $x^2 \equiv -a \pmod p$, and we can conclude that $a$ is a quadratic nonresidue. Thus $x = \pm a^{(p+1)/4}$ are the solutions, if the congruence has a solution. This takes care of roughly half the primes. As always with large exponents, the value of $a^{(p+1)/4} \pmod p$ is determined using the repeated squaring device discussed in Section 2.4. Hence the number of congruential multiplications required is only of the order of magnitude $\log p$.



Suppose now that $p \equiv 1 \pmod 4$. We have already considered the special case $x^2 \equiv -1 \pmod p$, and in proving Theorem 2.12 we gave a formula for the solutions, namely $x = \pm((p-1)/2)!$. However, this formula is useless for large $p$, as it involves $\approx p$ multiplications. On the other hand, if a quadratic nonresidue $z$ is known then we may take $x \equiv \pm z^{(p-1)/4} \pmod p$, since then $x^2 \equiv z^{(p-1)/2} \equiv -1 \pmod p$ by Euler's criterion. Thus in this special case it suffices to find a quadratic nonresidue. We can try small numbers in turn, or use a random number generator to provide "random" residue classes. In either case, since half the reduced residues are quadratic nonresidues, we may expect that the average number of trials is 2. (Here our interest is not in a deterministic algorithm of proven efficiency, but rather a calculational procedure that is quick in practice.)

We now develop these ideas to find the roots of the congruence $x^2 \equiv a \pmod p$ for arbitrary $a$ and $p$. We begin with a few general observations. Let $a$ and $b$ be relatively prime to $m$, and suppose that $a$ and $b$ both have order $h \pmod m$. Then $(ab)^h \equiv 1 \pmod m$, and hence the order of $ab$ is a divisor of $h$. In general nothing more can be said. It may be that $b$ is the inverse of $a$, so that $ab \equiv 1 \pmod m$, in which case the order of $ab$ is 1. On the other hand, the order of $ab$ may be as large as $h$. (Consider $3 \pmod{11}$, $5 \pmod{11}$, and $3 \cdot 5 \equiv 4 \pmod{11}$. All three of these numbers have order 5.) Nevertheless there is one particular situation in which a little more can be established.

**Theorem 2.45** *If $a$ and $b$ are relatively prime to a prime number $p$, and if $a$ and $b$ both have order $2^j \pmod p$ with $j > 0$, then $ab$ has order $2^{j'} \pmod p$ for some $j' < j$.*

*Proof* Since $a$ has order $2^j \pmod p$, it follows that $2^j \mid (p-1)$, and thus $p > 2$. Put $x = a^{2^{j-1}}$. Then $x \not\equiv 1 \pmod p$ but $x^2 = a^{2^j} \equiv 1 \pmod p$. Thus by Lemma 2.10 it follows that $x \equiv -1 \pmod p$. Similarly, $b^{2^{j-1}} \equiv -1 \pmod p$, and it follows that

$$(ab)^{2^{j-1}} = a^{2^{j-1}} b^{2^{j-1}} \equiv (-1)(-1) \equiv 1 \pmod p.$$

From this and Lemma 2.31 we deduce that the order of $ab$ is a divisor of $2^{j-1}$, that is, the order of $ab$ is $2^{j'}$ for some $j' < j$.

Neither Theorem 2.45 nor its proof involves primitive roots, but some further insight can be obtained by interpreting the situation in terms of powers of a given primitive root $g$. Write $a \equiv g^\alpha \pmod p$, where $0 \leqslant \alpha < p-1$. By Lemma 2.33, the order of $g^\alpha$ is $(p-1)/(p-1,\alpha)$. Write $p-1 = m2^k$ with $m$ odd. The hypothesis that $a$ has order $2^j$ is thus equivalent to the relation $(p-1,\alpha) = m2^{k-j}$. That is, $\alpha = \alpha_1 m2^{k-j}$ with $\alpha_1$ odd. Similarly, $b \equiv g^\beta \pmod p$ with $\beta = \beta_1 m2^{k-j}$, $\beta_1$ odd. But then $ab \equiv g^{\alpha+\beta} \pmod p$, and $\alpha + \beta = (\alpha_1 + \beta_1)m2^{k-j}$. Since $\alpha_1$ and $\beta_1$ are both odd, it follows that $\alpha_1 + \beta_1$ is even. Choose $i$ so that $(\alpha_1 + \beta_1, 2^i) = 2^i$. Since $j > 0$ by hypothesis, it follows that $i > 0$. Moreover, the order of $ab$ is $2^{j-i}$, so we have $j' = j - i < j$.

With these tools in hand, we describe the algorithm RESSOL (for RESidue SOLver), which locates $x$ such that $x^2 \equiv a \pmod p$. We begin by determining the power of 2 in $p-1$. Thus we find $k$ and $m$ with $m$ odd, so that $p - 1 = 2^k m$. We are supposing that $p > 2$, so that $k > 0$. Set $r \equiv a^{(m+1)/2} \pmod p$ and $n \equiv a^m \pmod p$. We note that

$$r^2 \equiv an \pmod p. \tag{2.13}$$

If $n \equiv 1 \pmod p$, then it suffices to take $x \equiv \pm r \pmod p$. If $n \not\equiv 1 \pmod p$, then we find a quadratic nonresidue $z$, and put $c \equiv z^m \pmod p$. We note that

$$c^{2^k} = z^{2^k m} = z^{p-1} \equiv 1 \pmod p.$$

Thus the order of $c$ is a divisor of $2^k$. Moreover,

$$c^{2^{k-1}} = z^{2^{k-1}m} = z^{(p-1)/2} \equiv -1 \pmod p$$

since $z$ is a quadratic nonresidue. Thus the order of $c$ is exactly $2^k$. Similarly,

$$n^{2^k} = a^{2^k m} = a^{p-1} \equiv 1 \pmod p,$$

so that the order of $n$ divides $2^k$. By repeatedly squaring $n$ we determine the exact order of $n$, say $2^{k'}$. Since

$$n^{2^{k-1}} = a^{2^{k-1}m} = a^{(p-1)/2},$$

we see that $a$ is a quadratic residue (mod $p$) if and only if

$$n^{2^{k-1}} \equiv 1 \pmod p,$$



which in turn is equivalent to the inequality $k' < k$. It is worth checking that this inequality holds, for otherwise $k' = k$, $a$ is a quadratic nonresidue and the proposed congruence has no solution. At this point of the algorithm, we begin a loop. Set $b \equiv c^{2^{k-k'-1}} \pmod p$. We put $r' \equiv br \pmod p$, $c' \equiv b^2 \pmod p$, $n' \equiv c'n \pmod p$. By multiplying both sides of (2.13) by $b^2$ we find that

$$r'^2 \equiv an' \pmod p. \tag{2.14}$$

The point of this construction is that $c'$ has order exactly $2^{k'}$. Since $n \not\equiv 1 \pmod p$ in the present case, it follows that $k' > 0$. Thus by Theorem 2.45, the order of $n' \equiv c'n$ is $2^{k''}$ where $k'' < k'$. (We determine the value of $k''$ by repeated squaring.) If $k'' = 0$, then $n' \equiv 1 \pmod p$, and we see from (2.14) that it suffices to take $x \equiv \pm r' \pmod p$. If $n' \not\equiv 1 \pmod p$, then $k'' > 1$, and the situation is the same as when the loop began, except that the numbers $c$ (of order $2^k$) and $n$ (of order $2^{k'}$) with $0 < k' < k$ have been replaced by $c'$ (of order $2^{k'}$) and $n'$ (of order $2^{k''}$) with $0 < k'' < k'$, while $r$ has been replaced by $r'$ and (2.13) has been replaced by (2.14). Since $k'' < k'$, some progress has been made. By executing this loop repeatedly, we eventually arrive at a set of these variables for which $n \equiv 1 \pmod p$, and then $x \equiv \pm r \pmod p$ is the desired solution.

As a numerical example of this algorithm, suppose we wish to find the roots of the congruence $x^2 \equiv 43 \pmod{97}$. Thus $p = 97$, and $p - 1 = 2^5 \cdot 3$. By using the method described in Section 2.4, we find that $r \equiv 43^{(3+1)/2} \equiv 6 \pmod{97}$, and that $n \equiv 43^3 \equiv 64 \pmod{97}$. Thus the congruence (2.13) is $6^2 \equiv 43 \cdot 64 \pmod{97}$. Since $n \not\equiv 1 \pmod{97}$, we must find a quadratic nonresidue. We note that $(p-1)/2 = 48$, and calculate that $2^{48} \equiv 1 \pmod{97}$. Thus 2 is a quadratic residue, by Euler's criterion. Similarly 3 is a quadratic residue, but 5 is a quadratic nonresidue. We set $z = 5$, $c \equiv 5^3 \equiv 28 \pmod{97}$. Thus $c$ has order $2^5 \pmod{97}$. By repeatedly squaring, we discover that $n$ has order $2^3 \pmod{97}$. That is, $k' = 3$, and we now begin the loop. Since $k - k' - 1 = 1$, we set $b \equiv c^2 \equiv 8 \pmod{97}$, and $c' \equiv b^2 \equiv 64 \pmod{97}$. On multiplying both sides of (2.13) by $b^2$ we obtain the congruence (2.14) with $r' \equiv 8 \cdot 6 \equiv 48 \pmod{97}$ and $n' \equiv 64 \cdot 64 \equiv 22 \pmod{97}$. That is, $48^2 \equiv 43 \cdot 22 \pmod{97}$. By repeated squaring, we discover that 22 has order $2^2 \pmod{97}$, so we take $k'' = 2$, and we are ready to begin the loop over. With the new values of the parameters, we now have $k - k' - 1 = 0$, so we set $b \equiv c \equiv 64 \pmod{97}$, $c' \equiv 64^2 \equiv 22 \pmod{97}$, and obtain the congruence $65^2 \equiv (64 \cdot 48)^2 \equiv 43 \cdot (22 \cdot 22)^2 \equiv 43 \cdot 96 \pmod{97}$. That is, $r' \equiv 65$, $n' \equiv 96 \pmod{97}$. Here 96 has order 2, so that $k'' = 1$. Since $n' \not\equiv 1 \pmod{97}$, we must execute the loop a third time. As $k - k' - 1 = 0$, we set $b \equiv c \equiv 22 \pmod{97}$, $c' \equiv b^2 \equiv 96 \pmod{97}$, and we obtain the congruence $72^2 \equiv (22 \cdot 65)^2 \equiv 43 \cdot (96 \cdot$

96) ≡ 43 (mod 97). Thus the solutions are $x \equiv \pm 72 \pmod{97}$. This example of the algorithm is unusually long because $p - 1$ is divisible by a high power of 2.

To gain further insight into this algorithm, let $g$ be a primitive root (mod $p$). Then $z \equiv g^n \pmod{p}$ for some $n$, and hence $c \equiv z^m \equiv g^{mn} \pmod{p}$. But $n$ is odd since $z$ is a quadratic nonresidue, and thus $(mn, p-1) = m$. Consequently by Lemma 2.33 the order of $c$ is $2^k$. In general, the order of $g^t$ is a power of 2 if and only if $m|t$. There are precisely $2^k$ such residue classes, namely $g^m, g^{2m}, g^{3m}, \cdots, g^{2^k m}$. On the other hand, the $2^k$ residue classes $c, c^2, c^3, \cdots, c^{2^k}$ are distinct, and each one has order a power of 2, so this latter sequence is simply a permutation of the former one. Thus the order of a residue class is a power of 2 if and only if it is a power of $c$. But $n \equiv a^m \pmod{p}$ has order that is a power of 2, and hence there is a non-negative integer $u$ such that $n \equiv c^u \pmod{p}$. A number $c^t$ is a quadratic residue or nonresidue according as $t$ is even or odd. Hence if $a$ is a quadratic residue, then $u$ is even, and the solutions sought are $x \equiv \pm c^{u/2} \pmod{p}$. Thus it suffices to determine the value of $u \pmod{2^k}$. As it stands, the algorithm does not do this, but it can be slightly modified to yield $u$. (See Problem 5 below.) If $n \not\equiv 1 \pmod{p}$, then $u \not\equiv 0 \pmod{2^k}$. Suppose that $0 < u < 2^k$. If the order of $n$ is $2^{k'}$ then $2^{k-k'} | u$ but $2^{k-k'+1} \nmid u$. Thus we obtain some information concerning the binary expansion of $u$. Repeated iterations of the loop (suitably modified) determine further coefficients in the binary expansion of $u$, and eventually $u$ is determined. Alternatively, the value of $u$ could be determined by calculating the successive powers of $c$ until $n$ is encountered, but that might require as many as $2^k$ multiplications. The algorithm given is much faster, as the loop is executed at most $k$ times.

## PROBLEMS

1. Reduce the following congruences to the form $(x - r)^2 \equiv k \pmod{p}$:
   (a) $4x^2 + 2x + 1 \equiv 0 \pmod{5}$; (b) $3x^2 - x + 5 \equiv 0 \pmod{7}$;
   (c) $2x^2 + 7x - 10 \equiv 0 \pmod{11}$; (d) $x^2 + x - 1 \equiv 0 \pmod{13}$.
2. Suppose that $f(x) = ax^2 + bx + c$, and that $D = b^2 - 4ac$. Show that if $p$ is an odd prime, $p \nmid a, p|D$, then $f(x) \equiv 0 \pmod{p}$ has exactly one solution. Show that if $p$ is an odd prime, $p \nmid a, p \nmid D$, then the congruence $f(x) \equiv 0 \pmod{p}$ has either 0 or 2 solutions, and that if $x$ is a solution then $f'(x) \not\equiv 0 \pmod{p}$.
*3. Let $f(x) = ax^2 + bx + c$, and let $p$ be an odd prime that does not divide all the coefficients $a, b, c$. Show that the congruence $f(x) \equiv 0 \pmod{p^2}$ has either 0, 1, 2, or $p$ solutions.

Head of the Department of Mathematics,
Mother Terasa College of Arts & Science,
Mettusalai, Illuppur, Pudukkottai Dt-622102.