PREFACE

These notes have been prepared by Dr Mike Canfell (with minor changes and extensions by Dr Gerd Schmalz) for use by the external students in the unit PMTH 338 Number Theory. This booklet covers all examinable material associated with the unit, and the lectures for internal students will be based on it. You should however be prepared to supplement these notes by reading and working from the appropriate parts of the text book as indicated in the course outline.

Dr Gerd Schmalz
Lecturer and Unit Coordinator
PMTH 338 Number Theory
School of Mathematics, Statistics and Computer Science
University of New England
Armidale NSW 2351

Printed at the University of New England
December, 2007

Contents

Chapter 1

Numbers

1.1 Algebraic properties of numbers

The main objective of this course are the integers. The set of integers consists of $0, \pm 1, \pm 2, \ldots$, and will be denoted by $\mathbb{Z}$. By $\mathbb{N}$ we denote the subset of non-negative integers $0, 1, 2, \ldots$, which are also known as natural numbers. For rational, real, or complex numbers the standard notations $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ will be used.

We recall some arithmetic properties of integers. They can be rigorously derived from Peano's axioms (see Appendix A).

1. There is a binary operation, called addition, that assigns to a pair of summands $a, b \in \mathbb{Z}$ the sum $a + b \in \mathbb{Z}$.

2. The addition is associative, i.e. $(a + b) + c = a + (b + c)$ for all $a, b, c \in \mathbb{Z}$.

3. There is a unique number, called zero (0), such that $a + 0 = 0 + a = a$ for any $a \in \mathbb{Z}$. Zero is the neutral element with respect to addition.

4. For any $a \in \mathbb{Z}$ there is a unique number b, such that $a + b = 0$. We write $b = -a$.

5. The addition is commutative, i.e. $a + b = b + a$ for all $a, b \in \mathbb{Z}$.

6. There is a second binary operation, called multiplication, that assigns to a pair of factors $a, b \in \mathbb{Z}$ the product $ab \in \mathbb{Z}$.

7. The multiplication is associative, i.e. $(ab)c = a(bc)$ for all $a, b, c \in \mathbb{Z}$.

8. There is a unique number, called one (1), such that $a \cdot 1 = 1 \cdot a = a$ for any $a \in \mathbb{Z}$. One is the neutral element with respect to multiplication.

9. The multiplication is commutative, i.e. $ab = ba$ for all $a, b \in \mathbb{Z}$.

10. For any $a, b, c \in \mathbb{Z}$ we have $(a + b)c = ac + bc$ (distributivity).

A set with one binary operation, such that properties 1.-5. hold is called *Abelian group*. A set with two binary operations, such that the arithmetic properties 1.-10. hold, is called a (commutative) ring. Thus, the set of integers is an Abelian group with respect to the addition and a commutative ring with respect to addition and multiplication. In this course we will encounter several groups and rings. We will see that these notions help to make statements about numbers more transparent and to simplify the proofs. This more general concept will provide results which are applicable in more general situations and it leads to a deeper understanding of the subject.

The rationals, real and complex numbers are commutative rings. After excluding zero (the neutral element with respect to the first operation) these sets form Abelian groups with respect to the second operation. This makes $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ into *fields* which are known from Linear algebra.

Example. The set $\mathcal{P}$ of polynomials $p = \sum_{k=0}^{n} a_k x^k$ with coefficients a_k in $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ or any other commutative ring will form a commutative ring with respect to the usual addition and multiplication of polynomials.

Example. $n \times n$- matrices with coefficients in a field (such as $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$) form a *non-commutative* ring. However, in this course we will only deal with commutative rings.

Another important arithmetic property of the integers that does not hold automatically is the so-called cancellation property:

If $ab = ac$ and $a \neq 0$ then $b = c$. If a had a multiplicative inverse a^{-1} this would just follow from multiplication by a^{-1} at both sides. We can derive the cancellation property for integers from the fact that an integer $a \neq 0$ has a multiplicative inverse in the bigger ring $\mathbb{Q}$ of rationals. Since the cancellation property holds for the rationals it must hold for the subset of integers.

We note that the cancellation property is equivalent to the statement: If $ab = 0$ then either $a = 0$ or $b = 0$. In fact, $ab = ac$ is equivalent to $a(b - c) = 0$ (we used the law of distributivity). Suppose $a \neq 0$. Then $b - c = 0$ is equivalent to $b = c$. In Chapter 2 we will encounter commutative rings in which the cancellation property does not hold.

1.2 Ordering of numbers

Before we introduce the concept of ordering in the set of integers we will take a much more general point of view.

We say a set M is equipped with an ordering by indicating pairs (a, b) of elements of M for which the relation $a < b$ holds. Thus an ordering is (similar to equivalence relations, which you might have encountered in earlier courses and will encounter again in Chapter 2) a binary relation. We can view a binary relation as the subset $\mathcal{O} \subset M \times M$ of the set of all pairs of elements of M, namely

$$\mathcal{O} = \{(a, b) \in M \times M : a < b\}.$$

The word *binary* says that two elements are involved in the relation. Of course not any binary relation satisfy our expectations of an ordering. In fact, only binary relations that

satisfy the following postulates (O1)-(O3) will be called a *partial ordering*.

$$a \not< a \quad \text{No element is (strongly) less than itself} \tag{O1}$$
$$a < b \Longrightarrow b \not< a \quad \text{antisymmetry} \tag{O2}$$
$$a < b \text{ and } b < c \Longrightarrow a < c \quad \text{transitivity} \tag{O3}$$

The main example is the following ordering for integers.

Definition 1.1. *An integer a is said to be less than an integer b (and b is said to be bigger than a) if there exists a positive integer c such that $a + c = b$. We write $a < b$ or $b > a$. If $a < b$ or $a = b$ we write $a \leq b$. If $a > b$ or $a = b$ we write $a \geq b$.*

The postulates (O1)-(O3) hold also in the following situation: Let M be the set of all subsets of $\mathbb{Z}$. Then we say $A \subset B$ if $A \subseteq M$ is a (proper) subset of $B \subseteq M$. Since there are subsets A and B such that neither of them is contained in the other, there are elements of M that cannot be compared with each other. To avoid this we add one more postulate

$$\forall a, b \in M \text{ either } a < b, \text{ or } b < a, \text{ or } a = b \tag{O4}$$

(O4) is obviously satisfied for integers since $c = b - a$ must be either positive, or negative or 0.

We will need to use a further principle which holds for the natural numbers (but not for the integers!).

We say that an element s_0 of a set of integers $S \subset \mathbb{Z}$ is called *smallest element* if $s \geq s_0$ for any $s \in S$. We postulate

The Well-Ordering Property (WOP): Every nonempty subset S of $\mathbb{N}$ has a smallest element s_0.

Though it seems to be intuitively clear that WOP holds for the ordering in $\mathbb{N}$, a formal proof requires *mathematical induction* and is actually equivalent to either of the following versions of mathematical induction (for a formal proof of this statement see Appendix B):

Principle of mathematical induction (FMI): If a subset S of $\mathbb{N}$ contains 0 and has the property that for any element a which belongs to S also $a + 1$ belongs to S then S must coincide with $\mathbb{N}$.

Second principle of mathematical induction (SMI): If a subset S of $\mathbb{N}$ contains 0 and has the property that $a + 1$ must belong to S if it was known that the consecutive elements $0, 1, \ldots, a$ belong to S then S in fact coincides with $\mathbb{N}$.

Remarks.

1. WOP does not hold neither for the set of all rational positive numbers nor for the set of real positive numbers. For whatever positive rational number x we choose, the number $x/2$ is a smaller positive rational number. Similarly, for whatever positive real number x we choose, the number $x/2$ is a smaller positive real number. The WOP is a distinctive property of $\mathbb{Z}$.

2. In using the WOP on a set S of non-negative integers, you will notice that we are always careful to check that S is nonempty. Notice that the Well-Ordering principle is equivalent to the statement: Any set S of non-negative integers that has no smallest element must be empty.

1.3 Divisibility

The proof of the following result illustrates a typical use of the Well-Ordering Property.

Theorem 1.2. (The Division Algorithm) *For each pair a, b with $a, b \in \mathbb{Z}$ and $b > 0$, there exists exactly one pair of integers q, r such that $a = qb + r$ and $0 \leq r < b$.*

Proof. We consider the set of non-negative integers of the form $a - nb$;

$$S = \{a - nb : n \in \mathbb{Z} \text{ and } a - nb \geq 0\}.$$

If $a \geq 0$, S contains the element $a = a - 0b$ while if $a < 0$, S contains the element $a - ab = -a(b - 1) \geq 0$. In either case S is nonempty. By the Well-Ordering Property, S has a smallest element which we denote by $r = a - qb$. By construction, $r \geq 0$ and $a = qb + r$.

Next, we show that $r < b$. Suppose on the contrary that $r \geq b$. Then

$$r - b = a - qb - b = a - (q + 1)b \geq 0.$$

Hence $r - b \in S$. But $r - b < r$ and this contradicts the choice of r as the smallest element of S. We conclude that $r \geq b$ is impossible and hence that $r < b$.

For uniqueness, suppose also that $a = q'b + r'$ with $0 \leq r' < b$. Then

$$\begin{aligned} 0 = a - a &= (q - q')b + (r - r') \\ r - r' &= (q' - q)b \end{aligned} \tag{1.1}$$

In order to show that $q = q'$, we eliminate the possibilities $q' > q$ and $q' < q$. Suppose that $q' > q$. Then from equation (1.1), $r - r' \geq b$ and so $r \geq b + r' \geq b$ which is impossible by the choice of r. Similarly we cannot have $q > q'$, so this leaves only $q = q'$. Finally, again from (1.1), we have $r = r'$. □

Remark. This is just the process of 'dividing a by b to get quotient q and remainder r', which we are familiar with from primary school work. Similar results are true in rings other than the ring of integers. The type of proof above carries over to those less familiar theorems.

Example 1.3. *(i) If $a = 77, b = 13$, we find that $q = 5, r = 12$ and $77 = 5 \cdot 13 + 12$.*

(ii) If $a = -77, b = 13$, we find that $q = -6, r = 1$ and $-77 = (-6) \cdot 13 + 1$.

When large numbers are involved it is convenient to use a calculator. We can find q by using the division button and dividing a by b.

Example 1.4. *Find q and r if $a = 1674823$ and $b = 23555$.*

Solution. Using a calculator we find that to 2 decimal places, $a/b = 71.10$. Dropping the terms after the decimal point we have $q = 71$. We calculate r from $r = a - qb$. Thus $r = 1674823 - 71 \cdot 23555 = 2418$. Hence $1674823 = 71 \cdot 23555 + 2418$.

Note that since the r above satisfies $0 \leq r < b$, it confirms (by the uniqueness part of the division algorithm) that we have the right q.

Since practically all of our work involves integers, it is convenient to make the following standing assumption.

Unless otherwise stated, lowercase letters in mathematics italics $a, b, c, \ldots$ denote integers.

Definition 1.5. *An integer $a \neq 0$ is a divisor of the integer b, (written $a \mid b$), if $b = ac$ for some integer c. Also $a \nmid b$ means that a does not divide b.*

Thus $a \mid b$ means that the remainder is 0 when b is divided by a.

If $a \mid b$ and $a \mid c$ then a is called a **common divisor** of b and c.

Definition 1.6. *If a and b are not both 0, the* **greatest common divisor** *(gcd) of a and b is the largest positive integer which divides both a and b.*

Notation. We use $\gcd(a, b)$ or simply (a, b) to denote the greatest common divisor of a and b.

Below we will prove that the gcd of two integers that are not both 0 exists and is unique.

Example 1.7. *The common divisors of 12 and 18 are $\pm 1, \pm 2, \pm 3, \pm 6$, and the greatest common divisor is 6.*

The gcd of a finite number of integers is defined similarly.

Theorem 1.8 (Euclid's Algorithm)**.** *If a and b are not both 0, they have a unique greatest common divisor d which can be written in the form $d = xa + yb$ for some integers x and y.*

Proof. If either (but not both) of the numbers a or b is zero, the result is obvious, so we assume that both are non-zero. Since $\gcd(a, b) = \gcd(-a, b)$, we can, if necessary, replace a or b by its negative. Hence we assume that both are positive integers, and by changing notation if necessary, we can assume that $a \geq b > 0$.

We construct a finite sequence of numbers $\{r_i\}$ by repeated use of the division algorithm. Begin by putting $r_0 = a$ and $r_1 = b$. The next term r_2 is the remainder obtained by dividing a by b.

$$r_0 = q_1 r_1 + r_2 \qquad 0 \le r_2 < r_1 \quad (1)$$

If $r_2 \neq 0$ we divide r_1 by r_2 to get

$$r_1 = q_2 r_2 + r_3 \qquad 0 \le r_3 < r_2 \quad (2)$$

We repeat the procedure so long as the remainder is non-zero to get

$$\begin{array}{ccccc} r_2 & = & q_3 r_3 + r_4 & & 0 \le r_4 < r_3 \quad (3) \\ \vdots & & \vdots & & \vdots \end{array}$$

But since $r_1 > r_2 > r_3 > r_4 > \cdots$ and each r_i is a non-negative integer, the process stops after a finite number of steps. Hence there is an integer n such that

$$\begin{array}{lcll} r_{n-3} & = & q_{n-2} r_{n-2} + r_{n-1} & 0 \le r_{n-1} < r_{n-2} \quad (4) \\ r_{n-2} & = & q_{n-1} r_{n-1} + r_n & 0 \le r_n < r_{n-1} \quad (5) \\ r_{n-1} & = & q_n r_n & (6) \end{array}$$

Define $d = r_n$ (the last non-zero remainder). By (6), $r_n \mid r_{n-1}$. From (5) we then see that $r_n \mid r_{n-2}$. Since $r_{n-3} = q_{n-2} r_{n-2} + r_{n-1}$ and $r_n \mid r_{n-1}, r_n \mid r_{n-2}$, we see that $r_n \mid r_{n-3}$. Working backwards through the set of equations, we eventually find that $r_n \mid r_1$ and $r_n \mid r_0$, that is, d is a common divisor of both a and b.

We next show that $d = xa + yb$ for some integers x and y. We do this by again working backwards through the set of equations. From (5) we see that $d = r_{n-2} - q_{n-1} r_{n-1}$ Now from (4) we have $r_{n-1} = r_{n-3} - q_{n-2} r_{n-2}$ and if we substitute for r_{n-1} in $d = r_{n-2} - q_{n-1} r_{n-1}$ and gather like terms, we find that d is a linear combination of r_{n-3} and r_{n-2}. Continuing backwards through the set of equations we eventually find that d is a linear combination of r_0 and r_1. That is,

$$d = xa + yb$$

for some integers x and y.

We need to show that d is the greatest common divisor of a and b. We know that $d \ge 1$. Suppose e is a positive common divisor of a and b, that is, $e \mid a$ and $e \mid b$. Since $d = xa + yb$ we see also that $e \mid d$. It follows that $e \le d$ and so d is the greatest common divisor of a and b.

Finally we need to show that $\gcd(a, b)$ is unique, i.e. that any other possible $\gcd(a, b) = d'$ must coincide with d. From $d = xa + yb$ and $d' \mid a$, $d' \mid b$ we get $d' \mid d$, i.e. $d = md'$ for some positive integer m. Then either $m > 1$ and $d > d'$, i.e. d' is not the gcd or $m = 1$ and $d = d'$ as required. □

Notice, that the numbers x, y in the representation $d = xa + yb$ are not unique! In fact, $d = xa + yb = (x - b)a + (y + a)b$.

The proof of the Theorem tells us how to compute the gcd for two given numbers and also how to write d in the form $xa + yb$. The process is illustrated in the following examples.

Example 1.9. *Find* $\gcd(91, 17)$ *and write it in the form* $91x + 17y$.

Solution.

$$\begin{aligned} 91 &= 5 \cdot 17 + 6 \\ 17 &= 2 \cdot 6 + 5 \\ 6 &= 5 + 1 \\ 5 &= 5 \cdot 1. \end{aligned}$$

Hence $\gcd(91, 17) = 1$.

Working backwards we find

$$\begin{aligned} 1 &= 6 - 5 \\ &= 6 - (17 - 2 \cdot 6) \\ &= -17 + 3 \cdot 6 \\ &= -17 + 3(91 - 5 \cdot 17) \\ &= 3 \cdot 91 - 16 \cdot 17 \end{aligned}$$

which is of the required form with $x = 3$ and $y = -16$.

Example 1.10. *Find* $\gcd(225, 85)$ *and write it in the form* $225x + 85y$.

Solution.

$$\begin{aligned} 225 &= 2 \cdot 85 + 55 \\ 85 &= 55 + 30 \\ 55 &= 30 + 25 \\ 30 &= 25 + 5 \\ 25 &= 5 \cdot 5. \end{aligned}$$

Hence $\gcd(225, 85) = 5$. Also

$$\begin{aligned} 5 &= 30 - 25 \\ &= 30 - (55 - 30) \\ &= 2 \cdot 30 - 55 \\ &= 2(85 - 55) - 55 \\ &= 2 \cdot 85 - 3 \cdot 55 \\ &= 2 \cdot 85 - 3(225 - 2 \cdot 85) \\ &= 8 \cdot 85 - 3 \cdot 225 \\ &= 225 \cdot (-3) + 85 \cdot 8. \end{aligned}$$

Hence we can take $x = -3$ and $y = 8$.

Example 1.11. *Find* $\gcd(-225, 85)$ *and write it in the form* $-225x + 85y$.

Solution. $(-225, 85) = (225, 85) = 5$. Also

$$5 = 225 \cdot (-3) + 85 \cdot 8 = -225 \cdot 3 + 85 \cdot 8.$$

Example 1.12. *Find* $\gcd(-225, 0)$ *and write it in the form* $-225x + 0y$.

Solution. $(-225, 0) = 225$ and $225 = (-225) \cdot (-1) + 0 \cdot 0$.

1.4 Prime Numbers

An element u of a ring R with 1 is called a **unit** if $uv = 1$ for some element v of the ring. In other words, units are the elements that have an inverse with respect to multiplication.

Proposition 1.13. *The only units of the ring* $\mathbb{Z}$ *are 1 and* -1.

Proof. $uv = 1$ implies $|u||v| = 1$ and therefore $|u| = |v| = 1$. Since $1 \cdot 1 = 1$ and $(-1) \cdot (-1) = 1$ the numbers ± 1 are the only units of $\mathbb{Z}$. □

The fact that only two elements in $\mathbb{Z}$ have an inverse with respect to multiplication makes division in $\mathbb{Z}$ difficult but at the same time an interesting subject of study. A central notion in this study is the notion of a prime which will be introduced now.

Definition 1.14. *We say that* a *is a* **proper divisor** *of* b *if* $b = ac$ *where neither* a *nor* c *is a unit.*

A positive integer p *is called a* **prime number** *if* $p \neq 1$ *and* p *has no proper divisors.*

A number $n \geq 4$ *is* **composite** *if it is not prime.*

Thus p is a prime if $p \geq 2$ and whenever $p = ab$ with $a, b \in \mathbb{Z}, a > 0, b > 0$, we must have either $a = 1$ or $b = 1$.

We can construct lists of primes using the method known as the Sieve of Eratosthenes. For example, to find all the prime numbers less than 20, begin by writing down all the integers from 2 to 20.

	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20

Now cross out all the multiples of 2 (but not 2 itself). Next, cross out the multiples of 3 (but not 3 itself). The numbers remaining are

$$2, 3, 5, 7, 11, 13, 17, 19.$$

These are the primes, in order, from 2 to 20.

Note that we did not have to go any further than multiples of 3. This is because if a number n is composite, say $n = ab$, then at least one of a and b must be less than or equal to $\sqrt{n}$. Any prime factor of n will be a prime factor of either a or b, so this n will have already been crossed out. (Actually, we are making use of a later result, the Corollary to Theorem 1.19, in saying this.)

Using this list of primes up to 19, we could if we wished, find all the primes less than 400. We would start with all the numbers from 2 to 400, cross out all the multiples of 2, 3, 5, 7, 11, 13, 17 and 19 (but not these numbers themselves) and the numbers remaining would be all prime.

There are a number of ways to make the process more efficient. We could just write down, say, the numbers from 21 to 400, omitting the even numbers and those ending in the digit 5. This gives 72 numbers to write down. In crossing out the multiples of 3, we can use the result (to be proved later), that if the sum of the digits of a number is divisible by 3, then the number itself is divisible by 3. Thus 171 gets crossed out since the sum of the digits $1 + 7 + 1 = 9$ is divisible by 3. In crossing out the multiples of 7, note that the first multiple of 7 which has not already been crossed out is 49, and similarly when it comes to crossing out multiples of 19, we start at $19^2 = 361$.

OPEN QUESTION (PRIMES) How plentiful are the primes and how are they distributed in relation to the non-primes? What are the properties of the function $\pi(n)$? Given an integer n, what is the best way to find the next prime after n?

Remark. Occasionally we will state an 'open question'. Often, the answer to the stated question is not fully known, and research into that question continues to this day. The purpose of stating such questions is to provide a guide for some of our subsequent work. You may like to keep track of how successful we are in answering these questions during the course.

We introduce a function which is useful in the description of the distribution of primes.

Definition 1.15. *$\pi(n)$ is the number of prime numbers less than or equal to n.*

Thus $\pi(20) = 8$ (the number of primes in our list above). Below we state (without proof) one famous result about $\pi(n)$ which is useful to know about:

The Prime Number Theorem.

$$\lim_{n\to\infty} \frac{\pi(n)}{n/\ln n} = 1.$$

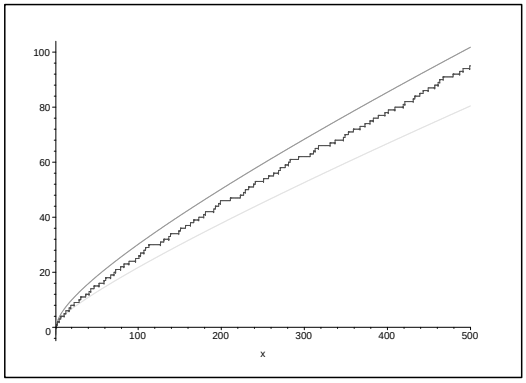
This says that there are approximately $\frac{n}{\ln n}$ primes which are less than or equal to n, that is,

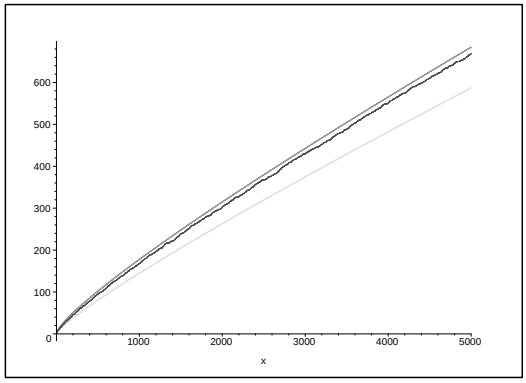
$$\pi(n) \sim \frac{n}{\ln n} \text{ as } n \to \infty.$$

Another approximation of π is given by the function $\mathrm{Li}(x) = \int_0^x \frac{dt}{\ln t}$. This (divergent) improper integral has to be interpreted as the limit

$$\lim_{\delta\to 0} \int_0^{1-\delta} \frac{dt}{\ln t} + \int_{1+\delta}^{x} \frac{dt}{\ln t}.$$

According to an (still unproved) conjecture Li gives a much better approximation of π. The picture below shows $\mathrm{Li}(x)$ (upper curve), π (middle curve) and $\frac{x}{\ln x}$ (lower curve) for $x \leq 500$ and for $x \leq 5000$.





Example 1.16. *Use the Prime Number Theorem to estimate the number of primes less than one million.*

Solution. Let $n = 10^6$. Then $\pi(10^6) \approx \frac{10^6}{\ln 10^6} \approx 72\ 382$. $\mathrm{Li}(10^6) \approx 78\ 628$.

The exact value of $\pi(10^6)$ is 78 498.

Example 1.17. *Suppose a computer is programmed to examine each of the primes from 1 to 10^{20} to see if it has a certain property. If it deals with one billion primes per second, approximately how long will it take to complete the search?*

Solution. By the Prime Number Theorem, there are approximately $\frac{10^{20}}{\ln 10^{20}} = \frac{10^{20}}{20 \ln 10}$ primes to examine. The time taken is $\sim 10^{-9} \frac{10^{20}}{20 \ln 10}$ seconds which is approximately

$$\frac{10^{10}}{2 \ln 10} \cdot \frac{1}{3600 \cdot 24 \cdot 365.25} \sim \text{ 69 years.}$$

Of course, having estimated the time needed to do the problem, you would no longer wish to proceed with the calculations.

Definition 1.18. *The integers a and b are relatively prime if $(a, b) = 1$. (The term* **co-prime** *is also used to mean the same thing).*

Theorem 1.19 (Dividing a product by a co-prime)**.** *If $a \mid (bc)$ and $(a, b) = 1$, then $a \mid c$.*

Proof. Since $(a, b) = 1$, by Theorem 1.8 (Euclid's algorithm) there exist x, y in $\mathbb{Z}$ such that $1 = xa + yb$. Multiplying through by c we get

$$c = xac + ybc.$$

Now since $a \mid xac$ and $a \mid y(bc)$, we see that a divides the RHS. Hence $a \mid c$. □

Corollary. If a prime p divides a product $b_1 b_2 \cdots b_k$ then it divides at least one of the factors $b_1, b_2, \ldots, b_k$.

Proof. If $p \mid b_1$ the conclusion holds. Otherwise, since p is prime, $(p, b_1) = 1$. By Theorem 1.19, $p \mid b_2 \cdots b_k$. Again, if $p \mid b_2$ the conclusion holds. Otherwise $(p, b_2) = 1$ and then $p \mid b_3 \cdots b_k$. Continuing, we find that if p does not divide any of $b_1, b_2, \ldots, b_{k-1}$, then it must divide b_k . □

Theorem 1.20 (Fundamental Theorem of Arithmetic)**.** *Every positive integer greater than* 1 *can be written as a product of primes. The factorisation is unique if the factors are written in order of non-decreasing size.*

Remark. When n itself is prime, we view $n = n$ as writing n as the product of one prime.

Proof. (a) (Existence of the factorisation). Let $\mathrm{P}(n)$ be the statement that n can be written as a product of primes. Clearly $\mathrm{P}(2)$ is true since 2 is prime. Assume $\mathrm{P}(k)$ is true for all integers k with $2 \leq k < n$. If n is prime, then $\mathrm{P}(n)$ is true. Otherwise $n = ab$ where a, b are positive integers with $2 \leq a < n$ and $2 \leq b < n$. By the inductive hypothesis, a and b can each be written as a product of primes. Thus $n = ab$ is also a product of (the same) primes, and again $\mathrm{P}(n)$ is true. By induction, $\mathrm{P}(n)$ is true for all $n \geq 2$.

(b) (Uniqueness of the Factorisation). Suppose we have two factorisations

$$n = p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_m$$

where p_i, q_j are positive primes and are possibly repeated. Assume without loss of generality that $k \geq m$. (We can just rename if necessary). Now $p_1 \mid n$ so by the Corollary to Theorem 1.19, p_1 divides at least one of the terms $q_1, q_2, \ldots, q_m$. After renaming if necessary we can assume $p_1 \mid q_1$. Since p_1 and q_1 are prime, this means $p_1 = q_1$. After cancelling we get

$$p_2 \cdots p_k = q_2 \cdots q_m.$$

Next $p_2 \mid q_2 \cdots q_m$ and again, by renaming if necessary we can assume $p_2 \mid q_2$ and so conclude $p_2 = q_2$. We continue in this way until all the q's are used up. If $k > m$ we end with a product of primes which is equal to 1, namely

$$p_{m+1} \cdots p_k = 1.$$

But this is impossible by the definition of a prime. Hence we must have $k = m$ and then the process ends with $p_k = q_m$. Hence the same primes occur in each factorisation and each prime occurs the same number of times in each factorisation.

Finally, we conclude that if the primes are arranged in order of non-decreasing size, then the factorisation is unique. □

Corollary. For each positive integer $n \geq 2$, there is a unique list of primes $p_1, p_2, \ldots, p_k$ with $p_1 < p_2 < \cdots < p_k$ and a unique list of integers $a_i > 0$ such that

$$n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}.$$

Example 1.21. *Find the prime factorisation of* 480.

Solution. We pull out all the 2's, then all the 3's (if any) and so on. Thus

$$\begin{aligned} 480 &= 2 \cdot 240 = 2 \cdot 2 \cdot 120 \\ &= 2 \cdot 2 \cdot 2 \cdot 60 = 2 \cdot 2 \cdot 2 \cdot 2 \cdot 30 \\ &= 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 15 \\ &= 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 3 \cdot 5 \\ &= 2^5 \cdot 3 \cdot 5. \end{aligned}$$

Remarks.

1. If we had allowed 1 to be counted as a prime number, then the uniqueness part of the prime factorisation would no longer hold. For example, we would have to allow $480 = 1 \cdot 2^5 \cdot 3 \cdot 5$ to count as another prime factorisation of 480.

2. The uniqueness part of the Fundamental Theorem is the more subtle of the two parts. To prove existence, we were able to give a direct proof based on Mathematical Induction. To prove the uniqueness, we made use of the Corollary to Theorem 1.19, which in turn was proved using preceding results.

Theorem 1.22 (Euclid's Second Theorem)**.** *There are infinitely many primes.*

Proof. Suppose there are only finitely many primes which can then be listed as $2, 3, 5, \ldots, p_k$ where p_k is the 'last' prime. Let

$$n = (2 \cdot 3 \cdot 5 \cdot \cdots \cdot p_k) + 1$$

Consider a prime factor p of n (such a p exists by the Fundamental Theorem). Now by assumption, p has to be one of the primes $2, 3, 5, \ldots, p_k$. Since

$$1 = n - (2 \cdot 3 \cdot 5 \cdot \cdots \cdot p_k)$$

and p divides both n and $2 \cdot 3 \cdot 5 \cdot \cdots \cdot p_k$, it follows that $p \mid 1$, an impossibility.

Since the assumption that there are only finitely many primes leads to a contradiction, we conclude that there must be infinitely many primes. □

Corollary. For each integer n there is a prime which is greater than n. Hence for each n, there is a 'next' prime $p > n$.

Apart from 2, every prime is odd and hence is either of the form $4k + 1$ or of the form $4k + 3$. One might ask which of the following four cases were true:

1. There are infinitely many primes of the form $4k + 1$ and infinitely many primes of the form $4k + 3$.

2. There are infinitely many primes of the form $4k+1$ and finitely many primes of the form $4k+3$.

3. There are finitely many primes of the form $4k+1$ and infinitely many primes of the form $4k+3$.

4. There are only finitely many primes of both types.

Actually the last case can be immediately ruled out, for if it were true there would only be finitely many primes altogether.

More generally we can divide a given integer n by a number q to get

$$n = qk + r, \quad 0 \leq r < n.$$

Clearly for n to be prime it is necessary that $(q, r) = 1$.
QUESTION (ON THE DISTRIBUTION OF PRIMES) Are there infinitely many primes of the form $qk + r$ for given q and r with $(q, r) = 1$ and $1 \leq r < q$?

The next theorem and a number of assignment questions will give an affirmative answer in some special cases. (It was proved by Dirichlet that the result is true in all cases, but since this theorem is not part of our course, we will not assume it in our discussions.)

Lemma 1.23. *If a and b are integers of the form $4n+1$ then ab is also of the same form.*

Proof. Write $a = 4r + 1$, $b = 4s + 1$. Then

$$\begin{aligned} ab &= (4r+1)(4s+1) \\ &= 16rs + 4s + 4r + 1 \\ &= 4(4rs + s + r) + 1 \end{aligned}$$

which is of the required form. □

Remark. This lemma is a special case of more general statements which will be obtained in the next chapter. At this stage you may convince yourself that the same argument shows that the statement remains true if you replace 4 by any positive integer.

Theorem 1.24. *There are infinitely many primes of the form $4n+3$.*

Proof. Suppose there are only finitely many primes of the form $4n + 3$, and list them as $p_0, p_1, \ldots, p_r$, where $p_0 = 3$, $p_1 = 7$, ..., and p_r is the last prime of the form $4n + 3$.

Let $Q = 4(p_1 p_2 \cdots p_r) + 3$, and consider its prime factorisation. If only primes of the form $4n + 1$ occurred then, by the lemma, Q itself would have to be of the form $4n + 1$ which is false. Hence Q must have at least one prime factor p of the form $p = 4n + 3$ and this p must be one of the primes $p_0, p_1, \ldots, p_r$.

Now $p_0 = 3$ cannot divide Q for if it did we would have $3 \mid (Q-3) = 4p_1p_2\cdots p_r$ whence $3 \mid p_i$ for some i, which is impossible. Also for $j \geq 1$, p_j cannot divide Q for if it did we would have

$$p_j \mid (Q - 4p_1p_2\cdots p_r),$$

that is, $p_j \mid 3$ which again is impossible.

Since the assumption that there are only finitely many primes of the form $4n+3$ leads to a contradiction, we conclude that there are infinitely many primes of that form. □

Although there are infinitely many primes, there are also some very large gaps between successive primes. For a given integer n, the n successive integers

$$(n+1)!+2, (n+1)!+3, (n+1)!+4, \ldots, (n+1)!+(n+1)$$

are all composite. (Can you see why?) So choosing $n = 10^9$ for example, we see that at some point there is a string of at least one billion integers to look through before we get to the next prime number.

1.5 Factorising an Integer

We can find the factors of a given integer n by using the method mentioned before, that is, obtain a list of the primes from 2 to n and try dividing them into n.

Example 1.25. *Factorise* 3127.

Solution. We find that the smallest prime which divides 3127 is 53. Dividing 3127 by 53 we get $3127 = 53 \cdot 59$.

For large numbers, this method involves a lot of work when the smallest prime factor is itself large, and even using the most powerful computer it is not practical to use this method for, say, a hundred digit number which is the product of two large primes.
OPEN QUESTION (FACTORISATION) Find an efficient method for factorising large composite integers.

The question has become of great importance in recent years because of the applications of number theory to cryptography. The following is another factorisation method which does not depend on finding all the primes between 1 and $\sqrt{n}$.

Fermat Factorisation

First note that if n is even it has 2 as a factor, so there is no difficulty in getting the first factor. If n is odd and composite then

$$\begin{aligned} n &= ab \quad (\text{for } a, b \text{ both odd and positive and } a \geq b), \\ &= \left(\frac{a+b}{2}\right)^2 - \left(\frac{a-b}{2}\right)^2 \end{aligned}$$

with $\frac{a+b}{2}$ and $\frac{a-b}{2}$ non-negative integers. Hence n is a difference of two squares.

Conversely, if n is a difference of squares, say $n = s^2 - t^2$ with s, t in $\mathbb{Z}$, then n has the factorisation $n = (s+t)(s-t)$. This becomes the trivial factorisation $n = n$ if $s = t + 1$.

These results are summarised in the following.

Theorem 1.26 (Fermat Factorisation)**.** *The odd positive number n is composite if and only if there are solutions for s and t of the equation*

$$n = s^2 - t^2$$

with $s > t + 1, t \geq 0$.

Note that $s^2 = n + t^2 \geq n$, so $s \geq \sqrt{n}$. Further, since $s > t + 1$, we see that

$$s^2 = n + t^2 \leq n + (s-1)^2 = n + s^2 - 2s + 1.$$

From this it follows that $2s \leq n + 1$ and hence that $s \leq (n+1)/2$. These two results together show that the s which we are seeking satisfies $\sqrt{n} \leq s \leq (n+1)/2$.

To implement Fermat's method, let x be the smallest integer greater than or equal to $\sqrt{n}$. Search for a square among the integers $x^2 - n, (x+1)^2 - n, (x+2)^2 - n, \ldots$, then use it to factorise n.

Example 1.27. *Use Fermat's method to factorise (a)* 121*, (b)* 3127*, (c)* 29*.*

Solution. (a) $\sqrt{121} = 11$, so $121 = 11^2$. In this example, $s = \sqrt{n}$.

(b) $\sqrt{3127} = 55.92$ (using a calculator to 2 places) so we take $x = 56$. Then $56^2 - 3127 = 9 = 3^2$. Hence $3127 = 56^2 - 3^2 = (56-3)(56+3) = 53 \cdot 59$.

(c) Here $x = 6$ and

$$\begin{aligned} 6^2 - 29 &= 7 \\ 7^2 - 29 &= 20 \\ 8^2 - 29 &= 35 \\ 9^2 - 29 &= 52 \\ 10^2 - 29 &= 71 \\ 11^2 - 29 &= 92 \\ 12^2 - 29 &= 115 \\ 13^2 - 29 &= 140 \\ 14^2 - 29 &= 167 \\ 15^2 - 29 &= 196 = 14^2. \end{aligned}$$

Hence $29 = 15^2 - 14^2 = 1 \cdot 29 = 29$. In this example, $s = t + 1 = (n+1)/2$.

This method works best when n is the product of two nearly equal numbers, and the process terminates with the trivial factorisation if we begin with a prime.

Note that both the methods we have discussed will, in principle, determine whether a given n is prime or composite, and will produce a factorisation when n is composite.

1.6 The Linear Diophantine Equation

The equation is

$$ax + by = c$$

where a, b, c are integer constants. A **solution** of this equation is a pair x, y of integers which satisfy the equation.

Note that there are some inherent difficulties in finding integral solutions of equations. Although the equation $6x + 9y = 2$ has plenty of solutions using real numbers (all the points lying on the line with equation $6x+9y = 2$), a little thought convinces us that there are no integral solutions to $6x + 9y = 2$. In fact, 3 must divide the LHS whatever the choice of integers x, y, but 3 can never divide the RHS. Hence whatever choice we make for x and y, the equation cannot be satisfied.

Notation. In the following discussion we let $d = \gcd(a, b)$.

Theorem 1.28 (The Linear Diophantine Equation)**.** *The equation $ax+by = c$ has solutions in integers if and only if $d \mid c$ (where $d = \gcd(a, b)$). Moreover, if x_0, y_0 is any solution, then all solutions are given by the expressions*

$$x = x_0 + \frac{b}{d}n, \quad y = y_0 - \frac{a}{d}n$$

where n is any integer.

Proof. Suppose first that there is a solution satisfying $ax + by = c$. Since $d \mid a$ and $d \mid b$ we see that $d \mid c$.

Conversely, suppose $d \mid c$. Then $c = de$ for some integer e. By Euclid's Algorithm there are integers s, t such that $d = as + bt$. Then multiplying through by e we get

$$c = a(se) + b(te)$$

so that there is at least one solution $se,\ te$.

For the last part, suppose x_0, y_0 is one solution and $x,\ y$ is any other solution. Then

$$\begin{aligned} ax + by &= c \\ ax_0 + by_0 &= c \\ \text{Hence} \quad a(x - x_0) + b(y - y_0) &= 0 \\ \frac{a}{d}(x - x_0) &= \frac{b}{d}(y_0 - y) \end{aligned}$$

Since $\left(\frac{a}{d},\ \frac{b}{d}\right) = 1$ and $\frac{b}{d}$ divides the LHS, we see that $\frac{b}{d}$ divides $x-x_0$. Hence $x-x_0 = \frac{b}{d}n$ for some integer n, that is, $x = x_0 + \frac{b}{d}n$. Also

$$b(y_0 - y) = a(x - x_0) = a\frac{b}{d}n.$$

Hence $y_0 - y = \frac{a}{d}n$, and therefore $y = y_0 - \frac{a}{d}n$.

Finally we can check by direct substitution that if n is any integer, then the pair $x = x_0 + \frac{b}{d}n$, $y = y_0 - \frac{a}{d}n$ is indeed a solution. □

Note that the proof of the theorem tells us how to construct solutions in particular cases. Also note that the complete solution $x = x_0 + \frac{b}{d}n$, $y = y_0 - \frac{a}{d}n$ is the sum of a particular solution $x = x_0$, $y = y_0$ and the complete solution $x = \frac{b}{d}n$, $y = \frac{a}{d}n$ of the homogeneous linear Diophantine equation $ax + by = 0$.

Example 1.29. *Solve the Diophantine equation* $3x + 5y = 1$.

Solution. Since $(3, 5) = 1$, solutions exist. By inspection (or by using Euclid's algorithm) we see that $x_0 = -3, y_0 = 2$ is one solution. The general solution is

$$x = -3 + 5n, \quad y = 2 - 3n \quad \text{for } n \in \mathbb{Z}.$$

Example 1.30. *Solve the Diophantine equation* $225x + 85y = 15$.

Solution. Referring to the working in Example 1.10, we have $(225, 85) = 5$, which is a divisor of 15. Hence solutions exist. Further, we found that $225 \cdot (-3) + 85 \cdot 8 = 5$. Hence on multiplying through by 3 we get $225 \cdot (-9) + 85 \cdot 24 = 15$ which shows that $x_0 = -9$, $y_0 = 24$ is one solution. The general solution is

$$x = x_0 + \frac{b}{d}n = -9 + \frac{85}{5}n = -9 + 17n$$
$$y = y_0 - \frac{a}{d}n = 24 - \frac{225}{5}n = 24 - 45n$$

for $n \in \mathbb{Z}$.

The problem of finding integral solutions to equations must have been considered at many different times in ancient civilisations. The name comes from Diophantus, who made a serious investigation into these types of equations.

If $f(x, y, z)$ is a polynomial in three variables with integral coefficients, the general Diophantine equation in three variables can be written $f(x, y, z) = 0$.

More examples of such equations include

$$x^2 + y^2 = z^2$$
$$x^3 + y^3 = z^3$$

The fact that $3^2 + 4^2 = 5^2$, that is, the triple 3, 4, 5 is a solution of $x^2 + y^2 = z^2$, has fascinated many people over the centuries.

More generally, we can consider Diophantine equations in any number of variables.
OPEN QUESTION (DIOPHANTINE EQUATIONS) Can we find all solutions of a given Diophantine equation, or else show that no solutions exist?

In general, the problem of solving Diophantine equations, or showing that they have no solutions, is incredibly difficult. Our Theorem 1.28 is considered a model answer to the question for the special case of a linear equation in two variables. We return to the problem in the last section of the course.

Chapter 2

CONGRUENCES

2.1 Residue classes

Some times we are interested in a particular property of a very complicated expression. It would be clever to detect the particular property without explicitly computing the whole expression. Congruences is such a clever method that allows us to answer questions like

- Is $203^5 + 1$ divisible by 17?
- What is the last digit of 203^5?

without computing 203^5.

The basic idea is to partition the set of all integers into finitely many classes (subsets) and to do the computations in terms of the classes rather than in terms of numbers. The simplest example is the partition of integers into *odd* and *even* numbers. If we want to know whether $132434 \cdot 6675567$ is odd or even we need not to compute the product. Instead we check whether the factors are odd or even and then apply the rule that a product is even if and only if at least one factor is even. (Hence, in our example the product is even!)

The parity of sums and products of numbers does not depend on the numbers themselves but only on their parity. We summarise this fact in the following tables.

$+$	even	odd
even	even	odd
odd	odd	even

$\cdot$	even	odd
even	even	even
odd	even	odd

Even numbers are exactly the integers with remainder 0 after division by 2, whereas odd numbers are the integers with remainder 1 after division by 2. Mimicking this for any divisor $m > 1$, we partition all integers into m classes (disjoint subsets), where each class consists of the numbers with the same remainder after division by m, according to the division algorithm.

Definition 2.1. *Fix an integer $m > 1$ which is called the* modulus. *Then an integer a is said to be* congruent *to an integer b modulo m, written $a \equiv b \pmod{m}$, if and only if $m \mid (b - a)$, i.e. a and b have the same remainder after division by m.*

Congruence modulo m as defined above is an *equivalence relation*. The set of integers splits into equivalence classes that are called *residue classes* modulo m. For any integer a we denote the corresponding residue class by $[a]$. Hence, $[a]$ is the set of all numbers that are congruent to a modulo m. a is called *representative* of $[a]$. Any number a' that is congruent to a would be a representative of the same residue class. Thus $[a] = [a']$. The m residue classes can be represented by $[0], [1], \ldots, [m-1]$. More general,

Definition 2.2. *A* complete system of residues modulo m *is a set of* m *integers* $\{a_1, a_2, \ldots, a_m\}$ *that are pairwise incongruent. Then all residue classes can be represented by* $[a_1], [a_2], \ldots, [a_m]$.

Examples for complete systems of residues for $m = 5$ are $\{0, 1, 2, 3, 4\}$, or $\{-2, -1, 0, 1, 2\}$, or $\{1, 2, 3, 4, 5\}$, but also $\{-10, -4, 7, 8, -1\}$.

Before we introduce addition and multiplication of the residue classes modulo m similar to addition and multiplication of "even" and "odd" we need to prove a lemma:

Lemma 2.3. *Let* $m > 1$ *be fixed modulus. Then the residue class of a sum* $[a+b]$ *depends only on the residue classes* $[a]$ *and* $[b]$*, but not on* a *and* b *themselves. The residue class of a product* $[a \cdot b]$ *depends only on the residue classes* $[a]$ *and* $[b]$*, but not on* a *and* b *themselves.*

Proof. Suppose $a \equiv a' \pmod{m}$ and $b \equiv b' \pmod{m}$, i.e. $m|(a-a')$ and $m|(b-b')$. Then $m|a+b-(a'+b')$, i.e. $a+b \equiv a'+b' \pmod{m}$. Hence $[a+b] = [a'+b']$.

For products we have $m|(a-a')b + a'(b-b')$, i.e. $m|ab - a'b'$. Hence, $[a \cdot b] = [a' \cdot b']$. □

Corollary. If $a \equiv a' \pmod{m}$, then $-a \equiv -a' \pmod{m}$

If $a \equiv a' \pmod{m}$, then $a^i \equiv (a')^i \pmod{m}$ for any integer $i \geq 1$.

Proof. If $m|(a-a')$ then, obviously $m|-a-(-a')$. The second statement can be proved by induction. □

Definition 2.4. *The sum* $[a]+[b]$ *of two residue classes* $[a]$ *and* $[b]$ *is defined as the residue class* $[a+b]$*. The product* $[a] \cdot [b]$ *of two residue classes* $[a]$ *and* $[b]$ *is defined as the residue class* $[a \cdot b]$.

The set of residue classes modulo m with the introduced above operations $+$ and $\cdot$ is another example of a commutative ring. In fact

1. $[a] + [b] = [a+b] = [b+a] = [b] + [a]$.
2. $([a] + [b]) + [c] = [a+b] + [c] = [a+b+c] = [a] + [b+c] = [a] + ([b] + [c])$
3. $[a] + [0] = [a+0] = [a]$ for any residue class $[a]$.
4. For any residue class $[a]$ the class $[-a]$ is an additive inverse, i.e. $[a] + [-a] = [-a] + [a] = [0]$.
5. $([a] \cdot [b]) \cdot [c] = [a \cdot b] \cdot [c] = [a \cdot b \cdot c] = [a] \cdot [b \cdot c] = [a] \cdot ([b] \cdot [c])$
6. $([a] + [b]) \cdot [c] = [a+b] \cdot [c] = [(a+b)c] = [ac+bc] = [ac] + [bc]$

7. $[a] \cdot [b] = [ab] = [ba] = [b] \cdot [a]$

8. $[1][a] = [a][1] = [a \cdot 1] = [a]$ for any $[a]$.

Consider the mapping that assigns to any number a its residue class $[a]$ modulo m:

$$\pi : \quad a \mapsto [a]$$

This mapping has the remarkable property $\pi(a+b) = \pi(a) + \pi(b)$ and $\pi(ab) = \pi(a)\pi(b)$. Such mapping is called a ring homomorphism.

Roughly speaking, any arithmetic relation between numbers must be true for the corresponding residue classes. This means that if a relation does not hold for the residue classes it cannot hold for integers.

There is one significant difference between the ring of integers and rings of residue classes. Whenever a product of two integers is zero, one of the factors must be zero. This is not true in general for rings of residue classes. For example, let $m = 6$ then $[2] \cdot [3] = [6] = [0]$, i.e. $[0]$ is the product of two factors which are different from zero. A consequence is that, in general, we are not allowed to cancel equal non-zero factors: the correct equation $[2] \cdot [3] = [2] \cdot [0]$ does not imply $[3] = [0]$. However we will prove that cancellation by "coprimes" of the modulus is allowed. In particular, cancellation of non-zero factors is always possible if the modulus is a prime.

Theorem 2.5 (Cancellation Modulo m)**.** *If* $ac \equiv bc \pmod{m}$ *and* $(c, m) = 1$*, then* $a \equiv b \pmod{m}$*. In other words,* $[a][c] = [b][c]$ *implies* $[a] = [b]$*.*

Proof. We are given that $ac \equiv bc \pmod{m}$. Hence

$$\begin{array}{lll} m & | & (ac - bc) \\ m & | & c(a-b) \\ m & | & (a-b) \quad \text{since} \quad (c,m) = 1. \end{array}$$

Hence $a \equiv b \pmod{m}$. □

The key point of using congruences can now be formulated as follows: Whenever we are interested to what residue class the result of an arithmetic expression involving addition, subtraction, and multiplication belongs we may substitute each summand or factor at each stage of the computation by another representative of the same residue class and thus simplify the computation.

Let us return to the questions at the beginning of this Section. Is $203^5 + 1$ divisible by 17? Thus we have to decide whether $203^5 + 1 \equiv 0 \pmod{17}$. This is equivalent to $203^5 \equiv -1 \pmod{17}$. Now we replace the factor 203 by the representative $-1 \equiv 203 \pmod{17}$. Obviously, $(-1)^5 \equiv -1 \pmod{17}$, hence $203^5 + 1$ is divisible by 17.

For the second question we realise that the last digit of a number (with respect to the decimal system) is the least non-negative remainder after division by 10. For our

computation we may replace 203 by 3 since $203 \equiv 3 \pmod{10}$. We have $3^5 = 3^3 \cdot 3^2 = 27 \cdot 9$. In this product we may replace 27 by 7 since $27 \equiv 7 \pmod{10}$. It follows that the last digit of 203^5 is the same as the last digit of $7 \cdot 9 = 63$, thus it is 3.

As another application of congruences we will prove some divisibility tests. These tests tell us whether an integer (the dividend) is divisible by another integer (the divisor) by just performing some simple operations with the digits of the dividend.

- An integer a is divisible by 2 if and only if its last digit is divisible by 2.
- An integer a is divisible by 5 if and only if its last digit is divisible by 5.
- An integer a is divisible by 3 if and only if the sum of its digits is divisible by 3.
- An integer a is divisible by 9 if and only if the sum of its digits is divisible by 9.
- An integer a is divisible by 11 if and only if the alternating sum (see below) of its digits is divisible by 11.

Recall that the decimal representation of an integer is a sum

$$a = \sum_{k=0}^{n} a_k 10^k.$$

Thus $a = f(10)$ where $f(x)$ is the polynomial $\sum_{k=0}^{n} a_k x^k$ whose coefficients are the digits of a (in reverse order).

Let us now prove the divisibility tests for 2 and 5. They are bases on the fact that any integer a is the sum $a_0 + a'$ where a_0 is the last digit of a and $a' = \sum_{k=1}^{n} a_k 10^k = 10 \sum_{k=0}^{n-1} a_{k+1} 10^k$ is divisible by 10, and hence by 2 and 5. Thus a is divisible by 2 or 5 if and only if the last digit a_0 is divisible by 2 or 5 respectively.

If we would choose a system based on a number N different from 10 we could derive similar divisibility tests for all divisors of N: The number $a = \sum_{k=0}^{n} a_k N^k$ is divisible by a divisor of N if and only if the last digit a_0 is divisible by that divisor.

An integer a will be divisible by 3 or 9 respectively, if and only it is congruent to 0 modulo 3 or 9 respectively. We have

$$[a] = \left[\sum_{k=1}^{n} a_k 10^k\right] = \left[\sum_{k=1}^{n} a_k [10]^k\right] = \left[\sum_{k=1}^{n} a_k\right]$$

because $10 \equiv 1 \pmod{3}$ as well as $10 \equiv 1 \pmod{9}$. It follows that the sum of digits of an integer a is in the same residue class modulo 3 or 9 as the number a itself.

For the test of divisibility by 11 we notice that even powers of 10, i.e. $1, 100, 1000, \ldots$ are congruent to 1 modulo 11 whereas odd powers $10, 1000, \ldots$ are congruent to -1 modulo 11. Hence

$$[a] = \left[\sum_{k=1}^{n} a_k 10^k\right] = \left[\sum_{k=1}^{n} a_k [10]^k\right] = \left[\sum_{k=1}^{n} (-1)^k a_k\right],$$

that is a belongs to the same residue class modulo 11 as the alternating sum of digits $a_0 - a_1 + a_2 - \cdots$.

Remark. Before computers were used in book-keeping, book-keepers often had to sum up long sequences of numbers by hands. In order to check their result they repeated the calculation modulo 9.

2.2 The Linear Congruence in One Variable

The following problem is called linear congruence in one variable: Given the modulus m and the integers a, b. Find all integers x such that

$$ax \equiv b \pmod{m}.$$

It is clear that all integers that belong to the same residue class as a solution x will also be solutions. We can reformulate the problem as follows: Find all residue classes $[x]$ such that $[a] \cdot [x] = [b]$. Hence, it is a division problem in the ring of residue classes modulo m.

Since every integer is congruent modulo m to one of $0, 1, 2, m-1$ we will be interested in finding all solutions x (if any) with $0 \leq x \leq m-1$, or equivalently all the solutions among a complete residue system. These are referred to as the incongruent solutions modulo m.

The congruence can be turned into an equation by noting that $ax \equiv b \pmod{m}$ is true if and only if there is an integer k such that $ax = b + km$. On replacing k by $-y$, we see that x is a solution of $ax \equiv b \pmod{m}$ if and only if there is an integer y such that the pair x, y is a solution of the Linear Diophantine Equation $ax + my = b$. According to the theory on the Linear Diophantine Equation, there is a solution if and only if $\gcd(a, m)$ is a divisor of b. Further, we can use the Euclidean Algorithm to find an x and a y which satisfy the equation. These observations prove the existence parts of the following theorem.

Theorem 2.6 (The Linear Congruence)**.** *(i) If $(a, m) = 1$, there is a unique solution modulo m of the linear congruence $ax \equiv b \pmod{m}$.*
(ii) If $(a, m) = d > 1$, there is a solution of the linear congruence $ax \equiv b \pmod{m}$ if and only if $d \mid b$. If $d \mid b$, there are exactly d incongruent solutions modulo m given by

$$x_0, x_0 + \frac{m}{d}, x_0 + 2\frac{m}{d}, \ldots, x_0 + (d-1)\frac{m}{d}$$

where x_0 is a particular solution of the linear congruence.

Proof. (i) Suppose that there are two solutions x and y of the congruence. Then

$$\begin{aligned} ax &\equiv b \pmod{m}, \text{ and} \\ ay &\equiv b \pmod{m}. \end{aligned}$$

On subtracting, we see that $a(x - y) \equiv 0 \pmod{m}$. Since $(a, m) = 1$, it follows from Theorem 2.5 that $x - y \equiv 0 \pmod{m}$ that is, $x \equiv y \pmod{m}$.

(ii) Suppose that $d \mid b$. The linear congruence is equivalent to the linear Diophantine equation $ax + my = b$. On dividing by d this becomes

$$\frac{a}{d}x + \frac{m}{d}y = \frac{b}{d}$$

which is equivalent to the linear congruence

$$\frac{a}{d}x \equiv \frac{b}{d} \pmod{\frac{m}{d}}.$$

Since $\left(\frac{a}{d}, \frac{m}{d}\right) = 1$, by part (i) above, this has a unique solution, x_0 say, modulo $\frac{m}{d}$.

This is the same as saying that there are d incongruent solutions modulo m given by

$$x_0, x_0 + \frac{m}{d}, x_0 + 2\frac{m}{d}, \ldots, x_0 + (d-1)\frac{m}{d}\,. \qquad \square$$

Example 2.7. *Solve* $17x \equiv 12 \pmod{91}$.

Solution. Since $(17, 91) = 1$, there is a unique solution modulo 91. To find x, we solve $17x + 91y = 12$.

We first use the Euclidean algorithm to get

$$91 = 5 \times 17 + 6, 17 = 2 \times 6 + 5, 6 = 5 + 1.$$

Then working backwards we get $1 = 6 - 5 = 6 - (17 - 2 \times 6) = 3 \times 6 - 17 = 3 \times (91 - 5 \times 17) - 17 = 3 \times 91 - 16 \times 17$. Hence $17(-16) \equiv 1 \pmod{91}$ and on multiplying by 12, we get $17(-192) \equiv 12 \pmod{91}$.

The solutions are given by $x \equiv -192 \pmod{91}$, that is, $x \equiv 81 \pmod{91}$.

Example 2.8. *Solve* $6x \equiv 3 \pmod{9}$.

Solution. Since $(6, 9) = 3$, and $3 \mid 3$, there are $\frac{9}{3} = 3$ distinct solutions modulo 9. By inspection, $6 \times 2 \equiv 3 \pmod{9}$ so we take $x_0 = 2$. The others are $2 + \frac{9}{3} = 5$ and $2 + 2 \times \frac{9}{3} = 8$.

The solutions are given by $x \equiv 2, 5, 8 \pmod{9}$.

Remark. Another way to do this problem is to write the congruence as a linear Diophantine equation $6x + 9y = 3$. After cancelling 3 from every term, this is equivalent to the equation $2x + 3y = 1$ which in turn is equivalent to the congruence $2x \equiv 1 \pmod{3}$. This has a unique solution modulo 3, which by inspection is $x \equiv 2 \pmod{3}$.

In the previous chapter we raised the problem of Diophantine equations. On replacing equality by congruence modulo m, we can ask a similar (simpler) question for congruences. It is clear that any solution of the original Diophantine equation gives rise to a solution of the corresponding equation modulo m. Hence, if one can show that there is no solution modulo m for some modulus m then it follows that there is no solution to the original problem.

Notice that a particular Diophantine equation modulo m can be solved by testing all residue classes for solutions. However, it is much more difficult to find a systematic solution.

For the linear case, we have a completely satisfactory answer given by our Theorem 2.6. For the case of a quadratic congruence ($n = 2$), the problem is already difficult, and we give a partial treatment of this case in Chapter 5 of the course.

Consider the linear congruence with $b = 1$.

Definition 2.9. *If* $ax \equiv 1 \pmod{m}$, *x is called an* **inverse of a modulo m**. *In this case, a is also an inverse of x modulo m.*

This definition is motivated by the fact that the residue class $[x]$ is inverse to $[a]$ in the ring of residue classes, i.e. $[a][x] = [1]$.

Theorem 2.10. *A number a has an inverse modulo m if and only if $(a, m) = 1$. The inverse is unique modulo m.*

Proof. If $(a, m) = 1$, Theorem 2.6 proves the existence of a unique unverse, since an inverse modulo m is the same as a solution of the linear congruence $ax \equiv 1 \pmod{m}$.

If $(a, m) = p > 1$ then $m = pq$ with $1 < p, q < m$. If there was an inverse a^{-1} we would have

$$aa^{-1}q \equiv 0 \pmod{m}$$

(since $m = pq|aq$) and at the othe hand

$$aa^{-1}q \equiv q \pmod{m}.$$

But $q \not\equiv 0 \pmod{m}$. The contradiction shows that a has no inverse. □

Note that if a has an inverse a^{-1}, modulo m, then the solution of the linear congruence

$$ax \equiv b \pmod{m}$$

can be found by multiplying both sides of the congruence by a^{-1} to give

$$x \equiv a^{-1}b \pmod{m}.$$

Example 2.11. *Find an inverse of* 2 *modulo* 23 *and use it to solve the linear congruence* $2x \equiv 7 \pmod{23}$.

Solution. We have to solve the linear congruence $2x \equiv 1 \pmod{23}$ and to do this we use one of the methods discussed before. By inspection, $2 \times 12 \equiv 1 \pmod{23}$ so 12 is the inverse.

If $2x \equiv 7 \pmod{23}$, then on multiplying through by 12 we get $x \equiv 7 \times 12 = 84 \equiv 15 \pmod{23}$.

Another application of residue classes are so-called "Selfcorrecting codes". During a transmission a message can be distorted. Therefore it is important to be able to check whether the received message coincides with the sent message and if not, to be able to correct it. A simple example of a selfcorrecting code is the ISBN number which identifies any published book worldwide. The ISBN consists of 10 characters: nine digits $x_1, \ldots, x_9$ that identify the book and a check symbol x_{10} which is either a digit or the letter X. The check symbol is chosen in such a way that the following congruence holds

$$\sum_{i=1}^{10} ix_i \equiv 0 \pmod{11},$$

where X stands for 10. Hence x_{10} is the least non-negative solution of the linear congruence

$$10x_{10} \equiv -\sum_{i=1}^{9} ix_i \pmod{11}$$

which is equivalent to saying that x_{10} is the least non-negative remainder modulo 11 of $\sum_{i=1}^{9} ix_i$.

The ISBN of the 3rd edition of the course textbook is 0 201 57889 1. We check

$$0 + 2 \cdot 2 + 0 + 4 \cdot 1 + 5 \cdot 5 + 6 \cdot 7 + 7 \cdot 8 + 8 \cdot 8 + 9 \cdot 9 + 10 \cdot 1 = 286 \equiv 0 \pmod{11}$$

If $\sum_{i=1}^{10} ix_i \not\equiv 0 \pmod{11}$ we conclude that one (or more) characters in the ISBN are wrong. We will see that the correct ISBN can be recovered if only one character at a known place was altered. Supposed $x_1, \cdots, x_{10}$ was the original valid ISBN and $y_1, \ldots, y_{10}$ the distorted (at one place) ISBN. Then

$$\sum_{i=1}^{10} iy_i = \sum_{i=1}^{10} ix_i + ja \equiv ja \equiv r \not\equiv 0 \pmod{11},$$

where j is the (known) place of the distortion and a is the (unknown) difference of the original and the altered value. We have to solve the linear congruence

$$ja \equiv r \pmod{11}$$

which has a unique solution since $(j, 11) = 1$ (The reason for the choice 11, not 10 for the modulus is that all $j = 1, \ldots, 10$ are coprimes of 11 and therefore have an inverse.)

2.3 The Chinese Remainder Theorem

In this section we study the question if there exist numbers x which satisfy a system of simultaneous congruences

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

A system of simultaneous congruences can be visualised as a mechanism of r cog-wheels with $m_1, \ldots, m_r$ cogs respectively, where the cogs at each wheel are marked. A solution will be a constellation where all marked cogs are at a certain prescribed place.

It is easy to see that for some simultaneous congruences no solution exists. Consider the system

$$\begin{aligned} x &\equiv 1 \pmod{2}, \\ x &\equiv 2 \pmod{4} \end{aligned}$$

The first congruence means precisely that x is odd. On the other hand, the second congruence implies that x must be even. Hence no solution exists. This is due to fact that the moduli 2 and 4 are not relatively prime and therefore the given informations about the residues are not independent.

Let us now assume that the moduli are relatively prime in pairs.

Theorem 2.12 (Chinese Remainder Theorem)**.** *If $m_1, m_2, \ldots, m_r$ are pairwise relatively prime positive integers, then the system of congruences has a solution x. The solution is unique modulo M where $M = m_1 m_2 \cdots m_r$.*

Proof. Let $M_k = \frac{M}{m_k}$, $1 \leq k \leq r$.

We first show that $(M_k, m_k) = 1$. Suppose that p is a prime which divides both M_k and m_k. Now M_k consists of a product of all but one of the terms m_i , so by the Corollary to Theorem 1.19, p divides m_i for some $i \neq k$. But this contradicts the condition that $(m_i, m_k) = 1$. We conclude that there cannot be a prime which divides both M_k and m_k, and hence that $(M_k, m_k) = 1$.

By Theorem 2.10, each M_k has an inverse modulo m_k . Thus, for $1 \leq k \leq r$ there exists y_k such that $M_k y_k \equiv 1 \pmod{m_k}$. Let

$$x = a_1 M_1 y_1 + a_2 M_2 y_2 + \cdots + a_r M_r y_r = \sum_{k=1}^{r} a_k M_k y_k.$$

We show that this x is a solution of each of the congruences. To do this, consider congruence modulo m_k. If $i \neq k$, then $m_k \mid M_i$ so that

$$M_i \equiv 0 \pmod{m_k}.$$

Hence in the expression for x, all but one of the terms (the k'th term) is congruent modulo m_k to 0, so

$$\begin{aligned} x &\equiv a_k M_k y_k \pmod{m_k}, \\ &\equiv a_k \pmod{m_k} \text{ since } M_k y_k \equiv 1 \pmod{m_k}. \end{aligned}$$

For the uniqueness part, suppose that both x and y are solutions of the congruences. Hence, for $1 \leq k \leq r$,

$$\begin{aligned} x &\equiv a_k \pmod{m_k}, \\ y &\equiv a_k \pmod{m_k}, \\ \text{and } x - y &\equiv 0 \pmod{m_k}. \end{aligned}$$

This means that $m_k \mid (x - y)$ for each k. Since $(m_1, m_2) = 1$, it follows that $m_1 m_2 \mid (x - y)$ (cp. Problem 8 in Assignment 1). Then, since $(m_1 m_2, m_3) = 1$, it again follows that $m_1 m_2 m_3 \mid (x - y)$. Continuing in this way, we find that $M = m_1 m_2 \cdots m_r \mid (x - y)$, that is, $x \equiv y \pmod{M}$. □

Remark. Coming back to the picture of the cog-wheel mechanism, the Chinese Remainder theorem states that any particular constellation will repeat itself exactly after M steps. Since there are exactly M constellations, all of them will be materialised during one cycle of M steps. The k-th cog-wheel will rotate M_k times during one cycle. If the numbers m_k are not pairwise relatively prime then the cycle will be smaller than M steps, thus some constellations will repeat within M steps and, consequently, some other constellations will not appear at all.

Example 2.13. *Solve the system*

$$\begin{aligned} x &\equiv 2 \pmod 3 \\ x &\equiv 4 \pmod 5 \\ x &\equiv 1 \pmod 4. \end{aligned}$$

Solution. Since $(3, 5) = (3, 4) = (4, 5) = 1$, there is a solution which is unique modulo $3 \times 5 \times 4 = 60$. We use the notation in the proof of the theorem:
$m_1 = 3, m_2 = 5, m_3 = 4$,
$M_1 = \frac{60}{3} = 20, M_2 = \frac{60}{5} = 12, M_3 = \frac{60}{4} = 15, M = 60$,
$a_1 = 2, a_2 = 4, a_3 = 1$.
A solution of $M_1 y \equiv 1 \pmod{m_1}$, that is, $20y \equiv 2y \equiv 1 \pmod 3$ is $y_1 = 2$.
A solution of $M_2 y \equiv 1 \pmod{m_2}$, that is, $12y \equiv 2y \equiv 1 \pmod 5$ is $y_2 = 3$.
A solution of $M_3 y \equiv 1 \pmod{m_3}$, that is, $15y \equiv 3y \equiv 1 \pmod 4$ is $y_3 = 3$.

One solution of the given system is

$$\begin{aligned} x &= a_1 M_1 y_1 + a_2 M_2 y_2 + a_3 M_3 y_3 \\ &= 2 \times 20 \times 2 + 4 \times 12 \times 3 + 1 \times 15 \times 3 = 269. \end{aligned}$$

Since $269 = 4 \times 60 + 29$, the smallest positive solution is 29.

ANS: $x \equiv 29 \pmod{60}$.

2.4 The Calendar

In the time it takes for the earth to go once around the sun (one year), the earth completes 365.2422 revolutions on its own axis. This number is obtained from astronomical observations and is the number of astronomical days in a year.

In the Julian calendar (instituted in the time of Julius Caesar), there was a leap year every four years. This made the length of a calendar year 365.25 days, a discrepancy of 0.0078 days per year. After a thousand years, the accumulated discrepancy was 1000×0.0078 days, or about 8 days. Thus Christmas in the northern hemisphere, which was known to have once occurred a few days after the winter solstice, was then occurring almost two weeks after the winter solstice. Similar discrepancies concerning the equinoxes were also apparent.

The calendar which we now use is the Gregorian calendar. In this calendar, 5 Oct 1582 (in the Julian calendar) became 15 Oct 1582 to correct the discrepancy, and new rules were introduced to calculate leap years. The main idea in the Gregorian calendar is to have 97 leap years in every cycle of 400 years so that the average length of a calendar year is (365 + 97/400) days, which to four decimal places is 365.2425 days. The accumulated discrepancy over a thousand year cycle in the Gregorian calendar is 1000×0.0003 days, or about 7 hours.

Many countries (including England) delayed changing to the new calendar, with the result that historians have to be careful to determine which calendar was in use when

interpreting some historical date. The Julian calendar is still in use by some churches to determine their dates of religious significance.

Further history and information about the calendar can be found in encyclopaedias. See also the science website http://www.science.org.au/nova/maths.htm where a number of articles on mathematics may be found. Articles on science in general are at http://www.science.org.au/nova/.

The rule for leap years is the following. Year N is a leap year if N is divisible by 4, except that the years divisible by 100 are leap years only if they are divisible by 400. Thus 1700, 1800 and 1900 were not leap years but 2000 was a leap year.

Problem. Given the date (in the Gregorian calendar), what day of the week is it?

We can use some elementary ideas on congruences to help answer this question, remembering that the days of the week work modulo 7, and that the months of the year work modulo 12.

We label the days of the week using a complete residue system modulo 7, and the months of the year using a complete residue system modulo 12. When a leap year arises, it is added on at the end of February. For this reason it is convenient to adopt a labeling system for the months which ends with February. We will use the following system of labels.

Day	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
W	0	1	2	3	4	5	6

Month	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb
m	1	2	3	4	5	6	7	8	9	10	11	12

We can get some idea of how to proceed by starting with a recent date and working forward. Let us choose 1 March 2000 as a starting point. We look up a recent calendar and find that it was Wednesday on this day, that is, $W = 3$ on 1 March 2000. The rest of the days in this month can be calculated from the formula

$$W \equiv k + 2 \pmod 7$$

where k denotes the day of the month and $1 \leq k \leq 31$.

For example, on 26 March 2000, we have $W \equiv 26 + 2 \equiv 0 \pmod 7$, so it was a Sunday on that day.

Because there are 31 days (four weeks plus three days) in March, the day for 1 April 2000 will be three days later than for March, so it will be on Saturday. For April 2000 we can use the formula

$$W \equiv k + 5 \pmod 7$$

where k denotes the day of the month and $1 \leq k \leq 30$.

Similarly the first of May is shifted two days later and is on a Sunday. We can continue working through the months and we get a formula of the type

$$W \equiv k + j(m) \pmod 7$$

where k denotes the day of the month and $j(m)$ is the initialising number for month m. The numbers $j(m)$ are given by the following table.

	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb
m	1	2	3	4	5	6	7	8	9	10	11	12
$j(m)$	2	5	0	3	5	1	4	6	2	4	0	3

The numbers in the table can be remembered from the mnemonic "It seems Charles has eaten a cold dinner; he eats nothing hot" and the length of the words in this sentence.

The formula works for January and February of 2001 as well, so for example on 14$^{\text{th}}$ February 2001 the day of the week is given by $W \equiv 14 + 3 \equiv 3 \pmod 7$ which translates to Wednesday.

Since there are $365 = 7 \times 52 + 1$ days in a normal year, the day of the week for 1 March 2001 is shifted by one day from what it was in 2000, so it will be on Thursday. This shift occurs throughout the next twelve months so from March 2001 to February 2002, the day of the week is given by the formula

$$W \equiv k + j(m) + 1 \pmod 7.$$

More generally for the year $2000 + x$ where $0 \leq x \leq 3$, the day of the week is given by

$$W \equiv k + j(m) + x \pmod 7$$

provided that we consider January and February as belonging to the previous year.

The year 2004 will be a leap year and a two day shift occurs in going from 2003 to 2004. The formula for March 2004 to February 2005 is

$$W \equiv k + j(m) + 5 \pmod 7.$$

What we have done is quite useful for determining dates over the next few years and has the advantage of being transparent. With a bit of practice you can amaze your friends by doing the calculations in your head. For a complete analysis it is better to use 1 March 0000 as a starting point, even though the Gregorian calendar was not in use then.

Our final formulation involves the greatest integer function. Recall that if x is a real number, then $[x]$ denotes the integer part of x. It is the unique integer n satisfying $n \leq x < n + 1$. For example, $[5] = 5$ and $[6.334] = 6$.

We first find the day of the week for 1 March 0000. In the 2000 years from 1 March 0000 to 1 March 2000 there are 5×97 leap years each of which causes a 2 day shift in the calendar. An ordinary year causes a 1 day shift. So altogether there are $2000 + 5 \times 97$ shifts. Since this number is divisible by 7, we conclude that 1 March 0000 was on the same day of the week as 1 March 2000, that is, on a Wednesday. In the period from 1 March 0000 to 28 February 0001, the same formula $W \equiv k + 2 \pmod 7$ applies.

For year N, we need to add on a shift of one day for each ordinary year and two days for each leap year. We get a formula of the type

$$W \equiv k + j(m) + N + L \pmod 7$$

where L is the number of leap years from 0000 to N. We include year N in this count for L but not 0000.

Example 2.14. *Show that there is at least one Friday 13th in each year.*

Solution. For any choice of N, the problem is to find m such that $5 \equiv 13 + j(m) + N + L \pmod 7$, that is, $j(m) \equiv 8 - N - L \pmod 7$. Looking at the table for $j(m)$, we see that the months from May to November have numbers 0, 3, 5, 1, 4, 6, 2 which form a complete set of residues modulo 7. Hence one of these months has a $j(m)$ which satisfies the last congruence. Thus, in any year, one of the months from May to November contains a Friday 13^{th}.

We will now get convenient expressions for $j(m)$ and $N + L$ and insert them into this formula.

First, there is an expression for $j(m)$ in terms of the greatest integer function. It is

$$j(m) = [2.6m - 0.2], \text{ where } 1 \leq m \leq 12$$

and you can check that it produces the desired set of numbers. The only significance of this formula is that it fits the set of numbers $j(m)$ and is a convenient way of generating the numbers.

Next, to calculate L, count the years divisible by four, take out the years divisible by 100, then add back in the years divisible by 400. This gives the expression

$$L = [N/4] - [N/100] + [N/400].$$

If we write $N = 100C + Y$, where $0 \leq Y \leq 99$, then

$$\begin{aligned} N &= 100C + Y \equiv 2C + Y \pmod 7 \\ [N/4] &= 25C + [Y/4] \equiv 4C + [Y/4] \pmod 7 \\ [N/100] &= C \\ [N/400] &= [C/4]. \end{aligned}$$

Hence

$$\begin{aligned} N + L &\equiv 2C + Y + 4C + [Y/4] - C + [C/4] \pmod 7 \\ &\equiv -2C + Y + [Y/4] + [C/4] \pmod 7. \end{aligned}$$

Finally we summarise our notation and give the final result. Remember that in calculating N, January and February are considered as belonging to the previous year.

k — the day of the month,
m — the month (using the above labeling),
N — the year,
C — the century,
Y — the year of the century.

Notice that here century means the number obtained by cancelling the last two digits of the year. Although we would say that 1995 was in the 20th century the corresponding number $C = 19$. Thus for 11 August 1995 we have $k = 11$, $m = 6$, $N = 1995$, $C = 19$, $Y = 95$. For 11 January 1995 we have $k = 11$, $m = 11$, $N = 1994$, $C = 19$, $Y = 94$.

Theorem 2.15. *Let W be the day of the week for day k of month m of year N. Then, with the above labeling, W satisfies the following formula:*

$$W \equiv k + [2.6m - 0.2] - 2C + Y + [Y/4] + [C/4] \pmod 7$$

provided that dates in January and February are treated as occurring in the previous year.

Example 2.16. *Find the day of the week for* 1 *February* 2100.

Solution. $k = 1, m = 12, Y = 99, C = 20$.

$$\begin{aligned} W &\equiv 1 + [2.6 \times 12 - 0.2] - 2 \times 20 + 99 + [99/4] + [20/4] \pmod 7 \\ &\equiv 1 + 31 - 40 + 99 + 24 + 5 \pmod 7 \\ &\equiv 120 \equiv 1 \pmod 7. \end{aligned}$$

Hence it will be a Monday on 1 February 2100.

At this stage you might wish to use the formula to calculate the day of the week on which you were born.

Using the formula in Theorem 2.15, it is possible to write a computer program which will produce a calendar for any given year. Such programs are often installed on computer systems. For example, on the student computer *metz* at UNE, the command **cal 1788** will display a calendar for the year 1788.

2.5 Scheduling a Round-Robin Tournament

Suppose N players (or teams) compete in a round-robin tournament so that they are to play each other exactly once.

If the pairings are done carelessly in the early rounds, it is possible to get into a situation where the pairings for the last few rounds cannot be done properly: it can become impossible to give every player a game against someone he or she has not already played[1]. A systematic procedure is needed.

[1]This phenomenom occurs if there are at least 5 players. If the first three rounds have been scheduled (1-4, 2-5, 3-bye), (1-5, 2-bye, 3-4), (1-bye, 2-4, 3-5) then in the next round two of the teams 1,2,3 have to play each other leaving the third without a partner. Moreover 4 and 5 have to play each other in round 4 and have no partner in round 5. A similar problem occurs with 6 players where the 6^{th} player replaces the bye.

Label the players

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

We need a procedure that assigns to any player and any round a partner in such a way that (1) nobody is assigned to himself (2) if a plays b in round k then b plays a in round k and (3) nobody is assigned to the same partner in different rounds.

We start with the case when N is an odd number. The best possible result is to have a draw in which only N rounds are required. In each round there is one player not playing. The basic idea of our pairing system is that in round k, the players are paired according to the following rule which immediately implies (2) and (3).

$$\text{In round } k, \text{ player } a \text{ meets player } k - a \pmod{N}.$$

Clearly, player a gets a different opponent in each round if we restrict k to $1 \leq k \leq N$ and $b = k - a \pmod{N}$ is assigned to a.

In fact, the condition can be stated symmetrically as follows:

Players a and b meet in round k, where $1 \leq k \leq N - 1$, if and only if

$$a + b \equiv k \pmod{N}.$$

Our rule does not imply (1). In fact, in any round k there is exactly one a modulo N with $a + a = k \pmod{N}$. Thus a is assigned to himself which fits to the fact that we need one player in each round who does not play.

If N was even, we distribute the first $N - 1$ players according to rule above. Player N will play in each round the player among the first $N - 1$ who was assigned to himself.

It is usually necessary to assign the home games in a tournament in such a way that each player gets a fair distribution of home and away games. (In a chess tournament, the 'home game' could be interpreted as having the white pieces).

A really fair distribution is only possible if N is an odd number. Otherwise there would be an odd number of games for each player which makes it impossible to have an equal number of home and away games. Let N be odd. Then for the players i, j ($1 \leq i, j < N$ and $i \neq j$), use the following rule.

$$\begin{gathered}\text{If } i + j \text{ is odd, the smaller of } i, j \text{ is the home team.} \\ \text{If } i + j \text{ is even, the larger of } i, j \text{ is the home team.}\end{gathered}$$

If N is even the following procedure is almost fair: Determine home games for the first $N - 1$ players as above. Then toss for N's first home game, and after that N alternates between home and away games.

As an illustration, suppose there is a round-robin tournament of 22 teams, and the above pairing rules are used with team 22 having its first game at home. Suppose team 14 wants to know whom it plays and where in round 8. Consider $8 - 14 = -6 \equiv 15 \pmod{21}$. Also $14 + 15$ is odd, so the smaller number has the home game. Thus team 14 has a home game against team 15 in round 8.

Chapter 3

SPECIAL CONGRUENCES and CRYPTOGRAPHY

3.1 Wilson's Theorem

If p be a prime number then $1, 2, \ldots, p-1$ are all relatively prime to p. According to Theorem 2.10 then all elements of the residue class ring except [0] have an inverse. This makes the residue class ring a field. Therefore prime modulus is special.

This section is devoted to congruences with respect to the prime modulus p.

Theorem 3.1. *Let p be prime. The only solutions of the congruence*

$$x^2 \equiv 1 \pmod{p}$$

are $x \equiv \pm 1 \pmod{p}$.

Proof. Clearly if $x \equiv \pm 1 \pmod{p}$, then $x^2 \equiv 1 \pmod{p}$. Conversely, suppose that x is a solution of $x^2 \equiv 1 \pmod{p}$. Then $p \mid x^2 - 1$, that is, $p \mid (x-1)(x+1)$. Since p is prime, this means that $p \mid (x-1)$ or $p \mid (x+1)$, that is, $x \equiv 1 \pmod{p}$ or $x \equiv -1 \pmod{p}$. □

Notice that when $p = 2$ we have $1 \equiv -1 \pmod{2}$, so there is only one incongruent solution in this case. We also notice from the definition of an inverse modulo p, that any x which satisfies $x^2 \equiv 1 \pmod{p}$ is actually its own inverse.

Theorem 3.2 (Wilson's Theorem)**.** *If p is prime, then $(p-1)! \equiv -1 \pmod{p}$.*

Proof. First, if $p = 2$, then $(p-1)! = 1 \equiv -1 \pmod{2}$, so the result is true. If p is any other prime, then

$$(p-1)! = 1 \times 2 \times \cdots \times (p-2) \times (p-1).$$

Each of $1, 2, \ldots, p-2, p-1$ is relatively prime to p, so each has an inverse modulo p. Now 1 and $p-1$ are their own inverses, since $1^2 \equiv 1 \pmod{p}$ and $(p-1)^2 \equiv (-1)^2 \equiv 1 \pmod{p}$. By Theorem 3.1 these are the only terms which are their own inverses.

Hence apart from 1 and $p-1$, the inverses occur in pairs, for if b is an inverse of a, then a is an inverse of b. Hence apart from 1 and $p-1$, each term is the inverse of one of the others, so the terms cancel in pairs modulo p. It follows that

$$(p-1)! \equiv 1 \times (p-1) \equiv -1 \pmod{p}. \qquad \square$$

Example 3.3. *Illustrate the proof of Wilson's Theorem with $p = 11$.*

Solution.

$$\begin{aligned} 10! &= 1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8 \times 9 \times 10 \\ &= 1 \times (2 \times 6) \times (3 \times 4) \times (5 \times 9) \times (7 \times 8) \times 10 \\ &\equiv 1 \times 1 \times 1 \times 1 \times 1 \times 10 \pmod{11} \\ &\equiv -1 \pmod{11}. \end{aligned}$$

Remark. An illustration of a proof using a numerical example does not itself constitute a proof. The purpose of a numerical illustration is to help pinpoint and understand the underlying reason why the proof works.

It is interesting that the converse of Wilson's theorem is also true:

Theorem 3.4. *If $(p-1)! \equiv -1 \pmod{p}$, then all numbers $1, \ldots p-1$ have an inverse modulo p and, therefore, p is prime.*

Proof. If $(p-1)! \equiv -1 \pmod{p}$ then for any k between 1 and $p-1$

$$k \cdot (-\prod_{\substack{j=1 \\ j \neq k}}^{p-1} j) \equiv 1 \pmod{p}.$$

Hence, $-\prod_{\substack{j=1 \\ j \neq k}}^{p-1} j$ is an inverse of k. But, according to Theorem 2.10 any proper divisor of p cannot have an inverse. Hence, p has no proper divisors. $\square$

The converse of Wilson's theorem allows us to detect whether p is prime or not without checking for divisibility by all primes that are $\leq \sqrt{p}$. Later on we will see other primality tests.

3.2 Euler's Theorem

Definition 3.5. *Let $m \geq 1$. The integer $\phi(m)$ is defined to be the number of positive integers less than or equal to m which are relatively prime to m.*

This function is referred to as the **Euler phi function**. It will play a crucial role in much of the remainder of our course. Its importance is related to the fact that the group of invertible residue classes modulo m has exactly $\phi(m)$ elements.

Theorem 3.6. *There are exactly $\phi(m)$ invertible residue classes modulo m. These residue classes are represented by positive integers less than or equal to m which are relatively prime to m.*

Proof. We prove this theorem by establishing a bijective map π between the set of positive integers less than or equal to m which are relatively prime to m and the set of invertible residue classes. The map π is defined by

$$\pi : r \mapsto [r],$$

i.e. it assigns a number r to its residue class $[r]$. It follows immediately from Theorem 2.10 that $[r]$ is invertible if $(r, m) = 1$. It is also clear from $0 < r \leq m$ that different r represent different $[r]$. Thus, π is injective. On the other hand, according to the division algorithm, any residue class is represented by a number r with $0 < r \leq m$. The class $[r]$ is invertible if and only if $(r, m) = 1$. Thus, π is surjective. □

Example 3.7. *Calculate $\phi(12)$.*

Solution. The positive integers less than or equal to 12 are 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12. The numbers 2, 3, 4, 6, 8, 9, 10, 12 have a prime factor in common with 12, so are not relatively prime to 12. If we delete these from the list, we are left with the four numbers 1, 5, 7, 11. Hence $\phi(12) = 4$.

This same method can in principle be used to evaluate $\phi(m)$ for any choice of m. For large m the method is inefficient, and we will develop more convenient methods later in this chapter.

Example 3.8. *Find a formula for $\phi(p)$ where p is a prime number.*

Solution. Consider the numbers $1, 2, \ldots, p-1, p$. The only one of these numbers which is not relatively prime to p is p itself. So on deleting p from the list we are left with $1, 2, \ldots, p-1$. It follows that $\phi(p) = p - 1$.

Definition 3.9. *A reduced system modulo m is a complete set of representatives of the invertible residue classes modulo m. In other words, a reduced residue system is a set of $\phi(m)$ integers each of which is relatively prime to m, and for which no two different elements of the set are congruent modulo m.*

Starting with a complete residue system modulo m (see Section 2.1), we can get a reduced residue system modulo m by removing those elements which are not relatively prime to m. For example, the set

$$\{-10, -4, 7, 8, -1\}$$

is a complete residue system modulo 5, and by removing the element -10, (the only one not relatively prime to 5), we are left with the reduced residue system modulo 5

$$\{-4, 7, 8, -1\}.$$

Let $r_1, r_2, \ldots, r_{\phi(m)}$ be the $\phi(m)$ positive integers which are less than m and which are relatively prime to m. These numbers form a reduced residue system modulo m, and this is the most frequently used example.

Theorem 3.10 (Euler's Theorem)**.** *If $m \geq 1$ and $(a, m) = 1$, then*

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

Using the notation of residue classes this can be restated as follows

$$[a]^{\phi(m)} = [1].$$

Proof. Let $r_1, r_2, \ldots, r_{\phi(m)}$ be the $\phi(m)$ positive integers which are less than m and which are relatively prime to m or any other reduced system. Then

$$[r_1], [r_2], \ldots, [r_{\phi(m)}]$$

is a list of all elements of the group G of invertible residue classes. The map L_a which is defined by

$$L_a : [r] \mapsto [a][r]$$

is bijective from G to G. In fact,

$$L_{a^{-1}} : [r] \mapsto [a]^{-1}[r]$$

is inverse to L_a. Thus,

$$[a][r_1], [a][r_2], \ldots, [a][r_{\phi(m)}]$$

is again a list of all elements of G (in different order). It follows (by commutativity) that the product of all elements of first list equals the product of all elements of the second list.

$$[r_1] \cdot [r_2] \cdots [r_{\phi(m)}] = [a][r_1] \cdot [a][r_2] \cdots [a][r_{\phi(m)}] = [a]^{\phi(m)}[r_1] \cdot [r_2] \cdots [r_{\phi(m)}].$$

Now, multiplying both sides with the inverse of

$$[r_1] \cdot [r_2] \cdots [r_{\phi(m)}]$$

yields

$$[a]^{\phi(m)} = [1].$$

□

Remark. Euler's Theorem is a special case of the following fact: If G is a finite group of n elements then $a^n = e$ for any element $a \in G$, where e is the neutral element of G. The proof from above can be literally carried over to the general case if G is Abelian. For non-Abelian groups a different proof is needed.

Example 3.11. *Illustrate the proof of Euler's Theorem with $m = 9, a = 7$.*

Solution. The numbers less than or equal to 9 and relatively prime to 9 are 1, 2, 4, 5, 7, 8. Hence $\phi(9) = 6$. Multiplying these numbers by 7 and reducing modulo 9 gives the following lists.

r_i	ar_i	s_i
1	$7 \times 1 = 7$	7
2	$7 \times 2 = 14$	5
4	$7 \times 4 = 28$	1
5	$7 \times 5 = 35$	8
7	$7 \times 7 = 49$	4
8	$7 \times 8 = 56$	2

The numbers in the second column are congruent mod 9 to the corresponding numbers in the third column. Thus

$$\begin{aligned}(7\times1)(7\times2)(7\times4)(7\times5)(7\times7)(7\times8) &\equiv 7\times5\times1\times8\times4\times2 \pmod 9\\ 7^6\times1\times2\times4\times5\times7\times8 &\equiv 1\times2\times4\times5\times7\times8 \pmod 9\\ 7^6 &\equiv 1 \pmod 9\end{aligned}$$

(after cancelling the common terms).

Example 3.12. *Find the last digit of* 137^{197}.

Solution. The problem is to find the non-negative integer less than 10 which is congruent mod 10 to 137^{197}.

First $137 \equiv 7 \pmod{10}$ so $137^{197} \equiv 7^{197} \pmod{10}$. Next, $(7, 10) = 1$, so we can use Euler's Theorem with $m = 10$. Now the non-negative integers which are less than 10 and relatively prime to 10 are 1, 3, 7, and 9 so $\phi(10) = 4$. Hence $7^4 \equiv 1 \pmod{10}$. Divide 197 by 4 to get $197 = 49 \times 4 + 1$. Then

$$137^{197} \equiv 7^{197} = 7^{49\times4+1} = (7^4)^{49} \times 7 \equiv 7 \pmod{10}.$$

The last digit is 7.

Remark. In the above examples we have been finding $\phi(m)$ the hard way. Later we will have a much better way using the prime factorisation of m.

Theorem 3.13 (Fermat's Little Theorem)**.** *If p is a prime and $p \nmid a$, then*

$$a^{p-1} \equiv 1 \pmod p.$$

Proof. Since p is prime, $\phi(p) = p-1$, by Example 3.8. Since $p \nmid a$, we must have $(a, p) = 1$. Euler's Theorem for this situation now says $a^{p-1} \equiv 1 \pmod p$ which is the required result. □

Corollary. If p is prime and a is any integer, then

$$a^p \equiv a \pmod p.$$

Proof. If $p \mid a$, then both sides are congruent to 0 mod p, so the result holds trivially. If $p \nmid a$, then $a^{p-1} \equiv 1 \pmod p$ by Fermat's Little Theorem, and on multiplying both sides by a, we again get $a^p \equiv a \pmod p$. □

Example 3.14. *The following statements are all consequences of Fermat's Little Theorem.*

(i) If 5^{10} is divided by 11, the remainder is 1 (on taking $p = 11, a = 5$).
(ii) $26^{100} \equiv 1 \pmod{101}$ (on taking $p = 101, a = 26$).
(iii) If 81^6 is divided by 13, the remainder is 1 (since $81^6 = 9^{12} \equiv 1 \pmod{13}$).

Example 3.15. *Suppose $(a, m) = 1$. Use Euler's Theorem to find a formula for the inverse of a modulo m.*

Solution. We know from Theorem 2.10 that the inverse exists and is unique modulo m. Now Euler's Theorem states that $a^{\phi(m)} \equiv 1 \pmod{m}$. The left hand side is a power of a. If we pull one of the a terms out to the front we get

$$aa^{\phi(m)-1} \equiv 1 \pmod{m}$$

which shows immediately that the inverse of a is $a^{\phi(m)-1}$.

Example 3.16. *Suppose $(a, m) = 1$. Use the result in the previous exercise to find a formula for the solution of the linear congruence $ax \equiv b \pmod{m}$.*

Solution. We know from Theorem 2.6(i) that the solution exists and is unique modulo m. Now an inverse of a is $a^{\phi(m)-1}$. Hence on multiplying both sides of the congruence by this inverse (see the comments after Theorem 2.10), we have the solution

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

Example 3.17. *Use the method in the previous exercise to solve the linear congruence*

$$17x \equiv 12 \pmod{91}$$

(which was considered in Example 2.7).

Solution. $91 = 7 \times 13$. To calculate $\phi(91)$ consider the list of integers from 1 to 91. By removing all multiples of 7 and all multiples of 13 and counting those removed, we find that $\phi(91) = 72$. Hence an inverse of 17 modulo 91 is 17^{71}. Hence the solution of the congruence can be written $x \equiv 12 \times 17^{71} \pmod{91}$.

The trouble with this method is that it does not produce the smallest positive solution, and this may have been what we really wanted. The number 17^{71} is too large to be handled exactly with a standard calculator (try it!). We saw in Example 2.7 that the solution can be written $x \equiv 81 \pmod{91}$.

3.3 Multiplicative Functions

Definition 3.18. *A function f defined on the positive integers is said to be* **multiplicative** *if $f(mn) = f(m)f(n)$ whenever $m, n > 0$ and $(m, n) = 1$.*

Our first objective in this subsection is to show that the Euler ϕ-function is multiplicative. Then we will see a couple of other functions which are also multiplicative.

Theorem 3.19 (ϕ is multiplicative)**.** *Let m and n be positive integers with $(m, n) = 1$. Then*

$$\phi(mn) = \phi(m)\phi(n).$$

Proof. Let G be the set of integers r such that $0 \leq r < mn$, $(r, mn) = 1$. Then G has $\phi(mn)$ elements. The set P of pairs (s, t) where $0 \leq s < m$, $(s, m) = 1$ and $0 \leq t < n$, $(t, n) = 1$ has $\phi(m)\phi(n)$ elements. We prove the theorem by establishing a bijective map between the two sets.

Let $g : G \to P$ be defined by

$$g : r \mapsto (s, t),$$

where s is the remainder of r on dividing by m and t is the remainder of r on dividing by n.

Since $(r, mn) = 1$ it follows that $(r, m) = 1$ and $(r, n) = 1$ and therefore $(s, m) = (t, n) = 1$. Thus the pair (s, t) belongs to P.

The map g is bijective if and only if it has an inverse. We construct the inverse map h of g as follows. Let $(s, t) \in P$. By the Chinese Remainder Theorem, there is a unique r such that $0 \leq r < mn$ and

$$r \equiv s \pmod{m}, \text{ and } r \equiv t \pmod{n}.$$

From $(s, m) = (t, n) = 1$ follows $(r, m) = (r, n) = 1$ and therefore $(r, mn) = 1$. Thus $r \in G$. Clearly, $h = g^{-1}$. □

Corollary. Let $m_1, m_2, \ldots, m_k$ be pairwise relatively prime positive integers. Then, by induction,

$$\phi(m_1 m_2 \cdots m_k) = \phi(m_1)\phi(m_2) \cdots \phi(m_k).$$

We cannot use this result when $(m, n) \neq 1$. For example, we cannot put $m = n = 3$ and expect that $\phi(9) = \phi(3)\phi(3) = 2 \times 2 = 4$. In fact, we know that $\phi(9) = 6$ which is different from 4.

We have seen earlier that $\phi(p) = p - 1$ when p is prime. It is useful to get the corresponding result for prime powers.

Theorem 3.20. *If p is a prime, then $\phi(p^a) = p^a - p^{a-1} = p^a(1 - \frac{1}{p})$ for any positive integer a.*

Proof. We consider the integers from 1 to p^a. Those numbers which are not relatively prime to p^a have p as a factor, so they are of the form kp where $1 \leq k \leq p^{a-1}$. There are

exactly p^{a-1} of these numbers. The others are relatively prime to p^a, so $\phi(p^a) = p^a - p^{a-1}$. □

Provided we know the prime factorisation of n, the last two results provide an extremely effective way of calculating $\phi(n)$, and replaces the methods we were using earlier on to calculate $\phi(n)$.

Example 3.21. *Evaluate (a) $\phi(101)$, (b) $\phi(91)$, (c) $\phi(36)$, (d) $\phi(9576)$.*

Solution. (a) 101 is prime so $\phi(101) = 101 - 1 = 100$.
(b) $91 = 7 \times 13$ so $\phi(91) = \phi(7)\phi(13) = (7-1)(13-1) = 6 \times 12 = 72$.
(c) $36 = 2^2 \times 3^2$ so $\phi(36) = \phi(2^2)\phi(3^2) = (2^2 - 2)(3^2 - 3) = 2 \times 6 = 12$.
(d) $9576 = 2^3 \times 3^2 \times 7 \times 19$ so

$$\begin{aligned} \phi(9576) &= \phi(2^3)\phi(3^2)\phi(7)\phi(19) \\ &= (2^3 - 2^2)(3^2 - 3)(7-1)(19-1) \\ &= 4 \times 6 \times 6 \times 18 = 2592. \end{aligned}$$

More generally, we can write down a formula for $\phi(n)$. Let $n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$ be the prime factorisation of n. Then

$$\begin{aligned} \phi(n) &= \phi(p_1^{a_1})\phi(p_2^{a_2}) \cdots \phi(p_k^{a_k}) \\ &= p_1^{a_1}(1 - \frac{1}{p_1})p_2^{a_2}(1 - \frac{1}{p_2}) \cdots p_k^{a_k}(1 - \frac{1}{p_k}) \\ &= n(1 - \frac{1}{p_1})(1 - \frac{1}{p_2}) \cdots (1 - \frac{1}{p_k}). \end{aligned}$$

Another area where the Euler ϕ-function occurs is in problems to do with reducing rational numbers to lowest terms. Let n be a positive integer, and consider the n rational numbers

$$\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}.$$

Reduce them all to lowest terms by cancelling out any common factor in the numerator and denominator. We will now analyse what happens in this process.

First of all, the rational number $\frac{m}{n}$ is already in lowest terms if $\gcd(m, n) = 1$, so there are $\phi(n)$ terms where no reduction is necessary. The denominators which occur after reduction to lowest terms are divisors of n, and we can use the Euler function to calculate how often each divisor appears. If $d \mid n$, then each of the rational numbers $\frac{1}{d}, \ldots, \frac{d}{d}$ is $\frac{m}{n}$ for some m and hence appears in the original list shown above. However, of the d fractions with denominator d just $\phi(d)$ are in lowest terms. Therefore the denominator d appears just $\phi(d)$ times when the fractions in the above list are written in lowest terms. But all the fractions have now been accounted for, so n, the number of fractions, is just the sum of the numbers $\phi(d)$ for d dividing n. Hence we have derived the following interesting result.

Theorem 3.22. *Let n be a positive integer. Then*

$$\sum_{d|n} \phi(d) = n$$

where the summation is taken over all the positive divisors of n, including 1 and n.

As an illustration, let us take $n = 12$ and write out the list of fractions. After reducing to lowest terms, we are left with

$$\frac{1}{12}, \frac{1}{6}, \frac{1}{4}, \frac{1}{3}, \frac{5}{12}, \frac{1}{2}, \frac{7}{12}, \frac{2}{3}, \frac{3}{4}, \frac{5}{6}, \frac{11}{12}, \frac{1}{1}.$$

We see that all the divisors of 12 occur in the denominators, and we can easily check that a divisor d occurs exactly $\phi(d)$ times.

We next consider briefly two other functions which turn out to be multiplicative.

Definition 3.23 (Sum of divisors function)**.** *For a positive integer n, the sum of all the positive divisors of n is denoted by $\sigma(n)$. Thus*

$$\sigma(n) = \sum_{d|n} d\,.$$

Definition 3.24 (Number of divisors function)**.** *For a positive integer n, the number of positive divisors of n is denoted by $\tau(n)$. Thus*

$$\tau(n) = \sum_{d|n} 1\,.$$

Example 3.25. *Find $\sigma(12)$ and $\tau(12)$.*

The positive divisors of 12 are 1, 2, 3, 4, 6, 12. Hence

$$\begin{aligned} \sigma(12) &= 1+2+3+4+6+12 = 28, \\ \tau(12) &= 6 \text{ (the number of terms in the above list).} \end{aligned}$$

We next derive formulae for the above functions in the case when n is a prime power.

Theorem 3.26. *Let p be a prime number and let a be a positive integer. Then*

$$\begin{aligned} \sigma(p^a) &= \frac{p^{a+1}-1}{p-1}, \\ \tau(p^a) &= a+1. \end{aligned}$$

Proof. The positive divisors of p^a are $1, p, p^2, \ldots, p^a$, so the sum of the divisors is

$$\sigma(p^a) = 1 + p + p^2 + \cdots + p^a = \frac{p^{a+1} - 1}{p - 1}$$

and the number of these divisors is $\tau(p^a) = a + 1$. □

Before we prove the multiplicativity of σ and τ we introduce a more general construction.

For a function f defined on the positive integers we define a new function F by

$$F(n) = \sum_{d|n} f(d).$$

We will show that F is multiplicative whenever f is multiplicative. (The converse is also true but we will not prove it.) We need a lemma.

Lemma 3.27. *If n is a positive integer such that $n = n_1 n_2$ with $(n_1, n_2) = 1$ and d is a positive integer such that $d|n$ then there are two uniquely determined positive integers d_1 and d_2 such that $d = d_1 d_2$ and $d_1|n_1$ and $d_2|n_2$.*

Proof. Let $d = p_1^{a_1} \cdots p_k^{a_k}$ the prime factorisation of d. Then any prime factor p_r of d divides n and hence precisely one of n_1 or n_2. Therefore d_1 can contain only the prime factors that divide n_1 and d_2 can contain only the prime factors that divide n_2. Let d_1 be the product of all $p_r^{a_r}$ that divide n_1 and d_2 be the product of all $p_r^{a_r}$ that divide n_2. Then d_1 and d_2 satisfy the requirements of the lemma. □

Theorem 3.28. *Let f be a multiplicative function. Then*

$$F(n) = \sum_{d|n} f(d)$$

is also multiplicative.

Proof. Let $(n_1, n_2) = 1$. Then

$$F(n_1 n_2) = \sum_{d|n_1 n_2} f(d).$$

Now, according to the lemma above, any divisor $d = d_1 d_2$ with $d_1|n_1$ and $d_2|n_2$. Hence

$$F(n_1 n_2) = \sum_{\substack{d_1|n_1 \\ d_2|n_2}} f(d_1 d_2) = \sum_{\substack{d_1|n_1 \\ d_2|n_2}} f(d_1) f(d_2)$$

$$= \left(\sum_{d_1|n_1} f(d_1) \right) \left(\sum_{d_2|n_2} f(d_2) \right) = F(n_1) F(n_2).$$

□

The constant function $f(n) = 1$ and the identical function $f(n) = n$ are obviously multiplicative. For $f(n) = 1$ we find $F(n) = \tau(n)$. Therefore τ is multiplicative. Analogously, for $f(n) = n$ we get $F(n) = \sigma(n)$. Hence, σ is multiplicative. Notice that Theorem 3.22 states that the identical function $F(n) = n$ can be obtained from the Euler function $f(n) = \phi(n)$ by the procedure above.

Example 3.29. *Find the sum of the positive divisors of* 9576 *and the number of positive divisors of* 9576*.*

Solution. $9576 = 2^3 \times 3^2 \times 7 \times 19$. Hence

$$\begin{aligned}
\sigma(9576) &= \sigma(2^3)\sigma(3^2)\sigma(7)\sigma(19) \\
&= \frac{2^{3+1}-1}{2-1}\frac{3^{2+1}-1}{3-1}(7+1)(19+1) \\
&= 31,200. \\
\tau(9576) &= \tau(2^3)\tau(3^2)\tau(7)\tau(19) \\
&= (3+1)(2+1)(1+1)(1+1) \\
&= 48.
\end{aligned}$$

3.4 Modular Exponentiation

The problem that we now consider is the following. Given numbers a, n, and m (where $m > 0$), what is a good way for finding the remainder when a^n is divided by m? For example, with $a = 17$, $n = 71$ and $m = 91$, we might wish to find the remainder when 17^{71} is divided by 91 (see Example 3.17). We outline below a few points which make this sort of problem easier to handle.

(1) **First reduce the base.**
If $a \equiv b \pmod{m}$, then $a^n \equiv b^n \pmod{m}$. Using the division algorithm, we can if necessary replace a by b, where $0 \leq |b| \leq m/2$.

Example 3.30. *Reduce the following expressions modulo* 11*. (i)* 13^6*, (ii)* 20^6*, (iii)* 10^6

Solution. (i) $13 \equiv 2 \pmod{11}$. Hence, $13^6 \equiv 2^6 = 64 \equiv 9 \pmod{11}$.
(ii) $20 \equiv -2 \pmod{11}$. Hence, $20^6 \equiv (-2)^6 = 64 \equiv 9 \pmod{11}$.
(iii) $10 \equiv -1 \pmod{11}$. Hence, $10^6 \equiv (-1)^6 = 1 \pmod{11}$.

(2) **Reduce the index using Fermat's Little Theorem, or Euler's Theorem.**
Assuming that $(a, m) = 1$, we can write $n = q\phi(m) + r$ where $0 \leq r < \phi(m)$, and then (using Euler's Theorem)

$$a^n = a^{q\phi(m)+r} = (a^{\phi(m)})^q a^r \equiv a^r \pmod{m}.$$

Example 3.31. *Find the last two digits of* 3^{84}*, that is, find the remainder when* 3^{84} *is divided by* 100*.*

Solution. $\phi(100) = \phi(2^2)\phi(5^2) = (2^2-2)(5^2-5) = 40$. Hence $3^{40} \equiv 1 \pmod{100}$. Hence

$$3^{84} = 3^{2\times 40+4} = (3^{40})^2 3^4 \equiv 3^4 \equiv 81 \pmod{100}.$$

The remainder is 81.

After using the above steps (if applicable), we are down to reducing a^n where $|a| \leq m/2$ and $0 \leq n < \phi(m)$.

(3) **Use Modular Exponentiation.**

We calculate the powers $a^1, a^2, a^4, a^8, \ldots$ modulo m and at each stage we reduce the result modulo m. We then use these calculations to evaluate a^n modulo m. We illustrate the process with an example.

Example 3.32. *Reduce* 2^{309} *modulo* 645*, that is, find the remainder when* 2^{309} *is divided by* 645*.*

Solution.

$$\begin{aligned} 2 &\equiv 2 \pmod{645} \\ 2^2 &\equiv 4 \pmod{645} \\ 2^4 &\equiv 16 \pmod{645} \\ 2^8 &\equiv 256 \pmod{645} \\ 2^{16} &\equiv 256^2 \pmod{645}. \end{aligned}$$

Now using a calculator and making use of the division button to find the quotient when using the division algorithm (see Example 1.4) we get $256^2 = 65536 \equiv 391 \pmod{645}$, since $65536 - 101 \times 645 = 391$.

We do this sort of calculation every time a number gets over 645. The table continues as below:

$$\begin{aligned} 2^{16} &\equiv 256^2 \equiv 391 \pmod{645} \\ 2^{32} &\equiv 391^2 \equiv 16 \pmod{645} \\ 2^{64} &\equiv 16^2 \equiv 256 \pmod{645} \\ 2^{128} &\equiv 256^2 \equiv 391 \pmod{645} \\ 2^{256} &\equiv 391^2 \equiv 16 \pmod{645}. \end{aligned}$$

The last few steps were easy to do because the pattern repeated. We do not need any power of 2 with the index greater than 309.

Now write $309 = 256 + 53 = 256 + 32 + 21 = 256 + 32 + 16 + 5 = 256 + 32 + 16 + 4 + 1$, that is, we write the index in binary form. Then using the results in the above table we have

$$\begin{aligned} 2^{309} &= 2^{256}2^{32}2^{16}2^4 2 \\ &\equiv 16 \times 16 \times 391 \times 16 \times 2 \pmod{645} \\ &\equiv 3,203,072 \equiv 2 \pmod{645}. \end{aligned}$$

3.5 Cryptography

Many groups in society, such as diplomats, spies, and the military, have had the need to send messages which are highly confidential. The purpose of cryptography is to encipher (or encrypt) these messages so that even if one of the messages should fall into enemy hands, it is still impossible for the enemy to decipher the message. More recently, other groups such as banks and networks of computer users, have also used cryptography in implementing the electronic transfer of funds and confidential files.

The simplest codes are based on a scrambled alphabet. An example is the following ancient code that was used by Julius Caesar and is called Caesar's cipher.

We first of all set up numerical equivalents of the letters of the alphabet. In our illustrations we will use the following table.

A	B	C	D	E	F	G	H	I	J	K	L	M
00	01	02	03	04	05	06	07	08	09	10	11	12
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
13	14	15	16	17	18	19	20	21	22	23	24	25

For Caesar's cipher one chooses a pair of integers (a, b), both between 0 and 25 and $(a, 26) = 1$ (thus, $a \neq 0, 2, 13$). Now, instead of the letter assigned to the number P one sends the letter assigned to

$$C \equiv aP + b \pmod{26}.$$

The original message P can be recovered by

$$P = a^{-1}(C - b) \pmod{26}.$$

The inverse a^{-1} modulo 26 of a exists since $(a, 26) = 1$.

The simplest version of Caesar's cipher is obtained for $a = 1$, thus

$$C = P + b \pmod{26}.$$

This means that instead of a letter P the letter C shifted by b places to the right will be sent. Decrypting is easily done by shifting by b places to the left.

Another simple system uses a randomly scrambled alphabet, such as the following.

A	B	C	D	E	F	G	H	I	J	K	L	M
F	J	D	L	M	S	A	Z	I	N	E	R	O
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
T	K	G	Y	U	X	B	H	Q	W	V	C	P

The table gives the enciphering transformation. Suppose the message to be sent is HAPPY BIRTHDAY. Then to encipher the message, each letter appearing in the message is replaced by the one below it in the above table. Hence the enciphered message will be ZFGGC JIUBZLFC. The receiver then deciphers the message by using the deciphering table below, which is constructed by reversing the substitution process above.

A	B	C	D	E	F	G	H	I	J	K	L	M
G	T	Y	C	K	A	P	U	I	B	O	D	E
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
J	M	Z	V	L	F	N	R	X	W	S	Q	H

The first letter Z of the enciphered message is replaced by the letter below it in this table, which is H, and so on. The original message HAPPY BIRTHDAY is then recovered.

There are two main problems with such simple systems:

(i) If the enciphering algorithm should fall into enemy hands, then the enemy can immediately work out how to decipher messages which have been sent using that algorithm.

(ii) If long messages are sent using such a system, then eavesdroppers will have a good chance of breaking the code using a method known as frequency analysis. The letter E is the most frequently used letter in the English alphabet, so the code-breaker would start by finding the most frequently letter occurring in the enciphered messages. There are also only a few one or two letter words in our language, and these can usually be quickly identified in long messages sent by the above system. Thus a good code-breaker would probably break the code if say, a one page message were sent using this system.

Many other systems have been devised for enciphering messages. The most remarkable of these systems are those described as public-key systems. The requirements for such systems are briefly as follows.

(a) Many long secret messages are routinely to be sent through public channels, such as telephones and satellite transmission. By the very nature of public channels, it is possible for eavesdroppers to listen in on messages sent. (The big advantage of using public channels is the low cost.)

(b) Each person in a network is to have an algorithm for enciphering and deciphering messages. The method for enciphering messages is made known to each member of the network. Should some unauthorised person obtain the enciphering algorithms, it is still impossible for him or her to decipher messages which have been sent using that algorithm.

(c) Messages can be 'signed' so that a receiver can tell whether the message is genuine, and not just junk mail sent by unauthorised persons who happen to have gained access to the enciphering algorithm.

It is very surprising that such systems exist. Surely if you know how a message is enciphered, then can't you decipher by just reversing the algorithm?

Our main purpose in this section is to describe the RSA Public-Key System, named after Rivest, Shamir and Adelman who proposed the system in 1978. We will consider very simplified versions of the system in our numerical examples.

A message is first of all translated into its numerical equivalents (according to the table above). For example, if the message to be sent is NO, then on replacing N by 13 and O by 14, the message becomes 1314. Usually, the numerical equivalents of letters are combined into blocks of a certain length (known to both sides). These blocks form (rather big) numbers themselves. These numbers are called the *plain-text*. The unciphered message consist of one or more plain-text blocks P.

3.6 RSA Cipher System

Let n be a number such that for all plain-text blocks $0 \leq P < n$. The RSA Cipher system considers the plain-text P and and the cipher-text C as representatives of their residue classes modulo n. Enciphering and deciphering use arithmetic operations with those residue classes.

Enciphering. First compute the Euler function $\phi(n)$. Then choose the *enciphering key* e which is a number with $(e, \phi(n)) = 1$. Hence, e has an inverse d modulo $\phi(n)$. The number d is called the *deciphering key* and will be used later for decryption. The cipher-text C is P^e modulo n.

Deciphering If the numbers d and n are known a cipher-text C which was encoded as described above can be deciphered by reducing C^d modulo n, provided the plain-text was relatively prime to n. In fact, according to Euler's theorem,

$$C^d = P^{ed} = P^{1+k\phi(n)} \equiv P \pmod{n}.$$

So far, this does not look like a cipher, because the message can only be deciphered if the pair (d, n) is known. On the other hand, anybody who knows the pair can decipher the message. However, by combining this simple idea with the following sophisticated procedure of key exchange one gets the ingenious RSA cipher system.

Suppose sender Alice wants to send a secret message to recipient Bob. First, Bob chooses the block size number n. He computes $\phi(n)$. This requires a factorisation of n (a difficult problem). But, Bob will take n as a product of two large primes p and q (testing p and q for primality is much simpler than factorising a large integer). So, $\phi(n) = (p-1)(q-1)$. Now Bob chooses the enciphering key e and computes the deciphering key d as above. Then he sends the pair (e, n) to Alice or to anyone else who might want to send him a secret message. The pair (e, n) is called (Bob's) public key. Many people create a public key and make it available through their website.

Now, Alice has all data necessary to encrypt her message, using Bob's encrypting key. Namely, she computes P^e and reduces the result modulo n, thus obtaining the cipher-text C. Then she sends the cipher-text to Bob. Only Bob knows the deciphering key d. So he can recover the plain-text P by computing C^d and reducing $\pmod{n}$. Of course, anybody who knows $\phi(n)$ would be able to compute d. But computing $\phi(n)$ is problem of the same difficulty as factorising n (see below) which is known to be a hard (time-consuming) problem for large numbers.

We give now an example.

Example 3.33. *Suppose Bob chose $n = 43 \cdot 59, e = 13$, and the message to be enciphered is NO. What is the corresponding cipher-text?*

Solution. First note that $n = 43 \times 59 = 2537$. This n is just large enough to encipher words of length 2, since the largest number in plain-text form which can occur is 2525 (corresponding to ZZ). Note also that $\phi(n) = (p-1)(q-1) = 42 \times 58$ and $(13, 42 \times 58) = 1$.

The plain-text corresponding to NO is 1314 (from the table above). Note that if we had a longer message to encipher, we could break it up into blocks of length 4, and encipher

each block separately. To get C we need to reduce 1314^{13} modulo 2537 and to do this we use modular exponentiation. The following calculations are done using a calculator and using the division button to get the quotients.

$$\begin{aligned}
1314^2 &\equiv 1436 \pmod{2537} \text{ since } 1314^2 - 680 \times 2537 = 1436,\\
1314^4 &\equiv 2052 \pmod{2537} \text{ since } 1436^2 - 812 \times 2537 = 2052,\\
1314^8 &\equiv 1821 \pmod{2537} \text{ since } 2052^2 - 1659 \times 2537 = 1821,\\
1314^{13} &= 1314^{8+4+1} \equiv 1821 \times 2052 \times 1314 \equiv 2431 \pmod{2537}.
\end{aligned}$$

Hence the cipher-text is 2431.

In the next expample we show that, if the key n is not big enough, not only Bob but any eavesdropper (who took Pmth 338) can decipher the message.

Example 3.34. *Decipher the cipher-text $C = 2431$, which was produced using the (public) enciphering key $(13, 2537)$.*

Solution. Since $n = 2537$ is small we (as well as any educated eavesdropper) easily find $2537 = 43 \times 59$.

Hence $\phi(2537) = 42 \times 58 = 2436$. We now find an inverse of 13 modulo 2436 making use of Euclid's Algorithm.

$$\begin{aligned}
2436 &= 187 \times 13 + 5\\
13 &= 2 \times 5 + 3\\
5 &= 3 + 2\\
3 &= 2 + 1
\end{aligned}$$

Working backwards, we get

$$\begin{aligned}
1 &= 3 - 2 = 3 - (5 - 3) = 2 \times 3 - 5\\
&= 2 \times (13 - 2 \times 5) - 5 = 2 \times 13 - 5 \times 5\\
&= 2 \times 13 - 5 \times (2436 - 187 \times 13)\\
&= (2 + 5 \times 187) \times 13 - 5 \times 2436.
\end{aligned}$$

Hence, we use $d = 2 + 5 \times 187 = 937$.

It follows that the private deciphering key is $(937, 2537)$. Now we can decipher the message $C = 2431$ as recipient Bob would do, namely we reduce 2431^{937} modulo 2537. This again can be done using modular exponentiation or using a computer program such as NUMBERS. The result is 1314 which, when we refer back to the table of numerical equivalents that we are using, translates to NO.

This example illustrates the fact that once the prime factorisation of n is known, it is easy to compute the deciphering key d.

The security of the system rests on the fact that for sufficiently large n, there is no good way known of finding the prime factors p and q which are needed to calculate $\phi(n)$

and hence an inverse of e modulo n. Example 1.17 indicates the amount of time needed to crack the code if inefficient methods are used.

Remark. There is a chance that the plain-text P might not be relatively prime to n, in which case the deciphering procedure might not work. For large n, this chance is small. We can avoid the possibility altogether by using blocks of small enough length so that P is smaller than both p and q.

Signing. Suppose that Alice wishes to sign her message to Bob. She enciphers her plain-text P first using the her own decrypting key (d_A, n_A). Then she enciphers the result P' once more, using Bob's public key (e_B, n_B). The result C of the second encryption will be sent to Bob. Bob deciphers C by means of his decrypting key (d_B, n_B) obtaining P'. Now he applies Alice's public encrypting key (e_A, n_A) which returns the original message P. Bob also concludes that P must have been encrypted by virtue of Alice's secret decrypting key, hence the message is genuine.

3.7 Postscript

(1) The RSA system is named after three mathematicians Ronald Rivest, Adi Shamir, and Leonard Adleman who proposed the system in 1978 while at the Massachusetts Institute of Technology. MIT patented the system and licensed the patent to a company RSA Data Security. This company is listed on the Nasdaq and is one of a number of companies competing for the world market in security applications. You can see what the company does by accessing its web site www.rsa.com.

(2) One of the recent uses of the RSA system is for digital certificates for computer users. This provides a way of verifying the signature of a user and after that less secure but more rapid communication takes place.

(3) There have been a lot of news reports over the last decade on the security of cryptography systems, and claims that someone has cracked the codes. The following report appeared in The Sydney Morning Herald on the 18$^{\text{th}}$ August, 1995.

A French student has cracked the most commonly used encryption system used to pass financial transactions over the Internet, threatening a business predicted to be worth billions worldwide. Market research companies forecast that money transmission over the Internet will be worth more than £30 billion ($65 billion) by 2005.

Damien Doligez, 27, a PhD student at the Inria research centre near Paris, has broken the software key used by the Netscape browsing program, which lets users navigate the World Wide Web. With Netscape, users can order goods by sending their credit card number and address over the network. To prevent anyone picking up these confidential details as they pass through the network, they are encrypted using a software key. Mr Doligez has compromised the system's security by decoding a test example of an encrypted transaction posted on Internet discussion groups in July. The transaction was scrambled using a digital key 40 bits long, which offers about 1,000 trillion possible combinations. Mr Doligez cracked it in eight days.

One of the reasons for such stories is that the USA passed laws preventing the export of

cryptography software. To get around this requirement, the manufacturers produced two versions of Netscape; one for use within the USA which was secure, and the other for use by the rest of the world which was less secure! The laws have changed now. Note that it is always possible to increase security by using larger p and q. Currently primes of about 100 digits in length are considered to provide adequate security.

(4) There is a do-it-yourself cryptography system known as PGP (pretty good privacy) which also uses the RSA system. The system is available free from the internet.

(5) You may have seen the movie Sneakers which came out a few years ago. The film is based on the idea that a mathematician at one of the Californian universities had discovered some powerful new mathematics which allowed him to break any encryption system in use by the US government. Presumably this meant that he could quickly factorise large integers. The mathematician met a speedy death in the film.

Chapter 4

PRIMITIVE ROOTS

4.1 Order and Primitive Roots

The group operation of a finite group can be described by a table (called Cayley table), similar to multiplication tables which are used at school. The Cayley table for the addition of residue classes shows a very simple pattern (see the Cayley table for addition modulo 6 below):

+	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

Similar to the addition of integers, where adding a number a causes a shift by a places to the right (for positive a) or by $|a|$ places to the left (for negative a), we can visualise addition modulo m as rotating a wheel with numbers 0 to $m-1$ by a places. Therefore the groups of residue classes modulo m with respect to addition (as well as the group of all integers with respect to addition) are called *cyclic groups*.

The group of invertible residue classes modulo 7 with respect to multiplication has also 6 elements, as the group considered above. However the Cayley table does not show any obvious cyclic pattern:

·	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	1	3	5
3	3	6	2	5	1	4
4	4	1	5	2	6	3
5	5	3	1	6	4	2
6	6	5	4	3	2	1

However, reordering the factors will reveal a cyclic pattern here as well

$\cdot$	$1 \equiv 3^0$	$3 \equiv 3^1$	$2 \equiv 3^2$	$6 \equiv 3^3$	$4 \equiv 3^4$	$5 \equiv 3^5$
1	1	3	2	6	4	5
3	3	2	6	4	5	1
2	2	6	4	5	1	3
6	6	4	5	1	3	2
4	4	5	1	3	2	6
5	5	1	3	2	6	4

The key to reveal the cyclic pattern was the element 3 whose powers generate all other elements. Motivated by this we will study what elements can be represented as powers of other elements.

Let $m > 0$ be a fixed modulus and a an integer that represents an invertible residue class modulo m, i.e. $(a, m) = 1$. Then, due to Euler's theorem, there exists a number $k = \phi(m)$ such that

$$[a]^k = [1]$$

or, in other words,

$$a^k \equiv 1 \pmod{m}.$$

There may be even smaller indices than $\phi(m)$ which have the same effect. The corresponding concept is captured in the following definition.

Definition 4.1. *The* **order of a modulo m** *is the* least *positive integer x such that*

$$a^x \equiv 1 \pmod{m}.$$

We denote this number by $\operatorname{ord}_m a$. By definition,

(i) $a^{\operatorname{ord}_m a} \equiv 1 \pmod{m}$,

(ii) if $a^s \equiv 1 \pmod{m}$ then $\operatorname{ord}_m a \leq s$.

In particular we know that $1 \leq \operatorname{ord}_m a \leq \phi(m)$. The following result improves on this observation.

Theorem 4.2. *If $s > 0$ and $a^s \equiv 1 \pmod{m}$, then $\operatorname{ord}_m a \mid s$.*

Proof. By the Division Algorithm we can write

$$s = q \operatorname{ord}_m a + r, \quad 0 \leq r < \operatorname{ord}_m a\,.$$

Now $a^s \equiv 1 \pmod{m}$ so

$$1 \equiv a^{q \operatorname{ord}_m a + r} = \left(a^{\operatorname{ord}_m a}\right)^q a^r \equiv a^r \pmod{m}$$

(using property (i) above).

Since $a^r \equiv 1 \pmod{m}$ and $\operatorname{ord}_m a$ is the least positive integer with this property, we must have $r = 0$. Hence $s = q \ \operatorname{ord}_m a$ and $\operatorname{ord}_m a \mid s$. □

Corollary. If $m > 0$ and $(a, m) = 1$ then $\operatorname{ord}_m a \mid \phi(m)$.

The concept of order often causes some difficulties when first encountered. In particular it is important to remember the minimal property in the definition. Using the last result, we can rewrite the definition of order as follows, and this form is usually the one that is used in solving problems or developing further results involving order.

Definition 4.1' (Alternate Form) The number x is the order of a modulo m if and only if satisfies the **two** properties

(i) $a^x \equiv 1 \pmod{m}$,

(ii) if $a^s \equiv 1 \pmod{m}$ then $x \mid s$.

The smallest and largest possible values of $\operatorname{ord}_m a$ are 1 and $\phi(m)$ respectively. When the second case occurs it is significant, and it leads us to the definition of a primitive root.

Definition 4.3. *If $r > 0, (r, m) = 1$, and $\operatorname{ord}_m r = \phi(m)$, the number r is called a* **primitive root** *for m.*

We will see that primitive roots are exactly the elements whose powers produce all invertible elements modulo m.

Example 4.4. *If $1 \leq a \leq 10$, what are the possible values of $\operatorname{ord}_{11} a$?*

Solution. $\phi(11) = 10$ and $\operatorname{ord}_{11} a \mid 10$. Hence $\operatorname{ord}_{11} a$ has to be a divisor of 10 and is one of the numbers $1, 2, 5, 10$.

Example 4.5. *Find* $\mathrm{ord}_{11}\,2$ *and* $\mathrm{ord}_{11}\,3$. *Which of 2 and 3 is a primitive root for 11?*

Solution. (a)

$$\begin{aligned} 2^1 &= 2 \\ 2^2 &= 4 \\ 2^5 &= 32 \equiv -1 \pmod{11} \\ 2^{10} &\equiv (-1)^2 \equiv 1 \pmod{11}. \end{aligned}$$

It follows that $\mathrm{ord}_{11}\,2 = 10$ and that 2 is a primitive root for 11.
(b)

$$\begin{aligned} 3^1 &= 3 \\ 3^2 &= 9 \\ 3^3 &= 27 \equiv 5 \pmod{11} \\ 3^5 &= 3^2 \cdot 3^3 \equiv 9 \cdot 5 \equiv 1 \pmod{11}. \end{aligned}$$

It follows that $\mathrm{ord}_{11}\,3 = 5$ and that 3 is not a primitive root for 11.
Remark. Observe how we only have to check whether $1, 2, 5, 10$ are the orders, since by the Corollary, $\mathrm{ord}_{11}\,a \mid \phi(11)$.
Remark. If $a \equiv b \pmod{m}$ then $\mathrm{ord}_m\,a = \mathrm{ord}_m\,b$. We are usually only interested in the numbers $1 \leq a \leq m-1$ as candidates for primitive roots.

Example 4.6. *Find* $\mathrm{ord}_m\,1$ *and* $\mathrm{ord}_m(m-1)$. *What happens if* $m = 2$*?*

Solution. $1^1 = 1 \pmod{m}$ so $\mathrm{ord}_m\,1 = 1$.
$(m-1)^2 \equiv (-1)^2 \equiv 1 \pmod{m}$ so $\mathrm{ord}_m(m-1) = \mathrm{ord}_m(-1) = 2$ when $m > 2$.
When $m = 2$, $1 \equiv -1 \pmod{2}$ so $\mathrm{ord}_2(-1) = 1$.

Example 4.7. *Show that 8 does not have any primitive roots.*

Solution. If $(a, 8) = 1$ then a is congruent to one of $1, 3, 5, 7 \pmod{8}$. Now $\phi(8) = \phi(2^3) = 2^3 - 2^2 = 4$. However, $1^2 \equiv 3^3 \equiv 5^2 \equiv 7^2 \equiv 1 \pmod{8}$ so none of these numbers has the required order 4. Hence there are no primitive roots for 8.

Theorem 4.8. *Suppose* $m > 0$ *and* $(a, m) = 1$. *Let* i, j *be integers such that* $i, j \geq 0$. *Then* $a^i \equiv a^j \pmod{m}$ *holds if and only if* $i \equiv j \pmod{\mathrm{ord}_m a}$.

Proof. Suppose $a^i \equiv a^j \pmod{m}$. Without loss of generality, we can assume that $i \geq j$. Then since $(a, m) = 1$, the term a can be cancelled j times to give

$$a^{i-j} \equiv 1 \pmod{m}.$$

By Theorem 4.2 it follows that $\mathrm{ord}_m a \mid (i - j)$ and so $i \equiv j \pmod{\mathrm{ord}_m\,a}$.

Conversely suppose that $i \equiv j \pmod{\operatorname{ord}_m a}$, and that $i \geq j \geq 0$. Then $i = j + k \operatorname{ord}_m a$ for some integer k and

$$a^i = a^j (a^{\operatorname{ord}_m a})^k \equiv a^j \pmod{m}. \qquad \square$$

Now we are ready to prove that the powers of a primitive root generate all invertible residue classes.

Theorem 4.9. *If r is a primitive root modulo m, then the list of integers*

$$r^1, r^2, \ldots, r^{\phi(m)}$$

form a reduced residue system modulo m.

Proof. Since $(r, m) = 1$ it follows that $(r^k, m) = 1$ for $k = 1, 2, \ldots, \phi(m)$.

Hence we have $\phi(m)$ integers each of which is relatively prime to m. It remains to show that no two integers in the list are congruent modulo m.

Suppose $r^i \equiv r^j \pmod{m}$ where $1 \leq i, j \leq \phi(m)$. By Theorem 4.8 we have $i \equiv j \pmod{\operatorname{ord}_m r}$. But $\operatorname{ord}_m r = \phi(m)$ since r is a primitive root so $i \equiv j \pmod{\phi(m)}$. Since both i, j lie between 1 and $\phi(m)$ it follows that $i = j$. Hence all the numbers in the list are distinct modulo m, and they form a reduced residue system modulo m. $\square$

Example 4.10. *Verify that 3 is a primitive root modulo 7 and that the powers $3^k, 1 \leq k \leq 6$, form a reduced residue system modulo 7.*

Solution.

$$\begin{aligned}
3^1 &= 3 \\
3^2 &= 9 \equiv 2 \pmod{7} \\
3^3 &\equiv 3 \cdot 2 = 6 \pmod{7} \\
3^4 &\equiv 3 \cdot 6 \equiv 4 \pmod{7} \\
3^5 &\equiv 3 \cdot 4 \equiv 5 \pmod{7} \\
3^6 &\equiv 3 \cdot 5 \equiv 1 \pmod{7}.
\end{aligned}$$

Thus $\operatorname{ord}_7 3 = 6 = \phi(7)$. Also the powers are congruent to $3, 2, 6, 4, 5, 1$ and form a reduced residue system modulo 7.

Finally in this section, we give a useful result about the order of powers which will be used later.

Theorem 4.11. *Suppose $(a, m) = 1$ and $\operatorname{ord}_m a = t$. Then a^k has order $\frac{t}{(k,t)}$.*

Proof. Suppose $(k, t) = l \geq 1$. Then we can write $k = k_0 l$, $t = t_0 l$ where $(k_0, t_0) = 1$. With this notation, we have to show that t_0 is the order of a^k, and to do this we will make repeated use of the criteria in the alternate form of the definition of order, Definition 4.1'.

Denote $s = \text{ord}_m(a^k)$. We prove $s = t_0$ by showing $s|t_0$ and $t_0|s$.

We have $a^t \equiv 1 \pmod{m}$ from the definition of t. Hence

$$(a^k)^{t_0} = a^{k_0 l t_0} = (a^t)^{k_0} \equiv 1 \pmod{m}.$$

From this it follows that $s \mid t_0$.

On the other hand $(a^k)^s \equiv 1 \pmod{m}$ from the definition of s. Hence $a^{ks} = (a^k)^s \equiv 1 \pmod{m}$. From this it follows that $t \mid ks$. Hence $t_0 l \mid k_0 l s$, that is, $t_0 \mid k_0 s$. Since $(t_0, k_0) = 1$ this implies $t_0 \mid s$.

From $s \mid t_0$ and $t_0 \mid s$ we conclude that $t_0 = s$ as required. □

Corollary. Suppose $(a, m) = 1$ and $\text{ord}_m a = s$. Then a^k has order s if and only if $(k, s) = 1$.

Corollary. If there is a primitive root a for modulus m then a^k is a primitive root if and only if $(k, \phi(m)) = 1$, i.e. there are exactly $\phi(\phi(m))$ primitive roots. In other words, either there is no primitive root at all or there are $\phi(\phi(m))$ primitive roots.

In fact, if r is a primitive root, then any other primitive root must be of the form r^k. But, according to the Corollary above, r^k is a primitive root if and only if k is a coprime of $\phi(m) = \text{ord}_m r$. Hence there are exactly $\phi(\phi(m))$ choices for k.

4.2 The Index

Suppose we have a primitive root r modulo m. If $(a, m) = 1$ then by Theorem 4.9, a must be congruent to one of the numbers $r, r^2, \ldots, r^{\phi(m)}$.

Definition 4.12. *If $(a, m) = 1$, and r is a primitive root modulo m, the* **index of a to the base r** *is the number x for which $a \equiv r^x \pmod{m}$ and $1 \le x \le \phi(m)$.*

Notation. We write $\text{ind}_r a$ for the number x in the above definition.

From Example 4.10, we can write down the following table of indices for the primitive root 3 modulo 7.

a	1	2	3	4	5	6
$\text{ind}_3 a$	6	2	1	4	5	3

To say that $\text{ind}_3 6 = 3$ is to say that $3^3 \equiv 6 \pmod{7}$ and you get this from the last column of the table.

The index has a number of properties which are analogous to those of the logarithm.

Note first of all that if $a \equiv b \pmod{m}$ then from the definition,

$$\text{ind}_r a \;=\; \text{ind}_r b.$$

Theorem 4.13 (Rules for Indices)**.** *Let r be a primitive root for $m > 0$, and suppose that $(a, m) = (b, m) = 1$. Then*

(i) $\operatorname{ind}_r 1 \equiv 0 \pmod{\phi(m)}$,

(ii) $\operatorname{ind}_r(ab) \equiv \operatorname{ind}_r a + \operatorname{ind}_r b \pmod{\phi(m)}$,

(iii) $\operatorname{ind}_r(a^k) \equiv k \operatorname{ind}_r a \pmod{\phi(m)}$ *where* $k > 0$.

Proof. (i) $r^{\phi(m)} \equiv 1 \pmod{m}$ by Euler's Theorem. Hence $\operatorname{ind}_r 1 = \phi(m) \equiv 0 \pmod{\phi(m)}$.

(ii) let $x = \operatorname{ind}_r a$, $y = \operatorname{ind}_r b$, $z = \operatorname{ind}_r(ab)$.

Then $r^x \equiv a \pmod{m}$ and $r^y \equiv b \pmod{m}$ so

$$r^x r^y = r^{x+y} \equiv ab \pmod{m}.$$

Also $r^z \equiv ab \pmod{m}$ so $r^z \equiv r^{x+y} \pmod{m}$. Using Theorem 4.8 we get

$$z \equiv x + y \pmod{\phi(m)}$$

which shows that (ii) holds.

(iii) Let $x = \operatorname{ind}_r a$ and $z = \operatorname{ind}_r(a^k)$. Then $r^x \equiv a \pmod{m}$ and raising both sides to the power k we get

$$r^{xk} \equiv a^k \pmod{m}.$$

But $r^z \equiv a^k \pmod{m}$ by definition so again by Theorem 4.8 we have

$$z \equiv xk \pmod{\phi(m)},$$

that is, $\operatorname{ind}_r(a^k) \equiv k \operatorname{ind}_r a \pmod{\phi(m)}$. □

Notice that indices work modulo $\phi(m)$ rather than modulo m, and that the above rules are like those for the logarithm.

Example 4.14. *Given that 3 is a primitive root for 17, construct a table of indices for $p = 17, r = 3$. Use the table and the index rules to solve the following congruences.*

(i) $5x^9 \equiv 10 \pmod{17}$,

(ii) $6x^{12} \equiv 11 \pmod{17}$,

(iii) $7^x \equiv 6 \pmod{17}$.

Solution. We calculate the powers $3^1, 3^2, \ldots$ and reduce them modulo 17 whenever a term is more than 17. This way, we keep the calculations down to simple ones which can

be done by mental arithmetic. After each line, we can enter the result in the table shown below.

$$\begin{aligned} 3^1 &= 3 && \text{so} && \operatorname{ind}_3 3 = 1 \\ 3^2 &= 9 && \text{so} && \operatorname{ind}_3 9 = 2. \\ 3^3 &= 27 \equiv 10 \pmod{17} && \text{so} && \operatorname{ind}_3 10 = 3. \end{aligned}$$

From now on we will just show the result and enter it into the table.

$$\begin{aligned} 3^4 &\equiv 3 \cdot 10 = 30 \equiv 13 && \pmod{17} \\ 3^5 &\equiv 39 \equiv 5 && \pmod{17} \\ 3^6 &\equiv 15 && \pmod{17} \\ 3^7 &\equiv 45 \equiv 11 && \pmod{17} \\ 3^8 &\equiv 33 \equiv 16 && \pmod{17} \\ 3^9 &\equiv 48 \equiv 14 && \pmod{17} \\ 3^{10} &\equiv 42 \equiv 8 && \pmod{17} \\ 3^{11} &\equiv 24 \equiv 7 && \pmod{17} \\ 3^{12} &\equiv 21 \equiv 4 && \pmod{17} \\ 3^{13} &\equiv 12 && \pmod{17} \\ 3^{14} &\equiv 36 \equiv 2 && \pmod{17} \\ 3^{15} &\equiv 6 && \pmod{17} \\ 3^{16} &\equiv 18 \equiv 1 && \pmod{17}. \end{aligned}$$

a	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\operatorname{ind}_3 a$	16	14	1	12	5	15	11	10	2	3	7	13	4	9	6	8

The last result of course follows from Fermat's Little Theorem. The fact that our table gets filled up with 1 as the last entry and no repetitions confirms that 3 is a primitive root for 17.

(i) To solve $5x^9 \equiv 10 \pmod{17}$, take indices of both sides to the base 3. Thus

$$\operatorname{ind}_3(5x^9) \equiv \operatorname{ind}_3 10 \pmod{16}.$$

Now using the index laws and reading from the table, we get

$$\begin{aligned} \operatorname{ind}_3 5 + 9 \operatorname{ind}_3 x &\equiv \operatorname{ind}_3 10 && \pmod{16} \\ 5 + 9 \operatorname{ind}_3 x &\equiv 3 && \pmod{16} \\ 9 \operatorname{ind}_3 x &\equiv -2 && \pmod{16}. \end{aligned}$$

This is a linear congruence where we treat $\text{ind}_3\, x$ as the unknown. Now $(9, 16) = 1$ so there is a unique solution modulo 16. Note that an inverse of 9 modulo 16 is 9 itself. On multiplying the congruence by 9 we get

$$\text{ind}_3\, x \equiv -18 \equiv 14 \pmod{16}.$$

Reading from the table, we see that $x = 2$ satisfies $\text{ind}_3\, x = 14$.

All solutions are given by

$$x \equiv 2 \pmod{17}.$$

(ii)

$$\begin{aligned} 6x^{12} &\equiv 11 && \pmod{17} \\ \text{ind}_3\, 6 + 12\,\text{ind}_3\, x &\equiv \text{ind}_3\, 11 && \pmod{16} \\ 15 + 12\ \text{ind}_3\, x &\equiv 7 && \pmod{16} \\ 12\ \text{ind}_3\, x &\equiv -8 \equiv 8 && \pmod{16}. \end{aligned}$$

This is a linear congruence where $\text{ind}_3\, x$ is the unknown.

Since $(12, 16) = 4$ and $4 \mid 8$, there are four incongruent solutions modulo 16. By inspection one solution for $\text{ind}_3\, x$ is $\text{ind}_3\, x = 2$. Then the other solutions are $6, 10, 14$. From the table the four corresponding solutions for x are $x = 9, x = 15, x = 8, x = 2$. All solutions of the original problem are given by

$$x \equiv 2, 8, 9, 15 \pmod{17}.$$

(iii) $7^x \equiv 6 \pmod{17}$.

As before, take indices of both sides to the base 3.

$$x\ \text{ind}_3\, 7 \equiv \text{ind}_3\, 6 \pmod{16}.$$

From the table, $\text{ind}_3\, 7 = 11$ and $\text{ind}_3\, 6 = 15$. Hence the congruence is

$$11x \equiv 15 \pmod{16}.$$

Now the inverse of 11 modulo 16 is 3. On multiplying through by 3 we get

$$x \equiv 45 \equiv 13 \pmod{16}.$$

4.3 Existence of primitive roots for prime modulus

The main goal of this section is to prove that there is always a primitive root for a prime modulus p. However some of our results apply for arbitrary modulus. For a given prime

p, we will count the number of elements of order d, if any. Such elements will be solutions of the polynomial congruence

$$x^d - 1 \equiv 0 \pmod{p}.$$

Therefore we first of all need to examine polynomial congruences with a prime modulus.

Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial with integral coefficients $a_i \in \mathbb{Z}$ and with $a_n \neq 0$. An integer c is called a **root of $f(x)$ modulo m** if

$$f(c) \equiv 0 \pmod{m}.$$

We know from Chapter 2 that if $b \equiv c \pmod{m}$ and if c is a root modulo m, then b is also a root modulo m. Hence the main interest is in determining the roots among a complete residue system. These will be referred to as the incongruent roots modulo m.

Example 4.15. *Let $m = 8, f(x) = x^2 - 1$. Then $f(x) \equiv 0 \pmod{m}$ has four incongruent roots $1, 3, 5, 7$ modulo 8.*

When the modulus is prime, the number of roots is at most the degree of the polynomial, as stated in the next result.

Theorem 4.16 (Lagranges Theorem)**.** *Suppose p is prime and $p \nmid a_n$. Then the integral polynomial $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ has at most n incongruent roots modulo p.*

Proof. The proof is by induction on n, the degree of the polynomial. For $n = 1$ the polynomial congruence is $a_1 x + a_0 \equiv 0 \pmod{p}$ where $p \nmid a_1$. This is the same as $a_1 x \equiv -a_0 \pmod{p}$ where $(a_1, p) = 1$. This is a linear congruence and since $(a_1, p) = 1$ there is a unique solution modulo p. Hence the result holds for $n = 1$.

Suppose now that the statement of the theorem holds for all polynomials of degree $n - 1$. Let

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

be a polynomial of degree n with $p \nmid a_n$. If $f(x)$ has no roots the result holds for $f(x)$. Otherwise $f(x)$ has a root, say c_0, modulo p. Then

$$\begin{aligned} f(x) - f(c_0) &= a_n(x^n - c_0^n) + a_{n-1}(x^{n-1} - c_0^{n-1}) + \cdots + a_1(x - c_0) \\ &= a_n(x - c_0)(x^{n-1} + x^{n-2}c_0 + \cdots + c_0^{n-1}) \\ &\qquad + a_{n-1}(x - c_0)(x^{n-2} + x^{n-3}c_0 + \cdots + c_0^{n-2}) \\ &\qquad\quad + \cdots + a_1(x - c_0) \\ &= (x - c_0)g(x) \end{aligned}$$

where $g(x)$ has degree $n - 1$ and leading coefficient a_n.

Suppose now that c is another root of $f(x)$ with $c \not\equiv c_0 \pmod{p}$. Then

$$f(c) - f(c_0) = (c - c_0)g(c) \equiv 0 \pmod{p}.$$

Since $c - c_0 \not\equiv 0 \pmod{p}$ it follows that $g(c) \equiv 0 \pmod{p}$. Hence the roots $f(x)$ are c_0 together with all the roots of $g(x)$. By the induction hypothesis, $g(x)$ has at most $n-1$ incongruent roots modulo p, hence $f(x)$ has at most n incongruent roots modulo p.

By induction the result holds for all $n \geq 1$. □

Theorem 4.17. *Suppose p is a prime and there is an element a which has order d modulo p. Then there are exactly $\phi(d)$ incongruent elements of order d.*

Proof. Any element of order d satisfies the congruence

$$x^d - 1 \equiv 0 \pmod{p}.$$

According to Lagrange's theorem it has at most d incongruent solutions. We show that the d integers

$$a, a^2, \ldots, a^d.$$

are these incongruent solutions. In fact, each integer in the list is a root of the polynomial congruence $x^d \equiv 1 \pmod{p}$ since

$$(a^k)^d \equiv (a^d)^k \equiv 1 \pmod{p}.$$

Moreover, the integers in the above list are incongruent modulo p. For if $a^i \equiv a^j \pmod{p}$ where $1 \leq i \leq j \leq d$, then $a^{j-i} \equiv 1 \pmod{p}$. Since the order of a is d and $0 \leq i - j < d$, this forces $j = i$.

Hence, all elements of order d must be contained in the list.

By the Corollary to Theorem 4.11, a^k has order d if and only if $(k, d) = 1$. There are $\phi(d)$ integers k with $1 \leq k \leq d$ and $(k, d) = 1$. Hence there are $\phi(d)$ incongruent elements of order d. □

Theorem 4.18. *Let p be a prime. Then modulo p, for each positive divisor d of $p-1$, there are exactly $\phi(d)$ incongruent integers of order d.*

Proof. For each positive divisor d of $p-1$ let $F(d)$ be the number of integers x such that $1 \leq x \leq p-1$ and $\text{ord}_p x = d$.

Then $p - 1 = \sum_{d|p-1} F(d)$ since each of the integers x with $1 \leq x \leq p-1$ has an order which divides $p-1$, so each is accounted for in the above sum.

We also know from Theorem 3.22 that

$$p - 1 = \sum_{d|p-1} \phi(d).$$

Hence

$$\sum_{d|p-1} F(d) = \sum_{d|p-1} \phi(d) \qquad (1)$$

Next we show that $F(d) \leq \phi(d)$ for each $d \mid p-1$. If $F(d) = 0$ the result holds trivially. If $F(d) \neq 0$ there is at least one element of order d and then by Theorem 4.18 there are exactly $\phi(d)$ incongruent integers of order d. In either case $F(d) \leq \phi(d)$.

Finally we must have $F(d) = \phi(d)$ for each positive divisor d of $p-1$, for if $F(d) = 0$ for some d and $F(d) \leq \phi(d)$ for all d we could not have (1) holding.

Hence for each positive divisor d of $p-1$ there are $F(d) = \phi(d)$ incongruent integers of order d. □

Theorem 4.19. *Each prime p has exactly $\phi(p-1)$ primitive roots.*

Proof. Let $d = p-1$. Then $d \mid p-1$ so by Theorem 4.18 there are $\phi(p-1)$ elements of order $p-1$. By definition, each of these $\phi(p-1)$ numbers is a primitive root, and modulo p, is incongruent to the others. □

The last two results not only tell us that there are primitive roots for primes but also tells us how many. Sometimes it is only the existence part of the above which is used so it is worth stating this as a separate result.

Corollary. (i) Each prime p has a primitive root.
(ii) For each divisor d of $p-1$ there is an integer of order d modulo p.

We saw earlier that 8 does not have a primitive root so the existence of primitive roots for primes is in fact quite a strong result.

The following Theorem which we cite without proof was obtained by Gauss and answers the question what numbers have primitive roots completely.

Theorem 4.20. *An integer $m > 1$ has a primitive root if and only if $m = 2$, or $m = 4$, or $m = p^n$, or $m = 2p^n$ where p is an odd prime and n is a positive integer.*

Example 4.21. *Suppose p is a prime of the form $p = 8k+1$. Show that there is an integer x such that $x^4 \equiv -1 \pmod{p}$.*

Solution. We observe that $p - 1 = 8k$ so $8 \mid p-1$. By theorem 4.18 there is an element x of order 8 modulo p. Thus $x^8 \equiv 1 \pmod{p}$ so $p \mid (x^8 - 1)$, that is, $p \mid (x^4+1)(x^4-1)$. Since p is prime, either $p \mid (x^4+1)$ or $p \mid (x^4-1)$.

But we cannot have $p \mid (x^4-1)$ for this would mean $x^4 \equiv 1 \pmod{p}$ which contradicts the fact that x has order 8.

It follows that $p \mid (x^4+1)$ and hence $x^4 \equiv -1 \pmod{p}$.

Example 4.22. *Let $p = 11$. The divisors of $p-1$ are 1, 2, 5, 10. Since $\phi(1) = 1$, $\phi(2) = 1$, $\phi(5) = 4, \phi(10) = 4$, the theorem tells us that modulo 11, there is one number of order 1, one number of order 2, 4 numbers of order 5, and 4 primitive roots for 11.*

Example 4.23. *How many primitive roots does 101 have?*

Solution. 101 is prime so there are $\phi(100)$ primitive roots. Now $100 = 2^2 \cdot 5^2$ so there are $\phi(2^2)\phi(5^2) = (2^2-2)(5^2-5) = 40$ primitive roots.

Chapter 5

QUADRATIC RESIDUES

5.1 Quadratic Congruences

We considered the linear congruence in Chapter 2 and determined a complete solution to that problem. The next simplest congruence is the quadratic congruence

$$ax^2 + bx + c \equiv 0 \pmod{m}$$

where a, b, c are integers, and $a \not\equiv 0 \pmod{m}$.

It turns out that this is a much more difficult problem than the linear congruence. The next example illustrates some of the issues involved in a simple case.

Example 5.1. *Write down all the quadratic congruences which are distinct modulo 2 and determine their solutions.*

Solution. Note that 0 and 1 form a complete residue system modulo 2, so the coefficient of x^2 has to be 1. The other coefficients can be either 0 or 1 so modulo 2 there are only four quadratic congruences. We can determine which of 0, 1 are solutions by trial and error. The results are as follows.

(i) $x^2 + x + 1 \equiv 0 \pmod{2}$ has no solutions.

(ii) $x^2 + x \equiv 0 \pmod{2}$ has two solutions $x \equiv 0 \pmod{2}$, and $x \equiv 1 \pmod{2}$.

(iii) $x^2 + 1 \equiv 0 \pmod{2}$ has one solution $x \equiv 1 \pmod{2}$.

(iv) $x^2 \equiv 0 \pmod{2}$ has one solution $x \equiv 0 \pmod{2}$.

In most cases of interest, we can use completion of squares to first simplify the quadratic. Multiply the general equation above by $4a$ to give

$$4a^2x^2 + 4abx + 4ac \equiv 0 \pmod{m}.$$

Provided $(4a, m) = 1$ we do not change the solutions. Now complete squares to get

$$(2ax + b)^2 + 4ac - b^2 \equiv 0 \pmod{m}.$$

Letting $u = 2ax + b$ and $k = b^2 - 4ac$, we then have to solve

$$u^2 \equiv k \pmod{m}.$$

If we can solve this for u, we can then go back and solve

$$2ax + b = u \pmod{m}$$

for x.

From now on we will only consider the problem of solving

$$x^2 \equiv a \pmod{m}$$

and our main interest is in determining whether or not there are solutions.

Definition 5.2. *Let $m > 1$. An integer a is called a* **quadratic residue of m** *if $(a, m) = 1$ and there is a solution of $x^2 \equiv a \pmod{m}$. Otherwise, if there is no solution, a is called a* **quadratic nonresidue of m**.

We note that if $a \equiv b \pmod{m}$, then b is a quadratic residue of m if and only if a is a quadratic residue of m.

Let $m = p_1^{\ell_1} \cdots p_k^{\ell_k}$ be the prime decomposition of m. The following theorem reduces the problem of quadratic residues to powers of primes.

Theorem 5.3. *a is a quadratic residue of m if and only if a is a quadratic residue for all $p_1^{\ell_1}, \ldots, p_k^{\ell_k}$.*

Proof. First we notice that $(a, m) = 1$ if and only if $(a, p_s^{\ell_s}) = 1$ for $s = 1, \ldots, k$. In fact, if there was a prime p that divides both a and m then p had to coincide with one of the p_s and would divide both a and $p_s^{\ell_s}$. Conversely, if p divides a and one of the $p_s^{\ell_s}$ then $p = p_s$ and $p|m$.

Now, $a \equiv x^2 \pmod{m}$ means $m|a - x^2$. But then $p_s^{\ell_s}|a - x^2$. Thus, if a is a quadratic residue of m it will be a quadratic residue for all $p_s^{\ell_s}$.

Conversely, suppose a is a quadratic residue for all s, i.e. there are x_s such that $a \equiv x_s^2 \pmod{p_s^{\ell_s}}$. Then, due to the Chinese remainder theorem there is an integer x such that

$$\begin{aligned} x &\equiv x_1 \pmod{p_1^{\ell_1}} \\ &\cdots \\ x &\equiv x_k \pmod{p_k^{\ell_k}} \end{aligned}$$

Now, $a \equiv x^2 \pmod{p_s^{\ell_s}}$, i.e. $p_s^{\ell_s}|a - x^2$. Then $m|a - x^2$, thus, a is a quadratic residue for m. □

A further reduction to quadratic residues of odd primes is established by the following theorem which we cite without proofs.

Theorem 5.4. *a is a quadratic residue for 2^ℓ with $\ell \geq 3$ if and only if $a \equiv 1 \pmod 8$.*

The missing powers 2 and $2^2 = 4$ are easily dealt with. For $m = 2$ only the odd numbers a satisfy $(a, 2) = 1$. They are quadratic residues since $a \equiv 1^2 \pmod 2$. For $m = 4$ the numbers a with $a \equiv \pm 1 \pmod 4$ satisfy $(a, 4) = 1$. Direct verification shows that $a \equiv 1^2 \pmod 4$ for $a \equiv 1 \pmod 4$ and that there is no x with $a \equiv x^2 \pmod 4$ for $a \equiv -1 \pmod 4$. Thus the quadratic residues for $m = 4$ are the integers a with $a \equiv 1 \pmod 4$.

For powers of odd primes we have

Theorem 5.5. *a is a quadratic residue for the odd power p^k if and only if it is a quadratic residue for the odd prime p.*

The theorems above justify that we may restrict our interest to the case of an odd prime modulus.

5.2 The Legendre Symbol

Definition 5.6. *Let p be an odd prime, let a be an integer such that $p \nmid a$. The* **Legendre Symbol** $\left(\dfrac{\boldsymbol{a}}{\boldsymbol{p}}\right)$ *is defined by*

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is a quadratic residue of } p, \\ -1 & \text{if } a \text{ is a quadratic nonresidue of } p. \end{cases}$$

Thus $\left(\dfrac{a}{p}\right)$ is ± 1 according as to whether there are, or are not, solutions of $x^2 \equiv a \pmod p$.

For small p there is only a small number of possibilities and we can proceed by examining each one. The process is illustrated in the next example.

Example 5.7. *Evaluate $\left(\dfrac{a}{13}\right)$ for $1 \leq a \leq 12$.*

Solution. We can reduce the number of calculations by doing two at a time, using $x^2 \equiv (13 - x)^2 \pmod{13}$. For example, $11^2 = (13 - 2)^2 \equiv 2^2 \equiv 4 \pmod{13}$. Now

$$\begin{aligned}
1^2 &\equiv 12^2 \equiv 1 && \pmod{13} \\
2^2 &\equiv 11^2 \equiv 4 && \pmod{13} \\
3^2 &\equiv 10^2 \equiv 9 && \pmod{13} \\
4^2 &\equiv 9^2 \equiv 16 \equiv 3 && \pmod{13} \\
5^2 &\equiv 8^2 \equiv 25 \equiv 12 && \pmod{13} \\
6^2 &\equiv 7^2 \equiv 36 \equiv 10 && \pmod{13}.
\end{aligned}$$

The numbers down the right hand side are $1, 3, 4, 9, 10, 12$ and these are the quadratic residues. The other numbers $2, 5, 6, 7, 8, 11$ are the quadratic nonresidues.

Using the Legendre symbol, the results can be expressed as

$$\begin{aligned}\left(\frac{1}{13}\right) = \left(\frac{3}{13}\right) = \left(\frac{4}{13}\right) = \left(\frac{9}{13}\right) = \left(\frac{10}{13}\right) = \left(\frac{12}{13}\right) &= 1,\\ \left(\frac{2}{13}\right) = \left(\frac{5}{13}\right) = \left(\frac{6}{13}\right) = \left(\frac{7}{13}\right) = \left(\frac{8}{13}\right) = \left(\frac{11}{13}\right) &= -1.\end{aligned}$$

For large p, this is a very inefficient method and we will develop better methods for evaluating the Legendre symbol.

First we will derive some information about the distribution of quadratic residues and non-residues.

Theorem 5.8. *If $m > 2$ then the number of quadratic residues between 1 and $m-1$ is not bigger than $\dfrac{\phi(m)}{2}$.*

Proof. The quadratic residues between 1 and $m-1$ are the least positive remainders modulo m of $x_1^2, x_2^2, \ldots x_{\phi(m)}^2$, where $x_1, \ldots, x_{\phi(m)}$ is a reduced residue system modulo m with $1 \leq x_s \leq m-1$. As observed above $x_s \equiv (m - x_s) \pmod{m}$. Thus, there are at most half as many quadratic residues than elements of the reduced system. □

Notice that the theorem above is not true for $m = 2$. There is one quadratic residue and 1 is bigger than $\dfrac{\phi(2)}{2} = \dfrac{1}{2}$.

Next we will show that if m is a prime there are exactly $\dfrac{\phi(m)}{2} = \dfrac{p-1}{2}$ quadratic residues and $\dfrac{p-1}{2}$ non-residues.

Theorem 5.9. *Let p be an odd prime. Then there are exactly $\dfrac{p-1}{2}$ quadratic residues of p and $\dfrac{p-1}{2}$ quadratic nonresidues of p among the integers $1, 2, \cdots, p-1$.*

Proof. Since p is a prime it has a primitive root a. Hence, the numbers $1, \ldots, p-1$ represent the same residue classes as $a^1, \ldots, a^{p-1}$. Now, the even powers of a are quadratic residues since

$$x^2 \equiv a^{2r} = (a^r)^2 \pmod{p}$$

has the solution a^r. For $p > 2$ there are exactly $\dfrac{p-1}{2}$ even powers of a. The remaining $\dfrac{p-1}{2}$ odd powers must be quadratic non-residues according to Theorem 5.8 □

Theorem 5.10 (Euler's Criterion)**.** *Let p be an odd prime, and suppose $p \nmid a$. Then*

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \pmod{p}.$$

Proof. $\left(\frac{a}{p}\right) = \pm 1$. We consider the two cases separately.

Suppose $\left(\frac{a}{p}\right) = 1$. Then there is an x_0 such that $x_0^2 \equiv a \pmod{p}$. Hence

$$\begin{aligned} a^{\frac{p-1}{2}} &\equiv (x_0^2)^{\frac{p-1}{2}} \pmod{p} \\ &\equiv x_0^{p-1} \pmod{p} \\ &\equiv 1 \pmod{p} \end{aligned}$$

(by Fermat's Little Theorem).

Hence, when $\left(\frac{a}{p}\right) = 1$, we get $a^{\frac{p-1}{2}} \equiv \left(\frac{a}{p}\right) \pmod{p}$.

Suppose now that $\left(\frac{a}{p}\right) = -1$. This means that there are no solutions of $x^2 \equiv a \pmod{p}$. Consider the integers i such that $1 \leq i \leq p-1$. Since $(i, p) = 1$ each congruence

$$ix \equiv a \pmod{p}$$

has a solution x with $1 \leq x \leq p-1$. This x has to be different to the i, that is $i \neq x$, for otherwise x would be a solution of $x^2 \equiv a \pmod{p}$. Hence we can group the integers $1, 2, \ldots, p-1$ into $\left(\frac{p-1}{2}\right)$ pairs with each pair having a product which is congruent to $a \pmod{p}$. Multiplying the pairs together we get

$$1 \cdot 2 \cdot 3 \cdot \cdots \cdot (p-1)! \equiv \underbrace{a \cdot a \cdot \cdots \cdot a}_{(p-1)/2 \text{ times}} \pmod{p}$$

that is, $(p-1)! \equiv a^{\frac{p-1}{2}} \pmod{p}$. By Wilson's Theorem, $(p-1)! \equiv -1 \pmod{p}$. Hence when $\left(\frac{a}{p}\right) = -1$ we also get $a^{\frac{p-1}{2}} \equiv \left(\frac{a}{p}\right) \pmod{p}$. □

Example 5.11. *Use Euler's Criterion to evaluate* $\left(\frac{8}{13}\right)$.

Solution. $\left(\frac{8}{13}\right) \equiv 8^{\frac{13-1}{2}} \equiv 8^6 \pmod{13}$.

Now $8^2 = 64 \equiv -1 \pmod{13}$.

Hence $8^6 \equiv (-1)^3 \equiv -1 \pmod{13}$, and therefore $\left(\frac{8}{13}\right) = -1$.

It follows that there are no solutions of $x^2 \equiv 8 \pmod{13}$.

We will rarely use Euler's Criterion to solve numerical problems like the last one. The criterion is more of a theoretical tool and is used in the next two theorems.

Theorem 5.12 (Properties of the Legendre Symbol)**.** *Let p be an odd prime, and let a and b be integers such that $p \nmid a, p \nmid b$. The following properties hold.*

(i) If $a \equiv b \pmod{p}$ then $\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$.

(ii) $\left(\frac{a}{p}\right)\left(\frac{b}{p}\right) = \left(\frac{ab}{p}\right)$.

(iii) $\left(\frac{a^2}{p}\right) = 1$.

Proof. (i) If $a \equiv b \pmod{p}$ then $x^2 \equiv a \pmod{p}$ if and only if $x^2 \equiv b \pmod{p}$. Hence $x^2 \equiv a \pmod{p}$ has a solution if and only if $x^2 \equiv b \pmod{p}$ has a solution and it follows that $\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$.

(ii)

$$\begin{aligned}
\left(\frac{ab}{p}\right) &\equiv (ab)^{\frac{p-1}{2}} && \pmod{p} && \text{by Euler's Criterion} \\
&\equiv a^{\frac{p-1}{2}} b^{\frac{p-1}{2}} && \pmod{p} && \\
&\equiv \left(\frac{a}{p}\right)\left(\frac{b}{p}\right) && \pmod{p} && \text{again using Euler's Criterion.}
\end{aligned}$$

(iii) $\left(\frac{a^2}{p}\right) \equiv \left(\frac{a}{p}\right)\left(\frac{a}{p}\right) = 1$ since $\left(\frac{a}{p}\right)$ is either 1 or -1. □

Example 5.13. *Evaluate* $\left(\frac{413}{101}\right)$.

Solution. 101 is prime and $413 \equiv 9 \pmod{101}$. Hence

$$\left(\frac{413}{101}\right) = \left(\frac{9}{101}\right) = \left(\frac{3^2}{101}\right) = 1$$

on using parts (i) and (iii) of the last theorem.

On the other hand, if we try the same technique to evaluate, say $\left(\frac{410}{101}\right)$, we get

$$\left(\frac{410}{101}\right) = \left(\frac{6}{101}\right) = \left(\frac{2 \cdot 3}{101}\right) = \left(\frac{2}{101}\right)\left(\frac{3}{101}\right)$$

on using part (ii) of the theorem.

In general, repeated use of Theorem 5.12 reduces the problem to evaluating $\left(\frac{-1}{p}\right)$ and $\left(\frac{q}{p}\right)$ for primes q with $q < p/2$. In fact, without loss of generality we may assume that $a < p$ by replacing a by its least positive remainder modulo p. If $a > p/2$ then we can replace a by $a - p = -(p - a)$ where $p - a < p/2$. Now let $a = \pm q_1^{k_1} \cdots q_\ell^{k_\ell}$ be the prime decomposition of a then, according to Theorem 5.12,

$$\left(\frac{a}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{q_1}{p}\right)^{k_1} \cdots \left(\frac{q_\ell}{p}\right)^{k_\ell}.$$

Here we can drop all factors with even exponent k_i and replace the odd exponents k_i by 1.

Hence we will need to be able to compute $\left(\frac{-1}{p}\right)$, $\left(\frac{2}{p}\right)$, $\left(\frac{q}{p}\right)$, where q is an odd prime. Below we solve the first two problems. The next section is devoted to the third problem.

Theorem 5.14. *Let p be an odd prime. Then*

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}} = \begin{cases} 1 \text{ if } p \equiv 1 \pmod 4, \\ -1 \text{ if } p \equiv -1 \pmod 4. \end{cases}$$

Proof. By Euler's Criterion,

$$\left(\frac{-1}{p}\right) \equiv (-1)^{\frac{p-1}{2}} \pmod p.$$

On the right hand side we get ± 1 according to whether $\frac{p-1}{2}$ is even or odd. If $p = 4k + 1$ for some $k \in \mathbb{Z}$ then $\frac{p-1}{2} = 2k$ which is even. If $p = 4k + 3$ for some $k \in \mathbb{Z}$ then $\frac{p-1}{2} = 2k + 1$ which is odd.

The stated result follows. □

Example 5.15. *Find* $\left(\frac{102}{103}\right)$.

Solution. 103 is prime and $103 \equiv 3 \pmod 4$. Hence $\left(\frac{102}{103}\right) = \left(\frac{-1}{103}\right) = -1$.

The next result tells us how to evaluate $\left(\frac{2}{p}\right)$. This is an important step in the whole theory of quadratic residues. The proof of this theorem is based on the so-called Gauss' Lemma.

Lemma 5.16 (Gauss)**.** *Let p be an odd prime and a an integer with $(a,p)=1$. If s is the number of least positive residues of the integers $a, 2a, \ldots, \frac{p-1}{2}a$ that are greater than $\frac{p}{2}$ then*

$$\left(\frac{a}{p}\right) = (-1)^s.$$

For the proofs of Gauss' lemma and Gauss' theorem below we refer to the textbook.

Theorem 5.17 (Gauss)**.** *Let p be an odd prime. Then*

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}.$$

Example 5.18. *Evaluate* $\left(\frac{18}{101}\right)$.

Solution. $18 = 2 \cdot 3^2$ so

$$\begin{aligned}\left(\frac{18}{101}\right) &= \left(\frac{2}{101}\right)\left(\frac{3^2}{101}\right)\\ &= \left(\frac{2}{101}\right)\\ &= (-1)^{\frac{101^2-1}{8}}\\ &= (-1).\end{aligned}$$

Rather than work out the index $\frac{p^2-1}{8}$ it is simpler to just determine whether it is even or odd, for this determines whether we get 1 or -1.

Note that if $p = 8k \pm 1$ then $p^2 = 64k^2 \pm 16k + 1$ so $\frac{p^2-1}{8} = 8k^2 \pm 2k = 2(4k^2 \pm 1)$ which is even. In this case $(-1)^{\frac{p^2-1}{8}} = 1$.

If $p = 8k \pm 3$ then $p^2 = 64k^2 \pm 48k + 9$ and $\frac{p^2-1}{8} = 8k^2 \pm 6k + 1 = 2(4k^2 \pm 3k) + 1$ which is odd. In this case $(-1)^{\frac{p^2-1}{8}} = -1$.

This leads us to an alternate formulation of the last theorem.

Theorem 5.17' (Alternate form) Let p be an odd prime. Then

$$\left(\frac{2}{p}\right) = \begin{cases} 1 & \text{if } p \equiv \pm 1 \pmod 8,\\ -1 & \text{if } p \equiv \pm 3 \pmod 8.\end{cases}$$

Hence, noting that $101 = 8 \cdot 13 - 3$ we have $101 \equiv -3 \pmod 8$ and $\left(\frac{2}{101}\right) = -1$.

5.3 Quadratic Reciprocity

There remains the problem of evaluating $\left(\frac{q}{p}\right)$ when both p and q are odd primes. This leads us to the topic of quadratic reciprocity which is one of the highlights of number theory. Quadratic reciprocity was conjectured by Legendre and Euler, but again it was Gauss who proved this result. Actually, Gauss found eight different proofs of this important and sophisticated result. One of the proofs is again based on Gauss' lemma. For the proof we refer to the exposition in the textbook.

Theorem 5.19 (Law of Quadratic Reciprocity)**.** *Let p and q be odd primes. Then*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}.$$

We will see how to use this result in examples. Essentially it allows you to tip a Legendre symbol upside down, with a change of sign in some cases.

Example 5.20. *Evaluate* $\left(\frac{3}{101}\right)$.

Solution. With $p = 3$ and $q = 101$, the law of quadratic reciprocity gives

$$\left(\frac{3}{101}\right)\left(\frac{101}{3}\right) = (-1)^{\frac{3-1}{2}\frac{101-1}{2}} = (-1)^{50} = 1.$$

$$\begin{aligned} \text{Hence } \left(\frac{3}{101}\right) &= \left(\frac{101}{3}\right) \\ &= \left(\frac{2}{3}\right) \text{ since } 101 \equiv 2 \pmod 3 \\ &= -1. \end{aligned}$$

Example 5.21. *Evaluate* $\left(\frac{54}{101}\right)$.

Solution. $54 = 2 \cdot 3 \cdot 3^2$. Hence

$$\left(\frac{54}{101}\right) = \left(\frac{2}{101}\right)\left(\frac{3}{101}\right)\left(\frac{3^2}{101}\right) = (-1)(-1)(1) = 1$$

using the results of previous exercises. It follows that there are solutions of the congruence $x^2 \equiv 54 \pmod{101}$.

Note that the term $\left(\frac{p-1}{2}\right)\left(\frac{q-1}{2}\right)$ is even if **either** $\frac{p-1}{2}$ **or** $\frac{q-1}{2}$ is even. This occurs if either $p \equiv 1 \pmod 4$ or $q \equiv 1 \pmod 4$. The term $\left(\frac{p-1}{2}\right)\left(\frac{q-1}{2}\right)$ is odd if and only if both $\frac{p-1}{2}$ and $\frac{q-1}{2}$ are odd and this occurs if and only if both $p \equiv 3 \pmod 4$ and $q \equiv 3 \pmod 4$.

This leads to the following alternate statement of the law of the law of quadratic reciprocity, and this is the one we will normally use in exercises.

Theorem 5.19' (Alternate Statement of the Law of Quadratic Reciprocity). Let p and q be odd primes. Then

(i) $\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right)$ if **either** $p \equiv 1 \pmod 4$ **or** $q \equiv 1 \pmod 4$

(ii) $\left(\frac{p}{q}\right) = -\left(\frac{q}{p}\right)$ if **both** $p \equiv 3 \pmod 4$ **and** $q \equiv 3 \mod 4$.

Example 5.22. *Evaluate* $\left(\frac{3}{43}\right)$.

Solution.

$$\begin{aligned}\left(\frac{3}{43}\right) &= -\left(\frac{43}{3}\right) \text{ since both } 43 \equiv 3 \pmod 4 \text{ and } 3 \equiv 3 \pmod 4 \\ &= -\left(\frac{1}{3}\right) \text{ since } 43 \equiv 1 \pmod 3 \\ &= -1 \text{ since } 1 = 1^2 \text{ is a square.}\end{aligned}$$

Example 5.23. *Determine whether or not there are solutions of*

$$x^2 \equiv 156 \pmod{199}.$$

Solution. $156 = 2^2 \cdot 3 \cdot 13$. Hence

$$\left(\frac{156}{199}\right) = \left(\frac{2^2}{199}\right)\left(\frac{3}{199}\right)\left(\frac{13}{199}\right).$$

$$\begin{aligned}
\text{Now } \left(\frac{2^2}{199}\right) &= 1 \text{ by Theorem 5.12 (iii).} \\
\text{Also } \left(\frac{3}{199}\right) &= -\left(\frac{199}{3}\right) \text{ since } 199, 3 \equiv 3 \pmod 4 \\
&= -\left(\frac{1}{3}\right) \text{ since } 199 \equiv 1 \pmod 3 \\
&= -1. \\
\text{Further, } \left(\frac{13}{199}\right) &= \left(\frac{199}{13}\right) \text{ since } 13 \equiv 1 \pmod 4 \\
&= \left(\frac{4}{13}\right) \text{ since } 199 \equiv 4 \pmod{13} \\
&= 1 \text{ since 4 is a square.} \\
\text{Hence } \left(\frac{156}{199}\right) &= (1)(-1)(1) = -1.
\end{aligned}$$

Hence there are no solutions of $x^2 \equiv 156 \pmod{199}$.

5.4 The Jacobi Symbol

Let n be an odd integer with prime factorisation

$$n = p_1^{t_1} p_2^{t_2} \cdots p_m^{t_m}.$$

Definition 5.24. *For $(a, n) = 1$, we define the Jacobi Symbol*

$$\left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right)^{t_1}\left(\frac{a}{p_2}\right)^{t_2}\cdots\left(\frac{a}{p_m}\right)^{t_m}.$$

Notes

(i) $\left(\frac{a}{p_i}\right)$ is a Legendre Symbol.

(ii) If $n = p$ is prime, then the Jacobi Symbol is the same as the Legendre Symbol.

(iii) When n is composite, the fact that $\left(\frac{a}{n}\right) = 1$ no longer signifies that $x^2 \equiv a \pmod n$ has a solution.

Example 5.25. *Evaluate (i)* $\left(\frac{61}{81}\right)$, *(ii)* $\left(\frac{5}{21}\right)$, *(iii)* $\left(\frac{5}{27}\right)$,

Solution.

(i) $\left(\frac{61}{81}\right) = \left(\frac{61}{3^4}\right) = \left(\frac{61}{3}\right)^4 = 1.$

(ii)

$$\begin{aligned}\left(\frac{5}{21}\right) &= \left(\frac{5}{3\cdot 7}\right) = \left(\frac{5}{3}\right)\left(\frac{5}{7}\right) = \left(\frac{2}{3}\right)\left(\frac{7}{5}\right) = \left(\frac{2}{3}\right)\left(\frac{2}{5}\right)\\ &= (-1)(-1) = 1.\end{aligned}$$

(iii) $\left(\frac{5}{27}\right) = \left(\frac{5}{3^3}\right) = \left(\frac{5}{3}\right)^3 = \left(\frac{5}{3}\right) = \left(\frac{2}{3}\right) = -1.$

It turns out the Jacobi symbol satisfies most of the properties satisfied by the Legendre symbol. We list these properties as a theorem. The proofs are omitted (they are in Rosen for those who are interested).

Theorem 5.26. (Properties of the Jacobi Symbol) *Let n be an odd positive integer and let a, b be integers with $(a, n) = (b, n) = 1$. Then*

(i) If $a \equiv b \pmod{n}$ then $\left(\frac{a}{n}\right) = \left(\frac{b}{n}\right)$.

(ii) $\left(\frac{ab}{n}\right) = \left(\frac{a}{n}\right)\left(\frac{b}{n}\right)$.

(iii) $\left(\frac{-1}{n}\right) = (-1)^{\frac{n-1}{2}}$

(iv) $\left(\frac{2}{n}\right) = (-1)^{\frac{n^2-1}{8}}$.

(v) Let m, n be odd positive integers with $(m, n) = 1$. Then

$$\left(\frac{m}{n}\right)\left(\frac{n}{m}\right) = (-1)^{\left(\frac{m-1}{2}\right)\left(\frac{n-1}{2}\right)}.$$

The effect of these results is that a Jacobi symbol can be evaluated using the same efficient techniques which we used for calculating Legendre symbols.

Example 5.27. *Evaluate* $\left(\frac{39}{55}\right)$.

Solution. Using the properties of the Jacobi symbol we get

$$\left(\frac{39}{55}\right) = -\left(\frac{55}{39}\right) = -\left(\frac{16}{39}\right) = -1.$$

Alternatively, reducing to Legendre symbol we get

$$\left(\frac{39}{55}\right) = \left(\frac{39}{5}\right)\left(\frac{39}{11}\right) = \left(\frac{4}{5}\right)\left(\frac{6}{11}\right)$$
$$= \left(\frac{2}{11}\right)\left(\frac{3}{11}\right) = -\left(\frac{3}{11}\right) = \left(\frac{11}{3}\right) = \left(\frac{2}{3}\right) = -1$$

5.5 Primality Tests

Wilson's theorem and its converse state that $(p-1)! \equiv -1 \pmod{p}$ if and only if p is prime. This allows us to test whether a number is a prime without checking for divisibility by all primes $\leq \sqrt{p}$. The importance of such primality tests follows from the role that primes play in cryptography.

We know more results which are known to be true for primes, such as Fermat's little theorem and its refinement by Euler. It turns out that their converse statements are "almost" true, i.e. if a number satisfies the conclusion of Little Fermat's or Euler's theorem it is very likely to be a prime. We introduce the following notation.

Definition 5.28. *Let $b > 1$. If n is a composite positive integer and*

$$b^n \equiv b \pmod{n},$$

then n is called a **pseudoprime** *to the base b.*

If p is a prime then $b^p \equiv b \pmod{p}$ by Fermat's Little Theorem. A pseudoprime to the base b is not a prime but nevertheless satisfies Fermat's Little Theorem to the base b.

Example 5.29. *We will show that $n = 341$ is a pseudoprime to the base 2.*

Working. First $n = 11 \cdot 31$ is composite. Next we have to show that $2^{341} \equiv 2 \pmod{341}$ and to do this we could use modular exponentiation. However, we can reduce the calculations by working with each of the factors and piecing two separate results together.

Now $2^{10} \equiv 1 \pmod{11}$ by Fermat's Little Theorem, so $2^{340} = (2^{10})^{34} \equiv 1 \pmod{11}$.

Also $2^5 = 32 \equiv 1 \pmod{31}$ so $2^{340} = (2^5)^{68} \equiv 1 \pmod{31}$.

We have shown that $11 \mid (2^{340} - 1)$ and $31 \mid (2^{340} - 1)$. Since $(11, 31) = 1$ we have $341 \mid (2^{340} - 1)$. Hence $2^{340} \equiv 1 \pmod{341}$ and on multiplying by 2 we get $2^{341} = 2 \pmod{341}$.

Definition 5.30. *Let $b > 1$. If n is a composite positive integer and a coprime of b such that*

$$b^{\frac{n-1}{2}} \equiv \left(\frac{b}{n}\right) \pmod{n},$$

then n is called an **Euler pseudoprime** *to the base b.*

If p is a prime and $p \nmid b$ then $b^{\frac{p-1}{2}} \equiv \left(\frac{b}{p}\right) \pmod{p}$ by Euler's Criterion. An Euler pseudoprime is not a prime number but shares with them the property of satisfying Euler's Criterion.

Note that since n is composite, the term $\left(\frac{b}{n}\right)$ is a Jacobi symbol.

Theorem 5.31. *Let $b > 1$ and let n be a composite positive integer. If n is an Euler pseudoprime to the base b, then n is a pseudoprime to the base b.*

Proof. Suppose n is an Euler pseudoprime to the base b. Then

$$b^{\frac{n-1}{2}} \equiv \left(\frac{b}{n}\right) \pmod{n}.$$

Square both sides noting that a Jacobi symbol is either 1 or -1.

$$b^{n-1} \equiv 1 \pmod{n}.$$

Hence $b^n \equiv b \pmod{n}$ and b is a pseudoprime to the base. □

Example 5.32. *Verify that $n = 1105$ is composite and is an Euler pseudoprime to the base 2.*

Solution. Note first that $1105 = 5 \cdot 221$ is composite. We have to verify that $2^{552} \equiv \left(\frac{2}{1105}\right) \pmod{1105}$.

Since $1105 = 8 \cdot 138 + 1$ we see that $1105 \equiv 1 \pmod 8$ and $\left(\frac{2}{1105}\right) = 1$.

Next, we could reduce $2^{552} \pmod{1105}$ by modular exponentiation and get 1. Alternatively you might like to use NUMBERS for this calculation.

We find that $2^{552} \equiv \left(\frac{2}{1105}\right) = 1 \pmod{1105}$ so 1105 is an Euler pseudoprime to the base 2.

The usefulness of pseudoprimes and Euler pseudoprimes is related to the fact that they are relatively rare. This is one area of number theory where it is useful to rely on computer calculations to generate lists of pseudoprimes and Euler pseudoprimes to various bases.

Primality Testing The following is a description in broad terms of how very large numbers can be tested to determine whether they are prime. By its very nature, the

material in this section relies on computers for its implementation in practical problems such as cryptography.

Given n (usually large), select a base b_1. Test whether $b_1^{\frac{n-1}{2}} \equiv \left(\frac{b_1}{n}\right) \pmod{n}$. If this relationship does not hold we know straight away that n is composite. If the relationship holds, we know that n is either a prime or an Euler pseudoprime. Because pseudoprimes are rather rare, we would know that it is much more likely that we have a prime than an Euler pseudoprime. If we have a list of known Euler pseudoprimes to the base b, we could check against the list and further increase the likelihood that we have a prime if it is not on the list of known pseudoprimes.

If the number n passes the test with base b_1, repeat the test with bases $b_2, \ldots, b_k$. The process stops if at any stage it is determined that n is composite.

If n passes all the tests the conclusion is that n is highly likely to be prime. This leads us into an area of number theory known as probabilistic number theory. The test can be organised so that you can be assured that you have a prime number to a very high degree of probability. The chance of a false positive can be made very low, say to the same chance of winning lotto three times in succession by buying a single ticket each time.

It turns out that all the calculations can be done very quickly on a computer.

Finding the Next Prime Once you have a way of testing quickly for a prime, you can organise a program to find the next prime. You might like to try out NUMBERS' capabilities in this direction. You input a very large number of say 50 digits in length. The program will successively test the numbers $n, n+1, \ldots$ until it finds a prime.

The process is used in cryptography. It gives a way of randomly choosing two large primes p and q so that the number $n = pq$ together with a number e with $(e, \phi(n)) = 1$ can be used as an enciphering key.

Chapter 6

DIOPHANTINE EQUATIONS

6.1 Pythagorean Triples

Definition 6.1. *A* **Pythagorean triple** *is a set of three positive integers* x, y, z *such that*

$$x^2 + y^2 = z^2.$$

The triple is said to be **primitive** *if* $\gcd(x, y, z) = 1$.

We are all familiar with examples such as

$$\begin{aligned} 3^2 + 4^2 &= 5^2 \\ 5^2 + 12^2 &= 13^2 \end{aligned}$$

which tell us that 3, 4, 5 and 5, 12, 13 are Pythagorean triples. Historically, interest was centred around finding right-angled triangles with integer sides. Given any Pythagorean triple, we can use it to construct a right-angled triangle with the sides having integer lengths.

We now discuss two ways of generating Pythagorean triples.

Procedure 1. Let m, n be positive integers. Now

$$\begin{aligned} (m^2 + n^2)^2 &= m^4 + 2m^2n^2 + n^4 \\ (m^2 - n^2)^2 &= m^4 - 2m^2n^2 + n^4 \end{aligned}$$

so on subtracting we get

$$(m^2 + n^2)^2 - (m^2 - n^2)^2 = 4m^2n^2 = (2mn)^2.$$

Hence

$$(m^2 - n^2)^2 + (2mn)^2 = (m^2 + n^2)^2$$

which has the form

$$x^2 + y^2 = z^2.$$

Hence any choice of m and n with $m > n > 0$ will produce a Pythagorean triple. Here a few examples of choices for m and n and the corresponding triple obtained by putting $x = m^2 - n^2, y = 2mn, z = m^2 + n^2$.

m	n	x	y	z
2	1	3	4	5
3	1	8	6	10
3	2	5	12	13
4	1	15	8	17
4	2	12	16	20
4	3	7	24	25

We notice that some of these triples are primitive and some are not.

Procedure 2. Once we have found a Pythagorean triple x, y, z, we can scale the triple by multiplying by a number k to give another triple kx, ky, kz which is Pythagorean since

$$(kx)^2 + (ky)^2 = k^2(x^2 + y^2) = k^2z^2 = (kz)^2.$$

Lemma 6.2. *Every Pythagorean triple is derived from a primitive Pythagorean triple by scaling.*

Proof. Let x, y, z be a Pythagorean triple, and let $d = \gcd(x, y, z)$. Write $x = dx_0, y = dy_0, z = dz_0$ with $\gcd(x_0, y_0, z_0) = 1$. Since $x^2+y^2 = z^2$ it follows immediately that x_0, y_0, z_0 is a primitive Pythagorean triple, and that x, y, z is obtained from x_0, y_0, z_0 by scaling. □

Notice that the Pythagorean triple 9, 12, 15 is obtained from 3, 4, 5 by trebling values. It cannot be obtained from the expressions $x = m^2 - n^2, y = m^2 - n^2, z = m^2 + n^2$ since there are no choices of m, n which will give $z = m^2 + n^2 = 15$.

Our program is to describe all Pythagorean triples. From the last result, it suffices to describe all primitive Pythagorean triples. We will show that Procedure 1 produces all primitive Pythagorean triples, and we will also describe the conditions necessary on m and n to produce a primitive triple.

Lemma 6.3. *If x, y, z is a primitive Pythagorean triple then $(x, y) = (x, z) = (y, z) = 1$.*

Proof. Suppose $(x, y) = d > 1$ and let p be a prime factor of d. Then $p \mid x$, $p \mid y$ and from $z^2 = x^2 + y^2$ we see that $p \mid z^2$. Since p is prime we have $p \mid z$. Hence $(x, y, z) \geq p$ contradicting the assumption that (x, y, z) is primitive. □

It is often useful to use congruences modulo 4 in discussing Pythagorean triples, and we recall the results about squaring even and odd integers.

If x is even then $x = 2k$ for some integer k, and $x^2 = 4k^2 \equiv 0 \pmod 4$. If x is odd then $x = 2k + 1$ for some integer k, and $x^2 = 4k^2 + 4k + 1 \equiv 1 \pmod 4$. We never have $x^2 \equiv 2 \pmod 4$ or $x^2 \equiv 3 \pmod 4$.

Lemma 6.4. *If (x, y, z) is a primitive Pythagorean triple, then either x is even and y is odd, or y is even and x is odd.*

Proof. Since $(x, y) = 1$, certainly x and y cannot both be even. Suppose x and y are both odd. We consider congruences modulo 4. Then $x^2 \equiv y^2 \equiv 1 \pmod 4$, and hence

$$z^2 = x^2 + y^2 \equiv 2 \pmod 4.$$

But this is impossible since 2 is a quadratic nonresidue of 4. It follows that x and y cannot both be odd. Hence one has to be even and the other odd. □

We will say that x and y have *opposite parity* if one of them is even while the other is odd. Hence, if x, y, z is a primitive Pythagorean triple, then x and y have opposite parity.

To simplify notation it is convenient to assume that in a primitive Pythagorean triple x, y, z that x is odd, y is even. It then follows that z is odd since $(y, z) = 1$.

Theorem 6.5. *If r, s, t are positive integers such that $(r, s) = 1$ and $rs = t^2$, then there are integers m, n such that*

$$r = m^2, \text{ and } s = n^2.$$

Proof. If $r = 1$ take $m = 1, n = t$. If $s = 1$ take $m = t, n = 1$. In the remainder of the proof, we can assume that $r > 1$ and $s > 1$.

Note that since $(r, s) = 1$, the same prime cannot occur in the factorisation of both r and s. Write

$$\begin{aligned} r &= p_1^{\alpha_1} p_2^{\alpha_2} \dots p_n^{\alpha_n} \\ s &= p_{n+1}^{\alpha_{n+1}} p_{n+2}^{\alpha_{n+2}} \dots p_v^{\alpha_v} \\ t &= q_1^{\beta_1} q_2^{\beta_2} \dots q_h^{\beta_h}. \end{aligned}$$

From $rs = t^2$ we have

$$p_1^{\alpha_1} p_2^{\alpha_2} \dots p_v^{\alpha_v} = q_1^{2\beta_1} q_2^{2\beta_2} \dots q_h^{2\beta_h}.$$

By unique factorisation, $v = h$ and each p_i occurs as some q_j with matching indices. Hence for each i we have $\alpha_i = 2\beta_j$ for some j. Hence α_i is even for each i. We can now define m and n by

$$\begin{aligned} m &= p_1^{\alpha_1/2} \dots p_n^{\alpha_n/2} \\ n &= p_{n+1}^{\alpha_{n+1}/2} \dots p_v^{\alpha_v/2.} \end{aligned}$$

Then $m^2 = r$ and $n^2 = s$ where m and n are integers. □

Remark. From the above, we see that an integer is a perfect square if and only if only even powers occur in its prime factorisation.

Theorem 6.6. (Primitive Pythagorean Triples) *Suppose x, y, z is a primitive Pythagorean triple with x odd and y even. Then there are positive integers m and n of opposite parity with $m > n$ and $(m, n) = 1$, such that*

$$x = m^2 - n^2, y = 2mn, z = m^2 + n^2.$$

Conversely, suppose that m and n are integers of opposite parity with $m > n$ and $(m, n) = 1$ and x, y, z are defined by $x = m^2 - n^2, y = 2mn, z = m^2 + n^2$. Then x, y, z is a primitive Pythagorean triple with x odd and y even.

Proof. Suppose x, y, z is a primitive Pythagorean triple with y even and x and z both odd. Now

$$\begin{aligned} y^2 &= z^2 - x^2 = (z + x)(z - x) \\ \left(\frac{y}{2}\right)^2 &= \left(\frac{z+x}{2}\right)\left(\frac{z-x}{2}\right) = rs \end{aligned}$$

where $r = \frac{z+x}{2}, s = \frac{z-x}{2}$. Since x and z are odd, r and s are integers which are also positive.

Let $d = (r, s)$. Since $r - s = x$ and $r + s = z$, we see that $d \mid x$ and $d \mid z$. Since $(x, z) = 1$ it follows that $d = 1$. This shows that $(r, s) = 1$.

By Theorem 6.5 there are integers m and n such that that $r = m^2$ and $s = n^2$. Thus $\left(\frac{y}{2}\right)^2 = m^2n^2$ and from this we see that $y = 2mn$. Also $x = r - s = m^2 - n^2$ and $z = r + s = m^2 + n^2$. Note that $m > n$ since $x = m^2 - n^2 > 0$. Let $(m, n) = d$. Then $d \mid m$ and $d \mid n$ and since $x = m^2 - n^2, z = m^2 + n^2$ we see that $d \mid x$ and $d \mid z$. Since $(x, z) = 1$ we must have $d = 1$ and this shows that $(m, n) = 1$.

Now m and n cannot both be even since $(m, n) = 1$. If m, n were both odd then $z = m^2 + n^2$ would be even, contradictory $(y, z) = 1$. It follows that one of m and n is even while the other is odd.

Conversely, suppose that m and n satisfy the stated conditions and that x, y, z are defined by $x = m^2 - n^2, y = 2mn, z = m^2 + n^2$. We have seen from Procedure 1 that x, y, z is a Pythagorean triple and it remains to show that the triple is primitive.

Since one of m, n is even and the other is odd, it follows from the equations for x, y, z that x and z are both odd. Suppose $(x, y, z) = d > 1$ and p is a prime divisor of d. Then $p \neq 2$ since x is odd and $p \mid x$. From the relations $p \mid x$ and $p \mid z$ we see that $p \mid (x + z)$ and $p \mid (x - z)$. Hence $p \mid 2m^2$ and $p \mid 2m^2$. Since $p \neq 2$ this means that $p \mid m$ and $p \mid n$, contradicting $(m, n) = 1$. It follows that $d = 1$ and x, y, z is a primitive triple. □

Example 6.7. *You can check that $341^2 + 420^2 = 541^2$ and that $(341, 420, 541) = 1$. Find m and n which satisfy the conditions of Theorem 2 and which produce this Pythagorean triple.*

Solution. We want $x = m^2 - n^2 = 341$ and $z = m^2 + n^2 = 541$. On adding these two expressions we get $2m^2 = 882$ where $m^2 = 441$ and $m = 21$. On subtracting we get

$2n^2 = 200$ where $n = 10$. Hence $m = 21$ and $n = 10$ is the required pair. We notice that $(21, 10) = 1$ and 21 is odd while 10 is even.

We also note that the formulae $x = m^2 - n^2 = 341$, $y = 2mn = 420$ and $z = m^2 + n^2 = 541$, produce the Pythagorean triple.

Example 6.8. *In a primitive Pythagorean triple x, y, z, verify that y is divisible by 4 and that $y + z$ is a perfect square.*

Solution. We use Theorem 6.5 to write $x = m^2 - n^2, y = mn, z = m^2 + n^2$ for integers m, n with one of m, n even. Then mn contains a factor 2 and so $y = 2mn$ is divisible by 4.

Also $y + z = 2mn + m^2 + n^2 = (m + n)^2$ is a perfect square.

Example 6.9. *Let x, y, z be a primitive Pythagorean triple. Show that exactly one of x and y is divisible by 3.*

Solution. We work directly with the equation $x^2 + y^2 = z^2$ and we will take congruences modulo 3. Note that 1 is the only quadratic residue modulo 3 while 2 is a quadratic non-residue.

If $3 \mid x$ or $3 \mid y$ there is nothing to prove, so suppose $3 \nmid x$ and $3 \nmid y$. Then $x^2 \equiv 1 \pmod 3$ and $y^2 \equiv 1 \pmod 3$. Hence $z^2 = x^2 + y^2 \equiv 2 \pmod 3$. But this is impossible since 2 is a quadratic nonresidue mod 3. We conclude that either $3 \mid x$ or $3 \mid y$.

Finally we cannot have both $3 \mid x$ and $3 \mid y$ since $(x, y) = 1$.

6.2 Some Other Diophantine Equations

We have completely analysed the two Diophantine equations.

$$\begin{aligned} ax &+ by = c \\ x^2 &+ y^2 = z^2 \end{aligned}$$

which are the linear and Pythagorean equations respectively. Our solutions provide a guide for the study of other equations. In particular, given a Diophantine equation we would like to be able to do the following:

(i) determine whether or not solutions exist,

(ii) describe all solutions when they exist.

Some Irrational Numbers

Another area where Diophantine equations arise is in discussions of irrational numbers. The ancient Greeks knew that $\sqrt{2}$ is irrational and we will now consider the arguments they used.

Theorem 6.10. *Let p be a prime number. Then the Diophantine equation*

$$x^2 - py^2 = 0$$

has no solutions in nonzero integers x and y.

Proof. Suppose x and y are nonzero integers such that $x^2 - py^2 = 0$. Let $d = (x, y)$ and write $x = dx_0$, $y = dy_0$ with $(x_0, y_0) = 1$. Then $(dx_0)^2 - p(dy_0)^2 = 0$ and from this we get

$$x_0^2 = py_0^2.$$

Now $p \mid x_0^2$ and since p is prime we get $p \mid x$. Write $x_0 = ps$. Then $p^2s^2 = py_0^2$ and therefore $ps^2 = y_0^2$. Again we see that $p \mid y_0^2$ and then $p \mid y_0$. But now we have $p \mid x_0$ and $p \mid y_0$ which contradicts $(x_0, y_0) = 1$. We conclude that there are no nonzero solutions of the equation. □

Corollary. Let p be a prime. Then $\sqrt{p}$ is irrational.

Proof. Suppose $\sqrt{p}$ is rational. Then $\sqrt{p} = \dfrac{x}{y}$ where x and y are non zero integers.

But then we have $x^2 - py^2 = 0$ which contradicts the theorem. We conclude that $\sqrt{p}$ cannot be written as a rational member so it must be irrational. □

Sums of Squares An integer n is said to be the sum of two squares if there are integers x and y such that

$$n = x^2 + y^2.$$

We allow one or both of x, y to be zero in this equation.

It is interesting to examine the situation for small n to get a feel for what happens. Here are some results.

$$\begin{aligned}
0 &= 0^2 + 0^2 \\
1 &= = 1^2 + 0^2 \\
2 &= 1^2 + 1^2 \\
3 &\quad \text{is not the sum of two squares} \\
4 &= 2^2 + 0^2 \\
5 &= 2^2 + 1^2 \\
6 &\quad \text{is not the sum of two squares} \\
7 &\quad \text{is not the sum of two squares} \\
8 &= 2^2 + 2^2 \\
9 &= 3^2 + 0^2 \\
10 &= 3^2 + 1^2 \\
11 &\quad \text{is not the sum of two squares.}
\end{aligned}$$

It is not clear what the pattern is and we will not do a complete analysis. However, it is possible to get one result without too much trouble, and this explains why numbers such as 3, 7, and 11 cannot be written as the sum of two squares.

Theorem 6.11. *If $n \equiv 3 \pmod 4$ then n cannot be written as the sum of two squares.*

Proof. Consider congruences modulo 4. Now

$$\begin{aligned} x^2 &\equiv 0 \text{ or } 1 \pmod 4 \\ y^2 &\equiv 0 \text{ or } 1 \pmod 4. \end{aligned}$$

Hence $x^2 + y^2$ is congruent to either 0, 1 or 2 (modulo 4). Hence if $n \equiv 3 \pmod 4$ we cannot have $n \equiv x^2 + y^2 \pmod 4$ and hence there are no solutions in x and y of $n = x^2 + y^2$. □

There is a famous result in number theory that if p is a prime and $p \equiv 1 \pmod 4$ then there are integers x and y such that $p = x^2 + y^2$.

For composites mn we have if m and n are sums of squares then mn is a sum of squares as well. (Tutorial problem)

We will not pursue this topic any further and refer you to the text book for further results.

6.3 The Equation $x^4 + y^4 = z^4$

Fermat knew that this equation had no solutions in nonzero integers, and we will now discuss what is essentially Fermat's proof of this result.

Since $x^4 + y^4 = (z^2)^2$, any solution x, y, z of $x^4 + y^4 = z^4$ also provides a solution of $x^4 + y^4 = z^2$. Hence it suffices to prove that there are no nonzero solutions of $x^4 + y^4 = z^2$.

Lemma 6.12. *Let a, b be positive integers such that $a^2 \mid b^2$. Then $a \mid b$.*

Proof. Let

$$\begin{aligned} a &= p^{\alpha_1}, \cdots p_n^{\alpha_n} \\ \text{and } b &= q_1^{\beta_1} \cdots q_m^{\beta_m} \end{aligned}$$

be the prime factorisations of a and b.

Then

$$\begin{aligned} a^2 &= p_1^{2\alpha_1} \cdots p_n^{2\alpha_n} \\ b^2 &= q_1^{2\beta_1} \cdots q_m^{2\beta_m} \end{aligned}$$

Since $a^2 \mid b^2$, each p_i occurs as a q_j with $p_i^{2\alpha_i} \mid q_j^{2\beta_j}$. It follows that $2\alpha_i \leq 2\beta_j$ and $\alpha_i \leq \beta_j$. From the prime factorisations of a and b we now see that $a \mid b$. □

Theorem 6.13. *The Diophantine equation $x^4 + y^4 = z^2$ has no solutions in nonzero integers x, y, z.*

Proof. Assume there is a solution with $x_0 \neq 0,\ y_0 \neq 0,\ z_0 > 0$. By the Well-Ordering Property we can choose this solution such that z_0 has the least possible positive value.

We first show that $(x_0, y_0) = 1$. For if $(x_0, y_0) = d$ we get $d^4 \mid x_0^4 + y_0^4 = z_0^2$. By the lemma, $d^2 \mid z_0$. Since $x_0^4 + y_0^4 = z_0^2$ we have

$$\left(\frac{x_0}{d}\right)^4 + \left(\frac{y_0}{d}\right)^4 = \left(\frac{z_0}{d^2}\right)^2$$

Hence we must have $d = 1$ for otherwise we would have a solution with a smaller value of z contradicting the choice of z_0.

Now $(x_0^2)^2 + (y_0^2)^2 = z_0^2$ and since $(x_0, y_0) = 1$ it follows that x_0^2, y_0^2, z_0 is a primitive Pythagorean triple.

Without loss of generality, assume y_0 is even. By Theorem 6.6 there are integers m, n of opposite parity with $m > n,\ (m, n) = 1$ such that

$$\begin{aligned} x_0^2 &= m^2 - n^2 \\ y_0^2 &= 2mn \\ z_0 &= m^2 + n^2. \end{aligned}$$

If m were even, then n would be odd and we would have $x_0^2 \equiv -n^2 \equiv -1 \equiv 3 \pmod 4$ which is impossible. It follows that m is odd, n even.

Since $x_0^2 + n^2 = m^2$ and $(m, n) = 1$ we have another primitive Pythagorean triple x_0, n, m with n even. Hence there exist integers r, s with $(r, s) = 1,\ r > s$ and one of r, s even and the other odd such that

$$\begin{aligned} x_0 &= r^2 - s^2 \\ n &= 2rs \\ m &= r^2 + s^2 \\ \text{Now } y_0^2 &= 2mn = 4rs(r^2 + s^2) \\ \left(\frac{y_0}{2}\right)^2 &= rs(r^2 + s^2). \end{aligned}$$

Since $(r, s) = 1$ we have $(r, r^2 + s^2) = 1$ and $(s, r^2 + s^2) = 1$. It follows from Theorem 6.5 that each of $r, s,\ r^2 + s^2$ are perfect squares.

Thus $r = \alpha^2\ s = \beta^2\ r^2 + s^2 = \gamma^2$ for integers α, β, γ. Since $r^2 + s^2 = \gamma^2$ we have $\alpha^4 + \beta^4 = \gamma^2$.

Note that $\alpha \neq 0$ since if $\alpha = 0$ we would get $r = 0,\ n = 0$ and $y_0 = 0$. Similarly $\beta \neq 0$

Now (the key point!)

$$\gamma < \gamma^2 = r^2 + s^2 = m < \sqrt{m^2 + n^2} = z_0.$$

Hence $\gamma < z_0$.

But this contradicts the assumption that z_0 was least. The assumption that $x^4+y^4 = z^2$ has solutions leads to a contradiction, and we conclude that there are no solutions in nonzero integers. □

Corollary. The equation $x^4 + y^4 = z^4$ has no solutions in non zero integers x, y, z.

Corollary. If n is a multiple of 4, there are no solutions of $x^n + y^n = z^n$ in non zero integers x, y, z.

Proof. Suppose $n = 4m$ and there is a solution x_0, y_0, z_0 in non zero integers with $x_0^n + y_0^n = z_0^n$. Then $(x_0^m)^4 + (y_0^m)^4 = (z_0^m)^4$, contradicting the fact that $x^4 + y^4 = z^4$ has no solutions. □

6.4 History of Fermat's Last Theorem

This section contains the material for the final lecture of PMTH 338. You are asked to read through these notes, in the expectation that you might like to see how the story ends, but this material will not specifically be examined.

Pythagoras' Theorem The square on the hypotenuse of a right-angled triangle is equal to the sum of the squares on the other two sides. Conversely, if the square on one side of a triangle is equal to the sum of the squares on the other two sides, then the triangle is right-angled. The fact that

$$3^2 + 4^2 = 5^2$$

allows us to construct a right-angled triangle with integer sides.

In ancient civilisations, this fact was used as a help in surveying. As we have seen, there are other possibilities for using integers to construct right-angled triangles, for example:

$$\begin{aligned} 5^2 + 12^2 &= 13^2, \\ 7^2 + 24^2 &= 25^2, \\ 8^2 + 15^2 &= 17^2. \end{aligned}$$

Diophantus of Alexandria ($\sim$ 250 AD) By the time of Diophantus, mathematics had reached an advanced stage in Greek civilisation. Diophantus recorded many results about finding integer solutions of various equations. In particular, he asked for a formula for generating solutions of the Pythagorean equation $x^2 + y^2 = z^2$. Our work on this equation constitutes the solution of Diophantus' problem.

The main body of Diophantus' works was lost when the library of Alexandria was burned in the fifth century AD. However, copies of some of his works survived in the Arab world and eventually these reached Europe in the 15'th century. They became generally available in Europe in about 1621, when Bachet published a translation.

Pierre de Fermat (1601 - 1665) Fermat was a magistrate in Toulouse and became interested in mathematics at a comparatively late age. He is known in optics for having

formulated the Principle of Least Time. He also did some work on finding tangents to curves using methods which we now recognise as belonging to differential calculus. However, his greatest work was in Number Theory, and he is recognised as being the founder of this subject in its modern form.

He had in his possession a copy of the works of Diophantus, and proceeded to build on what Diophantus had done. In particular, he became familiar with the methods used for finding solutions of the Pythagorean equation, and then found a proof showing that it is impossible to do a similar thing for fourth powers, that is, to find positive integers x, y, and z such that

$$x^4 + y^4 = z^4.$$

It is this proof which we studied in our course, and as you will have observed, it is highly non-trivial. Further, the proof is recognisable as being completely rigorous by the standards of modern mathematics. It is extraordinary to think that Fermat wrote the proof in the margin of his copy of Diophantus, with no intention of ever publishing it. He did however, challenge his friends to solve this problem and others like it. (Incidentally, the margins in Bachet's translation were quite large!)

Fermat discovered many new theorems and often made notes in the margins of his book, or wrote letters to friends. In many cases he included proofs but in others he did not.

The Last Theorem Fermat stated the 'Last Theorem' as an entry in the margin of his book, in 1637.

Cubum autem in duos cubos, aut quadratoquadratum in duos quadratoquadratos, et generaliter nullam in infinitum ultra quadratum potestatem in duos ejusdem nominis fas est dividere: cujus rei demonstrationem marabilem sane detexi. Hanc marginis exiguitas non caperet.

(It is impossible to separate a cube into two cubes, or a fourth power into two fourth powers, or in general, any power higher than the second into two like powers. I have discovered a truly marvelous proof of this, which this margin is too narrow to contain.)
The result is now usually stated in the following form:

Fermat's Last Theorem. The Diophantine equation

$$x^4 + y^4 = z^4$$

has no solutions in non-zero integers x, y, z when n is an integer with $n \geq 3$.

After his death, Fermat's son arranged for the publication of the main results which Fermat had discovered. Eventually, proofs were found to most of those results which had been stated without proof in the margin of the copy of Diophantus. One (the Second Last Theorem) proved very stubborn but a proof was eventually found over a hundred years later. It was because so many of Fermat's stated results had turned out to be correct that attention was drawn to the 'Last Theorem', which had so far eluded all attempts at proof.

Proofs of Special Cases As noted in our course material, the problem can be quickly reduced to the case $n = p$, where p is an odd prime. Quite a few special cases were then proved. Euler gave the proof for $n = 3$ in 1770, although his proof contained a gap which

was later fixed up. The case $n = 5$ fell in 1825 and soon after in 1839 the case $n = 7$ was proved. Some of the important advances during this period were made by an outstanding woman mathematician, Sophie Germain. However no general proof valid for all n was found. The arguments were becoming increasingly complex and it was clear that entirely different methods were required for the general case.

On two occasions during the nineteenth century, the Academie des Sciences de Paris offered a cash prize and a gold medal to the first French mathematician to prove Fermat's Last Theorem. The mystique surrounding the theorem continued to grow when several of the leading French mathematicians of the day submitted proofs which turned out to be faulty.

New Mathematics Created During the 19'th century, it was realised that new mathematical theories would have to be created in order to fully understand the Last Theorem. In 1843 the German mathematician Kummer made an important advance by inventing the notion of an *algebraic number* (an algebraic number is a complex number which satisfies a polynomial equation with integer coefficients and leading coefficient equal to 1), and used this idea to give a proof. It turned out that Kummer's proof of Fermat's Last Theorem contained a flaw. He had assumed that algebraic integers factorised uniquely into prime algebraic integers in a similar way to that stated in the Fundamental Theorem of Arithmetic, and this was soon shown to be false. (You will notice that our proof of uniqueness in the Fundamental Theorem required more thought than the proof of existence, indicating that the uniqueness assertion is the more subtle of the two). Nevertheless, Kummer's work has proved of immense importance, and the subject he founded, *algebraic number theory*, remains as one of the important branches of modern mathematics. The notion of an ideal arose in this work and this notion was later to become a key concept in the subject we now know as abstract algebra. Eventually, Kummer proved that the theorem held for a large class of prime numbers, those which he called the *regular* prime numbers. Kummer's work was followed up by others and by about 1990 it was known that the theorem is true for all primes p with $3 \leq p \leq 10^6$.

The Cash Prize In 1908 a large cash prize was bequeathed to the Academy of Science at Göttingen to be paid for the first complete proof of Fermat's Conjecture. This set in train one of the strange phenomena of twentieth century mathematics. A large number of amateur mathematicians tackled the problem and submitted thousands of manuscripts not only to the Academy of Science at Göttingen but to mathematics departments around the world. One line of reasoning of the amateurs is that Fermat did indeed discover a simple proof, and it just remains to rediscover it. Further, a professional mathematician is unlikely to discover a simple proof, since his very training makes him or her incapable of seeing the wood for the trees. A few years ago in Australia, the ABC's Quantum program did a feature on an amateur mathematician who took his proof into the University of Melbourne for the experts to pass judgment, and of course it turned out that the amateur had not proved the theorem.

The Current Generation Some of the leading mathematicians of the current generation have developed new mathematics which has finally led to a proof of Fermat's Last Theorem. In 1983, the German mathematician Gerd Faltings proved that the Fermat equation $x^n +$

$y^n = z^n$ can have at most a finite number of solutions in nonzero integers. It remained then to show that this finite number is in fact zero in each case. Further work by other leading number theorists led to the feeling that the problem was close to being settled once and for all. There was considerable publicity in 1988 when press reports suggested that Yoichi Miyaoka of Japan had proved the theorem. However it again turned out that the proof, which was immensely complicated, contained a number of flaws.

Andrew Wiles In about 1990, the English mathematician Andrew Wiles began to build on some of his earlier work and that of others to eventually construct a 200 page manuscript which he believed, amongst other things, gave a proof of Fermat's Last Theorem. Wiles gave his proof at the end of a series of lectures at Cambridge University in June 1993, and the event was widely reported in the media. The New York Times ran a front page story and the Sydney Morning Herald printed part of this story on 25 June 1993. Especially in the United States, there seemed to be extraordinary public interest in Wiles' proof, and the American Mathematical Society organised a number of public lectures on Fermat's Last Theorem and its proof.

An Australian Contribution There was a final twist to the story when a flaw was discovered in Wiles' proof at the end of 1993. John Coates, an old boy of Taree High School and now Professor of Mathematics at Cambridge University, was Wiles' Ph.D. thesis supervisor (at Cambridge), and has been a leading figure in Number Theory for the last thirty years. It was Coates who found the gap in Wiles' proof, and for a while it seemed that the proof had suffered the same fate as that of its predecessors. Wiles was unable to repair the gap but returned to an approach which he had tried earlier. (When Plan B failed he returned to Plan A.) Towards the end of 1994, he released two manuscripts, one of which was joint work with Richard Taylor. These two manuscripts have now appeared in the mathematical journals, and are considered by the experts (including Coates) to hold up. Together with Wiles' 1993 manuscript, they constitute a proof of Fermat's Last Theorem.

Further Information A number of books and articles have appeared recently which describe Wiles' feat in more detail. The BBC also interviewed Wiles and some of the other mathematicians mentioned above on the program Horizon. A transcript of the program is available at http://www.bbc.co.uk/horizon/95-96/960115.html

Epilogue Did Fermat really have a proof of the Last Theorem? We can never know, but a likely possibility is that he had proofs for both the cases $n = 3$ and $n = 4$, and at the time that he wrote the claim in the margin, he thought that the proofs generalised to all n. He may have realised later that the proofs did not generalise, as he made no further claims in letters to colleagues. He was not to know that his son would dig out all of his old notes and have them published!

Was it worth all the effort to find a proof of the Last Theorem? That depends on your point of view. Certainly the proof of Fermat's Last Theorem represents a marvelous intellectual achievement. But there is more to it than that. The problem was a good problem in the sense that it spurred mathematicians on to develop new theories in an effort to understand what was happening. The real value of the Last Theorem probably lies in this legacy.

Chapter 7

Appendices

7.1 Appendix A: Peano's axioms

According to Peano, all properties of the natural numbers $\mathbb{N}$ can be derived from the following 5 axioms:

1. 0 is a natural number.
2. Any natural number n has exactly one 'successor' n'.
3. 0 is not a successor of any natural number.
4. A natural number is the successor of at most one other number.
5. Any subset A of $\mathbb{N}$ that satisfies the properties: (a) 0 belongs to A, (b) If a number n belongs to A then its successor n' also belongs to A, coincides with $\mathbb{N}$.

It follows immediately that any natural number except 0 is the successor of exactly one number (called predecessor).

Proof. We have to show that the set $A = \{0\} \cup \{n : n = m'\} = \mathbb{N}$. In fact, $0 \in A$ and for any $n \in A$ the number $n' \in A$ since it is the successor of n.

Now we can define addition: For any natural number n the sum $n + 0 := n$. Assume $n + m$ is already defined, then $n + m' := (n + m)'$. (In fact, one has to prove that this defines the summ of all pairs n, m in a unique way.)

Prove: 1. $0 + n = n$ for all natural numbers n
2. $n + m = m + n$ for all natural numbers m, n (commutativity)

We prove associativity: (a+b)+c= a+(b+c) for all a, b, c. If $c = 0$ we have $(a+b)+0 = a+b = a+(b+0)$. Assume $(a + b) + c = a + (b + c)$ hold for some c. Then

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

Multiplication is defined as follows: For any natural number n the product $n \cdot 0 := 0$. Assume $n \cdot m$ is already defined, then $n \cdot m' = n \cdot m + n$.

The formal proofs of the arithmetic properties of numbers can be found in the book "Foundations of Analysis: The arithmetic of whole, rational, irrational and complex numbers" by Edmund Landau, Chelsea Publishing Company, New York.

7.2 Appendix B: Well-Ordering Property

Theorem 7.1. *The Well-Ordering Property is equivalent to either of the two principles of Mathematical induction.*

Proof. First we prove that the FMI implies the WOP. Let S be a set of non-negative numbers. We say that x is a lower bound of S if $a > x$ for all $a \in S$. Denote by X the set of all lower bounds of S. Then either $0 \in X$ or 0 is a smallest element of S. If b is a lower bound then $b + 1 \leq a$ for all $a \in S$. Now, either $b + 1$ belongs to S as its smallest element or $b+1$ belongs to X. It follows that either S has a smallest element or $0 \in X$ and for any $b \in X$ the number $b + 1 \in X$. But then FMI implies $X = \mathbb{N}$, i.e. all numbers are lower bounds of S which is only possible if S is empty.

Now we will show that the WOP implies the SMI. Let S be a set that contains 0 and for any $\{0, \ldots, a\} \subset S$ it can be shown that $a + 1 \in S$. Assume $X = \mathbb{N} \setminus S$. We use WOP to show (by contradiction) that X is empty. In fact, if X was not empty it would have a smallest element b according to WOP. We have $b \neq 0$ because $0 \in S$. On the other hand, b cannot be bigger than 0 because then $0, \ldots, b - 1 \in S$ and therefore $b \in S$ which contradicts $b \in X$.

It remains to show that SMI implies FMI. In order to do this we suppose that the assumption of the FMI is satisfied and we exploit the SMI to deduce the conclusion of FMI. According to the assumption of FMI the set S contains 0 and, if we assume $a \in S$, we are able to prove $a + 1 \in S$. But then we will be able to prove $a + 1 \in S$ under the stronger assumption $\{0, \ldots, a\} \subset S$. This means the assumption of SMI is fulfilled and, hence $S = \mathbb{N}$. But this is also the assertion of FMI.

(This short proof is much more difficult than it seems at the first glance. The logic difficulty is due to the fact that our assumptions are principles with their own assumptions which results in a nestled structure.) □

The following argument shows that WOP is, in fact, equivalent to the statement: For any natural number a there are only finitely many natural numbers between 0 and a.

Suppose there is a nonempty set S of positive integers which has no smallest element. Since S is not empty it has at least one element, say s_1. Since this cannot be the smallest element of S, there is another element $s_2 \in S$ with $s_2 < s_1$. Again s_2 cannot be the smallest element of S so we get another element $s_3 \in S$ with $s_3 < s_2$. Continuing in this way, we obtain an infinite string of positive integers $\{s_n\}_{n=1}^{\infty}$ satisfying $0 < \cdots < s_n < s_{n-1} < \cdots < s_3 < s_2 < s_1$. But this is impossible, since between any two given integers there can be only finitely many others. We conclude from this contradiction that there cannot be a nonempty set of positive integers which does not have a smallest element.

On the other hand WOP or the equivalent FMI can be used to prove the statement. There are finitely many (namely none) numbers between 0 and 0. If we know that there are only finitely many numbers between 0 and k then there are finitely many numbers between 0 and $k+1$. Thus, by induction, there are only finitely many numbers between 0 and a for any natural number a.