# A Course on Number Theory

Peter J. Cameron

# Preface

These are the notes of the course MTH6128, *Number Theory*, which I taught at Queen Mary, University of London, in the spring semester of 2009.

There is nothing original to me in the notes. The course was designed by Susan McKay, and developed by Stephen Donkin, Ian Chiswell, Charles Leedham-Green, and Thomas Müller; I have benefited greatly from Ian Chiswell's notes, which I have followed closely.

I am grateful to Mark Walters who stood in for me in the first six lectures of the course, and whose comments have been very helpful; also to the class tutors, markers, and most of all the students who took the course, for their comments and support.

The original course was largely based on continued fractions: this technique is very amenable to hand calculation, and can be used to solve Pell's equation, to write an integer as a sum of squares where this is possible, and to classify the indefinite binary quadratic forms. This is still the centrepiece of the course, but I have given alternate treatment of sums of squares.

The syllabus for the course reads

(a) Continued fractions: finite and infinite continued fractions, approximation by rationals, order of approximation.

(b) Continued fractions of quadratic surds: applications to the solution of Pell's equation and the sum of two squares.

(c) Binary quadratic forms: equivalence, unimodular transformations, reduced form, class number. Use of continued fractions in the indefinite case.

(d) Modular arithmetic: primitive roots, quadratic residues, Legendre symbol, quadratic reciprocity. Applications to quadratic forms.

The learning outcomes state

> Students will be able to use continued fractions to develop arbitrarily accurate rational approximations to rational and irrational numbers.

> They will be able to work with Diophantine equations, i.e. polynomial equations with integer solutions. They will know some of the famous classical theorems and conjectures in number theory, such as Fermat's Last Theorem and Goldbach's Conjecture, and be aware of some of the tools used to investigate such problems.

The recommended books are

[1] H Davenport, *The Higher Arithmetic*, Cambridge University Press (1999)

[2] Allenby & Redfern, *Introduction to Number Theory with Computing*, Edward Arnold (1989)

# Contents

# Chapter 1

# Overview and revision

In this section we will meet some of the concerns of Number Theory, and have a brief revision of some of the relevant material from Introduction to Algebra.

## 1.1 Overview

Number theory is about properties of the natural numbers, integers, or rational numbers, such as the following:

- Given a natural number $n$, is it prime or composite?
- If it is composite, how can we factorise it?
- How many solutions do equations like $x^2+y^2=n$ or $x^n+y^n=z^n$ have for fixed $n$, where the variables are required to be natural numbers?
- How closely can we approximate a given irrational number by rational numbers which are not too complicated?
- How many primes are there less than $10^{12}$ (or any other bound we might choose? Are more primes of the form $4k+1$ than $4k-1$, or vice versa?

Some of these questions are interesting because properties of numbers have fascinated humans for thousands of years. On the other hand, some of them (such as primality testing and factorisation) are of very great practical importance: the secret codes that keep internet commerce secure depend on properties of numbers such as primality, factorisation, and modular arithmetic.

Not all these questions will be covered in the course. But here are some problems, which turn out to be closely related to one another, which we will consider. Let $p$ be an odd prime number.

- Can we express $p$ in the form $x^2+y^2$ for some natural numbers $x$ and $y$? (For example, $13 = 3^2+2^2$, but 19 cannot be written in this form, as you can check.)

- Given a natural number $a$, is it congruent to the square of a number $x$ modulo $p$? How do we tell? (For example, $-1 \equiv 5^2 \bmod 13$, but there is no solution to $-1 \equiv x^2 \bmod 19$.)

- Does the equation $x^2-py^2 = 1$ have a solution? What about $x^2-py^2 = -1$? For example, $18^2-13\cdot 5^2 = -1$, but there is no solution to $x^2-19y^2 = -1$.

- How closely can $\sqrt{p}$ be approximated by a rational number? For example, $\sqrt{2}$ is approximately equal to $141421/100000$, but $1393/985$ is an even better approximation, and has much smaller numerator and denominator. How does one find such good approximations?

## 1.2 Euclid's algorithm

We will always count 0 as being a natural number.

We recall that, if $a$ and $b$ are natural numbers and $b > 0$, then there exist unique natural numbers $q$ and $r$ such that $a = bq+r$, with $0 \le r < b$. The numbers $q$ and $r$ are the quotient and remainder when $a$ is divided by $b$. We sometimes write $q = a \operatorname{div} b$ and $r = a \bmod b$. If $a \bmod b = 0$, we say that $b$ *divides* $a$ and write $b \mid a$. (Note: $a/b$ but $b \mid a$.) The division algorithm finds $q$ and $r$ from $a$ and $b$.

Euclid's algorithm is a procedure for finding the greatest common divisor of two natural numbers $a$ and $b$. It can be written as a function $\gcd(a,b)$, defined recursively as follows:

$$\gcd(a,b) = \begin{cases} a & \text{if } b = 0, \\ \gcd(b, a \bmod b) & \text{if } b \neq 0 \end{cases}.$$

The greatest common divisor $d = \gcd(a,b)$ is characterised by the following properties:

- $d \mid a$ and $d \mid b$;

- if $e$ is a natural number satisfying $e \mid a$ and $e \mid b$, then $e \mid d$.

**Example** Find $\gcd(225,157)$. Here is the calculation:

$$\begin{aligned} 225 &= 157\cdot 1+68 \\ 157 &= 68\cdot 2+21 \end{aligned}$$

$$\begin{aligned} 68 &= 21\cdot 3+5\\ 21 &= 5\cdot 4+1\\ 5 &= 1\cdot 5+0 \end{aligned}$$

So $\gcd(225,157)=1$.

The Euclidean algorithm also finds integers $u$ and $v$ such that

$$\gcd(a,b)=ua+vb.$$

In the above example, we can work back up the chain:

$$\begin{aligned} 1 &= 21-5\cdot 4\\ &= 21-(68-21\cdot 3)\cdot 4=21\cdot 13-68\cdot 4\\ &= (157-68\cdot 2)\cdot 13-68\cdot 4=157\cdot 13-68\cdot 30\\ &= 157\cdot 13-(225-157)\cdot 30=157\cdot 43-225\cdot 30. \end{aligned}$$

So we have $u=-30$, $v=43$.

**Remark** Sometimes we will modify the division algorithm as follows. Given integers $a$ and $b$, with $b>0$, there exist unique integers $q$ and $r$ such that $a=bq+r$ and $-b/2<r\leq b/2$. In other words, we allow the remainder to be negative, but choose it to have modulus at most $b/2$. For example, the classical division algorithm gives $12=7\cdot 1+5$, but the modified version gives $12=7\cdot 2-2$.

## 1.3 Primes and factorisation

A natural number $p$ is said to be *prime* if $p>1$ and, whenever $p=ab$ holds for some natural numbers $a$ and $b$, we have either $a=p$, $b=1$, or $a=1$, $b=p$. In other words, $p$ is prime if its only factors in the natural numbers are itself and 1, and these factors are different.

The fact that 1 is not counted as being prime is a convention, but is needed in order for unique factorisation to hold. (If we allowed 1 to be prime, then $6=2\cdot 3=1\cdot 2\cdot 3=1\cdot 1\cdot 2\cdot 3=\cdots$ would have infinitely many prime factorisations!

**Lemma 1.1** *Let $p$ be prime. If $p\mid ab$, then $p\mid a$ or $p\mid b$.*

**Proof** Suppose that $p$ does not divide $a$. Since the only divisors of $p$ are 1 and $p$, and $p$ doesn't divide $a$, we must have $\gcd(a,p)=1$, so there exist integers $u$ and $v$ with $ua+vp=1$. Now $b=uab+vpb$; and $p$ divides $uab$ (since it divides $ab$ by assumption) and $p$ divides $vpb$; so $p$ divides their sum, which is $b$. □

**Theorem 1.2** *Any natural number greater than* 1 *can be written as a product of prime numbers, and this expression is unique apart from re-ordering the factors.*

**Proof** We show the existence of a factorisation into primes by induction. Given a natural number $n$, if $n$ is prime, then it is the product of just one prime. (This starts the induction at $n = 2$, and is also part of the inductive step.) Otherwise, $n$ has a factorisation $n = ab$ with $a, b < n$. By the induction hypothesis (since both $a$ and $b$ are greater than 1 but smaller than $n$), they have factorisations into primes; putting them together we have a factorisation of $n$.

For the uniqueness, we use the lemma. Suppose that

$$n = p_1 p_2 \cdots p_r = q_1 q_2 \cdots q_s,$$

where $p_1, p_2, \ldots, p_r, q_1, q_2, \ldots, q_s$ are primes. Clearly $p_1$ divides $q_1 q_2 \cdots q_s$; by the lemma, either $p_1$ divides $q_1$ or $p_1$ divides $q_2 \cdots q_s$. Continuing, we find that $p_1$ divides one of the primes $q_1, \ldots, q_s$. By re-ordering them if necessary, we can assume that $p_1$ divides $q_1$, whence $p_1 = q_1$ since $q_1$ is prime. Now we can cancel off the first factor from both sides and continue the process, until we have shown that the two factorisations are the same. □

## 1.4 Congruences and modular arithmetic

Let $n$ be a natural number. We say that two integers $a$ and $b$ are *congruent modulo n* if $n$ divides $a - b$. We write this as

$$a \equiv b \bmod n.$$

Note that this is a slightly different use of the word "mod" from the one we used earlier to denote the remainder. But it is closely connected; two numbers are congruent modulo $n$ if and only if they leave the same remainder when they are divided by $n$.

Congruence modulo $n$ is an equivalence relation; the equivalence classes are called *congruence classes modulo n*. There are exactly $n$ congruence classes, corresponding to the $n$ possible remainders (0, 1, ..., $n - 1$) we could obtain when we divide a number by $n$.

We denote by $[a]_n$ the congruence class modulo $n$ containing $a$, and by $\mathbb{Z}_n$ the set of congruence classes modulo $n$. The set $\mathbb{Z}_n$ is a *ring*, in fact a *commutative ring with identity*; this means that congruence classes can be added or multiplied, by the rules

$$[a]_n + [b]_n = [a+b]_n, \qquad [a]_n \cdot [b]_n = [ab]_n,$$

and the usual laws (commutative, associative, distributive, identity, and additive inverse laws hold. See the Introduction to Algebra lecture notes if you need a reminder about this.

Here are the addition and multiplication tables of $\mathbb{Z}_4$. I have written the entries in the tables as $a$ rather than $[a]_4$ to save clutter.

| + | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | 0 | 1 | 2 | 3 |
| 1 | 1 | 2 | 3 | 0 |
| 2 | 2 | 3 | 0 | 1 |
| 3 | 3 | 0 | 1 | 2 |

| · | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 1 | 2 | 3 |
| 2 | 0 | 2 | 0 | 2 |
| 3 | 0 | 3 | 2 | 1 |

**Proposition 1.3** *If $p$ is prime, then $\mathbb{Z}_p$ is a* field*; that is, all non-zero elements have multiplicative inverses.*

**Proof** Suppose that $[a]_p$ is a non-zero element of $\mathbb{Z}_p$. This means $[a]_p \neq [0]_p$, so $p$ does not divide $a$. Since $p$ is prime, $\gcd(a,p) = 1$. By Euclid's algorithm, there are integers $u$ and $v$ satisfying $ua + vp = 1$. This means that $ua \equiv 1 \bmod p$, so that

$$[u]_p \cdot [a]_p = [1]_p.$$

So $[u]_p$ is the inverse of $[a]_p$. □

For example, take $p = 157$. What is the inverse of $[225]_{157}$? Our earlier calculation showed that $43 \cdot 157 - 30 \cdot 225 = 1$, so that the required inverse is $[-30]_{157} = [127]_{157}$.

As a consequence we prove *Fermat's Little Theorem*:

**Theorem 1.4** *Let $p$ be a prime number. Then $n^p \equiv n \bmod p$ for any natural number $n$.*

**Proof** If $n \equiv 0 \bmod p$, then the conclusion is certainly true; so suppose not. Then $[n]_p$ is an element of the multiplicative group of non-zero elements of $\mathbb{Z}_p$. By Lagrange's Theorem (see the Introduction to Algebra notes), the order of this element divides the order of the group, which is $p-1$. So $([n]_p)^{p-1} = [1]_p$, or in other words, $n^{p-1} \equiv 1 \bmod p$. Multiplying both sides by $n$ gives the result. □

**Exercise** Prove Fermat's Little Theorem by induction on $n$. (*Hint*: Use the Binomial Theorem and the fact (which you should prove) that the binomial coefficients $\binom{p}{k}$ are divisible by $p$ for $1 \leq k \leq p-1$.

Fermat's Little Theorem shows that it is possible to show that a number $n$ is composite without finding any factors of $n$. If we calculate $a^n \bmod n$ and the answer comes out to be different from $a$, then we know that $n$ is composite.

**Example** $3^{2047} \equiv 992 \bmod 2047$, so 2047 is not prime.

The computation is not as bad as it might appear. Since $2048 = 2^{11}$, we can work out $3^{2048} \bmod 2047$ by successive squaring. All congruences mod 2047; so no number occurring in the calculation is larger than $2046^2$, and the whole thing can easily be done on a calculator.

$$\begin{array}{rcl}
3^1 & = & 3 \\
3^2 & = & 9 \\
3^4 = 9^2 & = & 81 \\
3^8 = 81^2 & = & 420 \\
3^{16} = 420^2 & \equiv & 358 \\
3^{32} \equiv 358^2 & \equiv & 1250 \\
3^{64} \equiv 1250^2 & \equiv & 639 \\
3^{128} \equiv 639^2 & \equiv & 968 \\
3^{256} \equiv 968^2 & \equiv & 1545 \\
3^{512} \equiv 1545^2 & \equiv & 223 \\
3^{1024} \equiv 223^2 & \equiv & 601 \\
3^{2048} \equiv 601^2 & \equiv & 929
\end{array}$$

So $2^{2047} \equiv 992 \bmod 2047$, on dividing by 3 (equivalently, multiplying by the inverse of 3 in $\mathbb{Z}_{2047}$, which is 1365).

Note that the successive squaring method avoids having to compute very large numbers. We can evaluate $3^{2048}$ by just eleven squaring operations of numbers smaller than 2047 together with taking the remainder mod 2047.

Unfortunately, it doesn't always work. If we had used 2 rather than 3, we would have found that $2^{2047} \equiv 2 \bmod 2047$. The converse of Fermat's Little Theorem is false! (See the exercise on Carmichael numbers below.)

## 1.5 The Chinese Remainder Theorem

The Chinese Remainder Theorem is about solving simultaneous congruences to different moduli.

**Theorem 1.5** *Let a and b be natural numbers with* $\gcd(a,b) = 1$*, and let c and d be arbitrary integers. Then there is a solution to the simultaneous congruences*

$$x \equiv c \bmod a,$$
$$x \equiv d \bmod b.$$

*Moreover, the solution is unique modulo ab; that is, if $x_1$ and $x_2$ are two solutions, then $x_1 \equiv x_2 \bmod ab$.*

**Proof** Since $\gcd(a,b) = 1$, there are integers $u$ and $v$ with $ua+bv = 1$. Now let

$$x = dau+cbv.$$

We have $bv \equiv 1 \bmod a$, and $au \equiv 1 \bmod b$. So $x \equiv cbv \equiv c \bmod a$, and $x \equiv dau \equiv d \bmod b$, as required.

If $x_1$ and $x_2$ are two solutions, then $x_1 \equiv c \equiv x_2 \bmod a$ and $x_1 \equiv d \equiv x_2 \bmod b$. So both $a$ and $b$ divide $x_1 - x_2$. Since $a$ and $b$ are coprime, $ab$ divides $x_1 - x_2$, so that $x_1 \equiv x_2 \bmod ab$ as required. □

This can be extended to an arbitrary number of congruences to pairwise coprime moduli.

**Example** Find all numbers congruent to $2 \bmod 3$, $1 \bmod 4$ and $3 \bmod 5$.

The theorem shows that there is a unique solution mod 60, which can be found by trial and error, or systematically as in the proof, which we do here.

Since $-3+4 = 1$, the number $-3 \cdot 1+4 \cdot 2 = 5$ satisfies the first two congruences. Now we look for a number congruent to $5 \bmod 12$ and $3 \bmod 5$. We have $-2 \cdot 12+5 \cdot 5 = 1$, so the solution is $-2 \cdot 12 \cdot 3+5 \cdot 5 \cdot 5 = 53$. So the general solution is the congruence class $[53]_{60}$ (all numbers congruent to $53 \bmod 60$).

## 1.6 And finally . . .

Remember Euclid’s famous proof of the existence of infinitely many primes, which you will find in the Introduction to Algebra notes.

It is possible to adapt Euclid’s method for other purposes. Here is an example. Note that, apart from 2, all primes are odd, and so are of one or other of the forms $4k+1$ and $4k+3$ for some natural number $k$.

**Theorem 1.6** *There are infinitely many primes of the form $4k+3$ for natural numbers k.*

**Proof** Suppose that there are only finitely many such primes, say $p_1,\ldots,p_r$. Consider the number $n = 4p_1 \cdots p_r - 1$. Clearly $n$ is of the form $4k+3$, and so it must be divisible by some prime of this form. (A number with a factor 2 is even, while a product of factors of the form $4k+1$ is itself of this form, since

$(4k+1)(4l+1) = 4(4kl+k+l)+1$.) So one of the primes $p_1,\ldots,p_r$ must be a factor of $n$, since these are all primes congruent to 3 mod 4.

But by assumption, $n \equiv -1 \bmod p$ for $p = p_1,\ldots,p_r$, so none of $p_1,\ldots,p_r$ can divide $n$. So we have a contradiction to our assumption, and there must be infinitely many primes of this form. □

It is also true that there are infinitely many primes of the form $4k+1$ (and indeed, roughly equal numbers of the two forms below any given bound), but these things are more difficult to prove.

## Exercises

**1.1** Prove that a number with a periodic decimal expansion

$$a_1\ldots a_k.a_{k+1}\ldots a_{k+l}\overline{a_{k+l+1}\ldots a_{k+l+m}}$$

is rational. (This notation means that the digits from $a_{k+l+1}$ to $a_{k+l+m}$ repeat infinitely; for example, $1.2\overline{34} = 1.2343434\ldots$.)

**1.2** Find $\gcd(245,43)$ and express it in the form $245u+43v$.

**1.3** Find all integer solutions of the congruence $x^2 \equiv 2$ mod 17.

**1.4** Let $\mathbb{Z}_n$ denote the ring of integers modulo $n$.
How many solutions does the equation $x^2 = 1$ have

(a) in $\mathbb{Z}_8$,

(b) in $\mathbb{Z}_9$,

(c) in $\mathbb{Z}_{11}$?

**1.5** List the prime numbers less than 100. Which of them can be written in the form $x^2+y^2$ for integers $x$ and $y$?

**1.6** A natural number $q$ is said to be a *Carmichael number* if $q$ is not prime but satisfies the conclusion of Fermat's Little Theorem, that is, $n^q \equiv n$ mod $q$ for all integers $n$.

(a) Let $p$ be a prime number, and suppose that $p-1$ divides $q-1$. Show that $n^q \equiv n$ mod $p$.

(b) Hence show that, if $q$ is a product of distinct primes, and every prime $p$ which divides $q$ has the property that also $p-1$ divides $q-1$, then $q$ is a Carmichael number.

(c) Hence show that 561 is a Carmichael number.

**1.7** There is an unknown number of objects. When counted in threes, the remainder is 2; when counted in fives, the remainder is 3; when counted in sevens, the remainder is 2. How many objects are there?

This problem is taken from the fourth-century Chinese text *Sun Zi suanjing* (Master Sun's Arithmetic Manual). He gives the following formula for its solution:

> Not in every third person is there one aged three score and ten,
> On five plum trees only twenty-one boughs remain,
> The seven learned men meet every fifteen days,
> We get our answer by subtracting one hundred and five over and
> over again.

Can you explain this?

**Note:** Sun Zi (formerly transliterated as Sun Tsu), is not to be confused with the military strategist of the same name.

**1.8** (a) Let $p$ be a prime number, not 2 or 5. Show that there exists a positive integer $k$ such that $10^k \equiv 1 \bmod p$.

(b) Let $k_p$ be the smallest positive integer $k$ with this property. Show that $k_p$ divides $p-1$.

(c) Show that the digits in the infinite decimal expression for $1/p$ are periodic with period $k_p$.

(d) Find a prime $p$ for which $k_p = p-1$.

**1.9** What goes wrong with the argument in the last section if you try to prove that there are infinitely many primes congruent to 1 mod 4?

# Chapter 2

# Algebraic numbers

An algebraic number is one which satisfies a polynomial with integer coefficients. From Pythagoras to the present day, a lot of number theory hae been concerned with these numbers, and in particular in trying to decide whether particular numbers of interest to mathematics are algebraic or not.

## 2.1 Algebraic numbers and algebraic integers

Pythagoras and his school discovered that the square root of 2 is not a rational number. However, it is an easy number to describe geometrically: it is the ratio of the diagonal of a square to its side. The number $\pi$ has a more complicated geometric description: it is the ratio of the circumference of a circle to its diameter, but there is no simple method to construct a straight line which is equal to the circumference of a given circle. (We know now, for example, that such a line cannot be constructed with the traditional geometric instruments of "ruler and compass".)

We make a distinction between algebraic numbers (which are roots of polynomials with integer coefficients) and transcendental numbers (which are not):

**Definition** Let $u$ be a complex number. We say that $u$ is an *algebraic number* if there is a non-zero polynomial $f$ with integer coefficients such that $f(u) = 0$; and $u$ is a *transcendental number* otherwise. Moreover, $u$ is an *algebraic integer* if it is the root of a non-zero *monic* polynomial (one with leading coefficient 1) over the integers.

Note that, if we have any non-zero polynomial over the integers, we can divide by the leading coefficient to get a monic polynomial over the rationals. Conversely, given a monic polynomial over the rationals, we can multiply by the least common multiple of the denominators of the coefficients to obtain a non-zero polynomial over the integers. So an equivalent definition is: $u$ is an algebraic number if there is a non-zero monic polynomial $f$ with rational coefficients such that $f(u) = 0$.

For example, $u = \sqrt{2}$ is an algebraic integer since it satisfies the polynomial $u^2 - 2 = 0$. The complex number i is an algebraic integer; so is the *golden ratio* $\phi = (1+\sqrt{5})/2$ (it satisfies $\phi^2 - \phi - 1 = 0$). It is known that $\pi$ and e are transcendental numbers, but we will not give the proof here. (You can find these proofs in Ian Stewart's book *Galois Theory*.

Any integer is an algebraic integer; the integer $n$ satisfies the polynomial $x - n = 0$. Similarly, any rational number is an algebraic number. In the other direction, we have:

**Proposition 2.1** *A rational number is an algebraic integer if and only if it is an integer.*

For this reason, we sometimes call the ordinary integers "rational integers".

**Proof** We have seen that integers are algebraic integers; we have to prove that a rational number which is an algebraic integer is an integer.

Let $q = a/b$ be a rational number in its lowest terms (so that $\gcd(a,b) = 1$). Suppose that $q$ satisfies a monic polynomial with integer coefficients, say

$$f(q) = q^n + c_{n-1}q^{n-1} + \cdots + c_1 q + c_0 = 0.$$

Putting $q = a/b$, and multiplying this equation by $b^n$, we obtain

$$a^n + c_{n-1}a^{n-1}b + \cdots + c_1 ab^{n-1} + c_0 b^n = 0.$$

Now every term in this equation except the first is divisible by $b$, so $b$ divides $a^n$. Applying Lemma 1.1 repeatedly, we find that $b \mid a$. But also $b \mid b$, so $b \mid \gcd(a,b) = 1$. This means that $b = 1$, so that $q$ is an integer. □

There is a result, which I will not prove, which makes things like this *much* easier. This is known as Gauss's Lemma. It can be stated in many different ways. (You might like to look at p.258 in my book *Introduction to Algebra*.) But the following will do for our purposes.

We define the *minimal polynomial* of an algebraic number $\alpha$ to be the monic polynomial with rational coefficients *of smallest possible degree* satisfied by $\alpha$. (Any algebraic number satisfies a monic polynomial with rational coefficients, and we can certainly choose one of smallest degree. Why is it unique? Suppose that $f_1(x)$ and $f_2(x)$ were two different polynomials of the same (smallest) degree satisfied by $\alpha$, and let $g(x) = f_1(x) - f_2(x)$. Then $f_1(\alpha) = f_2(\alpha) = 0$, so $g(\alpha) = 0$, but $g$ has smaller degree than $f_1$ and $f_2$.)

**Theorem 2.2** *The algebraic number $\alpha$ is an algebraic integer if and only if its minimal polynomial has integer coefficients.*

Now let $q$ be a rational number. It satisfies the polynomial $x-q=0$, and clearly this is monic and has smallest possible degree, so it is the minimal polynomial of $q$. So $q$ is an algebraic integer if and only if the coefficients 1 and $-q$ of this polynomial are both integers, i.e. if and only if $q$ is an integer.

One of the most important properties of algebraic numbers is the following:

**Theorem 2.3** *(a) Let $a$ and $b$ be algebraic integers. Then $a-b$ and $ab$ are algebraic integers.*

*(b) Let $a$ and $b$ be algebraic numbers. Then $a-b$, $ab$, and (if $a \neq 0$) $1/a$ are algebraic numbers.*

I do not expect you to memorise the proof of this theorem. But it uses ideas from linear algebra, and may be useful revision of linear algebra for you. I have given the proof in an appendix to this section.

The theorem can be expressed in the language of algebraic structures as follows:

**Corollary 2.4** *(a) The algebraic numbers form a field.*

*(b) The algebraic integers form a commutative ring with identity.*

**Proof** The Theorem above says that they satisfy the conditions of the subfield and subring tests as subsets of the complex numbers. □

## 2.2 Quadratic irrationals

In this course we will be particularly interested in numbers of the form $a+b\sqrt{d}$, where $a$ and $b$ are rational numbers and $d$ is a squarefree integer not equal to 1. (An integer $d$ is *squarefree* if $n^2 \mid d$ implies $n=1$ for positive integer $n$. Clearly, if $d$ were not squarefree, we could write $d=cn^2$, and then $a+b\sqrt{d}=a+bn\sqrt{c}$.) A number of this form is called a *quadratic irrational*. There will be much more about quadratic irrationals later in the notes!

The number $u=a+b\sqrt{d}$ is an algebraic number, since it satisfies the quadratic equation $u^2-2au+(a^2-db^2)=0$. (This is a quadratic with rational coefficients; we obtain one with integer coefficients by multiplying up by the denominators of the coefficients.) In fact, the polynomial

$$f(x)=x^2-2ax+(a^2-db^2)$$

is the minimal polynomial of $u$. For it is a monic rational polynomial satisfied by $u$, and has degree 2; a polynomial of smaller degree would have to have degree 1, and have the form $x-q$, but if $u$ satisfied this polynomial, then $u=q$ would be a rational number.

Using Gauss's Lemma, we can now decide when a quadratic irrational is an algebraic integer.

**Proposition 2.5** *Let $a,b$ be rational numbers and $d$ a squarefree integer. Then $a+b\sqrt{d}$ is an algebraic integer if and only if either*

*(a) $a,b$ are integers; or*

*(b) $d\equiv 1 \bmod 4$ and $a-\frac{1}{2}$ and $b-\frac{1}{2}$ are integers.*

**Example** $(1+\sqrt{5})/2$ (the golden ratio) and $(-1+\sqrt{-3})/2$ (a complex cube root of unity) are algebraic integers but $(1+\sqrt{3})/2$ is not.

**Proof** By Gauss's Lemma, we just have to show that the monic quadratic equation satisfied by $u=a+b\sqrt{d}$ has integer coefficients precisely in the cases given.

The quadratic is $x^2-2ax+(a^2-db^2)$. So the question is, when is it true that $2a$ and $a^2-db^2$ are integers? If $2a$ is even then $a$ is an integer; if $2a$ is odd then $a-\frac{1}{2}$ is an integer.

Suppose that $a$ is an integer. Then $db^2$ is an integer; since $d$ is squarefree, this implies that $b$ is also an integer, since if $b=m/n$ with $\gcd(m,n)=1$ then necessarily $n^2\mid d$.

Suppose that $a=k+\frac{1}{2}$, with $k\in\mathbb{Z}$. Then $a^2=k^2+k+\frac{1}{4}$, so $db^2-\frac{1}{4}$ is an integer. This means that $b=l+\frac{1}{2}$ for $l\in\mathbb{Z}$ (so that $db^2$ has denominator 4). Then $db^2-\frac{1}{4}=(l^2+l)d+(d-1)/4$, so we must have $d\equiv 1 \bmod 4$. □

## 2.3 Appendix: Sums, products and quotients

In this section we prove Theorem 2.3. If $a$ and $b$ satisfy monic polynomials over the integers or rationals, we have to show that their difference and product do also. The direct approach is quite difficult; to convince yourself of this, try writing down a monic polynomial over the integers which has $\sqrt[3]{2}-\sqrt{3}$ as a root. So we need a different strategy.

First, we give an equivalent characterisation of algebraic numbers and algebraic integers, using the concept of eigenvalues from linear algebra.

**Proposition 2.6** *Let $u$ be a complex number. Then* *(a) u is an algebraic number if and only if it is an eigenvalue of a matrix over* $\mathbb{Q}$*;*

*(b) u is an algebraic integer if and only if it is an eigenvalue of a matrix over* $\mathbb{Z}$*.*

**Proof** The eigenvalues of a matrix $A$ are the roots of the *characteristic polynomial* of $A$, the monic polynomial $\det(xI-A)$. If $A$ is a rational (resp. integer) matrix, this polynomial has rational (resp. integer) coefficients.

Conversely, given any monic polynomial $f(x)$ of degree $n$, there is a matrix $C(f)$, called the *companion matrix* of $f$, whose characteristic polynomial is $f$ (and hence whose eigenvalues are the roots of $f$). It is the $n\times n$ matrix which has entries 1 immediately above the diagonal in the first $n-1$ rows, 0 in all other positions in these rowss, and the coefficients of $f$ (other than the coefficient of $x^n$) in reverse order with the signs changed in the $n$th row. So $C(f)$ is a matrix of integers (resp. rational numbers) if and only if the coefficients of $f$ are integers (resp. rational numbers). □

Now we define the *Kronecker product* of matrices. Let $A$ and $B$ be two matrices, possibly of different sizes. Suppose that $A$ is $n\times m$, with $(i,j)$ entry $a_{ij}$. Then the matrix $A\otimes B$ is defined to be the matrix in block form with $m\times n$ blocks, the $(i,j)$ block being $a_{ij}B$. For example, if

$$A=\begin{pmatrix}1&2\\3&4\end{pmatrix},\qquad B=\begin{pmatrix}5&6\\7&8\end{pmatrix}$$

then

$$A\otimes B=\begin{pmatrix}5&6&10&12\\7&8&14&16\\15&18&20&24\\21&24&28&32\end{pmatrix}.$$

The following result is what is needed. The proof is just a rather boring linear algebra argument involving playing with subscripts.

**Proposition 2.7** *Let $A,B,C,D$ be matrices of the appropriate sizes so that $AC$ and $BD$ are defined. Then $(A\otimes B)(C\otimes D)$ is defined, and*

$$(A\otimes B)(C\otimes D)=AC\otimes BD.$$

In particular, if $A$ and $B$ are matrices and $v,w$ are column vectors such that $Av$ and $Bw$ are defined, then

$$(A\otimes B)(v\otimes w)=Av\otimes Bw.$$

**Proposition 2.8** *Let $A$ and $B$ be square matrices of sizes $n \times n$ and $m \times m$ respectively, and suppose that $a$ and $b$ are eigenvalues of $A$ and $B$ respectively. Then*

- *$a-b$ is an eigenvalue of $A \otimes I_m - I_n \otimes B$;*
- *$ab$ is an eigenvalue of $A \otimes B$.*

**Proof** Let $u$ and $v$ be column vectors of lengths $n$ and $m$ which are eigenvectors of $A$ and $B$ respectively, so that $Au = au$ and $Bv = bv$. Then $u \otimes v$ is a column vector of length $mn$; and we have

$$\begin{aligned}(A \otimes I_m - I_n \otimes B)(u \otimes v) &= (a-b)(u \otimes v),\\ (A \otimes B)(u \otimes v) &= ab(u \otimes v).\end{aligned}$$

(For example, we have $(A \otimes B)(u \otimes v) = Au \otimes Bv$.) □

Now we can prove most of the theorem. If $a$ and $b$ are roots of monic polynomials over $\mathbb{Q}$ or $\mathbb{Z}$, then they are eigenvalues of matrices over $\mathbb{Q}$ or $\mathbb{Z}$, and hence so are their difference and product.

For the inverse, we proceed directly. Suppose that $a \neq 0$ and $a$ is a root of a rational polynomial $x^n + c_{n-1}x^{n-1} + \cdots + c_1x + c_0$; that is,

$$a^n + c_{n-1}a^{n-1} + \cdots + c_1a + c_0 = 0.$$

Dividing the polynomial by a suitable power of $x$, we can assume that $c_0 \neq 0$. Then dividing the whole expression by $c_0a^n$ and reversing the terms, we have

$$(1/a)^n + (c_1/c_0)(1/a)^{n-1} + \cdots + (c_{n-1}/c_0)(1/a) + c_n/c_0 = 0,$$

so that $1/a$ is also the root of a rational polynomial. □

## Exercises

**2.1** Find a polynomial with integer coefficients which has $\sqrt{2}+\sqrt{3}$ as a root.

**2.2** For each of the following numbers, say whether it is an algebraic number and whether it is an algebraic integer. If you answer "yes", justify your answer by giving a polynomial satisfied by the number in question.

(a) $\sqrt[3]{3}+1$

(b) $1+\frac{1}{2}\sqrt{5}$

(c) $(1+\sqrt{13})/2$.

**2.3** Let $\alpha$ satisfy the polynomial $x^3 + ax^2 + bx + c = 0$. Find a polynomial satisfied by $\alpha^2$.

# Chapter 3

# Finite continued fractions

Now we embark on a major theme of this course; a method for representing rational or irrational numbers by finite or infinite strings of integers, by means of continued fractions.

## 3.1 Introduction

Let us return to the calculation of $\gcd(225, 157)$ from the preceding chapter.

$$\begin{aligned}
225 &= 157 \cdot 1 + 68 \\
157 &= 68 \cdot 2 + 21 \\
68 &= 21 \cdot 3 + 5 \\
21 &= 5 \cdot 4 + 1 \\
5 &= 1 \cdot 5 + 0
\end{aligned}$$

We can express this in a different way.

$$\begin{aligned}
\frac{21}{5} &= 4 + \frac{1}{5} \\
\frac{68}{21} &= 3 + \cfrac{1}{4 + \cfrac{1}{5}} \\
\frac{157}{68} &= 2 + \cfrac{1}{3 + \cfrac{1}{4 + \cfrac{1}{5}}}
\end{aligned}$$

$$\frac{225}{157} = 1+\cfrac{1}{2+\cfrac{1}{3+\cfrac{1}{4+\cfrac{1}{5}}}}$$

These expressions are called *continued fractions*. Clearly we will cover a lot of paper writing out things like this in full, so we abbreviate them. The expression on the right of the last equation will be written $[1;2,3,4,5]$. Notice the semicolon ; in the notation. Later we will define $[1,2,3,4,5]$ to mean something different! The part after the semicolon represents the fractional part of $225/157$.

**Proposition 3.1** *Let $q = a/b$ be a rational number greater than $1$ in its lowest terms, so that $\gcd(a,b) = 1$. Then $q$ can be written as a continued fraction $[a_0;a_2,a_3,\dots,a_n]$ for some positive integers $a_0,\dots,a_n$ with $a_n > 1$.*

Note that this proposition includes a degenerate case: if $b = 1$ then the continued fraction is just $[a;]$.

**Proof** I will give two proofs of this. They are essentially the same, but the first is recursive, the second is more explicit. Both involve extracting the continued fraction from Euclid's algorithm, as we did in the example.

**First proof** We argue by induction on $b$. If $b = 1$ then, as we just remarked, the continued fraction is just $[a;]$ and the last integer is $a$, which is greater than 1 by assumption.

If $b > 1$ then, since the fraction is in its lowest terms, it is not an integer. If $a = bc + r$, then $r \neq 0$, and we have

$$\frac{a}{b} = c + \frac{r}{b} = a_0 + \frac{1}{b/r},$$

with $a_0 = c$, and $b/r > 1$ and in its lowest terms. By the induction hypothesis,

$$\frac{b}{r} = [a_1;a_2,\dots,a_n],$$

with $a_n > 1$. Then

$$\frac{a}{b} = a_0 + \frac{1}{[a_1;a_2,\dots,a_n]} = [a_0;a_1,\dots,a_n].$$

To see the last step, write it out as a continued fraction:

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots}}.$$

We can express this proof by a recurrence:

$$a_0 = \lfloor q \rfloor, \qquad \frac{1}{q-a_0} = [a_1, \ldots, a_n].$$

**Second proof** We run the Euclidean algorithm:

$$\begin{array}{rcl}
a = a_0 b + r_1 & \text{so} & a/b = a_0 + r_1/b \\
b = a_1 r_1 + r_2 & \text{so} & b/r_1 = a_1 + r_2/r_1 \\
r_1 = a_2 r_2 + r_3 & \text{so} & r_1/r_2 = a_2 + r_3/r_2 \\
\vdots & & \vdots \\
r_{n-2} = a_{n-1} r_{n-1} + 1 & \text{so} & r_{n-2}/r_{n-1} = a_{n-1} + 1/r_{n-1} \\
r_{n-1} = a_n \cdot 1 & \text{so} & r_{n-1} = a_n.
\end{array}$$

We have $1 < r_{n-1} < \cdots < r_1 < b < a$, so all the fractions on the left are greater than 1 and those on the right are less than 1; so the integers on the right are the integer parts of the fractions on the left. Putting it all together, we have

$$q = \frac{a}{b} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_{n-1} + \cfrac{1}{a_n}}}},$$

that is, $q = [a_0; a_1, \ldots, a_{n-1}, a_n]$. □

Conversely, any sequence of positive integers, the last greater than 1, defines a unique rational number greater than 1: Clearly we have

$$\begin{array}{rcl}
[a_0;] & = & a_0 \\
[a_0; a_1, \ldots, a_n] & = & a_0 + \dfrac{1}{[a_1; a_2, \ldots, a_n]} \quad \text{for } n > 0.
\end{array}$$

This also provides an inductive definition of the symbol $[a_0; a_1, \ldots, a_n]$ (by induction on $n$, the number of terms).

Now consider the second proof of the proposition. Let the fractions on the left of the equations be

$$q_0 = q = \frac{a}{b}, q_1 = \frac{b}{r_1}, q_2 = \frac{r_1}{r_2}, \ldots q_{n-1} = \frac{r_{n-2}}{r_{n-1}} \text{ and } q_n = r_{n-1} = a_n.$$

Then we can write the recurrence as

$$a_i = \lfloor q_i \rfloor, \qquad q_{i+1} = \frac{1}{q_i - a_i}$$

for $i = 0,\ldots,n-1$, and $a_n = q_n$.

This rule allows us to compute the continued fraction of a rational number without explicitly running the Euclidean algorithm. For example, let $q = 87/38$. We have

$$\begin{aligned} a_0 = \lfloor 87/38 \rfloor = 2, &\qquad q_1 = \frac{1}{(87/38-2)} = \frac{38}{11}, \\ a_1 = \lfloor 38/11 \rfloor = 3, &\qquad q_2 = \frac{1}{(38/11-2)} = \frac{11}{5}, \\ a_2 = \lfloor 11/5 \rfloor = 2, &\qquad q_3 = \frac{1}{(11/5)-2} = \frac{5}{1}, a_3 = q_3 = 5. \end{aligned}$$

So

$$\frac{87}{38} = [2;3,2,5].$$

**Remark 1** We assumed that $q > 1$. If we relax this assumption, the only difference is that the first term of the continued fraction may be zero or negative. For example, if $q = -3/5$, we have

$$\begin{aligned} a_0 = \langle -3/5 \rangle = -1, &\qquad q_1 = \frac{1}{-(3/5)+1} = \frac{5}{2}, \\ a_1 = \langle 5/2 = 2, &\qquad q_2 = \frac{1}{(5/2)-2} = \frac{2}{1}, \\ a_2 = q_2 = 2, & \end{aligned}$$

so $-5/3 = [-1;2,2]$.

**Remark 2** If we relax the condition that the last entry in the continued fraction is greater than one, then only one small change is necessary. Since $a_n = (a_n - 1) + \frac{1}{1}$, we see that

$$[a_0;a_1,a_2,\ldots,a_n] = [a_0;a_1,a_2,\ldots,a_n - 1, 1].$$

So $87/38 = [2;3,2,4,1]$.

We end this section by showing the uniqueness of the continued fraction for any rational number greater than 1 if we require that the last entry is greater than 1.

**Theorem 3.2** *If $q = [a_0;a_1,\ldots,a_n] = [b_0;b_1,\ldots,b_n]$, with $a_n, b_m > 1$, then $m = n$ and $a_i = b_i$ for $i = 0,\ldots,n$.*

**Proof** We prove this by induction on $n$. If $n = 0$, then $q = a_0$ is an integer. Now if $m > 0$ then

$$q = b_0 + \frac{1}{[b_1;b_2,\ldots,b_m]},$$

and the fraction is less than one (since the denominator is greater than 1), which is impossible. So $m = 0$ and $a_0 = b_0$. This starts the induction.

Suppose that the assertion is true with $n-1$ replacing $n$. Then we have

$$a_0 + \frac{1}{[a_1;\ldots,a_n]} = b_0 + \frac{1}{[b_1;\ldots,b_m]},$$

and again the fractions are less than one; so $a_0 = b_0 = \lfloor q \rfloor$. Then we have $[a_1;\ldots,a_n] = [b_1;\ldots,b_m]$, each expression having one fewer term in its continued fraction than $q$; so by the inductive hypothesis, $m-1 = n-1$ and $a_i = b_i$ for $i = 1,\ldots,n$. So we are done. □

## 3.2 The [ ] functions

In this section, we analyse continued fractions further by finding recurrence relations for the numerator and denominator of a given continued fraction. In this section, we will think of the numbers $a_0,\ldots,a_n$ which appear as arguments to these functions as being positive integers; but in fact everything is quite formal, and they could in fact be any real numbers.

**Definition** Let $n \geq 1$ and let $a_0,\ldots,a_n$ be positive real numbers. Define

$$\begin{aligned}
[a_0] &= a_0, \\
[a_0,a_1] &= a_0a_1+1, \\
[a_0,a_1,\ldots,a_k] &= a_0[a_1,\ldots,a_k]+[a_2,\ldots,a_k] \text{ for } 1 < k \leq n.
\end{aligned}$$

Note that the last clause expresses a function of $k$ variables in terms of functions of $k-1$ and $k-2$ variables; so the definition is good.

**Remark:** Often we will adopt the convention that [ ] (with no numbers in the square brackets) is equal to 1. If we do this, the induction gives the correct answer for $[a_0,a_1]$:

$$[a_0,a_1] = a_0[a_1]+[\,] = a_0a_1+1.$$

**Warning:** $[a_0,a_1,\ldots,a_n]$ is *not* the same as the continued fraction $[a_0;a_1,\ldots,a_n]$ defined in the last section. Be very careful to distinguish them!

**Example** Find $[1,2,3,4,5]$.

We calculate this by working from the back, since each expression only involves the last so many variables.

$$[5] = 5$$

$$\begin{aligned}
[4,5] &= 4\cdot 5+1=21\\
[3,4,5] &= 3[4,5]+[5]=3\cdot 21+5=68\\
[2,3,4,5] &= 2[3,4,5]+[4,5]=2\cdot 68+21=157\\
[1,2,3,4,5] &= 1[2,3,4,5]+[3,4,5]=157+68=225.
\end{aligned}$$

If you look back at our first example of a continued fraction, you will see a connection, which is expressed in the following theorem.

**Proposition 3.3** *Let $a_0,\dots,a_n$ be positive integers. Then*

*(a)* $\gcd([a_0,a-1,\dots,a_n],[a_1,\dots,a_n])=1$*;*

*(b) The continued fraction $[a_0;a_1,\dots,a_n]$ is equal to*

$$\frac{[a_0,a_1,\dots,a_n]}{[a_1,\dots,a_n]}.$$

**Remark:** With our convention that $[\,]=1$, this gives the correct answer for $n=0$ in part (b):

$$[a_0;]=\frac{[a_0]}{[\,]}=\frac{a_0}{1}=a_0.$$

**Proof** We prove both parts by induction on $n$.

(a) To start the induction, $[a_0,a_1]=a_0a_1+1$, and

$$\gcd(a_0a_1+1,a_1)=\gcd(a_1,1)=1.$$

So suppose that the result holds for $n-1$. Let $x=[a_0,\dots,a_n]$, $y=[a_1,\dots,a_n]$ and $z=[a_2,\dots,a_n]$. By the induction hypothesis, $\gcd(y,z)=1$; and $x=a_0y+z$, so

$$\gcd(a_0y+z,y)=\gcd(y,z)=1.$$

(b) By the remark, $[a_0;]=a_0=[a_0]$, so the induction starts. Suppose that it holds for $n-1$. With the same notation as in the previous part,

$$[a_1;a_2,\dots,a_n]=\frac{y}{z},$$

and so

$$\begin{aligned}
[a_0;a_1,\dots,a_n] &= a_0+\frac{1}{[a_1;\dots,a_n]}\\
&= a_0+\frac{z}{y}\\
&= \frac{a_0y+z}{y}\\
&= \frac{x}{y},
\end{aligned}$$

as required. □

**Example** Calculate $[3;1,4,1,6]$. We have

$$\begin{aligned}
[6] &= 6\\
[1,6] &= 1[6]+1=7\\
[4,1,6] &= 4[1,6]+[6]=34\\
[1,4,1,6] &= 1[4,1,6]+[1,6]=41\\
[3,1,4,1,6] &= 3[1,4,1,6]+[4,1,6]=157
\end{aligned}$$

so $[3;1,4,1,6]=157/41$.

The next theorem, due to Euler, gives a non-recursive way of computing these functions.

**Theorem 3.4** *Let $a_0,\dots,a_n$ be positive integers. Then $[a_0,a_1,\dots,a_n]$ can be found as follows: write the product $a_0a_1\cdots a_n$; in all possible ways, delete $k$ adjacent pairs of factors, where $k$ ranges from $0$ to $\lfloor(n+1)/2\rfloor$; add the resulting products. (By convention, if $n$ is even, then the term obtained by deleting everything has the value $1$.)*

**Example**

$$\begin{aligned}
[3,1,4,1,6] &= 3\cdot1\cdot4\cdot1\cdot6+\cancel{3}\cdot\cancel{1}\cdot4\cdot1\cdot6+3\cdot\cancel{1}\cdot\cancel{4}\cdot1\cdot6+3\cdot1\cdot\cancel{4}\cdot\cancel{1}\cdot6+\\
&\quad 3\cdot1\cdot4\cdot\cancel{1}\cdot\cancel{6}+\cancel{3}\cdot\cancel{1}\cdot\cancel{4}\cdot\cancel{1}\cdot6+\cancel{3}\cdot\cancel{1}\cdot4\cdot\cancel{1}\cdot\cancel{6}+3\cdot\cancel{1}\cdot\cancel{4}\cdot\cancel{1}\cdot\cancel{6}\\
&= 72+24+18+18+12+6+4+3\\
&= 157.
\end{aligned}$$

**Proof** Induction on $n$. When $n=1$, there is no way to delete any terms, and we just have the single term $a_0$, as required.

Suppose that the formula holds for $[a_0,\dots,a_m]$ with $m<n$, and consider $[a_0,\dots,a_n]$. We take all the terms in Euler's expression, and divide them into two types:

- Those for which $a_0$ is deleted. The only way this can happen is that $a_1$ is also deleted, and we delete all consecutive pairs of $a_2,\dots,a_n$. By the induction hypothesis, the sum of all these terms is $[a_2,\dots,a_n]$.

- Those for which $a_0$ is not deleted. Then every term has a factor $a_0$, and what remains is $a_1\cdots a_n$ with any number of consecutive pairs deleted; so the sum of all these terms is $a_0[a_1,\dots,a_n]$.

Putting the two pieces together and using the definition of $[a_0,\dots,a_n]$ gives the result. □

The result has a nice corollary:

**Corollary 3.5** *(a)* $[a_0, a_1, \dots, a_n] = [a_n, \dots, a_1, a_0]$.

*(b)* $[a_0, \dots, a_n] = [a_0, \dots, a_{n-1}]a_n + [a_0, \dots, a_{n-2}]$.

**Proof** (a) holds because in Euler's formula the reversed sequence obviously gives exactly the same result.

Then (b) is straightforward:

$$\begin{aligned} [a_0, \dots, a_n] &= [a_n, \dots, a_0] \\ &= a_n[a_{n-1}, \dots, a_0] + [a_{n-2}, \dots, a_0] \\ &= [a_0, \dots, a_{n-1}]a_n + [a_0, \dots, a_{n-2}]. \end{aligned}$$

□

This means that we can calculate by expanding "from the front" as well as "from the back":

$$\begin{aligned} [3] &= 3 \\ [3,1] &= [3]1 + 1 = 4 \\ [3,1,4] &= [3,1]4 + [3] = 19 \\ [3,1,4,1] &= [3,1,4]1 + [3,1] = 23 \\ [3,1,4,1,6] &= [3,1,4,1]6 + [3,1,4] = 157. \end{aligned}$$

## 3.3 The convergents of a finite continued fraction

Let $a_0, a_1, \dots, a_n$ be positive integers. We define the *convergents* of the continued fraction $[a_0; a_1, \dots, a_n]$ to be the numbers $c_k = [a_0; a_1, \dots, a_k]$ for $k = 0, 1, \dots, n$.

It doesn't make much sense yet to call them "convergents", but we will see when we turn to infinite continued fractions for irrational numbers that they do indeed converge!

**Proposition 3.6** *Given positive integers $a_0, \dots, a_n$, define*

$$p_k = [a_0, \dots, a_k], \qquad q_k = [a_1, \dots, a_k]$$

*for $k = 1, \dots, n$, with $p_0 = a_0$ and $q_0 = 1$; and let $c_k = [a_0; a_1, \dots, a_k]$ for $k \geq 0$. Then*

*(a)* $c_k = p_k/q_k$ *for* $k = 0, \dots, n$;

*(b)* $\gcd(p_k, q_k) = 1$*;*

*(c) for* $k > 1$*, we have*

$$\begin{aligned} p_k &= a_k p_{k-1} + p_{k-2} \\ q_k &= a_k q_{k-1} + q_{k-2} \end{aligned}$$

**Proof** Parts (a) and (b) are immediate from Proposition 3.3 applied to $a_0, \ldots, a_k$. Part (c) is just the second part of Corollary 3.5. □

**Example** What are the convergents to $[1; 1, \ldots, 1]$ (with an arbitrary number of ones)?

We have $p_0 = [1] = 1$, $p_1 = [1, 1] = 2$, and $q_0 = 1$, $q_1 = [1] = 1$; and

$$\begin{aligned} p_k &= p_{k-1} + p_{k-2} \\ q_k &= q_{k-1} + q_{k-2} \end{aligned}$$

for $k \geq 2$. These are the recurrence relations for the famous *Fibonacci numbers* $1, 2, 3, 5, 8, 13, 21, 34, \ldots$. Note that the $q$ sequence is just the $p$ sequence with 1 added at the front and all the other terms shifted along one place. (In fact, there are different conventions about the numbering of the Fibonacci numbers; some people say that the $k$th Fibonacci number is $p_k$, while others say that it is $q_k$. In any case, we see that $[1; 1, 1, \ldots 1]$ (with $n+1$ ones) is equal to $p_n/q_n$.) The convergents are

$$\frac{1}{1}, \frac{2}{1}, \frac{3}{2}, \frac{5}{3}, \frac{8}{5}, \frac{13}{8}, \ldots$$

You are encouraged to work out the first few of these fractions with a calculator. What pattern do you see? (The next theorem should confirm your guess.)

**Theorem 3.7** *With the above notation,*

*(a)* $p_k q_{k-1} - q_k p_{k-1} = (-1)^{k-1}$ *for* $k \geq 1$.

*(b)* $c_k - c_{k-1} = (-1)^{k-1}/q_{k-1}q_k$ *for* $k \geq 1$.

**Proof** (a) Induction on $k$. We have $p_0 = a_0$, $p_1 = a_0 a_1 + 1$, $q_0 = 1$, $q_1 = a_1$, and so $p_1 q_0 - q_1 p_0 = 1 = (-1)^0$, so the induction starts. If we assume that $p_{k-1}q_{k-2} - q_{k-1}p_{k-2} = (-1)^{k-2}$, then we have $p_k = a_k p_{k-1} + p_{k-2}$ and $q_k = a_k q_{k-1} = q_{k-2}$; so

$$\begin{aligned} p_k q_{k-1} - q_k p_{k-1} &= (a_k p_{k-1} + p_{k-2}) q_{k-1} - (a_k q_{k-1} + q_{k-2}) p_{k-1} \\ &= p_{k-2} q_{k-1} - q_{k-2} p_{k-1} \\ &= -(-1)^{k-2} \\ &= (-1)^{k-1}. \end{aligned}$$

(b) Divide both sides of (a) by $q_{k-1}q_k$ and use the fact that $c_k = p_k/q_k$ and $c_{k-1} = p_{k-1}/q_{k-1}$. □

**Corollary 3.8** *The convergents satisfy*

$$c_0 < c_2 < c_4 < \cdots < c_5 < c_3 < c_1.$$

In other words, the even-numbered convergents increase and the odd-numbered convergents decrease.

**Proof** Theorem 3.7 shows that the odd-numbered convergents are greater than the preceding even-numbered convergents. Also, the differences between consecutive convergents decrease; so, if $k$ is even, then $c_{k+1} - c_k < c_{k-1} - c_k$, so that $c_{k+1} < c_{k-1}$, with a similar argument if $k$ is odd. □

This is exactly the behaviour that you should have observed for the ratios of consecutive Fibonacci numbers.

## 3.4 A party trick

The continued fraction expansion of a rational number is the basis of a party trick (probably only for nerds and geeks) suggested to me by one of my colleagues.

Ask someone to think of two positive integers $r$ and $s$, to divide $r$ by $s$ using their calculator, and to tell you the result. You will find the numbers $r$ and $s$.

How? You simply calculate the continued fraction for the number $q = r/s$, and use this to express it as a fraction:

$$q = [a_0; a_1, \ldots, a_n] = \frac{[a_0, a_1, \ldots, a_n]}{[a_1, \ldots, a_n]}.$$

You tell them the numerator and denominator of this fraction.

There are a couple of traps, one theoretical, one practical. First of all, as we have seen, the greatest common divisor of the numerator and denominator of a continued fraction is 1. So if the original numbers $r$ and $s$ have greatest common divisor $d > 1$, then you will find $r/d$ and $s/d$ instead of $r$ and $s$. There is nothing that can be done about this; you have to bluff your way out of it as well as you can.

The practical problem is caused by rounding errors. Maybe, at some stage of the algorithm, the calculator will give you a number like 5.99999862, and you have to guess that this should really be 6, and terminate the algorithm at that point. There is no hard-and-fast rule for this.

**Example** Suppose the chosen numbers are 225 and 157, so that $q = 225/157 = 1.433121019$. Now we calculate as follows:

$$\begin{aligned} a_0 &= \lfloor q \rfloor = 1, & q_1 &= 1/(q-1) = 2.308823529 \\ a_1 &= \lfloor q_1 \rfloor = 2, & q_2 &= 1/(q_1-2) = 3.238095238 \\ a_2 &= \lfloor q_2 \rfloor = 3, & q_3 &= 1/(q_2-3) = 4.2 \\ a_3 &= \lfloor q_3 \rfloor = 4, & q_4 &= 1/(q_3-4) = 5 \end{aligned}$$

So $q = [1;2,3,4,5] = \frac{225}{157}$.

Of course $r = 450$ and $s = 314$ would have given the same result!

The chance of the gcd problem arising can be estimated rather precisely, by a surprising theorem which is not part of this course.

**Theorem 3.9** *Given a large positive integer $n$, let $p_n$ be the probability that two randomly chosen positive integers at most $n$ are coprime. Then* $\lim_{n\to\infty} p_n = \dfrac{6}{\pi^2}$.

For example, of the 1000000 pairs of positive integers not exceeding 1000, there are 608383 coprime pairs.

## Exercises

**3.1** Express $245/43$ as a continued fraction.

**3.2** (a) Let $\alpha = [a_0; a_1, a_2, \ldots, a_n]$, where $a_0, \ldots, a_n$ are positive integers.

(i) Show that $\alpha = a_0 + 1/[a_1; a_2, \ldots, a_n]$ if $n > 0$.

(ii) Show that $a_0 \le \alpha \le a_0 + 1$.

(b) Now let $\beta = [b_0; b_1, b_2, \ldots, b_m]$, where $b_0, \ldots, b_m$ are positive integers. Suppose that

$$a_i = b_i \text{ for } i = 0, \ldots, k-1 \text{ and } a_k < b_k.$$

Prove that

- if $k$ is even, then $\alpha \le \beta$;
- if $k$ is odd, then $\beta \le \alpha$.

[**Hint:** Induction on $k$.]

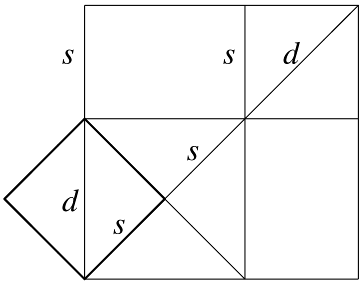
# Chapter 4

# Infinite continued fractions

Infinite continued fractions are not really continued fractions at all, but are limits of finite continued fractions. We show in this section that every real irrational number has an expression as an infinite continued fraction, and show that these provide "good" rational approximations to irrational numbers. For example, the famous approximations $22/7$ and $355/113$ to $\pi$ arise in this way.

## 4.1 An example

The Pythagoreans knew that the ratio of the diagonal to the side of a square is irrational. According to the historian of mathematics David Fowler, they may have reasoned something like this.



Let $s$ and $d$ be the side and diagonal lengths of a square. Rotate the square through 45 degrees. Prolong the diagonal by $s$ and draw a new square on this side, with side and diagonal lengths $S$ and $D$. We see from the figure that $S = s+d$, and $D = 2s$+d; so

$$\frac{S+D}{S} = \frac{3s+2d}{s+d} = 2+\frac{s}{s+d}.$$

Let $u=(s+d)/s$. Since any two squares are similar, we also have $u=(S+D)/S$, and so

$$u=2+\frac{1}{u}.$$

Substituting this expression for $u$ into the right-hand side of the expression repeatedly, we see that

$$\begin{aligned} u &= 2+\cfrac{1}{2+\cfrac{1}{u}} \\ &= 2+\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{u}}} \\ &= \cdots \end{aligned}$$

Now the Pythagoreans knew the essence of Euclid's algorithm, which as we have seen can be used to find the finite continued fraction of a rational number. Put another way, if $u$ were a rational number, then the procedure for finding the continued fraction for $u$ terminates, and the continued fraction is finite. The above argument shows that, if this algorithm is applied to our number $u$, it never terminates; the algorithm "spins its wheels" and the next number at each stage is always $u$.

You might guess that this means that $u$ can be expressed as an "infinite continued fraction"

$$u=2+\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{2+\cdots}}}.$$

In this chapter we will give a precise meaning to the notion of an infinite continued fraction, and verify that this is always the case. Morever, every real irrationanl number has a unique expression as an infinite continued fraction. We will also see that the finite continued fractions obtained by stopping after a finite number of steps give the "best possible" rational approximations to the number in question.

## 4.2 The definition

Our definition of an infinite continued fraction depends on the following theorem.

**Theorem 4.1** *Let $a_0,a_1,a_2$ be a sequence of positive integers, and define $c_n=[a_0;a_1,a_2,\dots,a_n]$ for $n\geq 0$. Then the sequence $c_0,c_1,c_2$ of rational numbers converges to a limit.*

**Remark** This explains why we called the numbers $c_0, c_1, \ldots$ "convergents".

**Proof** Since $c_0, c_1, \ldots, c_n$ are the convergents to the finite continued fraction $[a_0; a_1, \ldots, a_n]$, all the results of Chapter 2 apply here.

We have $c_n = p_n/q_n$, where $p_n = [a_0, a_1, \ldots, a_n]$ and $q_n = [a_1, \ldots, a_n]$. Now

$$c_0 < c_2 < c_4 < \cdots < c_5 < c_3 < c_1$$

and

$$c_k - c_{k-1} = \frac{(-1)^{k-1}}{q_{k-1}q_k}$$

for $k \geq 1$.

The even terms $c_0, c_2, c_4, \ldots$ form an increasing sequence which is bounded above (by $c_1$) and so tends to a limit $y$. Similarly, the odd terms $c_1, c_3, \ldots$ tend to a limit $z$, with $y \leq z$.

Now the recurrence relation $q_k = a_k q_{k-1} + q_{k-2}$ shows that the numbers $q_k$ increase strictly with $k$, so $c_k - c_{k-1} \to 0$ as $k \to \infty$. Hence $y = z$, and the whole sequence converges. □

We define the limit of the sequence of convergents to be the *value* of the infinite continued fraction $[a_0; a_1, a_2, \ldots]$.

For example, if $x_n = [2; 2, 2, \ldots, 2]$ (with $n+1$ 2s), then we have

$$x_n = 2 + \frac{1}{x_{n-1}}.$$

If $\lim_{n\to\infty} x_n = u$, then clearly

$$u = 2 + \frac{1}{u},$$

so $u^2 - 2u - 1 = 0$, or $u = 1 \pm \sqrt{2}$. But $u$ is obviously positive; so we have $u = 1 + \sqrt{2}$. This is exactly what we would expect for $u = (s+d)/s$, where $s$ are the side and diagonal of a square!

Now we show that any real number has a continued fraction expansion:

**Theorem 4.2** *For every irrational real number $y$ greater than* 1*, there is a sequence of positive integers $a_0, a_1, \ldots$ for which the limit of the sequence of convergents of $[a_0; a_1, \ldots]$ is $y$.*

**Proof** We take $a_0 = \lfloor y \rfloor$, so that $0 < y - a_0 < 1$. Then we put $y_1 = 1/(y - a_0)$, so that $y_1$ is an irrational nummber greater than 1, and continue the process:

$$a_i = \lfloor y_i \rfloor, \qquad y_{i+1} = \frac{1}{y_i - a_i}.$$

Then $a_0, a_1, a_2, \ldots$ are positive integers and $y_0 = y, y_1, y_2$ are irrational numbers greater than 1, so the process continues infinitely and produces an infinite continued fraction $[a_0; a_1, a_2]$. We have to show that the value of this continued fraction is $y$.

Let $c_0, c_1, \ldots$ be the convergents of $[a_0; a_1, a_2, \ldots]$. Then $c_n = p_n/q_n$, where $p_n = [a_0, a_1, \ldots, a_n]$ and $q_n = [a_1, \ldots, a_n]$.

Also, we have

$$y = a_0 + \frac{1}{y_1} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{y_2}} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{y_3}}} = \cdots.$$

In other words,

$$y = [a_0; a_1, a_2, \ldots, a_n, y_{n+1}] \text{ for all } n \geq 0.$$

So by Proposition 2.6 (and see the remark at the start of Section 2.2, this result is still valid even though the numbers $y$ and $y_{n+1}$ are not positive integers!), we have

$$y = \frac{y_{n+1}p_n + p_{n-1}}{y_{n+1}q_n + q_{n-1}} \text{ for all } n \geq 0.$$

So

$$\begin{aligned} |y - c_n| &= \left| \frac{y_{n+1}p_n + p_{n-1}}{y_{n+1}q_n + q_{n-1}} - \frac{p_n}{q_n} \right| \\ &= \frac{|p_{n-1}q_n - p_n q_{n-1}|}{(y_{n+1}q_n + q_{n-1})q_n} \\ &= \frac{1}{(y_{n+1}q_n + q_{n-1})q_n} \end{aligned}$$

Now $(y_{n+1}q_n + q_{n-1}) > q_n$, since $y_{n+1} > 1$ and $q_{n-1} > 0$; so $1/(y_{n+1}q_n + q_{n-1})q_n < 1/q_n^2$. Also the numbers $q_n$ are increasing positive integers, so $1/q_n^2 \to 0$ as $n \to \infty$. So finally $c_n \to y$ as $n \to \infty$, that is,

$$y = [a_0; a_1, a_2, \ldots],$$

as claimed. □

**Example** What is the continued fraction expansion of $\pi$?

A few minutes' work with a calculator shows that, to a few places of decimals,

$$\pi = 3.141592653589793,$$

$$\begin{aligned}
1/(0.141592653589793) &= 7.062513305931046,\\
1/(0.06251330593104577) &= 15.99659440668572,\\
1/(0.9965944066857205) &= 1.003417231013372,\\
1/(0.003417231013371963) &= 292.6345910144503,\\
1/(0.6345910144503185) &= 1.575818089492172
\end{aligned}$$

so the continued fraction begins $[3;7,15,1,292,1,\ldots]$.

The convergents are

$$3, \frac{22}{7}, \frac{333}{106}, \frac{355}{113}, \frac{103993}{33102}, \ldots$$

We recognise two famous approximations to $\pi$, namely $22/7$ and $355/113$. (The famous approximation $22/7$ was found by Archimedes, and the more accurate $355/113$ was given by Zu Chongzhi in the 5th century; in China it is called "Zu's ratio".) The best approximations occur when we stop just before a large number in the continued fraction. (Can you see why?) For example, the difference between $\pi$ and $355/113$ is only $-0.000000266764189\ldots$.

Finally we show that the continued fraction expression is unique:

**Theorem 4.3** *Every irrational number greater than* 1 *is the limit of a unique infinite continued fraction.*

**Proof** Suppose that $y$ is irrational and $y = [b_0; b_1, b_2, \ldots]$. Then $y = b_0 + 1/y_1$, where $y_1 > 1$; so $b_0 = \lfloor y \rfloor$ and $y_1 = 1/(y - b_0)$ are determined by $y$. Similarly $b_1 = \lfloor y_1 \rfloor$ and $y_2 = 1/(y_1 - b_1)$ are determined, and so on. □

**Remark** We have considered numbers greater than 1 so far. The results can be easily extended to arbitrary irrational numbers. We simply relax the condition that $a_0$ is a positive integer. For example,

$$\sqrt{2} - 1 = [0; 2, 2, 2, \ldots].$$

Here is a comparison between the representation of real numbers by infinite decimals and continued fractions. A sequence of positive integers may be finite, or recurring (that is, periodic after some point), or neither. Hidden in this table is an important theorem which we will prove later. A *quadratic irrational* is an irrational number which is a root of a quadratic equation with rational coefficients, in other words, a number of the form $a + b\sqrt{d}$ where $a, b \in \mathbb{Q}$ and $d$ is a squarefree integer greater than 1.

| | Decimal | Continued fraction |
|---|---|---|
| Finite | Rational with denominator $2^a 5^b$ | Rational |
| Recurring | Other rational | Quadratic irrational |
| Non-recurring | Irrational | Other irrational |

**Examples** To calculate the continued fraction for a real number $y$, set $y_0 = y$ and then $a_i = \lfloor y_i \rfloor$, $y_{i+1} = 1/(y_i - a_i)$.

(a) $u = 1 + \frac{1}{2}\sqrt{2}$. The

$$\begin{aligned} a_0 = \lfloor u \rfloor = 1, &\qquad y_1 = \frac{1}{\sqrt{2}/2} = \frac{2}{\sqrt{2}} = \sqrt{2} \\ a_1 = \lfloor \sqrt{2} \rfloor = 1, &\qquad y_2 = \frac{1}{\sqrt{2}-1} = \sqrt{2}+1, \\ a_2 = \lfloor \sqrt{2}+1 \rfloor = 2, &\qquad y_3 = \frac{1}{\sqrt{2}-1} = \sqrt{2}+1, \end{aligned}$$

and then the process simply repeats. So $1 + \frac{1}{2}\sqrt{2} = [1; 1, 2, 2, 2, \ldots]$.

(b) $u = \sqrt{15} - 3$. We have

$$\begin{aligned} a_0 = \lfloor \sqrt{15} - 3 \rfloor = 0, &\qquad y_1 = \frac{1}{\sqrt{15}-3} = \frac{\sqrt{15}+3}{6}, \\ a_1 = \left\lfloor \frac{\sqrt{15}+3}{6} \right\rfloor = 1, &\qquad y_2 = \frac{6}{\sqrt{15}-3} = \sqrt{15}+3, \\ a_2 = \lfloor \sqrt{15}+3 \rfloor = 6, &\qquad y_2 = \frac{1}{\sqrt{15}-3} = y_0, \end{aligned}$$

and the process repeats. So $u = [0; 1, 6, 1, 6, \ldots]$.

(c) Finally, let $\phi = (1+\sqrt{5})/2$ be the golden ratio. Then, as we noted before, $\phi = 1 + 1/\phi$, so the continued fraction is $\phi = [1; 1, 1, \ldots]$. We noted before that the convergents $[1; 1, 1, \ldots, 1]$ are ratios of consecutive Fibonacci numbers. So, if $F_n$ is the $n$th Fibonacci number, we have

$$\lim_{n\to\infty} \frac{F_{n+1}}{F_n} = \frac{1+\sqrt{5}}{2}.$$

## 4.3 Approximation by convergents

We have seen that, if $y = [a_0; a_1, a_2, \ldots]$, and $c_n = [a_0; a_1, \ldots, a_n]$ is the $n$th convergent to $y$, then the numbers $c_n$ are rational numbers which tend to the limit $y$. In this section, we will see that they give the best possible approximations to $y$.

What should a good rational approximation $p/q$ to $y$ be? First, of course, it should be close to $y$. Next, we want the denominator $q$ to be relatively small. In particular, there should be no rational number with smaller denominator which is closer to $y$. Finally, we should have a good estimate for $|y - p/q|$.

We will see that the convergents to the continued fraction for $y$ satisfy all these properties.

Let $c_n = p_n/q_n$, where $p_n = [a_0, \ldots, a_n]$ and $q_n = [a_1, \ldots, a_n]$ (so that $\gcd(p_n, q_n) = 1$). We know that $y$ lies between $c_n$ and $c_{n+1}$ for all $n$ – remember that we have

$$c_0 < c_2 < c_4 < \cdots < y < \cdots < c_5 < c_3 < c_1.$$

Thus we see that $|y - c_n| < |c_{n+1} - c_n|$. In x.x, we showed that $|c_{n+1} - c_n| < 1/q_n q_{n+1}$. In particular, $|y - c_n| < 1/q_n q_{n+1}$.

**Example** We showed that $\sqrt{15} - 3 = [0; 1, 6, 1, 6, \ldots]$. Let us compute the convergents, using the recurrence

$$p_{n+1} = a_n p_n + p_{n-1}, \qquad q_{n+1} = a_n q_n + q_{n-1}.$$

We have

$$\begin{aligned}
c_0 = \frac{0}{1}, &\qquad c_1 = \frac{1}{1}, \\
c_2 &= \frac{6 \cdot 1 + 0}{6 \cdot 1 + 1} = \frac{6}{7}, \\
c_3 &= \frac{1 \cdot 6 + 1}{1 \cdot 7 + 1} = \frac{7}{8}, \\
c_4 &= \frac{6 \cot 7 + 6}{6 \cdot 8 + 7} = \frac{48}{55}, \\
c_5 &= \frac{1 \cdot 48 + 7}{1 \cdot 55 + 8} = \frac{55}{63}, \\
c_6 &= \frac{6 \cdot 55 + 48}{6 \cdot 63 + 55} = \frac{378}{433},
\end{aligned}$$

and so on. How good an approximation is $c_6$? We know that

$$|y - c_6| \le \frac{1}{q_6 q_7} = \frac{1}{433(1 \cdot 433 + 63)} = \frac{1}{214768},$$

so the value is accurate to four places of decimals.

We know that $c_n$ is always a better approximation to $y$ than $c_{n-2}$ is. (If $n$ is even, we have $c_{n-2} < c_n < y$, and if $n$ is odd we have $y < c_n < c_{n-2}$.) What about $c_n$ and $c_{n-1}$? We show that, after the first step, the approximation always gets better as $n$ increases.

Let us stop and recall some notation. We have $y = [a_0; a_1, a_2, \ldots]$. The $n$th convergent is $c_n = p_n/q_n = [a_0; a_1, \ldots, a_n]$. As we saw on p. 4, $y = [a_0; a_1, \ldots, a_{n-1}, y_n]$, where $y_n = [a_n; a_{n+1}, \ldots]$. Also,

$$\left| y - \frac{p_n}{q_n} \right| = \frac{1}{q_n(y_{n+1}q_n + q_{n-1})} \leq \frac{1}{q_n q_{n+1}}.$$

**Proposition 4.4** *For all $n \geq 2$, we have*

*(a)* $|q_n y - p_n| < |q_{n-1} y - p_{n-1}|$;

*(b)* $|y - c_n| < |y - c_{n-1}|$.

**Proof** First we show that (a) implies (b). We have

$$q_n |y - c_n| = |q_n y - p_n|.$$

Also, $q_n > q_{n-1}$. So, if we show that $|q_n y - p_n| < |q_{n-1} y - p_{n-1}|$, then we will be able to conclude that

$$|y - c_n| = \frac{|q_n y - p_n|}{q_n} < \frac{|q_{n-1} y - p_{n-1}|}{q_{n-1}} = |y - c_{n-1}|.$$

So we only have to prove (a). For this, note first that

$$|q_n y - p_n| = \frac{1}{y_{n+1} q_n + q_{n-1}}, \qquad |q_{n-1} y - p_{n-1}| = \frac{1}{y_n q_{n-1} + q_{n-2}},$$

so it is enough to show that $y_{n+1} q_n + q_{n-1} > y_n q_{n-1} + q_{n-2}$.

Now $a_n = \lfloor y_n \rfloor$, and so $y_n < a_n + 1$, and

$$\begin{array}{rcl} y_n q_{n-1} + q_{n-2} & < & q_{n-1} + a_n q_{n-1} + q_{n-2} \\ & = & q_{n-1} + q_n \\ & < & y_{n+1} q_n + q_{n-1}, \end{array}$$

and we are done. □

Now we come to the main result.

**Theorem 4.5** *Let $[a_0;a_1,a_2,\ldots]$ be the continued fraction for the irrational number y, and let $[a_0;a_1,\ldots,a_n] = c_n = p_n/q_n$ be the nth convergent. Let $c = p/q$ be any rational number in its lowest terms. If $q < q_n$ with $n > 1$, then $|y - p/q| > |y - p_n/q_n|$.*

We say that a rational number $p/q$ is a *best approximation* to $y$ if $|y - p/q| < |y - a/b|$ for any rational number $a/b$ with $b < q$. We see that the convergents from $c_2$ on are best approximations to an irrational number.

The proof involves quite a bit of work, which we isolate in a preliminary lemma.

**Lemma 4.6** *Let $[a_0;a_1,a_2,\ldots]$ be the continued fraction for the irrational number y, and let $[a_0;a_1,\ldots,a_n] = c_n = p_n/q_n$ be the nth convergent. If $\gcd(p,q) = 1$ and $q \le q_n$, then*

$$|qy - p| \ge |q_{n-1}y - p_{n-1}|,$$

*with equality if and only if $p/q = p_{n-1}/q_{n-1}$.*

We will prove this in the appendix to this chapter.

**Proof of the Theorem** Suppose that $q < q_n$. By induction, if $q < q_{n-1}$, then $|y - p/q| > |y - p_{n-1}/q_{n-1}| > |y - p_n/q_n|$ (the last inequality by Proposition 4.4), so we can suppose that $q \ge q_{n-1}$. Then

$$\begin{aligned}
|y - p/q| &= \frac{1}{q}|qy - p| \\
&> \frac{1}{q}|q_{n-1}y - p_{n-1}| \qquad \text{by Lemma 4.6} \\
&> \frac{1}{q}|q_n y - p_n| \qquad \text{by Proposition 4.4(a)} \\
&> \frac{1}{q_n}|q_n y - p_n| \qquad \text{since } q < q_n \\
&= |y - p_n/q_n|.
\end{aligned}$$

□

## 4.4 Order of approximation

We say that an positive irrational number $y$ is *approximable by rationals to order n* if there exist a positive constant $c$ and infinitely many rationals $p/q$ with $q > 0$ such that

$$\left| y - \frac{p}{q} \right| \le \frac{c}{q^n}.$$

Note that is $y$ is approximable to order $n$, then it is approximable to any smaller order, since if $m < n$ then $c/q^n \leq c/q^m$ for positive integer $q$.

We will see that algebraic numbers (roots of polynomials over the integers) are not approximable to arbitrarily high order. Then, by writing down a number which is approximable to order $n$ for every $n$, we will have exhibited a transcendental number (one which is not a root of a polynomial over $\mathbb{Z}$).

**Theorem 4.7** *(a) Positive rational numbers are approximable to order* 1 *and no higher.*

*(b) Every positive irrational number is approximable to order* 2.

**Proof** (a) Let $y = a/b$ be a rational number, with $\gcd(a,b) = 1$. By Euclid's algorithm we can find $p_0, q_0$ such that $p_0 b - q_0 a = 1$. Then

$$\left|\frac{a}{b} - \frac{p_0}{q_0}\right| = \frac{1}{q_0 b} \leq \frac{1}{q_0}.$$

Similarly, if $p_m = p_0 + ma$ and $q_m = q_0 + mb$, then $p_m b - q_m a = 1$ (so $\gcd(p_m, q_m) = 1$), and exactly as before,

$$\left|\frac{a}{b} - \frac{p_m}{q_m}\right| \leq \frac{1}{q_m}.$$

So $a/b$ is approximable to order 1.

Now suppose that $a/b$ were approximable to order 2, so that there are infinitely many rationals $p/q$ such that $|a/b - p/q| \leq c/q^2$. We have

$$\left|\frac{a}{b} - \frac{p}{q}\right| = \frac{pb - qa}{qb} \geq \frac{1}{qb},$$

so that we have $1/(qb) < c/q^2$ or $q < cb$ for infinitely many $q$, which is clearly impossible.

(b) Let $y$ be irrational, and $c_n = p_n/q_n$ be any convergent. As we have seen,

$$\left|y - \frac{p_n}{q_n}\right| \leq \frac{1}{q_n q_{n+1}} \leq \frac{1}{q_n^2},$$

so $y$ is approximable to order 2. □

On the other hand, algebraic numbers are not approximable to arbitrary orders:

**Theorem 4.8** *Let the positive irrational number $y$ be the root of a polynomial of degree $n$ with integer coefficients. Then $y$ is not approximable to any order greater than $n$.*

**Proof** Suppose that $y$ is a root of the polynomial

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

with integer coefficients $a_0, \ldots, a_n$, not all zero. We can assume that $f(x)$ has no rational roots. (By the Remainder Theorem, if $f(p/q) = 0$, then $qx - p$ would be a factor of $f(x)$, and we could find a polynomial of smaller degree satisfied by $y$.) Hence

$$0 \neq f(p/q) = \frac{a_n p^n + a_{n-1} p^{n-1} q + \cdots + a_1 p q^{n-1} + a_0 q^n}{q^n}.$$

The numerator of this fraction is thus non-zero. But clearly it is an integer, so it is at least 1, and we have

$$|f(p/q)| \geq 1/q^n.$$

Now by the Mean Value Theorem, we have

$$-f(p/q) = f(y) - f(p/q) = (y - p/q) f'(z),$$

where $z$ is a number between $y$ and $p/q$. Now $f(p/q) \neq 0$, so $f'(z) \neq 0$. Choose a fixed interval containing $y$ and all the rationals that approximate it. The function $f'(x)$ is continuous on this interval, and so is bounded there, say by $M$; then $|f(p/q)| \leq M|y - p/q|$, and so

$$|y - p/q| \geq \frac{f(p/q)}{M} \geq \frac{1}{Mq^n}.$$

If $y$ were approximable to order $n+1$, then $|y - p/q| \leq c/q^{n+1}$ for infinitely many $q$. Combining these two bounds, we would have $q \leq Mc$ for infinitely many $q$, which is impossible. □

**Example** Let $y$ be *Liouville's number*

$$y = 0.110001000000000000000001 0\ldots$$

where the 1s appear in positions $n!$ for $n = 1, 2, 3, \ldots$. In other words,

$$y = \sum_{k=1}^{\infty} 10^{-k!}.$$

Let $c_n = p_n/q_n$ be the rational obtained by truncating the decimal after the $n$th occurrence of 1, so that

$$c_n = \sum_{k=1}^{n} 10{-}k!.$$

We have $q_n = 10^{n!}$, and

$$|y - p_n/q_n| \leq 2 \cdot 10^{-(n+1)!} = \frac{2}{q^{n+1}}.$$

(Since $(n+1)! = (n+1) \cdot n!$, we have $10^{(n+1)!} = (10^{n!})^{n+1}$.) So $y$ is approximable by rationals to every order, and is not algebraic; that is, *$y$ is transcendental.*

This was the first explicit example known of a transcendental number. Later Hermite and Lindemann showed that e and $\pi$ are transcendental. Later still, Cantor showed that *almost all real numbers are transcendental.* But there are still many mysteries: we don't know, for example, whether $\pi^{\mathrm{e}}$ is algebraic or transcendental!

## 4.5 Proof of Lemma 4.6

This ingenious argument is due to Lagrange.

Recall that $c_n = p_n/q_n$ is the $n$th convergent to the irrational number $y = [a_0; a_1, a_2, \ldots]$, and $c = p/q$ is any rational number in its lowest terms, with $q < q_n$ for some $n > 1$. We have to show that

$$|qy - p| \geq |q_{n-1}y - p_{n-1}|,$$

and that equality holds only if $c = c_{n-1}$.

We start by considering the equations

$$\begin{aligned} p_{n-1}u + p_n v &= p, \\ q_{n-1}u + q_n v &= q. \end{aligned}$$

Recalling that $p_{n-1}q_n - q_{n-1}p_n = (-1)^n \neq 0$, we see that these equations have a unique solution $(u, v)$, given by

$$\begin{aligned} u &= (-1)^n(pq_n - qp_n), \\ v &= (-1)^n(qp_{n-1} - pq_{n-1}). \end{aligned}$$

Note that $u$ and $v$ are integers.

If $u = 0$, then $p_n v = p$ and $q_n v = q$, so that $p/q = p_n/q_n$, contradicting our assumption that $q < q_n$. (Remember that all these fractions are in their lowest terms.) So $u \neq 0$.

Similarly, if $v = 0$, then $p_{n-1}u = p$, $q_{n-1}u = q$, so $p/q = p_{n-1}/q_{n-1}$, which is the extremal case. So we can also assume that $v \neq 0$, and we have to prove that strict inequality holds in the conclusion.

Now if $v < 0$, then $q_{n-1}u = q - q_n v$, so $u > 0$; and if $v > 0$, then $q_{n-1}u = q - q_n v < 0$ since $q < q_n$ and $v \geq 1$, so $u < 0$. Thus $u$ and $v$ have opposite signs.

Now $y$ lies between the two consecutive convergents $c_{n-1} = p_{n-1}/q_{n-1}$ and $c_n = p_n/q_n$. So $q_{n-1}y - p_{n-1}$ and $q_n y - p_n$ have opposite signs. So $u(q_{n-1}y - p_{n-1})$ and $v(q_n y - p_n)$ have the same sign. It follows that the absolute value of their sum is the sum of their absolute values:

$$\begin{aligned} |qy - p| &= |(q_{n-1}u + q_n v)y - (p_{n-1}u + p_n v)| \\ &= |u(q_{n-1}y - p_{n-1}) + v(q_n y - p_n)| \\ &= |u(q_{n-1}y - p_{n-1})| + |v(q_n y - p_n)| \\ &= |u| \cdot |q_{n-1}y - p_{n-1}| + |v| \cdot |q_n y - p_n| \\ &> |q_{n-1}y - p_{n-1}|, \end{aligned}$$

where in the last step we use the fact that $|u| \geq 1$ and $|v| \cdot |q_n y - p_n| > 0$. This completes the proof.

## Exercises

**4.1** Find the first six terms in the continued fraction for the number e, the root of natural logarithms. (You may use a calculator for this question.)

**Note:** You might spot a pattern here. The pattern really does continue!

**4.2** Find the continued fractions for the following numbers:

(a) $3 + 2\sqrt{2}$;

(b) $\sqrt{11} - 10$;

(c) $(1 + \sqrt{5})/4$.

**4.3** Why does the number $(-1 + \sqrt{-3})/2$ not have a continued fraction expansion?

**4.4** Let $F_n$ be the $n$th *Fibonacci number*, defined by the rules

$$F_1 = 1,\ F_2 = 2, \quad F_n = F_{n-1} + F_{n-2} \text{ for } n \geq 3.$$

(a) Show that $[1, 1, \ldots, 1] = F_n$ for $n \geq 1$ (where there are $n$ ones in the bracket).

(b) Find $\lim\limits_{n\to\infty} \dfrac{F_n}{F_{n-1}}$.

**Hint:** What is the number of which the ratios $F_n/F_{n-1}$ are the convergents?

# Chapter 5

# Periodic continued fractions

In this chapter, we see that the irrational numbers whose continued fraction expansion is periodic are precisely the (real) quadratic irrationals, and we determine which numbers have purely periodic expansion.

## 5.1 Periodic and purely periodic continued fractions

**Definition** The infinite continued fraction

$$[a_0; a_1, a_2, \ldots]$$

is *periodic* if there exist integers $k, l$ with $k > 0$ such that

$$a_{n+k} = a_n \text{ for all } n \geq l.$$

It is *purely periodic* if there exists $k > 0$ such that

$$a_{n+k} = a_n \text{ for all } n \geq 0.$$

If $a_{n+k} = a_n$ for all $n \geq l$, we write the continued fraction as

$$[a_0; a_1, a_2, \ldots, a_{l-1}, \overline{a_l, a_{l+1}, \ldots, a_{l+k-1}}].$$

This is a similar notation to the one used for periodic decimals. For example,

$$\begin{aligned} [2;1,2,1,2,1,2,1,\ldots] &= [\overline{2;1}] \\ [3;5,2,1,2,1,2,1,\ldots] &= [3;5,\overline{2,1}]. \end{aligned}$$

We now calculate these two continued fractions. Let $c = [\overline{2;1}]$. Then

$$\begin{aligned} c &= [2;1,c] \\ &= \frac{[2,1,c]}{[1,c]} \text{ (by Proposition 2.3(b))} \\ &= \frac{3c+2}{c+1}. \end{aligned}$$

So $c^2 + c = 3c + 2$, so that $c^2 - 2c - 2 = 0$, or $c = 1 \pm \sqrt{3}$. But $c > 2$, so we must take the plus sign; $c = 1 + \sqrt{3}$.

Now let $d = [3;5,\overline{2,1}]$. Then

$$\begin{aligned} d &= [3;5,c] \\ &= \frac{[3,5,c]}{[5,c]} \\ &= \frac{16c+3}{5c+1} \\ &= \frac{19+16\sqrt{3}}{6+5\sqrt{3}} \\ &= \frac{(19+16\sqrt{3})(6-5\sqrt{3})}{(6+5\sqrt{3})(6-5\sqrt{3})} \\ &= \frac{126-\sqrt{3}}{39}. \end{aligned}$$

Note that $d$, like $c$, is a “quadratic irrational”, an algebraic integer satisfying a quadratic equation. (We saw this already for $c$; and $d$ satisfies $(39x - 126)^2 = 3$.)

In this chapter we are going to show that the result suggested by these examples is true in general. A real number has a periodic continued fraction if and only if it is a quadratic irrational. We will also find which numbers have purely periodic continued fractions. We will apply these results to sums of squares and to a diophantine equation called *Pell’s equation* in the next chapter.

## 5.2 Quadratic irrationals

Recall that *quadratic irrational* is a real number of the form $a + b\sqrt{d}$, where $a$ and $b$ are rational numbers, $b \neq 0$, and $d$ is a squarefree positive integer.

These are precisely the roots of irreducible quadratics with rational coefficients. For the equation $Ax^2 + Bx + C = 0$ has the solutions

$$x = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}.$$

The roots are real if and only if $B^2-4AC>0$. Now $B^2-4AC$ is a rational number, and so can be written $u^2d/v^2$, where $u,v,d$ are integers with $d$ squarefree; then $x=a\pm b\sqrt{d}$, where $a=B/2A$ and $b=u/2Av$ are rational numbers. Conversely, the numbers $a\pm b\sqrt{d}$ have sum $2a$ and product $a^2-db^2$, so satisfy the equation $x^2-2ax+(a^2-db^2)=0$.

If $y=a+b\sqrt{d}$ is a quadratic irrational, we define its *algebraic conjugate* to be $y'=a-b\sqrt{d}$. Note that $y$ and $y'$ are the two roots of the same irreducible quadratic.

Now we define a *reduced quadratic irrational* to be a quadratic irreducible $y$ such that $y$ and its algebraic conjugate $y'$ satisfy

$$y>1 \quad \text{and} \quad -1<y'<0.$$

In our worked example, $c=1+\sqrt{3}>0$ and $-1<c'=1-\sqrt{3}<0$, so $c$ os a reduced quadratic irrational. On the other hand, $d=(126-\sqrt{3})/39>1$ but $d'=(126+\sqrt{3})/39$ is greater than $d$. So $d'$ is not reduced.

## 5.3 The main theorem

We are going to show that a number has a purely periodic continued fraction if and only if it is a reduced quadratic irrational. Here is the first part of the argument.

**Proposition 5.1** *Let $y$ be the value of a purely periodic continued fraction. Then $y$ is a reduced quadratic irrational.*

**Proof** Let $y=[\overline{a_0;a_1,\ldots,a_{k-1}}]$. We will suppose that $k\geq 3$; the argument for $k=1,2$ is easy to do directly, or we can simply pretend that the period is longer than it is (for example, $[\overline{2;1}]=[\overline{2;1,2,1}]$).

We know that $y$ is irrational, since the continued fraction for a rational number terminates. Also, just as we argued for the number $c=[\overline{2;1}]$, we have

$$\begin{aligned} y &= [a_0;a_1,\ldots,a_{k-1},y] \\ &= \frac{yp_{k-1}+p_{k-2}}{yq_{k-1}+q_{k-2}}, \end{aligned}$$

where $c_i=p_i/q_i$ is the $i$th convergent of $[a_0;a_1,\ldots,a_{k-1}]$.

Hence $y^2q_{k-1}+y(q_{k-2}-p_{k-1})-p_{k-2}=0$, so that $y$ is a quadratic irrational.

Also, $a_0=a_k\geq 1$ (remember that all terms except possibly the first in a continued fraction are positive), so $y>a_0\geq 1$. It remains to show that the algebraic conjugate $y'$ of $y$ satisfies $-1<y'<0$.

From the properties of quadratic equations, we have $yy' = -p_{k-2}/q_{k-1} < 0$. Also,

$$-y' = \frac{p_{k-2}}{yq_{k-1}} < \frac{p_{k-1}}{yq_{k-1}} = \frac{c_{k-1}}{y},$$

because $p_{k-2} < p_{k-1}$; and also

$$-y' = \frac{p_{k-2}}{yq_{k-1}} < \frac{p_{k-2}}{yq_{k-2}} = \frac{c_{k-2}}{y},$$

because $q_{k-2} < q_{k-1}$. One of $k-1$ and $k-2$, say $j$, is even; and we know from Corollary 2.8 that $c_j < y$. So $-y' < 1$, or $y' > -1$. Now we have verified all parts of the definition of a reduced quadratic irrational. □

**Proposition 5.2** *If $y$ is the value of a periodic continued fraction, then $y$ is a quadratic irrational.*

**Proof** Let $y = [a_0; a_1, \dots, a_m, \overline{a_{m+1}, \dots, a_{m+k}}]$. Let $z = [\overline{a_{m+1}; \dots, a_{m+k}}]$. By Proposition 5.1, $z$ is a (reduced) quadratic irrational, say $z = u + v\sqrt{d}$, where $u$ and $v$ are rational numbers and $d$ is a squarefree integer. We have

$$\begin{aligned} y &= [a_0; a_1, \dots, a_m, z] = \frac{[a_0, \dots, a_m, z]}{[a_1, \dots, a_m, z]} \\ &= \frac{[a_0, \dots, a_m]z + [a_0, \dots, a_{m-1}]}{[a_1, \dots, a_m]z + [a_1, \dots, a_{m-1}]}. \end{aligned}$$

Let $[a_0, \dots, a_m] = A$, $[a_0, \dots, a_{m-1}] = B$, $[a_1, \dots, a_m] = C$, $[a_1, \dots, a_{m-1}] = D$ (these are all positive integers). Then

$$\begin{aligned} y &= \frac{Az+B}{Cz+D} \\ &= \frac{Au+B+Av\sqrt{d}}{Cu+D+Cv\sqrt{d}} \\ &= \frac{(Au+B+Av\sqrt{d})(Cu+D-Cv\sqrt{d})}{(Cu+D)^2-(Cv)^2d}, \end{aligned}$$

which is a quadratic irrational since it has the form $x + y\sqrt{d}$ for some rational numbers $x$ and $y$. □

Now our goal is to prove the converse of the last two results: if $y$ is a (reduced) quadratic irrational, then its continued fraction is (purely) periodic. Let us begin with an example.

**Example** Find the continued fraction of $2+\sqrt{7}$. Note that $2+\sqrt{7}$ is reduced: it is greater than 1, and its algebraic conjugate $2-\sqrt{7}$ lies between $-1$ and 0.

$$\begin{aligned}
y_0 = 2+\sqrt{7}, &\qquad a_0 = \lfloor 2+\sqrt{7} \rfloor = 4 \\
y_1 = 1/(2+\sqrt{7}-4) = (2+\sqrt{7})/3, &\qquad a_1 = \lfloor (2+\sqrt{7})/3 \rfloor = 1 \\
y_2 = 3/(2+\sqrt{7}-3) = (1+\sqrt{7})/2, &\qquad a_2 = \lfloor (1+\sqrt{7})/2 \rfloor = 1 \\
y_3 = 2/(1+\sqrt{7}-2) = (1+\sqrt{7})/3 &\qquad a_3 = \lfloor (1+\sqrt{7})/3 \rfloor = 1 \\
y_4 = 3/(1+\sqrt{7}-3) = 2+\sqrt{7} = y_0 &
\end{aligned}$$

So

$$2+\sqrt{7} = [\overline{4;1,1,1}].$$

Note that all of $y_0, y_1, y_2, y_3$ are reduced quadratic irrationals, and we can read off their continued fractions: for example, $y_2 = [\overline{1;4,1,1}]$. Other observations which will be important in the proof are that, in each case, $y_i = (p_i+\sqrt{7})/q_i$, where $p_i$ and $q_i$ are integers (the $p_i$ are $2,2,1,1,\ldots$ and the $q_i$ are $1,3,2,3,\ldots$); and $0 < p_i < \sqrt{7}$, $0 < q_i < 2\sqrt{7}$.

We will see that all these properties hold quite generally.

Before we start the proofs, we introduce a slightly different way of writing quadratic irrationals.

**Lemma 5.3** *(a) A real quadratic irrational can be written as $y = (P+\sqrt{D})/Q$, where $P, Q$ are integers, $D$ is a positive integer which is not a square, and $Q$ divides $D-P^2$.*

*(b) If $y$ is reduced, then $0 < P < \sqrt{D}$ and $0 < Q < 2\sqrt{D}$.*

**Proof** (a) We know that $y = u + v\sqrt{d}$ where $u$ and $v$ are rationals and $d$ is squarefree. Suppose first that $v$ is positive. Let $q$ be the least common multiple of the denominators of $u$ and $v$, and $u = p/q$, $v = r/q$. Then

$$y = \frac{p+r\sqrt{d}}{q} = \frac{p+\sqrt{r^2d}}{q} = \frac{pq+\sqrt{q^2r^2d}}{q^2}.$$

Put $P = pq$, $Q = q^2$, and $D = q^2r^2d$, and note that $Q$ divides $P^2-D$.

If $u < 0$, then write $-y$ in the specified form and then replace $Q$ by $-Q$.

(b) Now suppose that $y$ is reduced; recall that this means $y > 1$ and $-1 < y' < 0$, where $y'$ is the conjugate of $y$ (so $y' = (P-\sqrt{D})/Q$). Then

- $y > 0 > y'$, so $(P+\sqrt{D})/Q > (P-\sqrt{D})/Q$. Hence $Q > 0$.

- $y > 1 > -y'$, so $(P+\sqrt{D})/Q > (-P+\sqrt{D})/Q$. Hence $P > 0$.
- $y' < 0$, so $P-\sqrt{D} < 0$. Hence $P < \sqrt{D}$.
- $y > 1$, so $(P+\sqrt{D})/Q > 1$. Hence $Q < P+\sqrt{D} < 2\sqrt{D}$. □

Suppose that $y$ is reduced. Now we start building the continued fraction for $y_0 = y$:

$$a_0 = \lfloor y_0 \rfloor, \qquad y_1 = \frac{1}{y_0 - a_0}.$$

**Claim:** $y_1$ is reduced.

Certainly $y_1 > 1$, since $y_1 = 1/(y_0 - a_0)$ and $y_0 - a_0 < 1$. We have to show that $-1 < y_1' < 0$.

Let $P^* = Qa_0 - P$. Then

$$\begin{aligned} y_1 &= \frac{1}{(P+\sqrt{D})/Q - a_0} \\ &= \frac{1}{(P+\sqrt{D}-Qa_0)/Q} \\ &= \frac{1}{(-P*+\sqrt{D})/Q} \\ &= \frac{P^*+\sqrt{D}}{(D-(P^*)^2)/Q} \\ &= \frac{P^*+\sqrt{D}}{Q^*}, \end{aligned}$$

where $Q^* = (D-(P^*)^2)/Q$. Now

$$Q^* = \frac{D-(Qa_0-P)^2}{Q} = -Qa_0^2 + 2Pa_0 + \frac{D-P^2}{Q},$$

so $Q^*$ is an integer. Moreover, $Q^*$ divides $D-(P^*)^2$ (the quotient is just $Q$). So we have written $y_1$ in the same form as $y$, with the same $D$, but maybe different $P$ and $Q$. Also, we have $y_1 = 1/(y-a_0)$, and so $y_1' = 1/(y'-a_0)$, so $-1 < y_1' < 0$ as required.

Our claim is proved.

Now calculate the continued fraction expansion of $y$. In the general step, we start with $y_n = (P_n+\sqrt{D})/Q_n$, where $Q_n$ divides $P_n^2 - D$; assume inductively that $y_n$ is reduced, so $0 < P_n < \sqrt{D}$, $0 < Q_n < 2\sqrt{D}$. Then we put

$$a_n = \lfloor y_n \rfloor, \qquad y_{n+1} = \frac{1}{y_n - a_n}.$$

By the above argument, we have $y_{n+1} = (P_{n+1} + \sqrt{D})/Q_{n+1}$, where the same conditions hold.

Now $P_n$ and $Q_n$ are integers satisfying $0 < P_n < \sqrt{D}$ and $0 < Q_n < 2\sqrt{D}$. There are only finitely many possible values of $P_n$ and $Q_n$, so after some number of steps, we must return to values we have seen before. Suppose this first happens when $y_m = y_{m+k}$. Clearly the sequence repeats after this point, that is, $y_n = y_{n+k}$ for all $n \geq m$.

We have to show that the repetition starts with $m = 0$. If not, then we have $y_{m-1} \neq y_{m+k-1}$. But

$$y_m = \frac{1}{y_{m-1} - a_{m-1}} = y_{m+k} = \frac{1}{y_{m+k-1} - a_{m+k-1}},$$

so

$$y'_m = \frac{1}{y'_{m-1} - a_{m-1}} = y'_{m+k} = \frac{1}{y'_{m+k-1} - a_{m+k-1}},$$

So $y'_{m-1} - a_{m-1} = y'_{m+k-1} - a_{m+k-1}$. Thus $y'_{m-1}$ and $y'_{m+k-1}$ differ by an integer. But they both lie between $-1$ and 0, so they are equal, whence $y_{m-1} = y_{m+k-1}$, contrary to assumption.

So, finally, we have proved the converse implication in the big theorem. (The forward implication was already proved in Proposition 5.1.)

**Theorem 5.4** *The irrational number y has a purely periodic continued fraction if and only if it is a reduced quadratic irrational.*

From this, it is not such a big step to the other main theorm of this chapter:

**Theorem 5.5** *The irrational number y has a periodic continued fraction if and only if it is a quadratic irrational.*

**Proof** We have proved the forward implication in Theorem 5.2. So suppose that $y$ is a quadratic irrational. Calculate its continued fraction: that is, put $y_0 = y$, and then

$$a_n = \lfloor y_n \rfloor, \qquad y_{n+1} = \frac{1}{y_n - a_n}$$

for $n \geq 0$. It is clear that all the $y_n$ are quadratic irrationals, since subtracting an integer from a quadratic irrational, and taking the reciprocal of one, gives again a quadratic irrational. We have to prove that, for some value of $n$, the number $y_n$ is a reduced quadratic irrational. Then by the preceding theorem, the continued fraction is periodic from that point on.

By construction, we have $y_n > 1$ for all $n > 1$ (it is the reciprocal of a number smaller than 1). Also, we have for any $k > 0$

$$y = [a_0; a_1, \dots, a_{k-1}, y_k] = \frac{y_k p_{k-1} + p_{k-2}}{y_k q_{k-1} + q_{k-2}}.$$

Hence

$$y' = \frac{y'_k p_{k-1} + p_{k-2}}{y'_k q_{k-1} + q_{k-2}},$$

so rearranging we obtain

$$y'_k = \frac{-y' q_{k-2} + p_{k-2}}{y' q_{k-1} - p_{k-1}} = -\frac{q_{k-2}}{q_{k-1}} \frac{y' - c_{k-2}}{y' - c_{k-1}},$$

where $c_n$ is the $n$th convergent. Since $c_n \to y$ as $n \to \infty$, we have

$$-\frac{q_{k-1}}{q_{k-2}} y'_k = \frac{y' - c_{k-2}}{y' - c_{k-1}} \to \frac{y' - y}{y' - y} = 1$$

as $k \to \infty$. So we can choose $n$ large enough that, for $k > n$, $|(y' - c_{k-2})/(y' - c_{k-1}) - 1| < 1$, and so this fraction is positive. Thus $-(q_{k-1}/q_{k-2})y'_k > 0$, so that $y'_k < 0$.

Also, we can ensure that $n$ is also large enough that $|c_k - y| < |y' - y|$ for $k > n$.

If $y' < y$, we use the fact that even-numbered convergents are smaller than $y$ and odd-numbered convergents are greater; choosing $k$ even, we have $y' < c_{k-2} < y < c_{k-1}$. If $y < y'$, then choosing $k$ odd we have $c_{k-1} < y < c_{k-2} < y'$. In either case, we have

$$y'_k = \frac{q_{k-2}}{q_{k-1}} \left| \frac{y' - c_{k-2}}{y' - c_{k-1}} \right| < 1,$$

so finally we conclude

$$-1 < y'_k < 0$$

and $y_k$ is a reduced quadratic irrational, as required. □

## Exercises

**5.1** Express each of the following periodic continued fractions in the form $u + v\sqrt{d}$, where $u$ and $v$ are rationals and $d$ is a squarefree integer greater than 1:

(a) $[\overline{1; 2, 3}]$,

(b) $[\overline{2; 3, 1}]$,

(c) $[1; \overline{1, 2, 3}]$,

**5.2** Express each of the following periodic continued fractions in the form $u + v\sqrt{d}$, where $u$ and $v$ are rationals and $d$ is a squarefree integer greater than 1:

(a) $[\overline{1;2,3}]$,

(b) $[\overline{2;3,1}]$,

(c) $[1;\overline{1,2,3}]$,

**5.3** Which of the following quadratic irrationals are reduced?

(a) $\sqrt{2} + (3/5)$

(b) $5 + \sqrt{101}/2$

(c) $(\sqrt{2} + \sqrt{3})^2$

(d) $(\sqrt{5} - 1)/2$

# Chapter 6

# Lagrange and Pell

The continued fraction expansion of $\sqrt{n}$, where $n$ is a positive integer which is not a square, has a very special form, which we derive in this chapter. We also use it to solve *Pell's equation* $x^2 - ny^2 = 1$, and to express primes congruent to 1 mod 4 as sums of two squares.

## 6.1 Introduction

A *diophantine equation* is an equation in more than one variable, where we are looking for integer solutions.

In this chapter we will look for solutions of the two diophantine equations:

Lagrange's equation: $x^2 + y^2 = n$,

Pell's equation: $x^2 - ny^2 = 1$, and the related equation $x^2 - ny^2 = -1$.

We will see that continued fractions give us constructive methods to solve these equations.

Pell's equation was given this name by Euler. According to mathematical legend (possibly apocryphal), Euler knew that an English mathematician had worked on it but couldn't remember which one (in fact it was Wallis). But many earlier mathematicians had studied this equation, notable among them the Indian mathematician Brahmagupta a thousand years earlier. Despite this, the name has stuck!

## 6.2 The continued fraction for $\sqrt{n}$

First, we need a bit more theory of continued fractions, which we introduce by way of an example. You know how to do these calculations now, so I will simply give the result.

**Example**

$$\sqrt{52} = [7;\overline{4,1,2,1,4,14}].$$

Here is the general statement:

**Theorem 6.1** *Let n be a positive integer, which is not a square. Then*

$$\sqrt{n} = [a_0;\overline{a_1,a_2,\ldots,a_{l-1},2a_0}],$$

*where* $a_1 = a_{l-1}$, $a_2 = a_{l-2}$, ... .

We begin with a lemma about purely periodic continued fractions.

**Lemma 6.2** *If* $y = [\overline{a_0;a_1,\ldots,a_{n-1},a_n}]$, *then* $-1/y' = [\overline{a_n;a_{n-1},\ldots,a_1,a_0}]$.

**Proof**

$$y = [a_0;a_1,\ldots,a_n,y] = \frac{yp_n + p_{n-1}}{yq_n + q_{n-1}},$$

so

$$q_ny^2 + (q_{n-1} - p_n)y + p_{n-1} = 0.$$

Let $z = [\overline{a_n;a_{n-1},\ldots,a_0}]$. Then

$$\begin{aligned} z = [a_n;a_{n-1},\ldots,a_0,z] &= \frac{z[a_n,\ldots,a_0] + [a_n,\ldots,a_1]}{z[a_{n-1},\ldots,a_0] + [a_{n-1},\ldots,a_1]} \\ &= \frac{zp_n + q_n}{zp_{n-1} + q_{n-1}}, \end{aligned}$$

where we use the fact that $[a_0,\ldots,a_n] = [a_n,\ldots,a_0]$. So

$$p_{n-1}z^2 + (q_{n-1} - p_n)z - q_n = 0.$$

In other words,

$$q_n(-1/z)^2 + (q_{n-1} - p_n)(-1/z) + p_{n-1} = 0.$$

Thus, $-1/z$ satisfies the same quadratic equation as $y$. But $z > 1$, so $-1 < -1/z < 0$, whereas $y > 1$; so $-1/z$ is the other root $y'$ of the quadratic equation. Thus, we have $z = -1/y'$, as required. □

**Proof of the theorem** Let $n$ be a positive integer which is not a perfect square, and $a_0 = \lfloor\sqrt{n}\rfloor$. Put $y = a_0 + \sqrt{n}$. Then $y > 1$, and $y' = a_0 - \sqrt{n}$ so $-1 < y' < 0$. Thus $y$ has purely periodic continued fraction. Since $\lfloor a_0 = \sqrt{n} = 2a_0$, we have

$$y = [\overline{2a_0; a_1, \dots, a_k}],$$

so

$$\sqrt{n} = [a_0; \overline{a_1, \dots, a_k, 2a_0}].$$

We have $\sqrt{n} = a_0 + 1/y_1$, where

$$y_1 = \frac{1}{\sqrt{n} - a_0} = [\overline{a_1; \dots, a_k, 2a_0}],$$

so

$$\frac{-1}{y_1'} = a_0 + \sqrt{n} = [\overline{2a_0; a_k, \dots, a_1}].$$

But we know that

$$a_0 + \sqrt{n} = [\overline{2a_0; a_1, \dots, a_k}],$$

so

$$a_1 = a_{k-1}, a_2 = a_{k-2}, \dots$$

as required.

In fact, any number $y$ which has a continued fraction of this form (that is, $y = [a_0, \overline{a_1, \dots, a_k, 2a_0}]$, where $a_1 = a_{k-1}$, $a_2 = a_{k-2}$, $\dots$) has the form $\sqrt{r}$ for some rational number $r$ (not necessarily an integer):

**Proposition 6.3** *Let $y = [a_0; \overline{a_1, \dots, a_l - 1, 2a_0}]$, where $a_1 = a_{l-1}$, $a_2 = a_{l-2}$, and so on. Then $y = \sqrt{r}$, where*

$$r = a_0^2 + \frac{[2a_0, a_1, \dots, a_{k-2}]}{[a_1, \dots, a_{k-1}]}.$$

**Proof** We have

$$\begin{aligned} y + a_0 &= [\overline{2a_0, a_1, \dots, a_{l-1}}] \\ &= [2a_0, a_1, \dots, a_{l-1}, y + a_0] \\ &= \frac{(y + a_0)p_{l-1} + p_{l-2}}{(y + a_0)q_{l-1} + q_{l-2}}, \end{aligned}$$

where $p_k/q_k$ are convergents to $y + a_0$. Now

$$\begin{aligned} 2a_0q_{l-1} + q_{l-2} &= 2a_0[a_1, \dots, a_{l-1}] + [a_1, \dots, a_{l-2}] \\ &= 2a_0[a_1, \dots, a_{l-1}] + [a_2, \dots, a_{l-1}] \\ &= p_{l-1}, \end{aligned}$$

using the fact that

$$[a_1,\dots,a_{l-2}] = [a_{l-1},\dots,a_2] = [a_2,\dots,a_{l-1},]$$

the first equality because $a_1 = a_{l-1}$, …, $a_{l-2} = a_2$, and the second by the symmetry of the square bracket function (see (2.5)(a)) in Notes 2). Hence

$$y + a_0 = \frac{(y+a_0)p_{l-1} + p_{l-2}}{(y-a_0)q_{l-1} + p_{l-1}}.$$

Multiplying up and cancelling, we obtain

$$(y^2 - a_0^2)q_{l-1} = p_{l-2},$$

ø$y = \sqrt{r}$ where $r = a_0^2 + p_{l-2}/q_{l-1}$, as claimed. □

**Example** Let $y = [4;\overline{2,1,3,1,2,8}]$.

This is of the above form, so $y = \sqrt{r}$, where

$$r = 4^2 + \frac{[8,2,1,3,1]}{[2,1,3,1,2]} = 16 + \frac{117}{39} = 19.$$

That is,

$$\sqrt{19} = [4;\overline{2,1,3,1,2,8}].$$

(Remember the rule for calculating the square bracket functions: delete consecutive pairs in all possible ways and take the product of the remaining terms, then add all these productss. Check that

$$[8,2,1,3,1] = 48+3+24+16+16+1+1+8 = 117,$$

and calculate $[2,1,3,1,2]$ yourself.)

## 6.3 Sums of two squares

We are going to investigate the question: Which positive integers can be written as the sum of two squares of integers? Of the numbers from 1 to 10, we see that

$$1 = 1^2+0^2,\ 2 = 1^2+1^2,\ 4 = 2^2+0^2,\ 5 = 2^2+1^2,\ 8 = 2^2+2^2,\ 9 = 3^2+0^2,\ 10 = 3^2+1^2,$$

while the other numbers 3, 6, 7 cannot be so written.

In this section, we are going to decide exactly which prime numbers can be written as the sum of two squares. Of course the prime 2 can be so written. For odd primes $p$, we will show that $p$ is the sum of two squares if and only if $p$ is congruent to 1 mod 4. On Coursework 1, you were asked to decide which primes less than 100 are sums of two squares, and which are not; you may have observed this pattern in your answer.

Before we begin, let us observe that any square is congruent to 0 or 1 mod 4, since Since $(2k)^2 = 4k^2$ and $(2k+1)^2 = 4k(k+1)+1$.

So we can very easily do one way round:

**Theorem 6.4** *Let p be a prime congruent to* 3 *mod* 4*. Then p cannot be expressed as a sum of two squares.*

**Proof** A square is congruent to 0 or 1 mod 4, and so a sum of two squares is congruent to 0, 1 or 2 mod 4. □

In the other direction, we have to show that a prime congruent to 1 mod 4 can be written as the sum of two squares, by actually constructing such a representation.

Here is Legendre's construction for expressing an integer as the sum of two squares.

**Proposition 6.5** *Suppose that*

$$\sqrt{n} = [a_0; \overline{a_1, \ldots, a_k, 2a_0}]$$

*where k is odd, say,* $k = 2m+1$*. Write* $y_{m+1} = (P_{m+1} + \sqrt{n})/Q_{m+1}$*. Then*

$$n = P_{m+1}^2 + Q_{m+1}^2.$$

**Proof** In this case, $y = [a_0; \overline{a_1, \ldots, a_m, a_m, \ldots, a_1, 2a_0}]$. We have

$$y_{m+1} = [\overline{a_{m+1}, \ldots, a_{2m}, 2a_0, a_1, \ldots, a_m}],$$

and

$$-1/y'_{m+1} = [\overline{a_m; \ldots, a_1, 2a_0, a_{2m}, \ldots, a_{m+1}}].$$

But by assumption, these two continued fractions are identical. So

$$y_{m+1} y'_{m+1} = -1.$$

Now $y_{m+1} = (P_{m+1} + \sqrt{n})/Q_{m+1}$ and $y'_{m+1} = (P_{m+1} - \sqrt{n})/Q_{m+1}$. So we have

$$\frac{P_{m+1}^2 - n}{Q_{m+1}^2} = -1,$$

so $n = P_{m+1}^2 + Q_{m+1}^2$, as required. □

**Example** Let $n = 41$. Put $y_0 = \sqrt{41}$. We have

$$\begin{aligned} a_0 &= \lfloor y_0 \rfloor = 6, & y_1 &= \frac{1}{y_0 - 6} = \frac{6+\sqrt{41}}{5} \\ a_1 &= \lfloor y_1 \rfloor = 2, & y_2 &= \frac{1}{y_1 - 2} = \frac{4+\sqrt{41}}{5} \\ a_2 &= \lfloor y_2 \rfloor = 2, & y_3 &= \frac{1}{y_2 - 2} = 6+\sqrt{41} = 6+y_0. \end{aligned}$$

So $\sqrt{41} = [6;\overline{2,2,12}]$. We have $m = 1$, $P_2 = 4$, $Q_2 = 5$, and so $41 = 4^2 + 5^2$.

**Remark** You do not have to completely work out the continued fraction for $\sqrt{n}$ in order to apply this method. At each step, you calculate $y_l = (P_l + \sqrt{n})/Q_l$; check whether $P_l^2 + Q_l^2 = n$. Stop when either this occurs, or you find a complete period of the continued fraction and it turns out to have even length (in which case the method has failed).

For example, $\sqrt{6} = [2;\overline{2,4}]$ (work this out for yourself!), and has even period, so the method fails. Indeed 6 is not the sum of two squares.

So we have to show that this construction works for any prime number congruent to 1 mod 4. This will take us the rest of the section.

Let us just make a checklist of what we need from the last section. Remember that a purely periodic continued fraction represents a reduced quadratic irrational $y$ (one satisfying $y > 1$ and $-1 < y' < 0$, where $y'$ is the algebraic conjugate of $y$). Any such number $y$ can be written in the form $(P+\sqrt{D})/Q$, where $P, Q, D$ are integers and $D$ is a positive non-square; we have $0 < P < \sqrt{D}$ and $0 < Q < 2\sqrt{D}$, and $Q$ divides $D - P^2$.

Suppose that $\sqrt{n} = [a_0;\overline{a_1,\ldots,a_{l-1},2a_0}]$. The numbers $y_k$ that arise in the calculation, given by

$$a_k = \lfloor y_k \rfloor, \qquad y_{k+1} = \frac{1}{y_k - a_k},$$

satisfy

$$y_k = \frac{P_k + \sqrt{n}}{Q_k},$$

where $0 < P_k < \sqrt{n}$, $0 < Q_k < 2\sqrt{n}$, and $Q_k$ divides $n - P_k^2$.

Let $p_k/q_k$ be the $k$th convergent to $\sqrt{n}$.

**Lemma 6.6** *With the above notation,* $p_k^2 - nq_k^2 = (-1)^{k-1} Q_{k+1}$ *for* $k \geq 0$.

**Proof** We have $p_0/q_0 = a_0/1$, so $p_0^2 - q_0^2 = a_0^2 - n$. Also,

$$y_1 = \frac{1}{\sqrt{n} - a_0} = \frac{a_0 + \sqrt{n}}{n - a_0^2},$$

so $P_1 = a_0$, $Q_1 = n - a_0^2$. Thus, $p_0^2 - q_0^2 = -Q_1$. So the result is true for $k = 0$.

For $k > 0$, we have

$$\begin{aligned} \sqrt{n} &= [a_0; a_1, \ldots, a_k, y_{k+1}] \\ &= \frac{y_{k+1}p_k + p_{k-1}}{y_{k+1}q_k + p_k} \\ &= \frac{P_{k+1}p_k + Q_{k+1}p_{k-1} + p_k\sqrt{n}}{P_{k+1}q_k + Q_{k+1}q_{k-1} + q_k\sqrt{n}}, \end{aligned}$$

using $y_{k+1} = (P_{k+1} + \sqrt{n})/Q_{k+1}$. Hence

$$\sqrt{n}(P_{k+1}q_k + Q_{k+1}q_{k-1} - p_k) = P_{k+1}p_k + Q_{k+1}p_{k-1} - q_k n.$$

Since $\sqrt{n}$ is irrational, and everything else in the equation is an integer, both sides must be zero. Eliminating $P_{k+1}$ from this equation, we find after some calculation that

$$p_k^2 - nq_k^2 = (p_k q_{k-1} - q_k p_{k_1})Q_{k+1} = (-1)^{k-1}Q_{k+1},$$

using (3.7)(a) from Chapter 3. □

**Proposition 6.7** *Suppose that $p_k/q_k$ is the kth convergent to $\sqrt{n}$. Then $p_k^2 - nq_k^2 = \pm 1$ if and only if $k$ is one less than a multiple of the period of the continued fraction for $\sqrt{n}$.*

**Proof** Suppose that $k$ is one less than a multiple of the period. Then $y_{k+1} = \sqrt{n} + a_0$, so $P_{k+1} = a_0$ and $Q_{k+1} = 1$. The lemma gives $p_k^2 - nq_k^2 = (-1)^{k-1}$.

Conversely, suppose that $p_k^2 - nq_k^2 = \pm 1$. Then the Lemma gives $Q_{k+1} = \pm 1$. But $0 < Q_{k+1} < 2\sqrt{n}$; so $Q_{k+1} = 1$. Thus, $y_{k+1} = P_{k+1} + \sqrt{n} = P_{k+1} + a_0 + 1/y_1$ (since $y_1 = 1/(\sqrt{n} - a_0)$, whence $a_{k+1} = \lfloor y_{k+1} \rfloor = P_{k+1} + a_0$, and $y_{k+2} = y_1$, so we have found a complete period of the continued fraction, which is an arbitrary multiple of the smallest period. □

Now here is the main theorem.

**Theorem 6.8** *Let $p$ be a prime congruent to 1 mod 4. Then $p$ can be expressed as a sum of two squares by Legendre's construction.*

**Proof** We know that Legendre's construction will succeed if $\sqrt{p} = [a_0; \overline{a_1, \ldots, a_l}]$ with $l$ odd, so we have to show that this does indeed happen. Clearly we can assume that $l$ is as small as possible, since any period is a multiple of the smallest one, so if the smallest period is even then all periods are even. What we have to show is that, if $p$ is an odd prime which has the above periodic continued fraction and $l$ is even, then $p$ is not congruent to 1 mod 4.

We can assume that $p > 3$. Let $l = 2m+2$, so that

$$\sqrt{p} = [a_0; \overline{a_1, \ldots, a_m, a_{m+1}, a_m, \ldots, a_1, 2a_0}].$$

Now $m+1 < l$, and so $\sqrt{p} \neq [a_0; \overline{a_1, \ldots, a_{m+1}}]$; by the preceding proposition, $p_m^2 - nq_m^2 \neq \pm 1$, so that $Q_{m+1} \neq 1$. Also, $Q_{m+1}$ divides $p - P_{m+1}^2$.

We have

$$y_{m+1} = [\overline{a_{m+1}; a_m, \ldots, a_1, 2a_0, a_1, \ldots, a_m}],$$

and so

$$-1/y'_{m+1} = [\overline{a_m; \ldots, a_1, 2a_0, a_1, \ldots, a_m, a_{m+1}}].$$

So $y_{m+1} = a_{m+1} + 1/(-1/y'_{m+1})$, that is, $y_{m+1} + y'_{m+1} = a_{m+1}$. But $y_{m+1} = (P_{m+1} + \sqrt{p})/Q_{m+1}$, and its algebraic conjugate is $y'_{m+1} = (p_{m+1} - \sqrt{p})/Q_{m+1}$; their sum is $2P_{m+1}/Q_{m+1}$. So $Q_{m+1}$ divides $2P_{m+1}$.

But we know that $Q_{m+1}$ divides $P_{m+1}^2 - p$; so $Q_{m+1}$ divides $2p$. The only divisors of $2p$ are $1, 2, p, 2p$. Since $0 < Q_{m+1} < 2\sqrt{p}$, we must have $Q_{m+1} = 1$ or $Q_{m+1} = 2$. But we know it isn't 1, so $Q_{m+1} = 2$.

So $p_m^2 - pq_m^2 = \pm 2$. Now any square is congruent to 0 or 1 mod 4 So

$$(0 \text{ or } 1) - p(0 \text{ or } 1) = \pm 2 \pmod 4.$$

This is impossible to satisfy if $p \equiv 1$ mod 4. □

## 6.4 The equations $x^2 - ny^2 = \pm 1$

**Example** Let $n = 2$. One can easily find the first few solutions of $x^2 - 2y^2 = \pm 1$ in positive integers:

$$\begin{aligned} 1^2 - 2 \cdot 1^2 &= -1, & (x,y) &= (1,1) \\ 3^2 - 2 \cdot 2^2 &= +1, & (x,y) &= (3,2) \\ 7^2 - 2 \cdot 5^2 &= -1, & (x,y) &= (7,5) \\ 17^2 - 2 \cdot 12^2 &= +1, & (x,y) &= (17,12) \end{aligned}$$

So the solutions of the two equations appear alternately. You might observe that, if $(x_k, y_k)$ is the $k$th solution, then

$$x_{k+1} = x_k + 2y_k, \qquad y_{k+1} = x_k + y_k;$$

so we can generate the solutions very easily. You might further observe that these equations imply

$$x_{k+1} + y_{k+1}\sqrt{2} = (x_k + y_k\sqrt{2})(1+\sqrt{2}),$$

so that $x_k + y_k\sqrt{2} = (1+\sqrt{2})^k$. This is the general pattern, as we will see. The only difference is that sometimes only the equation with the $+$ sign has solutions.

**Example** Let $n = 3$. The equation $x^2 - 3y^2 = -1$ has no solutions, since any square is congruent to 0 or $+1$ mod 3. The first few solutions of $x^2 - 3y^2 = +1$ are

$$\begin{array}{rl} 2^2 - 3\cdot 1^2 = 1, & (x,y) = (2,1) \\ 7^2 - 3\cdot 4^2 = 1, & (x,y) = (7,4) \\ 26^2 - 3\cdot 15^2 = 1. & (x,y) = (26,15) \end{array}$$

and the general solution is given by $x_k + y_k\sqrt{3} = (2+\sqrt{3})^k$.

We have seen that the continued fraction for $\sqrt{n}$ has the form

$$\sqrt{n} = [a_0; \overline{a_1, \ldots, a_{l-1}, 2a_0}].$$

We are going to show that, if $p_k/q_k$ denotes the $n$th convergent to $\sqrt{n}$, then $(p_l, q_l)$, $(p_{2l+1}, q_{2l+1})$, $(p_{3l+2}, q_{3l+2})$, $\ldots$ give all the solutions to the equation in the title of the section.

**Example** $\sqrt{2} = [1; \overline{2}]$. The successive convergents are $1/1$, $3/2$, $7/5$, $17/12, \ldots$

**Example** $\sqrt{3} = [1; \overline{1; 2}]$. The successive convergents are $1/1$, $2/1$, $5/3$, $7/4$, $19/11$, $26/15$, $\ldots$ . This time we see that only the odd-numbered convergents give solutions to Pell's equation.

We showed in Proposition 6.7 in the preceding section that convergents $p_k/q_k$ give solutions to $x^2 - ny^2 = \pm 1$ if and only if $k$ is one less than a multiple of the period of the continued fraction for $\sqrt{n}$.

Now we have to show that every solution to the equation arises from a convergent. We show that, if $x^2 - ny^2 = \pm 1$, then

$$(x + y\sqrt{n})(x - y\sqrt{n}) = \pm 1,$$

so

$$|\sqrt{n} - \frac{x}{y}| = \frac{1}{y(x + y\sqrt{n})},$$

and hence $x/y$ is a good rational approximation to $\sqrt{n}$; but we know that every good rational approximation is a convergent.

In detail: suppose that $u/v$ is another rational number in its smallest terms with $v < y$ and

$$|\sqrt{n} - \frac{u}{v}| < |\sqrt{n} - \frac{x}{y}|.$$

The difference $(x/y) - (u/v)$ is a non-zero rational number with denominator $yv$, so

$$\frac{1}{yv} \le \left|\frac{x}{y} - \frac{u}{v}\right| < \frac{2}{y(x+y\sqrt{n})},$$

which implies that $y > v > (x+y\sqrt{n})/2$, which is impossible if $\sqrt{n} > 2$. The cases $n = 2$ and $n = 3$ can be done directly; indeed, they were our introductory examples.

So we have proved:

**Theorem 6.9** *Let $n$ be a positive integer which is not a square, and suppose that $x^2 - ny^2 = \pm 1$. Then $x/y$ is a convergent to $\sqrt{n}$.*

**Example** We will find the continued fraction for $\sqrt{13}$ and use it both to express 13 as a sum of two squares and to solve Pell's equation.

$$\begin{array}{rl}
 & y_0 = \sqrt{13} \\
a_0 = \lfloor y_0 \rfloor = 3, & y_1 = 1/(\sqrt{13}-3) = (\sqrt{13}+3)/4 \\
a_1 = \lfloor y_1 \rfloor = 1, & y_2 = 4/(\sqrt{13}-1) = (\sqrt{13}+1)/3 \\
a_2 = \lfloor y_2 \rfloor = 1, & y_3 = 3/(\sqrt{13}-2) = (\sqrt{13}+2)/3 \\
a_3 = \lfloor y_3 \rfloor = 1, & y_4 = 3/(\sqrt{13}-1) = (\sqrt{13}+1)/4 \\
a_4 = \lfloor y_4 \rfloor = 1, & y_5 = 4/(\sqrt{13}-3) = \sqrt{13}+3 \\
a_5 = \lfloor y_5 \rfloor = 6, & y_6 = 1/(\sqrt{13}-3) = y_1
\end{array}$$

So $\sqrt{13} = [3; \overline{1,1,1,1,6}]$.

Since the period is 5, to write 13 as a sum of squares we look at

$$y_3 = \frac{2+\sqrt{13}}{3}, \qquad 13 = 2^2 + 3^2.$$

For Pell's equation, we have

$$[3;1,1,1,1] = \frac{[3,1,1,1,1]}{[1,1,1,1]} = \frac{18}{5},$$

and $18^2 - 13 \cdot 5^2 = -1$. So the smallest solution of Pell's equation is $(x,y)$, where

$$x + y\sqrt{13} = (18 + 5\sqrt{13})^2 = 649 + 180\sqrt{13},$$

that is, $(x,y) = (649, 180)$.

Here you can certainly find the sum of squares by trial and error, but solving Pell's equation without some theory would be more daunting!

**Definition** The solution to $x^2 - ny^2 = \pm 1$ in positive integers for which $x + y\sqrt{n}$ is smallest is called the *fundamental solution* of this equation.

**Theorem 6.10** *Let $(x_1, y_1)$ be the fundamental solution of $x^2 - ny^2 = \varepsilon$, where $\varepsilon = \pm 1$.*

*(a) If $\varepsilon = +1$, then there are no solutions to $x^2 - ny^2 = -1$, and all solutions $(x_k, y_k)$ of $x^2 - ny^2 = 1$ are given by*

$$(x_k + y_k\sqrt{n}) = (x_1 + y_1\sqrt{n})^k.$$

*(b) If $\varepsilon = -1$, then all solutions $(x_k, y_k)$ of $x^2 - ny^2 = \pm 1$ are given by*

$$(x_k + y_k\sqrt{n}) = (x_1 + y_1\sqrt{n})^k,$$

*where we get the plus sign if k is even and the minus sign if k is odd.*

**Proof** First let us see that these are all the solutions. This depends on the following:

> Let $(a,b)$ be a solution of $x^2 - ny^2 = s$, and $(c,d)$ be a solution of $x^2 - ny^2 = t$. Define $(e,f)$ by the rule
>
> $$e + f\sqrt{n} = (a + b\sqrt{n})(c + d\sqrt{n}).$$
>
> Then $(e,f)$ is a solution of $x^2 - ny^2 = st$.

For the given equation implies that

$$e - f\sqrt{n} = (a - b\sqrt{n})(c - d\sqrt{n}).$$

Multiplying these two equations together we find

$$e^2 - nf^2 = (a^2 - nb^2)(c^2 - nd^2 + = st,$$

as required.

This together with a short proof by induction shows that, if $x^2 - ny^2 = \varepsilon$, where $\varepsilon = \pm 1$, and $x_k + y_k\sqrt{n} = (x_1 + y_1\sqrt{n})^k$, then

$$x^2 - ny^2 = \varepsilon^k = \begin{cases} 1 & \text{if } \varepsilon = 1, \\ (-1)^k & \text{if } \varepsilon = -1. \end{cases}.$$

It remains to show that these are all the solutions. So suppose that

$$u^2 - nv^2 = \pm 1$$

and that $(u, v)$ is not of the form $(x_k, y_k)$ as in the theorem. We may assume that $|u + v\sqrt{n}|$ is minimal with this property; that is, all smaller solutions are of the appropriate form. Let $(x_1, y_1)$ be the fundamental solution. We see that there must exist $k$ such that

$$x_k + y_k\sqrt{n} < u + v\sqrt{n} < (x_{k+1} + y_{k+1}\sqrt{n}).$$

Now $(x_1 + y_1\sqrt{n})(x_1 - y_1\sqrt{n}) = x_1^2 - ny_1^2 = \delta$, say, where $\delta = \pm 1$. So

$$\delta(x_1 - y_1\sqrt{n}) \cdot (x_k + y_k\sqrt{n} = x_{k-1} + y_{k-1}\sqrt{n}.$$

We see that, if $u' + v'\sqrt{n} = \delta(x_1 - y_1\sqrt{n})(u + v\sqrt{n})$, then

$$x_{k-1} + y_{k-1}\sqrt{n} < u' + v'\sqrt{n} < (x_k + y_k\sqrt{n}).$$

By choice of the solution $(u, v)$, we must have $(u', v') = (x_m, y_m)$ for some $m$, whence $(u, v) = (x_{m+1}, y_{m+1})$, contrary to assumption. So the theorem is proved.
□

**Example** The fundamental solution of $x^2 - 3y^2 = 1$ is easily seen to be $(2, 1)$. So the next few solutions of this equation are given by

$$\begin{aligned} x_2 + y_2\sqrt{3} &= (2 + \sqrt{3})^2 = 7 + 4\sqrt{3}, & (x_2, y_2) &= (7, 4) \\ x_3 + y_3\sqrt{3} &= (2 + \sqrt{3})^3 = 26 + 15\sqrt{3}, & (x_3, y_3) &= (26, 15) \\ x_4 + y_4\sqrt{3} &= (2 + \sqrt{3})^4 = 97 + 56\sqrt{3}, & (x_4, y_4) &= (97, 56) \end{aligned}$$

and so on.

**Example** We have seen that $(32, 5)$ is the fundamental solution of $x^2 - 41y^2 = -1$. So the smallest solution of Pell's equation $x^2 - 41y^2 = +1$ is $(x_2, y_2)$, where

$$x_2 + y_2\sqrt{41} = (32 + 5\sqrt{41})^2 = 2049 + 320\sqrt{41},$$

in other words, $(2049, 320)$. By taking powers, we find infinitely solutions to the equation.

## Exercises

**6.1** We saw that $\sqrt{3} = [1; \overline{1,2}]$, and that the solutions $(x,y)$ in positive integers to $x^2 - 3y^2 = \pm 1$ are given by $x = p_n$, $y = q_n$, where $n$ is odd, and $p_n/q_n$ is the $n$th convergent to the continued fraction for $\sqrt{3}$. We saw further that all of these satisfy $x^2 - 3y^2 = 1$.

(a) Prove directly that the equation $x^2 - 3y^2 = -1$ has no solution in positive integers. (Hint: Congruence mod 3.)

(b) Let $a_m = [1,1,2,1,2,\ldots,1,2]$ (with $2m+1$ terms) and $b_m = [1,1,2,1,2,\ldots,1]$ (with $2m$ terms), so that $a_m = p_{2m}$ and $b_m = p_{2m-1}$. Prove that, for $m \geq 3$,

$$\begin{aligned} a_m &= 2b_m + a_{m-1}, \\ b_m &= a_{m-1} + b_{m-1}. \end{aligned}$$

Deduce that $b_m = 4b_{m-1} - b_{m-2}$ for $m \geq 3$, with $b_1 = 2$ and $b_2 = 7$.

(c) Similarly show that, if $c_m = [1,2,1,2,\ldots,1,2]$ (with $2m$ terms) and $d_m = [1,2,1,\ldots,1]$ (with $2m-1$ terms), so that $c_m = q_{2m}$ and $d_m = q_{2m-1}$, then $d_m = 4d_{m-1} - d_{m-2}$ for $m \geq 3$, with $d_1 = 1$ and $d_2 = 4$.

(d) Deduce that, if $(x_n, y_n)$ is the $n$th solution to $x^2 - 3y^2 = 1$ in positive integers, then

$$\begin{aligned} x_1 = 2, x_2 = 7, \qquad & x_n = 4x_{n-1} - x_{n-2} \text{ for } n \geq 3, \\ y_1 = 1, y_2 = 4, \qquad & y_n = 4y_{n-1} - y_{n-2} \text{ for } n \geq 3. \end{aligned}$$

(e) Hence find the first four solutions of this equation in positive integers.

# Chapter 7

# Euler's totient function

In this chapter, we look at Euler's totient function $\phi(n)$, and the existence of primitive roots modulo a prime number.

## 7.1 Euler's totient function

We say that non-negative integers $x$ and $y$ are *coprime* if $\gcd(x,y) = 1$.

*Euler's totient function*, or *euler's $\phi$-function*, is the function $\phi$ defined on the positive integers by the rule that $\phi(n)$ is the number of integers $x$ in $\{0,1,\ldots,n-1\}$ which are coprime to $n$.

**Example** If $p$ is prime, then $\phi(p) = p-1$: the integers $1,2,\ldots,p-1$ are all coprime to $p$.

**Example** $\phi(8) = 4$; the odd numbers $1,3,5,7$ are coprime to 8, while the even numbers are not.

Here is a more algebraic interpretation of Euler's function. Recall that, if $R$ is a commutative ring with identity, then an element $x \in R$ is a *unit* if there exists $y \in R$ such that $xy = 1$. The units in $R$ form a group (with the operation of multiplication).

**Proposition 7.1** *The number of elements in the group of units of $\mathbb{Z}_n$ is $\phi(n)$.*

**Proof** We show that $[x]_n$ is a unit in $\mathbb{Z}_n$ if and only if $x$ is coprime to $n$. Then the result follows.

The first part uses the same argument as we already saw for primes. Suppose that $\gcd(x,n) = 1$. Then there exist integers $y,z$ with $xy+nz = 1$, by Euclid's algorithm; thus $xy \equiv 1 \bmod n$, so $[x]_n[y]_n = [1]_n$, and $[x]_n$ is a unit.

Conversely, if $[x]_n$ is a unit, then by definition there exists $[y]_n$ so that $[x]_n[y]_n = [1]_n$, so that $xy \equiv 1 \bmod n$, or $xy + nz = 1$ for some integer $z$. Let $d = \gcd(x,n)$. Then $d$ divides $x$ and $d$ divides $n$, so $d$ divides $xy + nz = 1$; so $d = 1$, as required. □

From this we can deduce a theorem of Euler:

**Theorem 7.2** *Let n be a positive integer, and x an integer such that* $\gcd(x,n) = 1$. *Then* $x^{\phi(n)} \equiv 1$ *mod n.*

**Proof** There is a very simple proof using algebra. If $\gcd(x,n) = 1$, then $[x]_n$ is an element of the group of units of $\mathbb{Z}_n$. let $d$ be its order (the least positive integer such that $x^d \equiv 1 \bmod n$). By Lagrange's Theorem, $d$ divides the order of the group of units, which is $\phi(n)$, say $\phi(n) = de$. Then

$$x^{\phi(n)} = x^{de} = (x^d)^e \equiv 1^e = 1 \bmod n,$$

as required.

Here is a more direct proof. Let $y_1, \ldots, y_{\phi(n)}$ be the integers in $\{0, \ldots, n\}$ which are coprime to $n$. Then $xy_1, \ldots, xy_{\phi(n)}$ are all coprime to $n$; and no two of these are congruent mod $n$. (If $xy_i \equiv xy_j$, multiplying by the inverse of $x$ we find that $y_1 \equiv y_j$.) Thus $xy_1, \ldots, xy_{\phi(n)}$ are congruent to $y_1, \ldots, y_{\phi(n)}$ in some order, and so their products are congruent mod $n$:

$$x^{\phi(n)} y_1 \cdots y_{\phi(n)} \equiv y_1 \cdots y_{\phi(n)} \bmod n.$$

But again all the $y$s can be cancelled since they are coprime to $n$, leaving us with $x^{\phi(n)} \equiv 1 \bmod n$. □

As a corollary, we obtain *Fermat's Little Theorem:*

**Corollary 7.3** *Let p be prime. Then* $x^p \equiv x$ *mod p for any integer x.*

**Proof** If $p$ divides $x$, then $x \equiv 0 \bmod p$ and $x^p \equiv 0 \bmod p$, so the result is true. If $p$ does not divide $x$, then $\gcd(x,p) = 1$ and $\phi(p) = p - 1$, so Euler's theorem gives $x^{p-1} \equiv 1 \bmod p$. Multiplying both sides by $x$ gives $x^p \equiv x \bmod p$. □

The converse of this is not true. We saw in an exercise in Chapter 1 that there exist positive integers $n$ which are not prime but which satisfy $x^n \equiv x \bmod n$ for every integer $x$. Such integers are called *Carmichael numbers*; the smallest is 561.

## 7.2 Evaluation of $\phi(n)$

We now give a rule for calculating $\phi(n)$ for any integer $n$.

**Theorem 7.4** *(a) If $p$ is prime and $r > 0$, then $\phi(p^r) = p^{r-1}(p-1)$.*

*(b) If $\gcd(m,n) = 1$, then $\phi(mn) = \phi(m)\phi(n)$.*

*(c) Suppose that $n = p_1^{r_1} \cdots p_s^{r_s}$, where $p_1, \ldots, p_s$ are distinct primes and $r_1, \ldots, r_s > 0$. Then*

$$\phi(n) = \prod_{i=1}^{s} p_i^{r_i - 1}(p_i - 1) = n\prod_{i=1}^{s}(1 - 1/p_i).$$

**Example** $720 = 2^4 \cdot 3^2 \cdot 5$, so

$$\phi(720) = 2^3(2-1)3^1(3-1)5^0(5-1) = 8 \cdot 6 \cdot 4 = 192.$$

**Proof** (a) Let $n = p^r$, where $p$ is prime and $r > 0$, The numbers less than $n$ which are coprime to $n$ are precisely those which are not divisible by $p$. So, of the $p^r$ possibilities, we have to remove $p^{r-1}$ multiples of $p$, so that $\phi(p^r) = p^r - p^{r-1} = p^{r-1}(p-1)$.

(b) We use the *Chinese Remainder Theorem* (see Chapter 1). Suppose that $\gcd(m,n) = 1$. Given any $x$ and $y$, there exists $z$ such that

$$z \equiv x \bmod m, \qquad z \equiv y \bmod n;$$

and these congruences have a unique solution mod $mn$. We show that $\gcd(z,mn) = 1$ if and only if $\gcd(x,m) = 1$ and $\gcd(y,n) = 1$. This is true since any common factor of $z$ and $mn$ must divide either $m$ (and hence divides $\gcd(z,m) = \gcd(x,m)$) or $n$ (and hence divides $\gcd(z,n) = \gcd(y,n)$). Conversely, a common factor of $x$ and $m$ divides $z$ and $mn$.

So if $x_1, \ldots, x_{\phi(m)}$ are all the integers less than $m$ and coprime to $m$, and $y_1, \ldots, y_{\phi(n)}$ are all the integers less than $n$ and coprime to $n$, then for each pair $i, j$, the Chinese Remainder Theorem gives us a number $z_{ij}$ congruent to $x_i$ mod $m$ and to $y_j$ mod $n$; these $z_{ij}$ are all coprime to $mn$, and are all distinct mod $mn$, and every number less than $mn$ and coprime to $mn$ arises in this way. So $\phi(mn) = \phi(m)\phi(n)$.

(c) The result of (b) easily extends to the product of more than two pairwise coprime integers. So we can apply it to the prime powers $p_1^{r_1}, \ldots, p_s^{r_s}$ to obtain the first equality in (c). The second equality is a simple manipulation, since $p^{r-1}(p-1) = p^r(1-1/p)$. □

We need a technical result about Euler's function:

**Proposition 7.5** *Let $d$ be a divisor of $n$. Then the number of integers $x$ with $0 \leq x \leq n-1$ and $\gcd(x,n) = d$ is $\phi(n/d)$.*

**Proof** Let $n = dm$. We can write any such integer as $x = dy$, and when we cancel the factor $d$ from $x$ and $n$ there is no further common factor; so $\gcd(y,m) = 1$. There are $\phi(m)$ such numbers $y$; each of them, multiplied by $d$, gives a solution of $\gcd(x,n) = d$, and all solutions are obtained thus. □

## 7.3 Orders of elements

Let $n$ be a positive integer. The *order* of $x$ mod $n$ is the smallest positive integer $d$ such that $x^d \equiv 1 \bmod n$ (if such an integer $d$ exists).

**Proposition 7.6** *The integer $x$ has an order mod $n$ if and only if $\gcd(x,n) = 1$. If so, then the order of $x$ divides $\phi(n)$.*

**Proof** If $x$ has an order $d$, then $x^d \equiv 1 \bmod n$, so $\gcd(x^d,n) = 1$, and certainly $\gcd(x,n) = 1$. Conversely, if $\gcd(x,n) = 1$, then $x^{\phi(n)} \equiv 1 \bmod n$, so there certainly do exist such integers; the order $d$ is the smallest.

Write $\phi(n) = rq + r$, where $0 \leq r \leq d-1$, by the division algorithm. Then

$$1 \equiv x^{\phi(n)} = (x^d)^q \cdot x^r \equiv x^r \bmod n.$$

But $r < d$, and $d$ was the smallest positive integer with this property. So we must have $r = 0$, so that $d$ divides $\phi(n)$, as claimed. □

**Example** Let $n = 12$; we have $\phi(12) = 4$, and the four integers smaller than and coprime to 12 are $1,5,7,11$. Now we have

$$1^1 \equiv 1, \quad 5^2 \equiv 1, \quad 7^2 \equiv 1, \quad 11^2 \equiv 1$$

mod 12. So these four integers have orders $1,2,2,2$ respectively. This shows that not every divisor of $\phi(n)$ necessarily occurs as the order of an element mod $n$. In the next section, we consider one very important case where every divisor does indeed occur.

**Remark** How do we find the order of $x$ mod $n$? One way would be to calculate $x, x^2, x^3, \ldots$ mod $n$ until we first reach one which is congruent to 1. But the order must divide $\phi(n)$, so we only need test divisors of $\phi(n)$. For example, $\phi(10) = 4$, and $3^2 \not\equiv 1 \bmod 10$; so the order of 3 mod 10 must be 4.

## 7.4 Primitive roots

Let $p$ be a prime number. An integer $x$ is said to be a *primitive root* of $p$ if $x$ has order $p-1$ mod $p$. (This is the largest possible order, since $\phi(p) = p-1$.

**Example** Let $p = 17$. We find that $2^8 \equiv 1$ mod 17, so 2 is not a primitive root. But $3^8 \equiv 16$ mod 17. So the order of 3 is a divisor of 16 but is not a divisor of 8; thus 3 must be a primitive root of 17.

We show that primitive roots always exist. This depends on the following lemma together with some clever counting.

**Lemma 7.7** *Let p be prime, and let d be a divisor of* $p-1$*. Then the number of elements of order d mod p is either* 0 *or* $\phi(d)$*.*

**Proof** Suppose that the number of such elements is not zero; so there is at least one element of order $d$, say $a$. Then the numbers $1, a, a^2, \ldots, a^{d-1}$ are all distinct mod $p$. For if, say, $a^i \equiv a^j$, where $i < j$, then $a^{j-i} \equiv 1$, with $j-i < d$, contradicting the definition of the order of $a$.

Next we show that these numbers are all the solutions of $x^d = 1$ in $\mathbb{Z}_p$. For $\mathbb{Z}_p$ is a field, and the polynomial $x^d - 1$ of degree $d$ cannot have more than $d$ solutions; but we have found $d$ distinct solutions, so we have them all.

Finally, we show that $a^m$ has order $d$ if and only if $\gcd(m,d) = 1$. If $e$ divides $\gcd(m,d)$, then $(a^m)^{d/e} = (a^d)^{m/e} \equiv 1$, so $a^m$ has order at most $d/e$. Conversely, suppose that $\gcd(m,d) = 1$, and choose integers $u$ and $v$ with $mu + dv = 1$. Let $b = a^m$. Then $b^u = a^{mu} = a^{1-dv} \equiv a$, so each of $a$ and $b$ is a power of the other, and they must have the same order.

So the number of elements of order $d$ is the number of values of $m$ such that $\gcd(m,d) = 1$, which is just $\phi(d)$. □

Now we are ready to prove the existence of primitive roots.

**Theorem 7.8** *Let p be prime. Then, for every number d dividing* $p-1$*, the number of elements of order d mod p is equal to* $\phi(d)$*. In particular, there are* $\phi(p-1)$ *primitive roots of p.*

**Proof** Let $\psi(d)$ be the number of elements of order $d$ mod $p$. We show:

(a) $\displaystyle\sum_{d|p-1} \phi(d) = p-1.$

(b) $\displaystyle\sum_{d|p-1} \psi(d) = p-1.$

(c) For any $d$, we have $\psi(d) \leq \phi(d)$.

From these equations it clearly follows that $\psi(d) = \phi(d)$ for all $d$ dividing $p-1$, and the theorem is proved.

Proof of (a): By Proposition 7.5, the number of integers $x$ with $\gcd(x, p-1) = (p-1)/d$ is $\phi(d)$. But every non-negative integer $y < p-1$ satisfies $\gcd(y, p-1) = (p-1)/d$ for some $d$. So the equation just counts all these integers; we know that the total is $p-1$.

Proof of (b): Every non-zero integer less than $p$ has some order which divides $p-1$; so this equation just counts them by their order, and again there are $p-1$ of them.

Proof of (c): Lemma 7.7 showed the stronger result that, for each $d$ dividing $p-1$, we have either $\psi(d) = 0$ or $\psi(d) = \phi(d)$. □

## 7.5 The Möbius function

Another number-theoretic function which is closely connected with Euler's totient $\phi$ is the Möbius function $\mu$.

A positive integer $n$ is *squarefree* if it is not divisible by any square greater than 1; that is, $m^2 \mid n$ implies $m = 1$. So $n$ is squarefree if and only if $n$ is a product of distinct primes. Now we define

$$\mu(n) = \begin{cases} (-1)^r & \text{if } n = p_1 p_2 \cdots p_r \text{, where } p_1, \ldots, p_r \text{ are distinct primes,} \\ 0 & \text{otherwise.} \end{cases}$$

The most important property of the Möbius function is the following result which is known as *Möbius inversion.* The sums in each of the two parts are over all divisors of the positive integer $n$.

**Theorem 7.9** *Let $f$ and $g$ be functions defined on the set of positive integers. Then the following are equivalent:*

*(a)* $g(n) = \sum_{m \mid n} f(m)$;

*(b)* $f(n) = \sum_{m \mid n} g(m)\mu(n/m)$.

The proof of this theorem requires a lemma.

**Lemma 7.10** *For any positive integer n, we have*

$$\sum_{m|n} \mu(m) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{otherwise.} \end{cases}$$

**Proof** If $n = 1$, then the sum contains a single term $\mu(1) = 1$.

Suppose that $n > 1$; let $n = p_1^{a_1} \cdots p_r^{a_r}$, where $p_1, \ldots, p_r$ are distinct primes and $a_1, \ldots, a_r > 0$. Since the Möbius function is zero on non-squarefree arguments, the sum in the lemma is over the squarefree divisors of $n$, which are the products of some of the primes $p_1, \ldots, p_r$. Now, if $m$ is the product of $k$ of these primes, then $\mu(m) = (-1)^k$; and there are $\binom{r}{k}$ (the binomial coefficient) ways to choose $k$ primes from the set of $r$. So

$$\sum_{m|n} \mu(m) = \sum_{k=0}^{r} \binom{r}{k} (-1)^k = (1-1)^r = 0,$$

where we have used the Binomial Theorem. □

**Proof of the Theorem** Suppose first that (b) holds. Call the sum in part (a) $S$. Then

$$S = \sum_{m|n} f(m) = \sum_{m|n} \left( \sum_{k|m} g(k) \mu(m/k) \right).$$

The sum is over all pairs $(m,k)$ with $m \mid n$ and $k \mid m$. Putting $m/k = l$, we may sum over all pairs $(k,l)$ where $k \mid n$ and $l \mid (n/k)$, to get

$$S = \sum_{k|n} g(k) \left( \sum_{l|(n/k)} \mu(l) \right).$$

By the Lemma, the inner sum is 1 if $n/k = 1$ and 0 otherwise. So the only term in the outer sum is the one with $k = n$, and we conclude that

$$S = g(n).$$

Now assume that (a) holds, and call the sum in part (b) $T$. We have

$$T = \sum_{m|n} g(m) \mu(n/m) = \sum_{m|n} \mu(n/m) \left( \sum_{k|m} f(k) \right).$$

Again put $l = m/k$ and sum over pairs $(k,l)$ with $k \mid n$ and $l \mid= (n/k)$, to obtains

$$T = \sum_{k|n} f(k) \left( \sum_{l|(n/k)} \mu(n/kl) \right).$$

Again the inner sum is zero unless $n/k = 1$ in which case it is 1, so

$$T = f(n).$$

Using this we have a formula for Euler's function:

**Theorem 7.11** *For a positive integer n,*

$$\phi(n) = \sum_{m|n} m\mu(n/m).$$

**Proof** We saw in part (a) in the proof of Theorem 7.8 that $n = \sum_{m|n} \phi(m)$ (We observed there that the argument did not depend on the fact that $n$ is of the form $p-1$ with $p$ prime.) Now apply Möbius inversion. □

**Remark** In combinatorics there is a much more general Möbius function, associated with an arbitrary partially ordered set. The number-theorists' Möbius function is a special case. See my *Notes on Counting* on the Web for this.

## 7.6 Appendix: An algebraic view

If you are familiar with the language of algebraic structures, some of the results of this chapter can be re-written in more algebraic terminology.

Recall that $\mathbb{Z}_n$ is the ring of integers modulo $n$. The units in this ring form a group (with the operation of multiplication). We denote this group by $U(n)$; it is a group of order $\phi(n)$. Now we can re-write the second part of Theorem 7.4 as follows:

**Theorem 7.12** *Let m and n be positive integers with* $\gcd(m,n) = 1$. *Then*

$$U(mn) \cong U(m) \times U(n).$$

Here the notation $A \cong B$ means that the groups $A$ and $B$ are isomorphic, and $A \times B$ denotes the direct product of groups $A$ and $B$, whose elements are the ordered pairs $(a,b)$ with $a \in A$ and $b \in B$, with pointwise multiplication. The proof of the theorem, using the Chinese Remainder Theorem, gives a bijection between $U(mn)$ and $U(m) \times U(n)$, and it is straightforward to show that this bijection is an isomorphism.

A group $A$ is *cyclic* if it contains an element $g$ such that every element of $G$ is a power of $a$. Theorem 7.8, on the existence of primitive roots, can be stated as follows:

**Theorem 7.13** *Let $p$ be prime. Then the group $U(p)$ is cyclic.*

What about the converse? It can be shown that the following is true (but I will not do so here):

**Theorem 7.14** *Let $n$ be a positive integer which is not a prime number. Suppose that $U(n)$ is a cyclic group (that is, there exists a primitive root of $n$). Then $n = p^a$ or $n = 2p^a$ for some odd prime $p$ and integer $a > 1$, or $n = 4$. Conversely, for these values of $n$, the group $U(n)$ is cyclic.*

One can go on and determine the structure of the abelian group $U(n)$ for every positive integer $n$.

*Carmichael's lambda-function* $\lambda(n)$ is defined to be the largest order of any element of $U(n)$. So we have $\lambda(n) = \phi(n)$ if and only if $U(n)$ is cyclic. This function has applications in cryptography, but we do not discuss it further here.

## 7.7 Appendix: Cryptography

Let $g$ be a primitive root of $p$. Then, for any non-zero element $h$ of $\mathbb{Z}_p$, there is an exponent $m$ in the range $0 \leq m \leq p-2$ such that $g^m = h$. However, finding $m$ is difficult, especially for large primes. If I asked you to find the number $m$ such that $2^m \equiv 6 \bmod 11$, you would probably have to resort to calculating powers of 2 mod 11 until 6 occurs.

This problem is known as the *discrete logarithm problem*, since we are in essence finding the logarithm of $h$ to base $g$ in the finite field $\mathbb{Z}_p$. Its difficulty is the basis for some of the earliest *public-key cryptosystems*, such as Diffie–Hellman key exchange and the El-Gamal cryptosystem. I will briefly describe the former of these.

Alice and Bob, who have never met, need to exchange a secret message. If they both possessed some random information which nobody else knew, they could use this as a key to encrypt and decrypt the message (for example, using a one-time pad). But they can only communicate over an insecure line, and if Alice sent Bob the key, then all the world would know it.

So they choose a (large) prime $p$ and a primitive root $g$ of $p$, and share these – so everyone knows $p$ and $g$. Then

- Alice chooses a random number $a$ in the range $0 \leq a \leq p-2$, calculates $g^a$ mod $p$, and sends this to Bob.

- Bob chooses a random number $b$ in the range $0 \leq b \leq p-2$, calculates $g^b$ mod $p$, and sends this to Alice.

- Then both Alice and Bob can compute $(g^a)^b = (g^b)^a$; they use this as their secret key.

Any interceptor is faced with the job of calculating $g^{ab}$ from $g^a$ and $g^b$. The obvious approach (and nothing better has been found) is to solve the discrete logarithm problem to find, for example, $a$ from $g^a$, and then do Alice's calculation $(g^b)^a$.

Thus it is the difficulty of the discrete logarithm problem that keeps the secret secure!

## Exercises

**7.1** (a) Show that $\phi(n)$ is even if $n > 2$.

(b) Find all integers $n$ satisfying $\phi(n) = 4$.

(c) For any positive integer $d$, show that there are only finitely many integers $n$ satisfying $\phi(n) = d$.

**7.2** (a) How many primitive roots of 17 are there?

(b) Find them all.

# Chapter 8

# Quadratic residues and non-residues

Let $p$ be an odd prime. In this section we are going to show how to decide whether the congruence

$$x^2 \equiv a \bmod p$$

has integer solutions, for any integer $a$ not divisible by $p$.

## 8.1 Definition and basic properties

First, a definition: we say that $a$ is a *quadratic residue* mod $p$ if this congruence has a solution, and a *quadratic non-residue* mod $p$ if it does not. Clearly, if $a \equiv b \bmod p$, then $a$ is a quadratic residue mod $p$ if and only if $b$ is a quadratic residue mod $p$; so we may (and often will) restrict our attention to $a = 1, \dots, p-1$.

**Example** Let $p = 7$. The squares mod 7 are

$$1^2 = 1, \quad 2^2 = 4, \quad 3^2 = 2, \quad 4^2 = 2, \quad 5^2 = 4, \quad 6^2 = 1;$$

so $1, 2, 4$ are quadratic residues and $3, 5, 6$ are non-residues.

**Proposition 8.1** *Of the $p-1$ numbers $1, 2, \dots, p-1$, half of them are quadratic residues and half are quadratic non-residues.*

**Proof** Let $g$ be a primitive root of $p$. (Recall that this means that $g$ has order $p-1$ mod $p$.) Then the $p-1$ numbers $g^0, g^1, \dots, g^{p-2}$ are all distinct, and so must be congruent to $1, 2, \dots, p-1$ in some order. We claim that $g^i$ is a quadratic residue if and only if $i$ is even. The result obviously follows.

If $i$ is even, say $i = 2j$, then $g^i \equiv (g^j)^2$ is a quadratic residue.

Conversely, suppose that $a = g^i$ is a quadratic residue, say $a = b^2$. Let $b \equiv g^j$. Then $g^i = g^{2j}$, so $i \equiv 2j \bmod p-1$. But $p-1$ and $2j$ are even; so $i$ must also be even. □

We learn something very important from the proof:

**Proposition 8.2** *Let $g$ be a primitive root of the odd prime $p$. Let $a$ be an integer not divisible by $p$. Then $a$ is a quadratic residue if and only if it is an even power of $g$, and is a quadratic non-residue if and only if it is an odd power of $g$.*

However, this is not a practical method. For both finding a primitive root $g$ of $p$, and expressing an arbitrary element of $\mathbb{Z}_p$ as a power of $g$, are hard problems. The second of these problems is the *discrete logarithm problem*, which we met in the last chapter in connection with cryptography; it is the difficulty of this problem which keeps information secure!

**Example** For $p = 7$, it can be checked that 3 is a primitive root. The powers of 3 mod 7 are

$$3^0 = 1, \quad 3^1 = 3, \quad 3^2 = 2, \quad 3^3 = 6, \quad 3^4 = 4, \quad 3^5 = 5.$$

The even powers of 3 are thus $1, 2, 4$, agreeing with what we found earlier.

## 8.2 The Legendre symbol

We now introduce some notation. The *Legendre symbol* $\left(\dfrac{a}{p}\right)$ is defined by

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{if } a \text{ is divisible by } p, \\ +1 & \text{if } a \text{ is a quadratic residue mod } p, \\ -1 & \text{if } a \text{ is a quadratic non-residue mod } p. \end{cases}$$

We give four very important rules which enable us to calculate the value of the Legendre symbol. (This is equivalent to deciding whether the congruence $x^2 \equiv a \bmod p$ has a solution. Actually finding a solution is quite a different matter!) We saw that the “hard” method based on the discrete logarithm problem actually finds a solution; this “easy” method does not.

**Theorem 8.3** *For any odd prime $p$ and integers $a$ and $b$, we have*

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right).$$

**Theorem 8.4** *For any odd prime p,*

$$\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2} = \begin{cases} +1 & \textit{if } p \equiv 1 \textit{ mod } 4, \\ -1 & \textit{if } p \equiv 3 \textit{ mod } 4. \end{cases}$$

**Theorem 8.5** *For any odd prime p,*

$$\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8} = \begin{cases} +1 & \textit{if } p \equiv 1 \textit{ or } 7 \textit{ mod } 8, \\ -1 & \textit{if } p \equiv 3 \textit{ or } 5 \textit{ mod } 8. \end{cases}$$

**Theorem 8.6** *For any two distinct odd primes p and q,*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4} = \begin{cases} -1 & \textit{if } p \equiv q \equiv 3 \textit{ mod } 4, \\ +1 & \textit{otherwise.} \end{cases}$$

The fourth rule is known as the Theorem of Quadratic Reciprocity.

We stop here to see why the last inequality is true in each case. Of course, $(-1)^k$ is equal to $+1$ if $k$ is even and $-1$ if $k$ is odd. Now

- $(p-1)/2$ is odd if and only if $p \equiv 3$ mod 4;
- $(p^2-1)/8$ is odd if and only if $p \equiv 3$ or 5 mod 8;
- $(p-1)(q-1)/4 = ((p-1)/2)((q-1)/2)$ is odd if and only if both $(p-1)/2$ and $(q-1)/2$ are odd.

We will prove these four rules in the remainder of the section. But first we will have an example of their use, and an unexpected application of the second rule.

**Example** Is 38 a quadratic residue mod 43?

We could answer this by computing squares mod 43. But our four rules give us a much quicker method:

$$\begin{aligned}
\left(\frac{38}{43}\right) &= \left(\frac{2}{43}\right)\left(\frac{19}{43}\right) && \text{(Rule 1)}\\
&= -\left(\frac{19}{43}\right) && \text{(Rule 3)}\\
&= \left(\frac{43}{19}\right) && \text{(Rule 4)}\\
&= \left(\frac{5}{19}\right) && (43 \equiv 5 \bmod 19)\\
&= \left(\frac{19}{5}\right) && \text{(Rule 4)}\\
&= \left(\frac{4}{5}\right) && (19 \equiv 4 \bmod 5)\\
&= +1 && (4 \equiv 2^2 \bmod 19),
\end{aligned}$$

so 38 is a quadratic residue mod 43.

Let us examine these steps more closely. The first is straightforward. In the second, $43 \equiv 3 \bmod 8$, so $\left(\frac{2}{43}\right) = -1$. In the third, 43 and 19 are both congruent to 3 mod 4, so the product $\left(\frac{19}{43}\right)\left(\frac{43}{19}\right)$ is $-1$; so the two Legendre symbols have opposite signs. The fourth step is straightforward. In the fifth, 19 and 5 are not both congruent to 3 mod 4, so the Legendre symbols have the same sign. The next step is straightforward, and in the last step we observe that 4 is a square.

There are other ways we could proceed. For example, after the second step,

$$\begin{aligned}
-\left(\frac{19}{43}\right) &= -\left(\frac{-24}{43}\right) && (19 \equiv -24 \bmod 43)\\
&= -\left(\frac{-1}{43}\right)\left(\frac{2}{43}\right)^2\left(\frac{6}{43}\right) && \text{(Rule 1)}\\
&= \left(\frac{6}{43}\right) && \text{(Rule 2)}\\
&= \left(\frac{49}{43}\right) && (6 \equiv 49 \bmod 43)\\
&= +1.
\end{aligned}$$

## 8.3 A Euclid-type theorem

We showed in Chapter 1 that there are infinitely many primes congruent to 3 mod 4, and deferred the proof in the other case. Now is the time to fulfil that promise.

**Theorem 8.7** *There are infinitely many primes congruent to* 1 *mod* 4.

**Proof** Again we argue by contradiction. Suppose that $p_1,\ldots,p_r$ were all the primes congruent to 1 mod 4. Now let

$$x = 2p_1p_2\cdots p_r, \qquad N = x^2+1.$$

Let $q$ be a prime divisor of $N$. Then $q$ is odd. We have $x^2 \equiv -1 \bmod q$, so $-1$ is a quadratic residue mod $q$. By Theorem 8.4, $q \equiv 1 \bmod 4$. Hence by assumption, $q$ must be one of the primes $p_1,\ldots,p_r$. But this is a contradiction, since $N$ leaves remainder 1 when divisible by each of these primes. □

## 8.4 Proofs of the first two rules

In this section we prove the first two rules for the Legendre symbol (Theorems 8.3–8.4).

**Proof of Rule 1** Let $g$ be a primitive root of $p$. By Proposition 8.2, every integer not divisible by $p$ is congruent to a power of $g$, with the squares congruent to even powers and the non-squares congruent to odd powers. Now the addition table for exponents translates into the multiplication table for the integers as follows:

| $+$ | even | odd |
|---|---|---|
| even | even | odd |
| odd | odd | even |

| $\times$ | square | non-square |
|---|---|---|
| square | square | non-square |
| non-square | non-square | square |

□

**Proof of Rule 2** Again let $g$ be a primitive root of $p$, and consider $z = g^{(p-1)/2}$. We have $z^2 = g^{p-1} \equiv 1 \bmod p$, but $z$ is not congruent to 1 mod $p$ (since if it were, the order of $g$ would be at most $(p-1)/2$); so $z \equiv -1 \bmod p$. So $-1$ is a quadratic residue or not depending on whether $(p-1)/2$ is even or odd, that is, on whether $p \equiv 1$ or $p \equiv 3 \bmod 4$. □

We can prove something a little more general:

**Proposition 8.8** *Let a be an integer not divisible by p. Then*

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \textit{ mod } p.$$

**Proof** Let $g$ be a primitive root of $p$, and $a \equiv g^i$. Since $g^{p-1} \equiv 1$, we have

$$a^{(p-1)/2} \equiv \begin{cases} 1 & \text{if } i \text{ is even,} \\ g^{(p-1)/2} & \text{if } i \text{ is odd.} \end{cases}$$

But we saw in the proof of Rule 2 that $g^{(p-1)/2} \equiv -1 \bmod p$. □

**Exercise** Use this Proposition to give another proof of Rule 1.

## 8.5 Proofs of Rules 3 and 4

The proofs here depend on a method invented by Gauss. We fix an odd prime $p$. Let $S = \{1, 2, \ldots, (p-1)/2\}$. Noting that

$$-(p-1)/2, \ldots, -2, -1, 0, 1, 2, \ldots, (p-1)/2$$

is a complete set of residues mod $p$, we see that any integer coprime to $p$ is congruent to either an element of $S$, or the negative of one.

Now take any integer $a$ not divisible by $p$. Then for any $s \in S$, the integer $as$ is congruent to an element of $s$ or tne negative of one; we write

$$as = e(a,s)t(a,s),$$

where $e(a,s) = \pm 1$ and $t(a,s) \in S$. For example, let $p = 7$, $a = 4$, $s = 3$. Then $as = 12 \equiv -2 \bmod 7$, so $e(4,3) = -1$ and $t(4,3) = 2$.

For any fixed $a$, consider the map $s \mapsto t(a,s)$. This map takes $S$ to itself. We claim that it is injective. For suppose that $t(a,s_1) = t(a,s_2)$. Then $as_1 \equiv \pm as_2 \bmod p$, so $p$ divides $a(s_1 \pm s_2)$. But since $p$ does not divide $a$, and no element of $S$ is congruent to plus or minus another element, we must have $s_1 = s_2$. Now an injective map of a finite set is bijective. So for fixed $a$, the elements $t(a,s)$ run through $S$ as $s$ does.

The heart of Gauss's method is the following result.

**Proposition 8.9** *With the above notation,*

$$\left(\frac{a}{p}\right) = \prod_{s \in S} e(a,s).$$

**Proof** We have

$$a^{(p-1)/2} \prod_{s \in S} s = \prod_{s \in S} as$$

$$\begin{aligned} &\equiv \left(\prod_{s\in S} e(a,s)\right)\left(\prod_{s\in S} t(a,s)\right) \\ &= \left(\prod_{s\in S} e(a,s)\right)\left(\prod_{s\in S} s\right). \end{aligned}$$

(In the last line we use the fact that the elements $t(a,s)$ run through all of $S$ as $s$ does.) Now $\prod_{s\in S} s$ is coprime to $p$ and can be cancelled; so we get

$$a^{(p-1)/2} \equiv \prod_{s\in S} e(a,s) \bmod p.$$

Now the result follows because

$$a^{(p-1)/2} \equiv \left(\frac{a}{p}\right) \bmod p$$

by Proposition 8.8. □

**Example** Let $p = 11$ and $a = 3$. Then

$$\begin{array}{lll} 3\cdot 1 = +3 & \text{so} & e(3,1) = +1 \\ 3\cdot 2 = -5 & \text{so} & e(3,2) = -1 \\ 3\cdot 3 = -2 & \text{so} & e(3,3) = -1 \\ 3\cdot 4 = +1 & \text{so} & e(3,4) = +1 \\ 3\cdot 5 = +4 & \text{so} & e(3,5) = +1 \end{array}$$

So $\left(\frac{3}{11}\right) = +1$. Indeed, $3 \equiv 5^2 \bmod 11$.

Before going on to the proofs, let us note that Rule 2 follows very easily from this. Multiplying $S$ by $-1$ takes each element to its negative. So $e(-1,s) = -1$ for all $s \in S$, and

$$\left(\frac{-1}{p}\right) = (-1)^{|S|} = (-1)^{(p-1)/2}.$$

**Proof of Rule 3** We split into cases depending on the congruence of $p$ mod 8:

$p = 8k+1$: Multiplying $s$ by 2 gives

$$2,4,\ldots,4k,4k+2 = -(4k-1),\ldots,8k = -1;$$

there are $2k$ positive and $2k$ negative terms, so the product is $+1$.

$p = 8k+3$: Multiplying $S$ by 2 gives

$$2,4,\dots,4k,4k+2 = -(4k+1),\dots,8k+2 = -1;$$

there are $2k$ positive and $2k+1$ negative terms, so the product is $-1$.

$p = 8k+5$: Multiplying $S$ by 2 gives

$$2,4,\dots,4k+2,4k+4 = -(4k+1),\dots,8k+4 = -1;$$

there are $2k+1$ positive and $2k+1$ negative terms, so the product is $-1$.

$p = 8k+7$: Multiplying $S$ by 2 gives

$$2,4,\dots,4k+2,4k+4 = -(4k+3),\dots,8k+6 = -1;$$

there are $2k+1$ positive and $2k+2$ negative terms so the product is 1. □

Gauss gave many different proofs of Rule 4, the Law of Quadratic Reciprocity; we will make do with just one. We need a lemma:

**Lemma 8.10** *Let $q$ be an odd integer not divisible by the odd prime $p$. Then*

$$e(q,s) = (-1)^{\lfloor t(2,s)q/p \rfloor}.$$

**Proof** We have $sq \equiv e(q,s)t(q,s) \bmod p$, so

$$sq = kp + e(q,s)t(q,s)$$

for some integer $k$. Hence

$$\frac{2sq}{p} = 2k + e(q,s)\frac{2t(q,s)}{p}.$$

Now $t(q,s) \in S = \{1,\dots,(p-1)/2\}$; so $0 < 2t(q,s) < p$. So the second term in the above equation is a fraction between 0 and 1. Hence

$$\left\lfloor \frac{2sq}{p} \right\rfloor = \begin{cases} 2k & \text{if } e(q,s) = +1, \\ 2k-1 & \text{if } e(q,s) = -1, \end{cases}$$

while

$$\left\lfloor -\frac{2sq}{p} \right\rfloor = \begin{cases} -2k-1 & \text{if } e(q,s) = +1, \\ -2k & \text{if } e(q,s) = -1. \end{cases}$$

Also, $2s \in \{2,\dots,p-1\}$, so either $2s = t(2,s)$ or $2s = p - t(2,s)$. We treat the two cases separately.

If $2s = t(2,s)$, then we see that $\lfloor t(2,s)q/p \rfloor$ is even if $e(q,s) = +1$ and odd if $e(q,s) = -1$, so the conclusion of the lemma is true.

If $2s = p - t(2,s)$, then

$$\frac{2sq}{p} = q - \frac{t(2,s)q}{p},$$

so the argument applies with the parities reversed (here we use that $q$ is odd). □

**Proof of Rule 4** Let $p$ and $q$ be distinct odd primes. Put $S_p = \{1,\ldots,(p-1)/2\}$ and $S_q = \{1,\ldots,(q-1)/2\}$. We have

$$\begin{aligned}\left(\frac{p}{q}\right) &= \prod_{s\in S_q} e(p,s)\\ &= \prod_{s\in S_q}(-1)^{\lfloor t(2,s)p/q\rfloor}\\ &= (-1)^{\sum\limits_{s\in S_q}\lfloor t(2,s)p/q\rfloor}.\end{aligned}$$

Now, as $s$ takes the $(q-1)/2$ distinct values in $S_q$, then $t(2,s)$ also takes these values once each; so

$$\left(\frac{p}{q}\right) = (-1)^{\sum\limits_{s\in S_q}\lfloor sp/q\rfloor}.$$

Similarly

$$\left(\frac{q}{p}\right) = (-1)^{\sum\limits_{t\in S_p}\lfloor tq/p\rfloor}.$$

So we have

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\Sigma_1+\Sigma_2},$$

where $\Sigma_1$ and $\Sigma_2$ are the summations in the two preceding formulae.

We are going to show that $\Sigma_1+\Sigma_2 = (p-1)(q-1)/4$.

Let

$$T = \{(sp,tq) : 1 \le s \le (q-1)/2, 1 \le t \le (p-1)/2\}.$$

Clearly $|X| = (p-1)(q-1)/4$. Also we can divide $T$ into two parts $T_1$ and $T_2$ by

$$T_1 = \{(sp,tq) \in T : sp > tq\}, \qquad T_2 = \{(sp,tq) \in T : sp < tq\}.$$

(We cannot have $sp = tq$, since this would imply that $q$ divides $s$, contradicting $1 \le s \le (q-1)/2$.)

Suppose that $(sp,tq) \in T_1$, so that $sp > tq$. Then $1 \le t \le \lfloor sp/q\rfloor$, so there are $\lfloor sp/q\rfloor$ points in $T_1$ with a given first coordinate $sp$. So the sum $\Sigma_1$ is precisely the number of points in $T_1$.

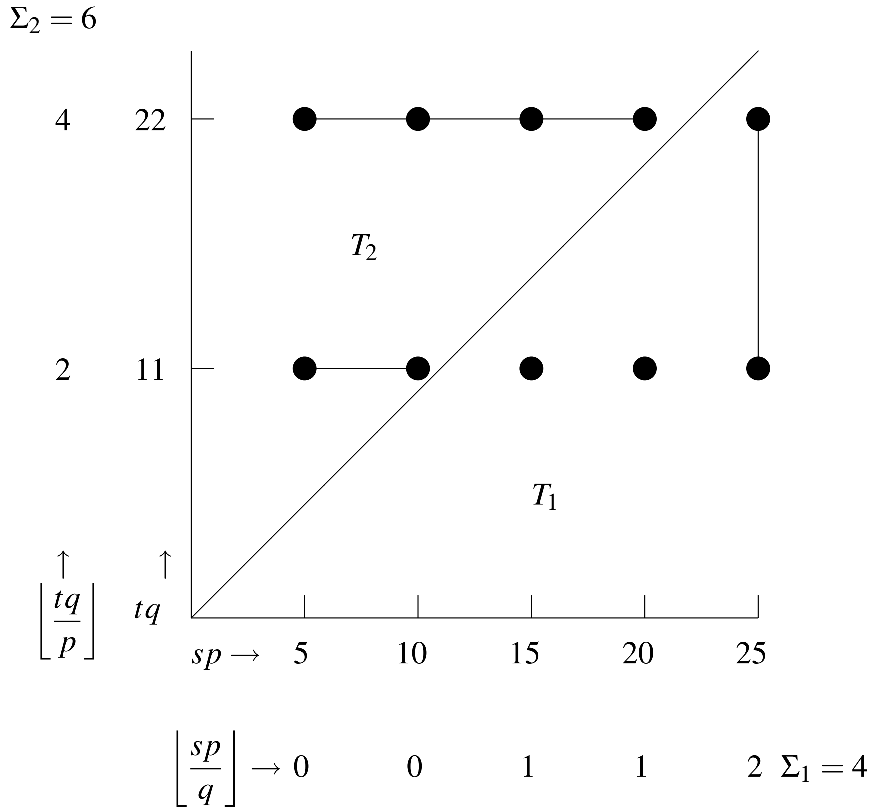
Similarly $\Sigma_2$ is the number of points in $T_2$. So altogether we have $\Sigma_1+\Sigma_2 = |T| = (p-1)(q-1)/4$.

We conclude that

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4},$$

and the theorem is proved. □

Here is a diagram illustrating the case $p = 5$, $q = 11$; so $1 \le s \le (11-1)/2 = 5$ and $1 \le t \le (5-1)/2 = 2$.


$\Sigma_2 = 6$
4
22
2
11
$T_2$
$T_1$
↑
$\left\lfloor \frac{tq}{p} \right\rfloor$
↑
$tq$
$sp \to$
5
10
15
20
25
$\left\lfloor \frac{sp}{q} \right\rfloor \to$ 0
0
1
1
2
$\Sigma_1 = 4$

We have $|T_1| = 4$, $|T_2| = 6$, and

$$|T| = 4 + 6 = 10 = \frac{5-1}{2} \times \frac{11-1}{2}.$$

## Exercises

**8.1** Calculate the following Legendre symbols:

(a) $\left(\frac{36}{109}\right)$

(b) $\left(\frac{26}{109}\right)$

(c) $\left(\dfrac{7}{103}\right)$

**8.2** (a) Let $p$ be an odd prime. Show that there is an integer $x$ such that $\left(\dfrac{x}{p}\right) = +1$ and $\left(\dfrac{x+1}{p}\right) = -1$.

(b) Let $p = 71$. Find the smallest positive integer which is a quadratic non-residue mod $p$.

**8.3** Let $p_1, \ldots, p_r$ be odd primes, and let $q$ be a prime divisor of $(p_1 \cdots p_r)^2 - 2$. Show that $\left(\dfrac{2}{q}\right) = +1$, and deduce that $q \equiv \pm 1 \bmod 8$.

Hence show that there are infinitely many primes $p$ satisfying $p \equiv \pm 1 \bmod 8$.

# Chapter 9

# Sums of squares

In this chapter we are going to decide which integers can be written as the sum of two squares, or the sum of four squares.

## 9.1 Sums of two squares

In Chapter 6, we found which primes can be written as the sum of two integer squares. Now we will extend this to arbitrary positive integers.

Let $n$ be any positive integer. Then we can write $n = a^2b$ where $a, b$ are positive integers and $b$ is squarefree. (Write down the prime factorisation of $n$. Let $b$ be the product of all the primes which occur to an odd power in the factorisation. Then $b$ is squarefree, and $n/b$ has all its prime factors occurring to an even power, so $n/b$ is a square.) For example,

$$1440 = 2^5 3^2 5 = 12^2 \cdot 10,$$

where 10 is squarefree.

In the above representation, we call $b$ the *squarefree part* of $n$.

**Theorem 9.1** *The positive integer $n$ is the sum of two squares of integers if and only if the squarefree part of $n$ has no prime factors congruent to* 3 *mod* 4.

In the example, $10 = 2 \cdot 5$ has no prime factor congruent to 3 mod 4, so 1440 is the sum of two squares. Indeed

$$1440 = 12^2 \cdot 10 = 12^2(3^2 + 1^2) = 36^2 + 12^2.$$

**Proof** First we have to show that every number which satisfies this condition can be written as the sum of two squares. This uses the following fact. Suppose that two numbers $a$ and $b$ are each the sum of two squares. Then so is their product. For, if $a = x^2 + y^2$ and $b = u^2 + v^2$, then

$$ab = (x^2+y^2)(u^2+v^2) = (xu-yv)^2 + (xv+yu)^2,$$

as is easily verified. (The fact that

$$(x^2+y^2)(u^2+v^2) = (xu-yv)^2 + (xv+yu)^2$$

is called the *two-squares identity*.)

Now any number satisfying the conditions of the theorem is a product of factors of the following types: a square; the prime 2; and primes congruent to 1 mod 4. Now all of these are sums of two squares: $a^2 = a^2 + 0^2$; $2 = 1^2 + 1^2$; and the conclusion for primes congruent to 1 mod 4 was shown in Theorem 6.8. So the product of such numbers is the sum of two squares.

For example, $340 = 2^2 \cdot 5 \cdot 17$, and $5 = 1^2 + 2^2$, $17 = 1^2 + 4^2$. We have

$$85 = (1^2+2^2)(1^2+4^2) = 9^2 + 2^2,$$

and

$$340 = (2^2+0^2)(9^2+2^2) = 18^2 + 4^2.$$

Now we turn to the converse. Suppose that $n = x^2 + y^2$. We have to show that no prime congruent to 3 mod 4 divides the squarefree part of $n$; in other words, if $p$ is a prime congruent to 3 mod 4, then the power of $p$ which divides $n$ is even. Our proof will be by induction on $n$. Clearly $n = 1$ has no prime divisors at all, so the induction starts. So suppose that the result is true for all numbers less than $n$.

Suppose that $p$ divides $n$, where $p \equiv 3 \bmod 4$. We claim that $p$ divides both $x$ and $y$. For suppose not. Then $x^2 + y^2 \equiv 0 \bmod p$. If $p$ does not divide $x$, then there is an inverse $z$ of $x$ mod $p$; and $(xz)^2 + (yz)^2 \equiv 0 \bmod p$. But $xz \equiv 1 \bmod p$; so $(yz)^2 \equiv -1 \bmod p$, whence $\left(\dfrac{-1}{p}\right) = +1$, contradicting Rule 2. So the claim is proved.

Now write $x = pu$ and $y = pv$; then $n = p^2(u^2 + v^2)$, so $p^2$ divides $n$, and $n = p^2 m$, where $m = u^2 + v^2$. By the induction hypothesis, the power of $p$ dividing $m$ is even; so the same is true for $n$, and we are done. □

## 9.2 Sums of four squares

In this section, as a companion piece, we will prove a theorem of Lagrange:

**Theorem 9.2** *Every positive integer can be written as the sum of four squares of integers.*

**Proof** In the two-squares theorem we used the identity

$$(a^2+b^2)(x^2+y^2) = (ax-by)^2+(ay+bx)^2.$$

There is a similar identity for four squares:

$$\begin{aligned}&(a^2+b^2+c^2+d^2)(w^2+x^2+y^2+z^2)\\ &\quad = (aw-bx-cy-dz)^2+(ax+bw+cz-dy)^2\\ &\qquad\quad +(ay-bz+cw+dx)^2+(az+by-cx+dw)^2.\end{aligned}$$

This can be proved just by multiplying it out.

Because of the four-squares identity, we see that if two numbers can be written as the sum of four squares, then so can their product. So to prove that every positive integer can be so written, it is enough to show that every prime number is the sum of four squares. We do this by dividing the primes into three classes:

**The prime** 2 We have $2 = 1^2+1^2+0^2+0^2$.

**Primes congruent to** 1 **mod** 4 By Theorem 6.8, these primes can be written as the sum of two squares, say $p = x^2+y^2$. So we simply put

$$p = x^2+y^2+0^2+0^2.$$

**Primes congruent to** 3 **mod** 4 This case is the hardest. We proceed in two steps. Let $p$ be a prime with $p \equiv 3 \bmod 4$.

**Step 1:** We can find a positive integer $r$ such that $rp$ is a sum of four squares.

Of the numbers $1,\ldots,p-1$, half are quadratic residues (including 1) and half are non-resudues. Let $a+1$ be the smallest quadratic non-residue. Since $-1$ is also a quadratic non-residue, we see that $a$ and $-(a+1)$ are quadratic residues: say $x^2 \equiv a$, $y^2 \equiv -(a+1) \bmod p$. Then $0^2+1^2+x^2+y^2 \equiv 0 \bmod p$; so $0^2+1^2+x^2+y^2 = rp$ for some positive integer $r$.

**Step 2:** If $rp$ is the sum of four squares and $r > 1$, then there is an integer $s$ with $0 < s < r$ such that $sp$ is a sum of four squares.

For suppose that $rp = a^2+b^2+c^2+d^2$ with $r > 1$, and suppose for a contradiction that no smaller multiple of $p$ is the sum of four squares. We may assume that $-p/2 < a,b,c,d < p/2$. For changing $a$ by a multiple of $p$ does not change the fact that the sum of squares is a multiple of $p$; and there is a unique element $a'$ of the congruence class of $a$ with smallest modulus, satisfying $-p/2 < a' < p/2$.

Replacing $a$ by $a'$ would reduce the sum of squares, which by assumption is not possible. Similarly for $b,c,d$. Then

$$rp = a^2+b^2+c^2+d^2 < 4(p/2)^2 = p^2,$$

so $r < p$.

Now there is a unique number $w$ with $-r/2 < w \leq r/2$ and $w \equiv a \bmod r$. Define $x,y,z$ similarly but with a twist: for example, $-r/2 < x \leq r/2$ and $x \equiv -b \bmod 4$. Then

$$w^2+x^2+y^2+z^2 \equiv a^2+b^2+c^2+d^2 \equiv 0 \bmod r,$$

so $w^2+x^2+y^2+z^2 = rs$ for some integer $s$. The inequalities on $w,x,y,z$ show that

$$rs = w^2+x^2+y^2+z^2 \leq 4(r/2)^2 = r^2,$$

so $s \leq r$.

Could equality hold? If so then $w = x = y = x = r/2$, and so $a,b,c,d$ are all congruent to $r/2 \bmod r$. But then

$$rp = a^2+b^2+c^2+d^2 \equiv 4(r^2/4) \equiv 0 \bmod r^2,$$

so $r^2$ divides $rp$, and $r$ divides $p$; a contradiction since $p$ is prime and $r < p$. So in fact $s < r$.

Now we have $rp$ and $sr$ both written as the sum of four squares. The four-squares identity now expresses their product $r^2sp$ as the sum of four squares. Looking more closely at the identity, we observe that each term is a multiple of $r$. For recall that $w \equiv a$, $x \equiv -b$, $y \equiv -c$, and $z \equiv -d \bmod r$; so

$$wa - xb - yc - zd \equiv a^2+b^2+c^2+d^2 \equiv 0 \bmod r,$$

and

$$wb + xa + yd - zc \equiv ab - ab - cd + cd \equiv 0 \bmod r,$$

with a similar calculation for the other two terms. So

$$sr^2p = (re)^2 + (rf)^2 + (rg)^2 + (rh)^2,$$

and cancelling $r^2$ we have written $sp$ as the sum of four squares.

This completes the proof. □

**Example** Let us apply the method of proof to the prime $p = 7$. The smallest quadratic non-residue is 3, and $-3 \equiv 2^2$, $2 \equiv 3^2$. We find that

$$0^2 + 1^2 + 2^2 + 3^2 = 14 = 2 \cdot 7.$$

Reducing mod 2, we have

$$0^2 + 1^2 + 0^2 + 1^2 = 2 = 1 \cdot 2.$$

The four-squares identity gives (ignoring minus signs)

$$4^2 + 2^2 + 2^2 + 2^2 = 1 \cdot 2^2 \cdot 7,$$

and cancelling $2^2$ gives $2^2 + 1^2 + 1^2 + 1^2 = 7$.

## 9.3 Two squares revisited

We used Legendre's continued fraction method to express a prime congruent to 1 mod 4 as a sum of two squares. This can also be done by Lagrange's method, similar to the argument in the four-squares proof. The two steps are almost the same as in that proof.

**Step 1:** There is a positive number $r < p$ such that $rp$ is the sum of two squares. For the congruence of $p$ mod 4 shows that $\left(\dfrac{-1}{p}\right) = +1$, so that there is a positive integer $x$ (less than $p/2$) satisfying $x^2 \equiv 1$ mod $p$; that is, $rp = x^2 + 1^2$.

**Step 2:** If $rp$ is the sum of two squares, with $r > 1$, then there exists $s < r$ such that $sp$ is the sum of two squares. For take the equation

$$rp = a^2 + b^2$$

and reduce mod $r$ to get

$$rs = x^2 + y^2,$$

where $x, y$ are congruent to $a$ and $-b$ respectively mod $r$, and $|x|, |y| \leq r/2$ (so that $s < r$). Use the two-squares identity to get

$$r^2 sp = (a^2 + b^2)(x^2 + y^2) = (ax - by)^2 + (ay + bx)^2.$$

Now both $ax - by$ and $ay + bx$ are divisible by $r$. (For, modulo $r$, we have $x \equiv a$ and $y \equiv -b$, and so $ax - by \equiv a^2 + b^2 \equiv 0$ and $ay + bx \equiv -ab + ab = 0$). Say

$ax-by=ru$ and $ay+bx=rv$; then we can divide the last displayed equation by $r^2$ to get

$$sp=u^2+v^2.$$

Now repeating this process, we must eventually reach an expression for $p$ as the sum of two squares.

**Example:** $p=29$.

First let's remember how we expressed $p$ as the sum of two squares. We calculated the continued fraction for $\sqrt{p}$, expressing the intermediate quantities in the calculation as $y_m=(P_m+\sqrt{p})/Q_m$. For the assumed congruence on $p$, the period of the continued fraction will be odd, say $2k+1$, and when we reach the point just past half-way, we will have the required expression: $P_{k+1}^2+Q_{k+1}^2=p$.

For $p=29$, we have

$$\begin{aligned} a_0=\lfloor\sqrt{29}\rfloor=5, &\qquad y_1=\frac{1}{\sqrt{29}-5}=\frac{\sqrt{29}+5}{4}\\ a_1=\lfloor y_1\rfloor=2, &\qquad y_2=\frac{4}{\sqrt{29}-3}=\frac{\sqrt{29}+3}{5}\\ a_2=\lfloor y_2\rfloor=1, &\qquad y_3=\frac{5}{\sqrt{29}-2}=\frac{\sqrt{29}+2}{5} \end{aligned}$$

and we have $29=2^2+5^2$.

You should complete the calculation to show that $\sqrt{29}=[5;\overline{2,1,1,2,10}]$.

Now let us apply Lagrange's method. First, since $29\equiv 1 \bmod 4$, we know that $\left(\frac{-1}{29}\right)=+1$, so there is a solution of $x^2+1\equiv 0 \bmod 29$. By trial, we find that

$$12^2+1^2=5\cdot 29.$$

Reducing this equation mod 5, we have

$$2^2+1^2=5\cdot 1.$$

By the two-squares identity, multiplying the two expressions gives

$$5^2\cdot 1\cdot 29=(12^2+1^2)(2^2+1^2)=25^2+10^2,$$

so $29=5^2+2^2$.

## 9.4 Sums of three squares

You might wonder: do we really need four squares? Two are not enough, what about three? Since squares are congruent to 0, 1 or 4 mod 8, no sum of three squares can be congruent to 7 mod 8. Moreover, if $4n$ is the sum of three squares, then each of the squares must be even (three or fewer odd squares cannot add up to a multiple of 4, since squares are congruent to 0 or 1 mod 4), so $n$ is also such a sum. This proves the easy direction in the following theorem:

**Theorem 9.3** *Every positive integer can be written a the sum of three squares of integers except for those of the form $4^a(8b+7)$ for $a,b \geq 0$.*

But this is more difficult to prove, and we will not give the proof.

## 9.5 Where do these identities come from?

You might recognise that the two-squares identity has something to do with the complex numbers. We have

$$|a+b\mathrm{i}|^2 = a^2+b^2,$$

and the two-squares identity

$$(a^2+b^2)(x^2+y^2) = (ax-by)^2+(ay+bx)^2$$

just says that $|z_1|^2|z_2|^2 = |z_1z_2|^2$, since if $z_1 = a+b\mathrm{i}$ and $z_2 = x+y\mathrm{i}$ then

$$z_1z_2 = (ax-by)+(ay+bx)\mathrm{i}$$

(using the fact that $\mathrm{i}^2 = -1$).

The four-squares identity comes in the same way from another number system, the *quaternions*. A quaternion has the form

$$a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k},$$

where $a,b,c,d$ are real numbers; the units satisfy the multiplication rules

$$\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = -1$$

and

$$\mathrm{ij} = -\mathrm{ji} = \mathrm{k}, \quad \mathrm{jk} = -\mathrm{kj} = \mathrm{i}, \quad \mathrm{ki} = -\mathrm{ik} = \mathrm{j}.$$

The norm of the quaternion $z = a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k}$ is given by

$$|z|^2 = a^2+b^2+c^2+d^2,$$

and it is an exercise to show that the four-squares identity once again expresses the fact that the norm is multiplicative, that is, $|z_1|^2|z_2|^2 = |z_1z_2|^2$.

There is a similar eight-squares identity, derived from the multiplication on an eight-dimensional number system called the *octonions* or *Cayley numbers* (actually invented by Graves). More surprisingly, the pattern does not continue; there is no similar sixteen-squares identity, and indeed no identity for any other number of squares except one, two, four and eight.

## 9.6 Pythagoras and Fermat

Every perfect square is the sum of two squares: $x^2 = x^2+0^2$. Indeed, this was one of the "base cases" that we used in proving Theorem 9.1 determining those positive integers which are sums of two squares. But things are very different if we ask which perfect squares are the sum of two squares of positive integers. The smallest such is 25: $3^2+4^2 = 5^2$.

The equation $x^2+y^2 = z^2$ is associated with Pythagoras. Not only did he prove his famous theorem asserting that this holds if $z$ is the hypotenuse of a right-angled triangle and $x$ and $y$ are the other two sides, but he also gave a rule for finding all the solutions of this equation in positive integers. (The famous solution $3^2+4^2 = 5^2$ gives a right-angled triangle which had been used by surveyors since before the time of Pythagoras. Take a loop of string with twelve equally-spaced knots. Taking hold of the appropriate knots and pulling the string tight gives a right angle.)

**Theorem 9.4** *Let $x$, $y$ and $z$ be positive integers satisfying $x^2+y^2 = z^2$. Then there are positive integers $d,s,t$ with* $\gcd(s,t) = 1$*, such that, after interchanging $x$ and $y$ if necessary, we have*

$$x = 2std, \quad y = (s^2-t^2)d, \quad z = (s^2+t^2)d.$$

Note that $s = 2$, $t = 1$, $d = 1$ gives the solution $(x,y,z) = (3,4,5)$.

**Proof** We use the following principle: if $ab = c^2$ and $\gcd(a,b) = 1$, then each of $a$ and $b$ is a square. For any prime divisor of $ab$ occurs to an even power, and must occur in one of $a$ and $b$ and not the other; so each of $a$ and $b$ is a product of even powers of primes, and so is a square. More generally, if the product of any number of pairwise coprime factors is a square, then each factor is a square.

In a similar way, if $ab = c^2$ where $\gcd(a,b) = 2$, then $a = 2m^2$ and $b = 2n^2$ for some integers $m,n$. (Just replace $a$ and $b$ by $a/2$ and $b/2$; these are coprime and their product is $(c/2)^2$, so each is a square.)

Now suppose that $x^2 + y^2 = z^2$.

- any common factor of two of $x,y,z$ must divide the third, and we can divide through by it and get a smaller solution. So we can assume that $x,y,z$ are pairwise coprime, by dividing by their gcd (say $d$).

- Since squares are congruent to 0 or 1 mod 4, the only solutions to our equation mod 4 are $0+0 = 0$ and $0+1 = 1$. Since the variables are pairwise coprime, we can assume the latter: that is, (swapping $x$ and $y$ if necessary) $x$ is even, $y$ and $z$ are odd.

- Now $x^2 = z^2 - y^2 = (z+y)(z-y)$, and $\gcd(z+y)(z-y) = 2$. (For both $z+y$ and $z-y$ are even, so there is a common factor 2; and if $d$ is the gcd, then $d$ divides both $(z+y)+(z-y) = 2z$ and $(z+y)-(z-y) = 2y$, so $d = 2$.) So $z-y = 2s^2$ and $z+y = 2t^2$ for some (coprime) integers $s$ and $t$. Now

- 
  - ○ $x^2 = 4s^2t^2$, so $x = 2st$;
  - ○ $y = ((z+y)-(z-y))/2 = s^2 - t^2$;
  - ○ $z = ((z+y)+(z-y))/2 = s^2 + t^2$. □

In the seventeenth century, Pierre de Fermat wrote a note in the margin of his copy of the book on number theory by the Greek mathematician Diophantus. The note was opposite the place where Diophantus gave the preceding theorem of Pythagoras. Fermat claimed that he had a "truly wonderful" proof that, for any $n > 2$, the equation $x^n + y^n = z^n$ has no solution in positive integers, but the margin where he was writing was too small to contain it.

Mathematicians took up the challenge of trying to find the proof of what became known, ironically, as Fermat's Last Theorem. Finally in the 1990s, Andrew Wiles succeeded in finding a proof. But his proof was very long and complicated, and used many concepts which had not been invented in Fermat's time. Moreover, no evidence of such a proof was ever found in Fermat's papers. It is generally believed now that he didn't have a proof; perhaps he thought he had one but it contained a mistake.

We certainly cannot prove Wiles' Theorem here. But as an illustration, we prove a simple case:

**Theorem 9.5** *The equation $x^4 + y^4 = z^4$ has no solution in positive integers $x,y,z$.*

**Proof** We actually consider a slightly different equation, namely $x^4+y^4=z^2$. If we show that this equation has no solution, then the equation of the theorem has no solution either, since if $x,y,z$ satisfy the equation in the theorem, then $x,y,z^2$ satisfy the modified equation.

Suppose that $x^4+y^4=z^2$, where $x,y,z$ are positive integers. We may suppose that this is the solution with the smallest possible value of $z$. Then $x,y,z$ are pairwise coprime, since if there were a common prime factor $p$ of any two of them it would divide the third and we could replace the solution $(x,y,z)$ by $(x/p,y/p,z/p^2)$. So one of $x^2$ and $y^2$ is even; without loss of generality, $x$ is even.

Since $(x^2)^2+(y^2)^2=z^2$, we can apply Pythagoras to conclude that

$$x^2=2st, \quad y^2=s^2-t^2, \quad z=s^2+t^2,$$

where $\gcd(s,t)=1$.

Applying Pythagoras to the equation $t^2+y^2=s^2$ (remembering that $y$ is odd), we have

$$t=2uv, \quad y=u^2-v^2, \quad s=u^2+v^2,$$

where $\gcd(u,v)=1$. It follows that $\gcd(u,u^2+v^2)=\gcd(v,u^2+v^2)=1$. Then $x^2=2st=4uv(u^2+v^2)$, so that $uv(u^2+v^2)$ is a square. Since the factors are pairwise coprime, we have $u=m^2$, $v=n^2$, and $u^2+v^2=r^2$. Thus

$$m^4+n^4=r^2.$$

But $r\le u^2+v^2=s<s^2+t^2=z$, so we have $(m,n,r)$ is a solution of the original equation smaller than the solution $(x,y,z)$, which we assumed to be the smallest. This contradiction shows that no solution can exist. □

## 9.7 Open problems

Just to show that we don’t know everything, here are three problems which are still unsolved despite a lot of effort from many mathematicians:

**Goldbach’s Conjecture:** Every even number greater than 2 is the sum of two prime numbers.

The conjecture is known to be true for all “small” even numbers (less than $10^{18}$).

Note that we can decide whether $n$ is the sum of two primes in a finite amount of time: we only have to check the numbers $a$ with $1\le a\le n/2$ to see whether $a$ and $n-a$ are prime. By contrast, if you conjecture instead that any even number is the difference of two primes, you do not know *a priori* how long it will take to check a given value. Of course, 2 is the difference of two primes: $2=5-3$. But a famous related problem is currently unsolved:

**The twin-primes conjecture:** There are infinitely many pairs of primes differing by 2.

The last of the three problems has a different flavour.

**The congruent number problem:** Decide for which positive integers $n$ there exists a right-angled triangle of area $n$ with all sides rational.

The numbers $1,2,3,4$ are not congruent, but $5,6,7$ are. 157 is a congruent number, but the "simplest" right-angled triangle with rational sides and area 157 has hypotenuse

$$\frac{224403517704336969924557513090674863160948472041}{8912332268928859588025535178967163570016480830}.$$

Andrew Wiles, who proved Fermat's Last Theorem, has said that the congruent number problem is even harder!

Here is what Wiles had to say about doing mathematical research, in an interview with Simon Singh for the Horizon program about Fermat's Last Theorem:

> Perhaps I can best describe my experience of doing mathematics in terms of a journey through a dark unexplored mansion. You enter the first room of the mansion and it's completely dark. You stumble around bumping into furniture, but gradually you learn where each piece of furniture is. Finally after six months or so, you find the light switch, you turn it on, and suddenly it's all illuminated. You can see exactly where you were. Then you move into the next room and spend another six months in the dark. So each of these breakthroughs, while sometimes they're momentary, sometimes over a period of a day or two, they are the culmination of – and couldn't exist without – the many months of stumbling around in the dark that preceded them.

Look at `http://www.maths.qmul.ac.uk/~pjc/comb/quotes.html#work` for more quotes by mathematicians about how they make their discoveries.

## 9.8 Appendix: an algebraic proof

The fact that a prime congruent to 1 mod 4 is a sum of two squares can be proved in many different ways; we have already seen two. Here is a third, which depends on algebraic properties of a certain ring.

A *Gaussian integer* is a number of the form $a+b\mathrm{i}$, where $a,b \in \mathbb{Z}$. The Gaussian integers form a ring $R$; you may have learnt in Algebraic Structures I that this ring is a *principal ideal domain*. You don't need to know the definition of this; but it implies that, if $p$ is prime and $p$ divides $ab$, then either $p$ divides $a$ or $p$ divides $b$.

Let $p$ be a prime congruent to 1 mod 4. We know that $\left(\frac{-1}{p}\right) = +1$, so there is an integer $x$ such that $p$ divides $x^2 + 1 = (x+\mathrm{i})(x-\mathrm{i})$.

Suppose that $p$ is prime in $R$. Then $p$ must divide one of the factors $x \pm \mathrm{i}$, which is impossible, since the quotient $x/p \pm \mathrm{i}/p$ is not a Gaussian integer.

So $p$ is composite in $R$, say $p = (a+b\mathrm{i})(c+d\mathrm{i})$. Taking the complex conjugate gives $p = (a-b\mathrm{i})(c-d\mathrm{i})$. Multiplying these two equations, we obtain the equation $p^2 = (a^2+b^2)(c^2+d^2)$. This is an equation in the integers; since $p$ is prime and neither factor on the right is equal to 1, we must have $p = a^2 + b^2$ (and also $p = c^2 + d^2$) – that is, $p$ is the sum of two squares.

## Exercises

**9.1** Which of the following numbers can be written as the sum of two squares? Give such an expression if it exists, and explain why not if not.

(a) 120

(b) 720

(c) 8633

**9.2** A *triangular number* is a number of the form $n(n+1)/2$, for $n \geq 0$. Express each integer between 10 and 20 inclusive as a sum of three triangular numbers.

# Chapter 10

# Quadratic forms

A *quadratic form* over $\mathbb{Z}$ in the variables $x_1,\dots,x_n$ is an expression of the form

$$f(x_1,\dots,x_d) = \sum_{1\le i\le j\le d} a_{ij}x_ix_j,$$

where the coefficients $a_{ij}$ are integers. If there are $d$ variables, we call it a $d$-ary quadratic form. For $d = 2,3,4$ we use the terms *binary*, *ternary* and *quaternary*. In this chapter we are only concerned with binary quadratic forms.

Given an $d$-ary quadratic form $f$ and an integer $n$, do there exist integers $x_1,\dots,x_d$ such that $f(x_1,\dots,x_d) = n$? If so, we say that the integer $n$ is *represented* by the form $f$. We are interested in the question:

> Which integers are represented by a given quadratic form?

Note that $f(0,0,\dots,0) = 0$, so any quadratic form represents 0.

We solved this question for the quadratic forms $x_1^2+x_2^2$ and $x_1^2+x_2^2+x_3^2+x_4^2$ in the last chapter.

A quadratic form $f$ is called

- *positive definite* if it only represents positive integers apart from $f(0,0,\dots,0) = 0$;
- *negative definite* if it only represents negative integers apart from $f(0,0,\dots,0) = 0$;
- *indefinite* if it represents both positive and some negative integers.

In this chapter, we only consider binary quadratic forms (forms in two variables), and write such a form as $f(x,y) = ax^2+bxy+cy^2$.

**Example** The quadratic form $x^2+y^2$ is positive definite, and represents precisely zero and those positive integers whose squarefree part has no prime divisor congruent to 3 mod 4.

**Example** The quadratic form $x^2 - y^2$ is indefinite. It represents precisely those integers (positive or negative) which are not congruent to 2 mod 4. For, if $n$ is odd, say $n = 2k+1$, then $n = (k+1)^2 - k^2$, while if $n$ is a multiple of 4, say $n = 4k$, then $n = (k+1)^2 - (k-1)^2$. But $x^2 - y^2$ cannot be congruent to 2 mod 4, since squares are congruent to 0 or 1 mod 4.

**Example** The quadratic form $ax^2 + bxy + xy^2$ represents the integer $a$ (since $f(1,0) = a$).

**Example** The quadratic form $4x^2 + 12xy + 9y^2$ satisfies none of our three conditions, since $f(3,-2) = 0$.

## 10.1 Linear forms and degenerate quadratic forms

Let us take a step back and solve an easier question: Which integers are represented by linear forms?

**Proposition 10.1** *The equation $ax + by = n$ has a solution in integers $x$ and $y$ if and only if $\gcd(a,b)$ divides $n$.*

**Proof** Let $d = \gcd(a,b)$. If the equation has a solution, then $d \mid a$ and $d \mid b$, so $d \mid ax+by = n$. Conversely, suppose that $d \mid n$, say $n = md$. By Euclid's algorithm, we can find integers $u, v$ such that $au + bv = d$; then $ax + by = n$, with $x = mu$, $y = mv$. □

As an exercise, you should find all solutions to the equation $ax + by = n$.

A binary quadratic form is called *degenerate* if it is a multiple of a square of a linear form, say $k(ax+by)^2$.

**Corollary 10.2** *The degenerate form $k(ax+by)^2$ represents the integers of the form $k(md)^2$ for $m \in \mathbb{Z}$, where $d = \gcd(a,b)$.*

For example, the form $4x^2 + 12xy + 9y^2 = (2x+3y)^2$ is degenerate, and represents precisely the perfect squares.

**Remark** A degenerate quadratic form falls into none of the three classes we described earlier: for if $f(x,y) = k(ax+by)^2$, then $f(b,-a) = 0$. Conversely, a form which falls into none of these classes is degenerate.

## 10.2 Matrix, discriminant, equivalence

Let $f(x,y) = ax^2 + bxy + cy^2$ be a quadratic form. We define the *matrix* of the form to be $M = \begin{pmatrix} 2a & b \\ b & 2c \end{pmatrix}$, and the *discriminant* of the form to be $b^2 - 4ac = -\det(M)$.

Note that

$$f(x,y) = \tfrac{1}{2}\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} 2a & b \\ b & 2c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Following the notation from Linear Algebra I, we write this using column vectors as

$$f(x,y)) = \tfrac{1}{2}v^\top Mv,$$

where $v = \begin{pmatrix} x \\ y \end{pmatrix}$.

**Proposition 10.3** *A quadratic form is*

- *indefinite if its discriminant is positive;*
- *positive definite if its discriminant is negative and $a,c > 0$;*
- *negative definite if its discriminant is negative and $a,c < 0$;*
- *degenerate if its discriminant is zero.*

**Proof** Assume that $a \neq 0$. Then calculation shows that

$$ax^2 + bxy + cy^2 = \frac{1}{4a}\left((2ax+by)^2 - (b^2-4ac)y^2\right).$$

If $b^2 - 4ac > 0$, then clearly this takes both positive and negative values. (If $y = 0$, the quantity in brackets is positive, while if $y = 2a, x = -b$, then it is negative.) If $b^2 - 4ac < 0$, then $4ac < 0$ and $a,c$ have the same sign; and the values taken by the form have the same sign as $a$, since the quantity in brackets is a sum of squares with positive coefficients. Finally, if $b^2 - 4ac = 0$, then the form is $(1/4a)(2ax+by)^2$, which is degenerate.

If $c \neq 0$, then the same argument applies with $a$ and $c$ reversed.

If $a = c = 0$, then $f(x,y) = bxy$. If $b \neq 0$, the form is indefinite – putting $x = 1$, $y = \pm 1$, we get the values $\pm b$. Its discriminant is $b^2$, which is positive. If $b = 0$, the form is (very) degenerate! □

**Example** The form $x^2+3xy+5y^2$ has discriminant $9-20<0$, so is positive definite. The form $x^2+3xy+y^2$ has discriminant $9-4=5$, so is indefinite.

What are the possible discriminants of quadratic forms?

**Proposition 10.4** *An integer $d$ is the discriminant of a quadratic form if and only if $d \equiv 0$ or $1$ mod $4$.*

**Proof** We have $d=b^2-4ac$; and $b^2 \equiv 0$ or 1 mod 4, while $-4ac \equiv 0$ mod 4.

Conversely, if $d \equiv 0$ mod 4, then $d=4e$, and $x^2-ey^2$ has discriminant $d$; and if $d \equiv 1$ mod 4, then $d=4e+1$, and $x^2+xy-ey^2$ has discriminant $d$. □

We can break the representation problem into two, hopefully more tractable, parts:

- Which quadratic forms have given discriminant $d$?
- Which integers are represented by forms of discriminant $d$?

First we define an equivalence relation on quadratic forms, so that equivalent forms have the same discriminant and represent the same integers.

Let $P$ be a $2\times 2$ matrix with integer entries. If $\det(P)=1$, then $P$ has an inverse which is also a matrix with integer entries. For if $ps-qr \neq 0$, then

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix}^{-1} = \frac{1}{ps-qr}\begin{pmatrix} s & -q \\ -r & p \end{pmatrix}.$$

We call the matrix $P$ *unimodular* if its determinant is 1.

Now let $f$ and $f'$ be quadratic forms with matrices $M$ and $M'$ respectively. We say that $f$ and $f'$ are *equivalent* if there is a unimodular matrix $P$ such that

$$M' = P^\top MP.$$

**Proposition 10.5** *(a) Equivalence of quadratic forms is an equivalence relation.*

*(b) Equivalent forms have the same discriminant and represent the same integers.*

**Proof** (a) is straightforward using the fact that the identity is unimodular and products and inverses of unimodular matrices are unimodular.

[If you have taken the course Algebraic Structures I, there is another way to view this theorem. The unimodular matrices form a group, called the *special linear group* and denoted $\mathrm{SL}(2,\mathbb{Z})$; this group acts on the set of quadratic forms, and two forms are equivalent if and only if they lie in the same orbit of the group.]

(b) Calculate. If $\det(P) = 1$ and $M' = P^\top MP$ then

$$\det(M') = \det(P)\det(M)\det(P) = \det(M),$$

since $\det(P^\top) = \det(P)$. Also, if $n$ is represented by $f'$, then there exist $x$ and $y$ such that

$$(x \quad y)P^\top MP \begin{pmatrix} x \\ y \end{pmatrix} = n.$$

Now put

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = P \begin{pmatrix} x \\ y \end{pmatrix};$$

then

$$(x' \quad y')M \begin{pmatrix} x' \\ y' \end{pmatrix} = n,$$

so $n$ is represented by $f$. The converse follows using the fact that the relation of equivalence is symmetric (or by using the inverse of the matrix $P$). □

Suppose that the quadratic form $f$ is represented by the matrix $M$, and $P = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$ is unimodular. Then the equivalent form $f'$ is represented by $M' = P^\top MP$; we have

$$f'(x,y) = f(px+qy, rx+sy).$$

So, for example, the quadratic forms $x^2 + y^2$ and

$$(3x+4y)^2 + (2x+3y)^2 = 13x^2 + 36xy + 25y^2$$

are equivalent; they have the same discriminant (namely $-4$) and represent the same integers (namely, the positive integers whose squarefree part has no prime divisor congruent to 3 mod 4).

**Remark** Let $M$ represent $f(x,y) = ax^2 + bxy + cy^2$. If $M' = P^\top MP$, where $P = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$, then $M'$ represents $a'x^2 + b'xy + c'y^2$, where $a' = f(p,r)$, $c' = f(q,s)$, and

$$b' = (p \quad r)M \begin{pmatrix} q \\ s \end{pmatrix}.$$

**Remark** Equivalent forms have the same discriminant. But the converse is not true. The forms $x^2 + 6y^2$ and $2x^2 + 3y^2$ have the same discriminant (namely, $-24$), and are both positive definite; but the first represents the integer 1, while the second obviously does not. So they are not equivalent.

## 10.3 Positive definite forms

To understand an equivalence relation, a common strategy is to pick one element out of each equivalence class to use as a representative. (Such a representative is often called a "canonical form".) We now do this for positive definite quadratic forms:

- we define "reduced" forms, and give a simple test for recognising them;
- we give an algorithm for finding the unique reduced form equivalent to a given form;
- we show that there are only a finite number of equivalence classes of positive definite forms with given discriminant.

This gives us a fairly satisfactory classification of positive definite forms.

**Definition** Let $f(x,y) = ax^2 + bxy + cy^2$ be a positive definite quadratic form (so that $b^2 - 4ac < 0$, $a > 0$, $c > 0$). We say that $f$ is *reduced* if either

- $c > a$, $-a < b \leq a$; or
- $c = a$, $0 \leq b \leq a$.

We are going to show that any positive definite quadratic form is equivalent to a unique reduced form. This requires two steps: first we show that there is a reduced form equivalent to any given form; then we show that if two reduced forms are equivalent then they are equal.

We begin with a simple observation.

**Proposition 10.6** *let $f = ax^2 + \cdots$ be a reduced positive definite form. Then $a$ is the smallest positive integer represented by $f$.*

**Proof** The form $f = ax^2 + bxy + cy^2$ does represent the integer $a$ (just put $(x,y) = (1,0)$). So it suffices to assume that $f$ is reduced (so that, in particular, $|b| \leq a \leq c$), and that $ax^2 + bxy + cy^2 < a$ for some integers $x,y$ (not both zero), and derive a contradiction.

We divide into four cases, and reach a contradiction in each case.

- If $y=0$, then $a>ax^2\geq a$ (since $x^2\geq 1$).
- If $x=0$, then $a>cy^2\geq c\geq a$ (since $y^2\geq 1$).
- If $x,y\neq 0$ and $|x|\leq|y|$, then $|bxy|\leq cy^2$ (since $|b|\leq c$), and $a>ax^2+bxy+cy^2\geq ax^2\geq a$.
- If $x,y\neq 0$ and $|y|\leq|x|$, then $|bxy|\leq ax^2$, and $a>ax^2+bxy+cy^2\geq cy^2\geq c\geq a$.

First we look at a particular kind of equivalence. The matrix

$$P=\begin{pmatrix}0 & 1\\ -1 & k\end{pmatrix}$$

is unimodular. Our calculations above show that if $f(x,y)=ax^2+bxy+cy^2$ is represented by $M$, then the equivalent form represented by $P^\top MP$ is

$$f(y,ky-x)=cx^2-(b+2k)xy+(a+bk+ck^2)y^2.$$

We call this the *right neighbour* of $f$ via $k$. This is a formula to which we shall return very often! Note that the right neighbour of $f$ via 0 is $cx^2-bxy+ay^2$. Any right neighbour of a form $\cdots+cy^2$ looks like $cx^2+\cdots$. Note also that the coefficient of $y^2$ in the right neighbour of $f$ by $k$ is $f(1,k)$.

**Theorem 10.7** *Any positive definite quadratic form is equivalent to a reduced form.*

**Proof** We will see that the method of proof gives us a construction for finding the reduced form equivalent to a given positive definite form.

Let our form be

$$f_0=a_0x^2+b_0xy+a_1y^2.$$

(You will see in a minute why we write it this way.)

Define $q$ and $b_1$ by $b_0=2a_1q-b_1$, with $-a_1<b_1\leq a_1$. (In other words, divide $b_0$ by $2a_1$, so that the remainder is between $-a_1$ and $a_1$.)

Now take the right neighbouring form by $-q$. This form is

$$f_1=a_1x^2-(b_0-2a_1q)xy+(a_0-b_0q+a_1q^2)y^2=a_1x^2+b_1xy+a_2y^2,$$

where $a_2=f_0(1,-q)=a_0-b_0q+a_1q^2$.

If $a_2\geq a_1$, then stop; otherwise repeat to obtain

$$f_2=a_2x^2+b_2xy+a_3y^2,$$

with $-a_2 < b_2 \leq a_2$. Continue the process, obtaining forms $a_n x^2 + b_n xy + a_{n+1} y^2$ for $n = 1, 2, \ldots$; stop when $a_{n+1} \geq a_n$.

We have $a_1 > a_2 > a_3 > \cdots$, and this sequence cannot continue for ever since $a_i > 0$. When it terminates we have a form

$$f_n = a_n x^2 + b_n xy + a_{n+1} y^2,$$

with $-a_n < b_n \leq a_n$ and $a_n \leq a_{n+1}$.

This form is reduced unless $a_{n+1} = a_n$ and $b_n < 0$; in this case take the right neighbour by 0 to change the sign of $b_n$.

**Example** Find a reduced form equivalent to the positive definite form $31x^2 + 22xy + 4y^2$.

First we put $22 = 8q - b_1$ with $-4 < b_1 \leq 4$. The solution is $q = 3$, $b_1 = 2$, and $f_1 = 4^2 + 2xy + a_2 y^2$, where $a_2 = 31 - 22 \cdot 3 + 4 \cdot 9 = 1$, that is, $f_1 = 4x^2 + 2xy + y^2$.

Now we put $2 = 2q - b_2$ with $-1 < b_2 \leq 1$, with solution $q = 1$, $b_2 = 0$. We get $f_2 = x^2 + (4 - 2 + 1)y^2 = x^2 + 3y^2$, which is reduced.

**Theorem 10.8** *If the two reduced positive definite forms $f = ax^2 + bxy + cy^2$ and $g = a'x^2 + b'xy + c'y^2$ are equivalent, then they are equal: $a = a'$, $b = b'$, $c = c'$.*

The proof of this theorem is a calculation which we omit. Note that the fact that $a = a'$ follows from Proposition 10.6.

Now the theorem gives us an important conclusion. We have seen that equivalent forms have the same discriminant but that the converse is false. But the following holds:

**Theorem 10.9** *There are only finitely many equivalence classes of positive definite forms with given discriminant.*

**Proof** Since there is a unique reduced form in any given equivalence class, it is enough to show that there are only finitely many reduced forms with any given discriminant.

Let $f = ax^2 + bxy + cy^2$ be a reduced form, so that $c \geq a \geq |b|$, and

$$-d = -b^2 + 4ac \geq 3ac \geq 3c.$$

So we have

$$-a \leq b \leq a \leq c \leq |d|/3,$$

and for given $d$ there are only finitely many choices of $(a, b, c)$.

**Example** Discriminant $-11$. We have $b^2 \leq ac \leq 11/3 < 4$, so $b = -1$, 0 or 1. But $b^2 - 4ac$ is odd, so $b = \pm 1$. Now $4ac = 12$ so $ac = 3$. Now $a \leq c$, so $a = 1$, $c = 3$, and $-1 < b \leq 1$ gives $b = 1$. So the unique reduced form is $x^2 + xy + 3y^2$. This means that all positive definite forms with discriminant 11 are equivalent.

**Example** Discriminant $-12$. We have $b^2 \leq ac \leq 12/3 = 4$, and $b$ is even, so $b = -2$, 0 or 2.

If $b = \pm 2$ then $b^2 - 4ac = -12$ gives $ac = 4$, and $-a < b \leq a \leq c$ gives $b = a = c = 2$.

If $b = 0$, then $b^2 - 4ac = -12$ gives $ac = 3$; and $0 < a \leq c$ gives $a = 1$, $c = 3$.

Thus there are two reduced forms of discriminant 12, namely $2x^2 + 2xy + 2y^2$ and $x^2 + 3y^2$, and hence two equivalence classes of such forms. Note that the first form represents only even numbers while the second represents both even and odd numbers. So the form $31x^2 + 22xy + 4y^2$, which is positive definite with discriminant $-12$, is equivalent to the second of these (since $f(1,0) = 43$). This agrees with our calculation in the earlier example.

In general, if we know all the reduced forms of discriminant $-d$, and are given an arbitrary form with this discriminant, we can decide which reduced form it is equivalent to, by simply applying the method of Theorem 10.7.

The second part of the program, having classified forms up to equivalence, is to decide which integers are represented by any given reduced form. But this would take more time than we can afford!

## 10.4 Indefinite quadratic forms

Things work rather differently for indefinite quadratic forms. Recall that $f(x,y) = ax^2 + bxy + cy^2$ is indefinite if its discriminant $b^2 - 4ac$ is positive. Such a quadratic form may factorise over the integers: for example, $5x^2 + 12xy + 7y^2 = (5x + 7y)(x + y)$; we exclude these forms from our consideration. In particular, this implies that $a$ and $c$ are non-zero (if $c = 0$, then $f(x,y) = x(ax + by)$, and if $a = 0$, then $f(x,y) = y(bx + cy)$).

We will see that the theory of indefinite forms is more difficult than that of positive definite forms, but links up with the theory of continued fractions for quadratic irrationals. Recall that a quadratic irrational $s$ is said to be *reduced* if $s > 1$ and $-1 < s' < 0$, where $s'$ is the algebraic conjugate of $s$. Recall also that an irrational number has purely periodic continued fraction if and only if it is a reduced quadratic irrational.

One conclusions are:

- we give a definition of reduced forms, more complicated than in the positive definite case, and show that every form is equivalent to a reduced form;
- we show that the number of reduced forms in a given equivalence class is the least common multiple of 2 and $k$, where $k$ is the period of the continued fraction of a certain quadratic irrational associated with the form;
- we show that there are only finitely many reduced forms of any given discriminant.

This is not as satisfactory as for positive definite forms, since we do not have a unique reduced form in each equivalence class.

Consider the equation $f(x,y)=0$. Putting $y=1$, we have

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

so that

$$x=\frac{-b\pm\sqrt{d}}{2a}.$$

Note that $d$ is not a square (if it was, the quadratic would have rational roots, and $f$ would factorise), and the two solutions are conjugate quadratic irrationals. We call the root $t$ with the $+$ sign the *first root* of $f$. For technical reasons we need a related quadratic irrational: $s=|b+\sqrt{d}/2c|$.

**Proposition 10.10** *Let $s$ and $t$ be defined as above. Then $s=1/|t|$.*

**Proof**

$$\frac{1}{t}=\frac{2a}{\sqrt{d}-b}=\frac{2a(b+\sqrt{d})}{d-b^2}=-\frac{b+\sqrt{d}}{2c}$$

since $d=b^2-4ac$. Taking the modulus gives the result. □

Recall that we said a quadratic irrational $s$ is *reduced* if $s>1$ and $-1<s'<0$, where $s'$ is the algebraic conjugate of $s$. If $s$ is reduced and $u=-1/s$, then $u$ is also a quadratic irrational and satisfies $u>1$ and $-1<u'<0$; so $u$ is also reduced. It follows from Proposition 10.10 that, if $s$ is reduced, then so is either $t$ or $-t$.

We say that the indefinite quadratic form $ax^2+bxy+cy^2$ is *reduced* if $s=|(b+\sqrt{d})/2c|$ is a reduced quadratic irrational, where $d$ is the discriminant. (Note that this is quite different from the definition we used in the positive definite case!)

**Proposition 10.11** *(a) If the indefinite quadratic form $f(x,y)=ax^2+bxy+cy^2$ with discriminant $d$ is reduced, then $0<b<\sqrt{d}$ and $0<|c|<\sqrt{d}$.*

*(b) There are only finitely many reduced indefinite quadratic forms of given discriminant d.*

**Proof** (a) We have

$$s = \left|\frac{b+\sqrt{d}}{c}\right| = \frac{b+\sqrt{d}}{2\varepsilon c},$$

where $\varepsilon = \pm 1$ and $d - b^2 = -4ac$ is divisible by $2\varepsilon c$. Assume that $\varepsilon c > 0$ (so that $\varepsilon c = |c|$); the case $\varepsilon c < 0$ is similar, and we cannot have $c = 0$ since then $f$ would factorise. Assuming that $s$ is reduced, we have

$$\begin{aligned} s > 0, &\quad \text{so} \quad b + \sqrt{d} > 2|c|; \\ 0 > s' > -1, &\quad \text{so} \quad 0 > b - \sqrt{d} > -2|c|. \end{aligned}$$

It follows that $b > 0$, and then the second inequality gives $b < \sqrt{d}$. Subtracting the inequalities gives $2\sqrt{d} > 2|c|$.

(b) There are only finitely many values for $b$ and $c$, by (a); and $a$ is determined by $b$ and $c$, since $a = (b^2 - d)/4c$. □

**Example** Find all the reduced forms of discriminant 13.

We have $b^2 - 4ac = 13$, so $b$ is odd, and $0 < b < \sqrt{13}$, so $b = 1$ or $b = 3$. Also, $-3 \le c \le 3$.

If $b = 1$, then $1 - 4ac = 13$, so $ac = -3$, and $c = \pm 1$ or $\pm 3$. But then

$$s = \frac{1+\sqrt{13}}{2|c|} > 1, \qquad -1 < s' = \frac{1-\sqrt{13}}{2|c|} < 0,$$

giving $-1 + \sqrt{13} < 2|c| < 1 + \sqrt{13}$. The only even number in this range is 4, so $|c| = 2$, and we have a contradiction.

If $b = 3$, then $9 - 4ac = 13$, so $ac = -1$, and so $c = \pm 1$, $a = -c$. So the possible forms are

$$f(x,y) = x^2 + 3xy - y^2, \qquad g(x,y) = -x^2 + 3xy + y^2.$$

Since $s = (3 + \sqrt{13})/2$ is a reduced quadratic irrational, both of these quadratic forms really are reduced.

In this case, we see that the right neighbour of $f$ by 3 is

$$f(y, 3y - x) = y^2 + 3y(3y - x) - (3y - x)^2 = y^2 + 3xy - x^2 = g(x,y),$$

so the forms $f$ and $g$ are equivalent. (Remember that for positive definite forms, no two reduced forms are equivalent.)

We are going to decide when it happens that two reduced forms are equivalent. First we have to look more closely at reduced forms.

Let $f(x,y) = ax^2 + bxy + cy^2$ have discriminant $d = b^2 - 4ac > 0$, where $d$ is not a square. We defined the *first root* of $f$ to be $t = (-b+\sqrt{d})/2a$, a root of the quadratic $ax^2 + bx + c = 0$.

**Proposition 10.12** *Let $f(x,y) = ax^2 + bxy + cy^2$ be an indefinite form with discriminant $d > 0$ (d a non-square) with first root t. Then*

*(a) f is reduced if and only if $1/|t|$ is a reduced quadratic irrational;*

*(b) if g is the right neighbour of f by k, then g has first root $k - 1/t$.*

**Proof** (a) Immediate from the definition and Proposition 10.10.

(b) The right neighbour of $f$ by $k$ is

$$g(x,y) = cx^2 - (b+2ck)xy + (f(1,k))y^2,$$

with first root

$$\frac{b+2ck+\sqrt{d}}{2c} = k + \frac{b+\sqrt{d}}{2c} = k - 1/t$$

by (a). □

(Note that $f$ and $g$, being equivalent, have equal discriminants.)

Now we give an algorithm to show that any indefinite form with non-square discriminant is equivalent to a reduced form.

We start with such a form, say $f_0(x,y) = a_0x^2 + b_0xy + c_0y^2$.

Suppose that we have constructed $f_i(x,y) = a_ix^2 + b_ixy + c_iy^2$. If $i = 0$, or if $|a_i| > |c_i|$, then we write $b_i = (2c_i)q_i - b_{i+1}$, where $-|c_i| < b_{i+1} \leq |c_i|$. Let $f_{i+1}$ be the right neighbour of $f_i$ by $-q_i$. Then

$$f_{i+1} = c_ix^2 - (b_i - 2c_iq_i)xy + (a_i - b_iq_i + c_iq_i^2)y^2 = a_{i+1}x^2 + b_{i+1}xy + c_{i+1}y^2.$$

If $i > 0$ and $|a_i| \leq |c_i|$, then put $i = n+1$ and stop.

Now return to $f_n$, the penultimate form in the sequence. Put $b_n = (2c_n)q - b$, where $\sqrt{d} > b > \sqrt{d} - 2|c_n|$. Let $g$ be the right neighbour of $f_n$ by $-q$.

**Proposition 10.13** *(a) The above algorithm terminates.*

*(b) The form g is reduced and equivalent to f.*

**Proof** (a) We have $a_i = c_{i-1}$ and $b_i = r_{i-1}$. So if the algorithm fails to terminate we would have

$$|a_1| > |c_1| = |a_2| > |c_2| = \dots,$$

which is impossible.

(b) A fairly long calculation shows that $g$ is reduced. It is clearly equivalent to $f$. □

**Example** Let $f(x,y) = 3x^2 + 7xy + 3y^2$, with discriminant $7^2 - 4 \cdot 9 = 13$. We have $a_0 = 3$, $b_0 = 7$, $c_0 = 3$.

FIrst we put $7 = 6q_0 - b_1$, with $-3 < b_1 \le 3$, giving $q_0 = 1$ and $b_1 = -1$. The right neighbour of $f$ by $-1$ is $3x^2 - xy - y^2$.

Next we solve $-1 = 2q_1 - b_2$, with $-1 < b_2 \le 1$, giving $q_2 = 0$, $b_2 = 1$. The right neighbour of $f_1$ by 0 is $-x^2 + xy + y^2$, and now $|c_2| = |a_3| = 3 > |a_2| = 1$, so we stop.

We return to $f_1 = 3x^2 - xy - y^2$, and solve $-1 = -2q - b$ with $\sqrt{13} > b > \sqrt{13} - 2$, giving $q = -1$, $b = 3$. The right neighbour of $f_1$ by $+1$ is

$$g(x,y) = -x^2 + 3xy + y^2,$$

which is reduced according to the proof of the Proposition. Indeed this is one of the two reduced forms of discriminant 13 we found earlier.

Now what is the connection with continued fractions?

We have associated a quadratic irrational $t$ with each indefinite form $f$, namely the first root of $f$. Now $t$ has a continued fraction expansion (which is ultimately periodic, as we saw). Indeed, if $f$ is reduced, then either $t$ or $-t$ is a reduced quadratic irrational, and so has purely periodic continued fraction.

**Proposition 10.14** *Let $f(x,y) = ax^2 + xy + cy^2$ be a reduced indefinite quadratic form with first root $t$, and let $|t| = \varepsilon t$ (with $\varepsilon = \pm 1$). Let*

$$\frac{1}{|t|} = [a_0; a_1, a_2, \ldots].$$

*Then*

*(a) If $g$ is the right neighbour of $f$ by $\varepsilon a_0$, and $g$ has first root $T$, then $g$ is reduced, $|T| = -\varepsilon T$, and*

$$\frac{1}{|T|} = [a_1; a_2, \ldots].$$

*(b) $g$ is the only right neighbour of $f$ which is reduced.*

**Proof** (a) We have

$$\begin{aligned} T &= \varepsilon a_0 - 1/t \\ &= \varepsilon(a_0 - 1/|t|) \\ &= \varepsilon(a_0 - [a_0; a_1, a_2, \ldots]) \\ &= -\varepsilon \frac{1}{[a_1; a_2, \ldots]} \end{aligned}$$

since $[a_0;a_1,a_2,\ldots] = a_0 + 1/[a_1;a_2,\ldots]$.Hence $T = -\varepsilon|T|$ and

$$\frac{1}{|T|} = [a_1;a_2,\ldots].$$

Since $f$ is reduced, $1/|t|$ is reduced, so $[a_0;a_,a_2,\ldots]$ is purely periodic. So $[a_1;a_2,\ldots]$ is periodic (we have just moved one place along in the period), so $1/|T|$ is reduced, whence $g$ is reduced.

(b) A calculation, which we omit. □

Now any reduced indefinite form $f$ has associated with it a sign $\varepsilon = \pm 1$ (where $|t| = \varepsilon t$), and a purely periodic continued fraction

$$1/|t| = [\overline{a_0;a_1,a_2,\ldots,a_{k-1}}]$$

of period $k$, say. We construct a chain in which one step is to take the right neighbour by $\varepsilon a_0$. This has the effect of changing the sign of $\varepsilon$ and shifting the continued fraction one place along:

$$\varepsilon \mapsto -\varepsilon, \qquad [\overline{a_0;a_1,a_2,\ldots,a_{k-1}}] \mapsto [\overline{a_1;a_2,\ldots,a_{k-1},a_0}].$$

How many steps does it take to return to our starting point? After $k$ steps, the continued fraction has cycled right around and returned to its starting value, but $\varepsilon$ has been multiplied by $(-1)^k$. If $k$ is even, then everything is the same as when we started out; but if $k$ is odd, then the sign of $\varepsilon$ has changed, and we have to go round the cycle one more time to reach our starting point. So the number of steps we take to return is $k$ if $k$ is even, or $2k$ if $k$ is odd. This can be written more succinctly as $\mathrm{lcm}(2,k)$. This explains why the number of reduced forms in an equivalence class is $\mathrm{lcm}(2,k)$.

Note that, even if the period of the continued fraction is 1, we still need two steps. So each equivalence class contains at least two reduced forms (unlike the positive definite case where there was a unique reduced form in each class).

**Example** Consider the form $x^2+3xy-y^2$, with discriminant 13. We have seen that there are only two reduced forms of discriminant 13, so they must be equivalent. But even without knowing this, we could find another form equivalent to $f$, using the method of proof of the last propositon.

We have $t = (-3+\sqrt{13})/2$, so

$$\frac{1}{t} = \frac{2}{-3+\sqrt{13}} = \frac{3+\sqrt{13}}{2} > 0,$$

so $\varepsilon = +1$. The continued fraction of $1/t$ is given by

$$a_0 = \left\lfloor \frac{3+\sqrt{13}}{2} \right\rfloor = 3, \qquad y_1 = \frac{2}{\sqrt{13}-3} = \frac{1}{t},$$

so the continued fraction is $[3;3,\ldots]$.

The right neighbour of $f$ by 3 is $-x^2+3xy+y^2$ (which, as we have seen, is the other reduced form of discriminant 13). The first root of this form is $T = (-3+\sqrt{13})/(-2) = (3-\sqrt{13})/2$, which is negative; and $-T = t$, asit should be. The continued fraction of $-1/T$ is the same as that of $1/t$, namely $[3;3,\ldots]$.

If we continue the method of the Proposition, we take the right neighbour of $g$ by $-3$, which is $x^2+3xy-y^2$, that is, $f$. So the sequence of reduced forms is $f,g,f,g,\ldots$.

So we finally come to the description of all the reduced forms equivalent to a given one:

**Theorem 10.15** *Let $f(x,y)$ be an indefinite reduced form with discriminant $d$ and first root $t$, where $|t| = \varepsilon t$. Suppose that*

$$\frac{1}{|t|} = [\overline{a_0;a_1,\ldots,a_{k-1}}],$$

*where $k$ is even. (That is, if the period of $1/|t|$ is even, we take one period; if the period is odd, we take two periods.)*

*Let $f_0 = f$ and for $i = 1,2,\ldots$ let $f_i$ e the right neighbour of $f_{i-1}$ by $(-1)^{i-1}\varepsilon a_{i-1}$. Then $f_0,f_1,\ldots,f_{k-1}$ are all the reduced forms equivalent to $f$, and $f_k = f$.*

We call the sequence $(f_0,f_1,\ldots,f_{k-1})$ the chain of reduced forms equivalent to $f$. We see that every reduced form equivalent to $f$ is contained in the chain beginning at $f$. Note that $k$ is either the period of the continued fraction of $1/|t|$ or twice this period, where $t$ is the first root of $f$, and the first root of $f_i$ is the reciprocal of $(-1)^i\varepsilon[\overline{a_i;a_{i+1},\ldots,a_{i-1}}]$.

We forego a detailed proof of the Theorem, and conclude with two examples.

**Example** Find all the reduced quadratic forms of discriminant 17, and partition them into equivalence classes (that is, chains).

A reduced form $ax^2+bxy+cy^2$ of discriminant 17 has $0 < b < \sqrt{17}$, $0 < |c| < \sqrt{17}$, and $b$ odd (since $b^2-4ac = 17$). So $b = 1$ or $b = 3$.

If $b = 1$, we have $-4ac = 16$, so $ac = -4$; thus $c = \pm 1$, $\pm 2$ or $\pm 4$. Now $|(b+\sqrt{d})/2c| = (1+\sqrt{17})/2|c|$ is a reduced quadratic irrational, so

$$1+\sqrt{17} > 2|c|, \qquad 1-\sqrt{17} > -2|c|,$$

giving $|c| = 2$, $c = \pm 2$. So we get two forms:

$$2x^2 + xy - 2y^2 \text{ and } -2x^2 + xy + 2y^2.$$

If $b = 3$, we have $-4ac = 8$, so $ac = 2$, giving $c = \pm 1$ or $\pm 2$. This gives four further forms,

$$x^2 + 3xy - 2y^2,\ -x^2 + 3xy + 2y^2,\ 2x^2 + 3xy - y^2,\ -2x^2 + 3xy + y^2,$$

all of which can be checked to be reduced. So there are six reduced forms in all.

The chains have even length, though we don't yet know how many there are or what their lengths are.

Take $f_0 = 2x^2 + xy - 2y^2$. Its first root is $t = (-1 + \sqrt{17})/4$, with $1/t = (1 + \sqrt{17})/4 > 0$. Find the continued fraction of $1/t$:

$$\begin{aligned} a_0 &= \left\lfloor \frac{1}{t} \right\rfloor = 1, & y_1 &= \frac{4}{\sqrt{17} - 3} = \frac{\sqrt{17} + 3}{2} \\ a_1 &= \lfloor y_1 \rfloor = 3, & y_2 &= \frac{2}{\sqrt{17} - 3} = \frac{\sqrt{17} + 3}{4} \\ a_2 &= \lfloor y_2 \rfloor = 1, & y_3 &= \frac{4}{\sqrt{17} - 1} = \frac{1}{t}. \end{aligned}$$

So

$$\frac{1}{t} = [\overline{1;3,1}] = [\overline{1;3,1,1,3,1}].$$

We see that the chain contains six forms, so that the six reduced forms we found are all equivalent and form a single chain.

You should check for yourself that the procedure of taking successive right neighbours of $f$ does indeed produce all six reduced forms.

**Example** Do the same for discriminant 12.

A reduced form $ax^2 + bxy + cy^2$ of discriminant 12 has $0 < b < \sqrt{12}$, $0 < |c| < \sqrt{12}$, and $b^2 - 4ac = 12$, so $b$ is even. We must have $b = 2$. Then $-4ac = 8$, so $ac = -2$, and there are just four forms,

$$x^2 + 2xy - 2y^2,\ -x^2 + 2xy + 2y^2,\ 2x^2 + 2xy - y^2,\ -2x^2 + 2xy + y^2.$$

All are reduced.

Since the only possible first neighbour of the first form is the last, and *vice versa* (since a right neighbour of $\cdots + cy^2$ is $cx^2 + \cdots$), we see that there must be two chains each containing two forms; so there are two equivalence classes of forms of discriminant 12.

You should calculate the first roots of these forms and the appropriate continued fractions to check out this conclusion!

## Exercises

**10.1** Suppose that $f(x,y) = ax^2 + bxy + cy^2$ is an indefinite quadratic form: that is, it takes both positive and negative values for suitable integers $x$ and $y$.

(a) Show that there are real numbers $u$ and $v$, not both zero, such that $f(u,v) = 0$.

(b) Do there necessarily exist integers $u$ and $v$, not both zero, such that $f(u,v) = 0$?

**10.2** For each of the following quadratic forms, say whether it is positive definite, negative definite, or indefinite:

(a) $5x^2 + 12xy + 7y^2$

(b) $13x^2 + 36xy + 25y^2$.

**10.3** Find the continued fraction expansions associated with the four reduced quadratic forms of discriminant 12, and verify that there are two equivalence classes of such forms.

**10.4** Find all reduced positive definite quadratic forms with discriminant $-15$.

**10.5** Find a reduced quadratic form equivalent to the form $76x^2 + 249xy + 204y^2$.

**10.6** Find all reduced quadratic forms with discriminant 5, and classify them into chains.

**10.7** Find all reduced quadratic forms equivalent to the form $19x^2 + 29xy + 11y^2$.

**10.8** Suppose that the prime $p > 3$ is represented by the quadratic form $x^2 - xy + y^2$: say $u^2 - uv + v^2 = p$, where $u, v \in \mathbb{Z}$.

(a) Show that $p$ does not divide either $u$ or $v$.

(b) Show that $u^3 \equiv -v^3 \bmod p$ but $u \not\equiv -v \bmod p$.

(c) Show that $uv^{-1}$ has order 6 in $\mathbb{Z}_p$.

(d) Deduce that $p \equiv 1 \bmod 6$.

# Chapter 11

# Revision problems and solutions

## 11.1 Problems

**1** (a) Find $\gcd(131,52)$ and express it in the form $131x+52y$ for integers $x,y$.

(b) Does 52 have an inverse mod 131? If so, what is it? If not, why not?

(c) State the *Chinese Remainder Theorem.*

(d) Use your result to part (a) to find an explicit formula for the solution to the two simultaneous congruences

$$\begin{aligned} x &\equiv a \bmod 131, \\ x &\equiv b \bmod 52, \end{aligned}$$

in terms of $a$ and $b$.

(e) True or false? $52^{130} \equiv 1 \bmod 131$. Give reasons for your answer.

**2** (a) Express $\dfrac{131}{52}$ as a continued fraction.

(b) Is this expression unique? If so, why? If not, give another expression.

(c) Define Euler's "square bracket" function $[a_0,a_1,\ldots,a_n]$, and prove that

$$\gcd([a_0,a_1,\ldots,a_n],[a_0,a_1,\ldots,a_{n-1}]) = 1.$$

(d) Let $x_n = [2,2,2,\ldots,2]$, with $n$ terms in the bracket. Show that

$$x_0 = 1, \quad x_1 = 2, \qquad x_n = 2x_{n-1} + x_{n-2} \text{ for } n \geq 2,$$

and find $\lim_{n\to\infty} x_{n+1}/x_n$.

**3** (a) What is an *algebraic number*? What is an *algebraic integer*?

(b) Which of the following are algebraic numbers and/or algebraic integers?

$$\text{(i) } -273, \qquad \text{(ii) } (3+\sqrt{5})/2, \qquad \text{(iii) } \sqrt[3]{3}+1, \qquad (iv)\pi$$

You should prove any positive assertions but are not required to prove negative assertions. Standard results may be used if clearly stated.

(c) What does it mean to say that a real number $y$ is *approximable to order n*? State a theorem about the approximability of algebraic numbers.

**4** (a) Let $a_0, a_1, a_2, \ldots$ be integers, with $a_n > 0$ for $n > 0$. Let $c_n$ be the continued fraction $[a_0; a_1, a_2, \ldots, a_n]$. State a theorem about the ordering of the numbers $c_n$. Do they have a limit as $n \longrightarrow \infty$?

(b) Define the infinite continued fraction $[a_0; a_1, a_2, \ldots]$.

(c) Which numbers have a representation as infinite continued fractions? Is the representation unique? (Proof not required.)

(d) Find a continued fraction for $\sqrt{7}$.

**5** (a) Prove that the value of a periodic continued fraction is a *quadratic irrational.* (You should define this term.)

(b) Is the converse true? (No proof required.)

(c) What is meant by saying that a quadratic irrational is *reduced*? Give a characterisation of the continued fractions of reduced quadratic irrationals.

(d) Show that any quadratic irrational can be written in the form $y = (P + \sqrt{D})/Q$, where $P$ and $Q$ are integers, and $D$ is a positive integer which is not a square, such that $Q$ divides $D - P^2$. Show further that, if $y$ is reduced, then $0 < P < \sqrt{D}$ and $0 < Q < 2\sqrt{D}$.

**6** (a) You are given that $\sqrt{19} = [4; \overline{2, 1, 3, 1, 2, 8}]$.

- Explain how to find all solutions of the equation $x^2 - 19y^2 = \pm 1$, and find the smallest solution of this equation.
- Can 19 be written as a sum of two squares? Why or why not?

(b) You are given that $\sqrt{29} = [5; \overline{2, 1, 1, 2, 10}]$.

- Explain Legendre's method for expressing 29 as the sum of two squares.
- By carrying out the appropriate steps to find the continued fraction, find an expression for 29 as the sum of two squares.

**7** (a) Define *Euler's totient function* $\phi(n)$.

(b) From the definition, calculate $\phi(20)$.

(c) Prove that, if $p$ is prime and $a > 1$, then $\phi(p^a) = p^{a-1}(p-1)$.

(d) Find all positive integers $n$ such that $\phi(n) = 2$. State carefully any result you need in your argument.

(e) What is a *primitive root* of a prime $p$? How many primitive roots of $p$ are there? (Proof not required.)

(f) Find a primitive root of 13.

**8** Let $p$ be an odd prime.

(a) Define the terms *quadratic residue*, *quadratic non-residue* mod $p$, and the *Legendre symbol* $\left(\dfrac{a}{p}\right)$.

(b) Show that there are equally many (namely $(p-1)/2$) quadratic residues and non-residues among the set $\{1, 2, \ldots, p-1\}$. (You may assume the existence of a primitive root of $p$.)

(c) Calculate the following Legendre symbols, explaining carefully the results you use in your argument:

$$\text{(i)} \left(\frac{8}{11}\right), \quad \text{(ii)} \left(\frac{39}{23}\right), \quad \text{(iii)} \left(\frac{24}{41}\right).$$

(d) Show that, if $p$ does not divide $a$, then $a^{(p-1)/2} \equiv \left(\dfrac{a}{p}\right) \bmod p$.

**9** Let $p$ be a prime congruent to 1 mod 4.

(a) Show that there are positive integers $x$ and $r$, with $r < p$, such that $x^2 + 1 = rp$. (You may assume that $\left(\frac{-1}{p}\right) = +1$.)

(b) Show that, if $rp = x^2 + y^2$ and $1 < r < p$, then there exists a positive integer $s < r$ and integers $u, v$ such that $sp = u^2 + v^2$.

(c) Deduce that $p$ is the sum of two squares.

(d) Outline a proof, using the above fact, that if $n$ is an integer whose squarefree part has no prime factors congruent to 3 mod 4, then $n$ is the sum of two squares.

**10** (a) Explain what is meant by a *quadratic form* over $\mathbb{Z}$ in two variables $x$ and $y$. What is meant by saying that the quadratic form $f(x,y)$ *represents* the integer $n$?

(b) What is meant by saying that a quadratic form is (i) *positive definite*, (ii) *negative definite*, (iii) *indefinite*.

(c) For each of the quadratic forms below, state (with reasons) whether it is positive definite, negative definite or indefinite:

(i) $7x^2 + xy + y^2$,

(ii) $3x^2 - 2xy - 8y^2$.

(d) What is meant by saying that two quadratic forms are *equivalent*? Show that equivalent forms represent the same integers.

(e) Find the reduced form equivalent to $7x^2 + xy + y^2$.

## 11.2 Solutions

Extra comments or alternative answers are given in [square brackets].

**1** (a)

$$\begin{aligned}
131 &= 2 \cdot 52 + 27 \\
52 &= 1 \cdot 27 + 25 \\
27 &= 1 \cdot 25 + 2 \\
25 &= 12 \cdot 2 + 1 \\
2 &= 2 \cdot 1
\end{aligned}$$

So $\gcd(131,52)=1$.

$$\begin{aligned}
1 &= 25-12\cdot 2\\
&= 25-12(27-25)=13\cdot 25-12\cdot 27\\
&= 13\cdot(52-27)-12\cdot 27=13\cdot 52-25\cdot 27\\
&= 13\cdot 52-25\cdot(131-2\cdot 52)=63\cdot 52-25\cdot 131
\end{aligned}$$

So $x=-25$, $y=63$.

(b) Yes. We have $63\cdot 52\equiv 1 \bmod 131$, so the inverse of 52 mod 131 is 63.

(c) If $\gcd(m,n)=1$, then for any integers $a,b$, the simultaneous congruences

$$x\equiv a \bmod m, \qquad x\equiv b \bmod n$$

have a solution, which is unique mod $mn$.

(d) We have $63\cdot 52\equiv 1 \bmod 131$, $63\cdot 52\equiv 0 \bmod 52$, and $-25\cdot 131\equiv 0 \bmod 131$, $-25\cdot 131\equiv 1 \bmod 52$.

So the congruences have a solution

$$x=63\cdot 52\cdot a-25\cdot 131\cdot b,$$

and the general solution is the congruence class $[x]_{52\cdot 131}$, that is, all integers congruent to $x$ mod $52\cdot 131$.

(e) True: 131 is prime, so this follows from Fermat's Little Theorem. (If $p$ is prime and $p$ does not divide $a$, then $a^{p-1}\equiv 1 \bmod p$.)

**2** (a) From the application of Euclid's algorithm in the preceding part, we see that $\frac{131}{52}=[2;1,1,12,2]$. In more detail:

$$\begin{aligned}
\frac{2}{1} &= 2\\
\frac{25}{2} &= 12+\frac{1}{1/2}=12+\frac{1}{2}\\
\frac{27}{25} &= 1+\frac{1}{25/2}=1+\cfrac{1}{12+\cfrac{1}{2}}\\
\frac{52}{27} &= 1+\frac{1}{27/25}=1+\cfrac{1}{1+\cfrac{1}{12+\cfrac{1}{2}}}\\
\frac{131}{52} &= 2+\frac{1}{52/27}=2+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{12+\cfrac{1}{2}}}}
\end{aligned}$$

(b) It is not unique; it can also be written as $[2;1,1,12,1,1]$, since $2 = 1+\frac{1}{1}$. [In fact these are the only possible representations.]

(c) We define the function by induction:

$$\begin{aligned}[] &= 1\\ [a_0] &= a_0\\ [a_0,a_1,\ldots,a_n] &= a_0[a_1,\ldots,a_n]+[a_2,\ldots,a_n]\end{aligned}$$

for $n \geq 1$.

[If you don't like empty brackets you can start the induction one place later, by saying $[a_0] = a_0$, $[a_0,a_1] = a_0a_1+1$.]

We prove the assertion about gcd by induction on $n$. For $n = 0$, we have $\gcd([a_0],[]) = \gcd(a_0,1) = 1$. [Again if you prefer you can start the induction at $n = 1$: $\gcd([a_0,a_1],[a_0]) = \gcd(a_0a_1+1,a_0) = 1$.] Assuming the result for $n$, put $x = [a_0,\ldots,a_n]$ and $y = [a_0,\ldots,a_{n-1}]$. Then $\gcd(x,y) = 1$ by the induction hypothesis. Using the fact that we can expand the bracket function from the back as well as from the front, we get

$$\gcd([a_0,\ldots,a_{n+1}],[a_0,\ldots,a_n]) = \gcd(a_{n+1}x+y,x) = \gcd(y,x) = 1.$$

(d) Clearly $x_0 = [] = 1$ and $x_1 = [2] = 2$. By (c), $x_n = 2x_{n-1}+x_{n-2}$.

We try a solution of the recurrence relation of the form $x_n = \alpha^n$. This satisfies the relation if $\alpha^n = 2\alpha^{n-1}+\alpha^{n-2}$ for all $n$, which holds if $\alpha^2-2\alpha-1 = 0$, or $\alpha = 1\pm\sqrt{2}$. Since the recurrence is linear, the general solution is

$$x_n = a(1+\sqrt{2})^n+b(1-\sqrt{2})^n.$$

The values $a$ and $b$ can be found from the initial values: we have $a+b = 1$, $a(1+\sqrt{2})+b(1-\sqrt{2}) = 2$. Clearly $a \neq 0$, and it follows (since $1+\sqrt{2} > |1-\sqrt{2}|$) that $\lim_{n\to\infty} x_{n+1}/x_n = 1+\sqrt{2}$.

[The fractions $x_{n+1}/x_n$ are the convergents to the continued fraction $[2;2,2,\ldots] = [\overline{2;}]$; so the value of this continued fraction is $1+\sqrt{2}$, as you can easily establish directly. Note that $1+\sqrt{2}$ is a reduced quadratic irrational.]

**3** (a) An *algebraic number* $u$ is a complex number which satisfies an equation of the form $a_nun+a_{n-1}u^{n-1}+\cdots+a_0 = 0$, where $a_n,a_{n-1},\ldots,a_0$ are integers, not all zero. An *algebraic integer* satisfies an equation of the above form where $a_n,\ldots,a_0$ are integers and $a_n = 1$.

[It is also correct to say "an algebraic number $u$ is a complex number which satisfies an equation of the form $a_nu^n+a_{n-1}u^{n-1}+\cdots+a_0 = 0$, where $a_n,\ldots,a_0$ are rational numbers and $a_n = 0$."]

(b) We use *Gauss's Lemma*: $u$ is an algebraic integer if and only if its minimal polynomial (the monic polynomial of least degree satisfied by $u$) has integer coefficients.

(i) $u = -273$ satisfies $u + 273 = 0$, so it's an algebraic integer (and an algebraic number).

(ii) Let $u = (3+\sqrt{5})/2$. Then $u^2 = (7+3\sqrt{5})/2 = 3\alpha - 1$, so $u^2 - 3u + 1 = 0$. So it's an algebraic integer (and an algebraic number).

(iii) Let $u = \sqrt[3]{3} + 1$. Then $(u-1)^3 = 3$, so $u^3 - 3u^2 + 3u - 4 = 0$. So it's an algebraic integer (and an algebraic number).

(iv) $\pi$ is known to be transcendental (i.e. satisfies no polynomial equation with rational coefficients). So it is not an algebraic number (and not an algebraic integer).

[You may instead quote the known results that an integer is an algebraic integer, and that a quadratic irrational $a + b\sqrt{d}$ (with $d$ squarefree) is an algebraic integer if and only if either $a, b$ are integers, or $d \equiv 1 \bmod 4$ and $a - \frac{1}{2}, b - \frac{1}{2}$ are integers.]

(c) The real number $y$ is *approximable to order n* if there are infinitely many rational numbers $p/q$ such that $|y - p/q| < c/q^n$, for some constant $c$.

The theorem of Liouville says that, if $y$ is an algebraic number whose minimal polynomial has degree $n$, then $y$ is not approximable to any order greater than $n$.

**4** (a) The ordering of the convergents $c_n$ is

$$c_0 < c_2 < c_4 < \cdots < c_5 < c_3 < c_1.$$

The sequence of convergents $c_n$ does tend to a limit as $n \to \infty$.

[It is not asked why, but this follows immediately from this ordering and the fact that $|c_{n+1} - c_n| = 1/q_n q_{n+1} \to 0$ as $n \to \infty$.]

(b) The infinite continued fraction $[a_0; a_1, a_2, \ldots]$ is defined to be the limit of the continued fractions in (a).

(c) Theorem: Every irrational real number has a unique representation as an infinite continued fraction.

(d) Let $y_0 = \sqrt{7}$. We carry out the standard algorithm to find the continued fraction: successively put $a_n = \lfloor y_n \rfloor$ and $y_{n+1} = 1/(y_n - a_n)$.

$$\begin{aligned}
a_0 &= \lfloor y_0 \rfloor = 2, & y_1 &= \frac{1}{\sqrt{7}-2} = \frac{\sqrt{7}+2}{3}, \\
a_1 &= \lfloor y_1 \rfloor = 1, & y_2 &= \frac{3}{\sqrt{7}-1} = \frac{\sqrt{7}+1}{2}, \\
a_2 &= \lfloor y_2 \rfloor = 1, & y_3 &= \frac{2}{\sqrt{7}-1} = \frac{\sqrt{7}+1}{3},
\end{aligned}$$

$$a_3 = \lfloor y_3 \rfloor = 1, \qquad y_4 = \frac{3}{\sqrt{7}-2} = \sqrt{7}+2$$
$$a_4 = \lfloor y_4 \rfloor = 4, \qquad y_5 = \frac{1}{\sqrt{7}-2} = y_1.$$

So the continued fraction repeats:

$$\sqrt{7} = [2;1,1,1,4,1,1,1,4,1,1,1,4,\ldots] = [2;\overline{1,1,1,4}].$$

**5** (a) A quadratic irrational is a number of the form $a+b\sqrt{d}$, where $a$ and $b$ are rational, $b \neq 0$, and $d$ is squarefree.

We first show that the value of a purely periodic continued fraction is a quadratic irrational. Let $y$ be a real number whose continued fraction is purely periodic; say

$$y = [\overline{a_0;a_1,\ldots,a_{k-1}}] = [a_0;a_1,\ldots,a_{k-1},a_0,a_1,\ldots].$$

Then

$$\begin{aligned} y &= [a_0;a_1,\ldots,a_{k-1},y] \\ &= \frac{[a_0,a_1,\ldots,a_{k-1},y]}{[a_1,\ldots,a_{k-1},y]} \\ &= \frac{Ay+B}{Cy+D}, \end{aligned}$$

where $A = [a_0,\ldots,a_{k-1}]$, $B = [a_0,\ldots,a_{k-2}]$, $C = [a_1,\ldots,a_{k-1}]$, $D = [a_1,\ldots,a_{k-2}]$. So $Cy^2+(D-A)y-B = 0$. The solution to this quadratic equation is a quadratic irrational.

Now suppose that $y$ is periodic but not purely periodic, say

$$y = [a_0;a_1,\ldots,a_{k-1},\overline{b_0,\ldots,b_{l-1}}].$$

Let $u = [\overline{b_0;b_1,\ldots,b_{l-1}}]$. Then $u$ is a quadratic irrational; and

$$\begin{aligned} y &= [a_0;a_1,\ldots,a_{k-1},u] \\ &= \frac{[a_0,\ldots,a_{k-1},u]}{[a_1,\ldots,a_{k-1},u]} \\ &= \frac{Pu+Q}{Ru+S} \end{aligned}$$

as above. Hence $y$ is a quadratic irrational too.

(b) The converse is almost true (values of continued fractions are real!) Any real quadratic irrational is the value of a periodic continued fraction.

(c) A quadratic irrational $u$ is *reduced* if $u > 1$ and $-1 < u' < 0$, where $u'$ is the algebraic conjugate of $u$ (so, if $u = a + b\sqrt{d}$, then $u' = a - b\sqrt{d}$).

A quadratic irrational is reduced if and only if it is the value of a purely periodic continued fraction.

(d) This part is Lemma 5.3 from the lecture notes: I have copied it out here.

We know that $y = a + b\sqrt{d}$ where $a$ and $b$ are rationals and $d$ is squarefree. Suppose first that $b$ is positive. Let $q$ be the least common multiple of the denominators of $a$ and $b$, and $a = p/q$, $b = r/q$. Then

$$y = \frac{p + r\sqrt{d}}{q} = \frac{p + \sqrt{r^2 d}}{q} = \frac{pq + \sqrt{q^2 r^2 d}}{q^2}.$$

Put $P = pq$, $Q = q^2$, and $D = q^2 r^2 d$, and note that $Q$ divides $P^2 - D$.

If $u < 0$, then write $-y$ in the specified form and then replace $Q$ by $-Q$.

Now suppose that $y$ is reduced; recall that this means $y > 1$ and $-1 < y' < 0$, where $y'$ is the conjugate of $y$ (so $y' = (P - \sqrt{D})/Q$). Then

- $y > 0 > y'$, so $(P + \sqrt{D})/Q > (P - \sqrt{D})/Q$. Hence $Q > 0$.
- $y > 1 > -y'$, so $(P + \sqrt{D})/Q > (-P + \sqrt{D})/Q$. Hence $P > 0$.
- $y' < 0$, so $P - \sqrt{D} < 0$. Hence $P < \sqrt{D}$.
- $y > 1$, so $(P + \sqrt{D})/Q > 1$. Hence $Q < P + \sqrt{D} < 2\sqrt{D}$.

**6** (a) We know that solutions to $x^2 - ny^2 = \pm 1$ are given by $x = p_n$, $y = q_n$, where $p_n/q_n$ is the $n$th convergent to $\sqrt{n}$, and $n$ is one less than a multiple of the period of the continued fraction.

So the smallest solution of $x^2 - 19y^2 = \pm 1$ is given by

$$\begin{aligned} x &= [4,2,1,3,1,2] = 48 + 6 + 24 + 16 + 16 + 24 + 2 + 2 + 3 + 8 + 12 + 8 + 1 = 170, \\ y &= [2,1,3,1,2] = 12 + 6 + 4 + 4 + 6 + 2 + 3 + 2 = 39. \end{aligned}$$

[The question does not ask whether we have a solution to the equation with the plus or minus sign. Now $170^2 - 19 \cdot 29^2 = +1$ (either by direct calculation, or noticing that it is congruent to 1 mod 10, or by using the fact that $-1$ is a non-square mod 19). This shows that the method gives only solutions to $x^2 - 19y^2 = +1$.]

Since the period of the continued fraction is even, 19 cannot be written as the sum of two squares. [You could also say: Since $19 \equiv 3 \bmod 4$, it cannot be written as the sum of two squares.]

(b) The method involves expressing the quantity $y_k$ obtained at stage $k$ of the continued fraction, where $k$ is half of one more than the period, in the form $(\sqrt{29}+P)/Q$; then $P^2+Q^2=29$. Here are the calculations:

$$\begin{aligned} a_0 &= \lfloor\sqrt{29}\rfloor = 5, & y_1 &= \frac{1}{\sqrt{29}-5} = \frac{\sqrt{29}+5}{4}, \\ a_1 &= \lfloor y_1\rfloor = 2, & y_2 &= \frac{4}{\sqrt{29}-3} = \frac{\sqrt{29}+3}{5}, \\ a_2 &= \lfloor y_2\rfloor = 1, & y_3 &= \frac{5}{\sqrt{29}-2} = \frac{\sqrt{29}+2}{5}. \end{aligned}$$

So $29 = 2^2+5^2$.

[You could just calculate $P^2+Q^2$ at each stage obtaining $5^2+4^2=41$, $3^2+5^2=34$, $2^2+5^2=29$: stop when the right value is obtained.]

**7** (a) $\phi(n)$ is the number of members $x$ of the set $\{0,1,2,\ldots,n-1\}$ which satisfy $\gcd(x,n)=1$.

(b) We have to exclude all the even numbers, and the multiples of 5, leaving $\{1,3,7,9,11,13,17,19\}$; so $\phi(20)=8$.

(c) Of the $p^a$ numbers $0,1,,\ldots,p^a$, we have to exclude just the $p^{a-1}$ multiples of $p$; so $\phi(p^a)=p^a-p^{a-1}=p^{a-1}(p-1)$.

(d) Can we have $\phi(p^a)=2$, where $p$ is prime? From the expression above this can happen only if either $p^{a-1}=2$, $p-1=1$ (so $p^a=4$) or $p^{a-1}=1$, $p-1=2$ (so $p^a=3$).

Now we use the fact that if $n=p_1^{a_1}\cdots p_r^{a_r}$, where $p_1,\ldots,p_r$ are distinct primes, then $\phi(n)=\phi(p_1^{a_1})\cdots\phi(p_r^{a_r})$. So each prime power factor of $n$ must have either $\phi(p^a)=2$ (whence $p^a=3$ or 4), or $\phi(p^a)=1$ (whence clearly $p^a=2$). Since the primes must be distinct, the only new value of $n$ we obtain is $3\cdot 2=6$.

So the values of $n$ are 3, 4, and 6.

(e) A *primitive root* of $p$ is an integer $u$ such that every integer not divisible by $p$ is congruent to a power of $u$ mod $p$. [You can say: an integer $u$ such that the order of $u$ mod $p$ is $p-1$.]

The number of primitive roots of $p$ is $\phi(p-1)$.

(f) We have mod 13:

$$2^1=2, 2^2=4, 2^3=8, 2^4=3, 2^5=6, 2^6=12.$$

So the order of 2 mod 13 divides 12 (by Fermat's little theorem) but is not 1, 2, 3, 4 or 6; so it must be 12, that is, 2 is a primitive root of 12.

[You may if you wish continue calculating powers of 2,

$$2^7 = 11, 2^8 = 9, 2^9 = 5, 2^{10} = 10, 2^{11} = 7, 2^{12} = 1,$$

at which point you can observe that all non-zero integers mod 13 have been obtained.]

**8** (a) The integer $a$, not divisible by $p$, is a *quadratic residue* mod $p$ if the congruence $x^2 \equiv a \bmod p$ has a solution; it is a *quadratic non-residue* mod $p$ otherwise.

The *Legendre symbol* $\left(\dfrac{a}{p}\right)$ is defined for integer $a$ and odd prime $p$ to be

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{if } p \text{ divides } a; \\ +1 & \text{if } a \text{ is a quadratic residue mod } p; \\ -1 & \text{if } a \text{ is a quadratic non-residue mod } p. \end{cases}$$

(b) Let $u$ be a primitive root of $p$, an element whose order mod $p$ is $p-1$. Then all $p-1$ non-zero integers mod $p$ are powers of $u$. Clearly the even powers are quadratic residues and the odd powers are non-residues.

(c) (i)

$$\left(\frac{8}{11}\right) = \left(\frac{2}{11}\right)^3 = \left(\frac{2}{11}\right) = -1,$$

using the multiplicative property (Rule 1) and the fact that $\left(\dfrac{2}{p}\right)$ is $-1$ if $p \equiv 3$ mod 8 (Rule 3).

(ii)

$$\left(\frac{39}{23}\right) = \left(\frac{16}{23}\right) = +1,$$

using the fact that $\left(\dfrac{a}{p}\right) = \left(\dfrac{b}{p}\right)$ if $a$ and $b$ are congruent mod $p$ and the fact that 16 is clearly a square modulo any prime.

(iii)

$$\left(\frac{24}{41}\right) = \left(\frac{2}{41}\right)\left(\frac{3}{41}\right) = \left(\frac{41}{3}\right) = \left(\frac{2}{3}\right) = -1,$$

using the multiplicative property, the fact that 4 is a square, the value $\left(\dfrac{2}{p}\right) = +1$ if $p \equiv 1$ mod 8, and the law of quadratic reciprocity (Rule 4).

(d) Let $u$ be a primitive root of $p$. Then $u^{p-1} \equiv 1 \bmod p$. So if $x = u^{(p-1)/2}$, we have $x^2 \equiv 1 \bmod p$, so $x \equiv 1$ or $x \equiv -1$. The first is impossible since the order of $u$ is $p-1$. So $u^{(p-1)/2} \equiv -1 \bmod p$. Now, if $a = u^k$, then

$$a^{(p-1)/2} = u^{k(p-1)/2} \equiv (-1)^k = \begin{cases} +1 & \text{if } k \text{ is even,} \\ -1 & \text{if } k \text{ is odd.} \end{cases}$$

But we saw that $a$ is a residue or non-residue according as it is an even or odd power of $u$ in part (b); so $a^{(p-1)/2} \equiv \left(\dfrac{a}{p}\right) \bmod p$.

**9** (a) The given value of the Legendre symbol shows that there is a solution of $x^2 \equiv -1 \bmod p$, that is, $x^2+1$ is a multiple of $p$. Choosing $x$ so that $x < p$ (since we can replace it by its remainder on dividing by $p$), we have $rp = x^2+1 < (p-1)^2+1 < p^2$, so $r < p$.

(b) Choose $a, b$ so that $|a|, |b| < r/2$ and $a \equiv -x$, $b \equiv y \bmod r$. Then $a^2+b^2 \equiv x^2+y^2 \equiv 0 \bmod r$, so $a^2+b^2 = rs$, with $s < r$. The two-squares identity gives

$$rp \cdot rs = (x^2+y^2)(a^2+b^2) = (xa-yb)^2+(xb+ya)^2,$$

and $xa-yb \equiv x^2+y^2 \equiv 0 \bmod r$, $xb+ya \equiv xy-yx \equiv 0 \bmod r$. Put $xa-yb = ru$, $xb+ya = rv$. Then

$$r^2sp = (ru)^2+(rv)^2,$$

so $sp = u^2+v^2$.

(c) This process can be continued until we obtain $p = x^2+y^2$ for some $x, y$, since the multiple of $p$ which is a sum of two squares cannot decrease for ever.

(d) If the squarefree part of $n$ has no prime factor congruent to 3 mod 4, then $n = m^2 p_1 \cdots p_r$, where $p_1, \ldots, p_r$ are primes congruent to 1 mod 4 or the prime 2. Each factor is the sum of two squares: $m^2 = m^2+0^2$, $2 = 1^2+1^2$, and the result for odd primes is (c) above. By the two-squares identity, $n$ is the sum of two squares.

**10** (a) A *quadratic form* in $x$ and $y$ is an expression $f(x,y) = ax^2+bxy+cy^2$, where $a, b, c$ are integers. We say that it represents the integer $n$ if there are integer values $x$ and $y$ such that $f(x,y) = n$.

(b) The form $f$ is positive definite if $f(x,y) > 0$ for all integers $x, y$ not both zero; negative definite if $f(x,y) < 0$ for all integers $x, y$ not both zero; and indefinite if it takes both positive and negative values.

(c) (i) This form can be written as $(27/4)x^2+(x/2+y)^2$, so clearly its values are $\geq 0$. If it is zero, then $x = (x/2+y) = 0$, so $x = y = 0$. So the form is positive definite.

(ii) $F(1,0) = 3 > 0$ and $f(0,1) = -8 < 0$, so the form is indefinite.

[You could also use the standard tests here. For the first form, the discriminant is $1^2 - 4 \cdot 7 \cdot 1 = -27$ which is negative, and $a = 7 > 0$; so the form is positive definite. For the second, the discriminant is $2^2 - 4 \cdot 3 \cdot (-8) = 100$, which is positive; so the form is indefinite.]

(d) Two forms $f(x,y)$, $g(x,y)$ are equivalent if $g(x,y) = f(px+qy, rx+sy)$ for some $p,q,r,s$ satisfying $ps-qr=1$ (that is, such that the matrix $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$ is unimodular).

If $f$ and $g$ satisfy this relation and $f(x,y)=n$, then $g(px+qy, rx+sy)=n$. Conversely, since the inverse of the matrix $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$ is $\begin{pmatrix} s & -q \\ -r & p \end{pmatrix}$, we have $f(x,y) = g(sx-qy, -rx+py)$, so any integer represented by $f$ is also represented by $g$. So they represent the same integers.

(e) Remember that the positive definite form $ax^2+bxy+cy^2$ is *reduced* if either $c>a$ and $-a<b\le a$, or $c=a$ and $0\le b\le a$.

Apply the algorithm in Chapter 10 of the notes. We begin with the form $f_0 = a_0x^2+b_0xy+a_1y^2$, with $a_0=7$, $b_0=a_1=1$.

First we put $b_0 = 2a_1q-b_1$ with $-a_1<b_1\le a_1$; in other words, $1=2q-b_1$ with $-1<b_1\le 1$. Clearly $q=1$ and $b_1=1$. The right neighbour of $f_0$ by $-1$ is $a_1x^2+b_1xy+a_2y^2$, with $a_1=1$, $b_1=1$, $a_2=f_0(1,-1)=7$; that is, $f_1 = x^2+xy+7y^2$.

Now $a_2=7>a_1=1$ and $b_1=1$ satisfies $-1<b_1\le 1$; so this form is reduced.

# Index