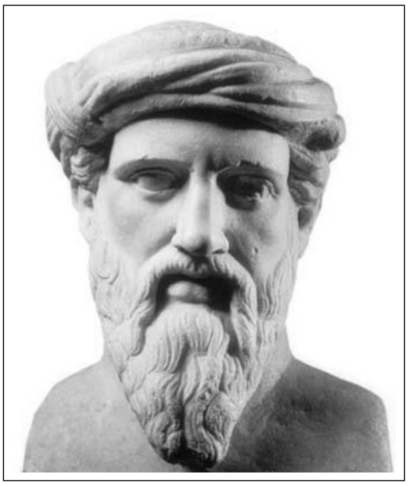
UNIVERSITY OF DUBLIN

TRINITY COLLEGE

SCHOOL OF MATHEMATICS

2015 Course 2316
Introduction to Number Theory

Timothy Murphy



ἁπας εἰναι αριθμὸς

All is number

Pythagoras of Samos

(570–495 BC)

Contents

Chapter 0

Prerequisites

0.1 The number sets

We follow the standard (or Bourbaki) notation for the number sets $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$.

Thus $\mathbb{N}$ is the set of natural numbers $0, 1, 2, \ldots$; $\mathbb{Z}$ is the set of integers $0, \pm 1, \pm 2, \ldots$; $\mathbb{Q}$ is the set of rational numbers n/d, where $n, d \in \mathbb{Z}$ with $d \neq 0$; $\mathbb{R}$ is the set of real numbers, and $\mathbb{C}$ the set of complex numbers $x + iy$, where $x, y \in \mathbb{R}$.

Note that $\mathbb{Z}$ is an *integral domain*, ie a commutative ring with 1 having no zero divisors:

$$xy = 0 \implies x = 0 \text{ or } y = 0.$$

Also $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ are all *fields*, ie integral domains in which every non-zero element has a multiplicative inverse.

All 5 sets are *totally ordered*, ie given 2 elements x, y of any of these sets we have either $x < y$, $x = y$ or $x > y$. Also the orderings are compatible (in the obvious sense) with addition and multiplication, eg

$$x \geq 0,\ y \geq 0 \implies x + y \geq 0,\ xy \geq 0.$$

0.2 The natural numbers

According to Kronecker, "God gave us the integers, the rest is Man's".
("Gott hat die Zahlen gemacht, alles andere ist Menschenwerk.")

We follow this philosophy in assuming the basic properties of $\mathbb{N}$.

In particular, we assume that $\mathbb{N}$ is *well-ordered*, ie a decreasing sequence of natural numbers

$$a_0 \geq a_1 \geq a_2 \ldots$$

is necessarily stationary: for some n,

$$a_n = a_{n+1} = \cdots .)$$

We also assume that we can "divide with remainder"; that is, given $n, d \in \mathbb{N}$ with $d \neq 0$ we can find $q, r \in \mathbb{N}$ such that

$$n = qd + r,$$

with remainder

$$0 \le r < d.$$

If we wanted to prove these results, we would have to start from an axiomatic definition of $\mathbb{N}$ such as the Zermelo-Fraenkel, or ZF, axioms. But we don't want to get into that, and assume as 'given' the basic properties of $\mathbb{N}$.

0.3 Divisibility

If $a, b \in \mathbb{Z}$, we say that a *divides* b, written $a \mid b$, or a is a *factor* of b, if

$$b = ac$$

for some $c \in \mathbb{Z}$.

Thus every integer divides 0; but the only integer divisible by 0 is 0 itself.

Chapter 1

The Fundamental Theorem of Arithmetic

1.1 Primes

Definition 1.1. *We say that $p \in \mathbb{N}$ is* prime *if it has just two factors in $\mathbb{N}$, 1 and p itself.*

Number theory might be described as the study of the sequence of primes

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

Definition 1.2. *1. We denote the nth prime by p_n.*

2. If $x \in \mathbb{R}$ then we denote the number of primes $\le x$ by $\pi(x)$.

Thus

$$p_1 = 2,\ p_2 = 3,\ p_3 = 5, \ldots,$$

while

$$\pi(-2) = 0,\ \pi(2) = 1,\ \pi(\pi) = 2, \ldots .$$

1.2 The fundamental theorem

Theorem 1.1. *Every non-zero natural number $n \in \mathbb{N}$ can be expressed as a product of primes*

$$n = p_1 \cdots p_r;$$

and this expression is unique up to order.

By convention, an empty sum has value 0 and an empty product has value 1. Thus $n = 1$ is the product of 0 primes.

Another way of putting the theorem is that each non-zero $n \in \mathbb{N}$ is uniquely expressible in the form

$$n = 2^{e_2} 3^{e_3} 5^{e_5} \cdots$$

where each $e_p \in \mathbb{N}$ with $e_p = 0$ for all but a finite number of primes p.

The proof of the theorem, which we shall give later in this chapter, is non-trivial. It is easy to lose sight of this, since the theorem is normally met long before the concept of *proof* is encountered.

1.3 Euclid's Algorithm

Definition 1.3. *Suppose $m, n \in \mathbb{Z}$. We say that $d \in \mathbb{N}$ is the* greatest common divisor *of m and n, and write*

$$d = \gcd(m, n),$$

if

$$d \mid m, \; d \mid n,$$

and if $e \in \mathbb{N}$ then

$$e \mid m, \; e \mid n \implies e \mid d.$$

The term *highest common factor* (or hcf), is often used in schools; but we shall always refer to it as the gcd.

Note that at this point we do not know that $\gcd(m, n)$ exists. This follows easily from the Fundamental Theorem; but we want to use it in proving the theorem, so that is not relevant.

It is however clear that if $\gcd(m, n)$ exists then it is unique. For if $d, d' \in \mathbb{N}$ both satisfy the criteria then

$$d \mid d', \; d' \mid d \implies d = d'.$$

Theorem 1.2. *Any two integers m, n have a greatest common divisor*

$$d = \gcd(m, n).$$

Moreover, we can find integers x, y such that

$$d = mx + ny.$$

Proof. We may assume that $m > 0$; for if $m = 0$ then it is clear that

$$\gcd(m, n) = |n|,$$

while if $m < 0$ then we can replace m by $-m$.

Now we follow the Euclidean Algorithm. Divide n by m:

$$n = q_0 m + r_0 \quad (0 \le r_0 < m).$$

If $r_0 \neq 0$, divide m by r_0:

$$m = q_1 r_0 + r_1 \quad (0 \le r_1 < r_0).$$

If $r_1 \neq 0$, divide r_0 by r_1:

$$r_0 = q_2 r_1 + r_2 \quad (0 \le r_2 < r_1).$$

Continue in this way.

Since the remainders are strictly decreasing:

$$r_0 > r_1 > r_2 > \cdots,$$

the sequence must end with remainder 0, say

$$r_{s+1} = 0.$$

We assert that

$$d = \gcd(m, n) = r_s,$$

ie the gcd is the last non-zero remainder.

For

$$d =| r_{s-1} \text{ since } r_{s-1} = q_{s+1} r_s.$$

Now

$$\begin{aligned} d \mid r_s, r_{s-1} &\implies d \mid r_{s-2} \text{ since } r_{s-2} = r_s - q_s r_{s-1}; \\ d \mid r_{s-1}, r_{s-2} &\implies d \mid r_{s-3} \text{ since } r_{s-3} = r_{s-1} - q_{s-1} r_{s-2}; \\ \ldots\ldots & \\ d \mid r_2, r_1 &\implies d \mid m; \\ d \mid r_1, m &\implies d \mid n. \end{aligned}$$

Thus

$$d \mid m, n.$$

Conversely, if $e \mid m, n$ then

$$\begin{aligned} & e \mid r_0 \text{ since } r_0 = n - q_0 m; \\ & e \mid r_1 \text{ since } r_1 = m - q_1 r_0; \\ & \quad \ldots\ldots \\ & e \mid r_s \text{ since } r_s = r_{s-1} - q_s r_{s-1}. \end{aligned}$$

Thus

$$e \mid m, n \implies e \mid d.$$

We have proved therefore that $\gcd(m, n)$ exists and

$$\gcd(m, n) = d = r_s.$$

To prove the second part of the theorem, which states that d is a linear combination of m and n (with integer coefficients), we note that if a, b are linear combinations of m, n then a linear combination of a, b is also a linear combination of m, n.

Now r_1 is a linear combination of m, n, from the first step in the algorithm; r_2 is a linear combination of m, r_1, and so of m, n, from the second step; and so on, until finally $d = r_s$ is a linear combination of m, n:

$$d = mx + ny.$$

□

We say that m, n are *coprime* if

$$\gcd(m, n) = 1.$$

Corollary 1.1. *If m, n are coprime then there exist integers x, y such that*

$$mx + ny = 1.$$

1.4 Speeding up the algorithm

Note that if we allow *negative* remainders then given $m, n \in \mathbb{Z}$ we can find $q, r \in \mathbb{Z}$ such that

$$n = qm + r,$$

where $|r| \le |m|/2$.

If we follow the Euclidean Algorithm allowing negative remainders then the remainder is at least halved at each step. It follows that if

$$2^r \le n < 2^{r+1}$$

then the algorithm will complete in $\le r$ steps.

Another way to put this is to say that if n is written to base 2 then it contains at most r bits (each bit being 0 or 1).

When talking of the efficiency of algorithms we measure the input in terms of the number of bits. In particular, we define the *length* $\ell(n)$ to be the number of bits in n. We say that an algorithm completes in polynomial time, or that it is in class P, if the number of steps it takes to complete its task is $\le P(r)$, where $P(x)$ is a polynomial and r is the number of bits in the input.

Evidently the Euclidian algorithm (allowing negative remainders) is a polynomial-time algorithm for computing $\gcd(m, n)$.

1.5 Example

Let us determine

$$\gcd(1075, 2468).$$

The algorithm goes:

$$\begin{aligned}
2468 &= 2 \cdot 1075 + 318, \\
1075 &= 3 \cdot 318 + 121, \\
318 &= 3 \cdot 121 - 45, \\
121 &= 3 \cdot 45 - 14, \\
45 &= 3 \cdot 14 + 3, \\
14 &= 5 \cdot 3 - 1, \\
3 &= 3 \cdot 1.
\end{aligned}$$

Thus

$$\gcd(1075, 2468) = 1;$$

the numbers are coprime.

To solve

$$1075x + 2468y = 1,$$

we start at the end:

$$\begin{aligned}
1 &= 5 \cdot 3 - 14 \\
&= 5(45 - 3 \cdot 14) - 14 = 5 \cdot 45 - 16 \cdot 14 \\
&= 5 \cdot 45 - 16(3 \cdot 45 - 121) = 16 \cdot 121 - 43 \cdot 45 \\
&= 16 \cdot 121 - 43(3 \cdot 121 - 318) = 43 \cdot 318 - 113 \cdot 121 \\
&= 43 \cdot 318 - 113(1075 - 3 \cdot 318) = 382 \cdot 318 - 113 \cdot 1075 \\
&= 382(2468 - 2 \cdot 1075) - 113 \cdot 1075 = 382 \cdot 2468 - 877 \cdot 1075.
\end{aligned}$$

Note that this solution is not unique; we could add any multiple $1075t$ to x, and subtract $2468t$ from y, eg

$$\begin{aligned}
1 &= (382 - 1075) \cdot 2468 + (2468 - 877) \cdot 1075 \\
&= 1591 \cdot 2468 - 693 \cdot 1075.
\end{aligned}$$

We shall return to this point later.

1.6 An alternative proof

There is an apparently simpler way of establishing the result.

Proof. We may suppose that x, y are not both 0, since in that case it is evident that $\gcd(m, n) = 0$.

Consider the set S of all numbers of the form

$$mx + ny \qquad (x, y \in \mathbb{Z}).$$

There are evidently numbers > 0 in this set. Let d be the smallest such integer; say

$$d = ma + nb.$$

We assert that

$$d = \gcd(m, n).$$

For suppose $d \nmid m$. Divide m by d:

$$m = qd + r,$$

where $0 < r < d$. Then

$$r = m - qd = m(1 - qa) - nqd,$$

Thus $r \in S$, contradicting the minimality of d.

Hence $d \mid m$, and similarly $d \mid n$.

On the other hand

$$d' \mid m, n \implies d' \mid ma + nb = d.$$

We conclude that

$$d = \gcd(m, n).$$

□

The trouble with this proof is that it gives no idea of how to determine $\gcd(m,n)$. It appears to be *non-constructive*.

Actually, that is not technically correct. It is evident from the discussion above that there is a solution to

$$mx + ny = d$$

with

$$|x| \leq |n|, \ |y| \leq |m|.$$

So it would be theoretically possible to test all numbers (x,y) in this range, and find which minimises $mx+ny$.

However, if x,y are very large, say 100 digits, this is completely impractical.

1.7 Euclid's Lemma

Proposition 1.1. *Suppose p is prime; and suppose $m,n \in \mathbb{Z}$. Then*

$$p \mid mn \implies p \mid m \text{ or } p \mid n.$$

Proof. Suppose

$$p \nmid m.$$

Then p,m are coprime, and so there exist $a,b \in \mathbb{Z}$ such that

$$pa + mb = 1.$$

Multiplying by n,

$$pna + mnb = n.$$

Now

$$p \mid pna, \ p \mid mnb \implies p \mid n.$$

□

1.8 Proof of the Fundamental Theorem

Proof.

Lemma 1.1. *n is a product of primes.*

Proof. We argue by induction on n If n is *composite*, ie not prime, then

$$n = rs,$$

with

$$1 < r,s < n.$$

By our inductive hypothesis, r,s are products of primes. Hence so is n. □

To complete the proof, we argue again by induction. Suppose

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

are two expressions for n as a product of primes.

Then

$$\begin{aligned} p_1 \mid n &\implies p_1 \mid q_1 \cdots q_s \\ &\implies p_1 \mid q_j \end{aligned}$$

for some j.

But since q_j is prime this implies that

$$q_j = p_1.$$

Let us re-number the q's so that q_j becomes q_1. Then we have

$$n/p_1 = p_2 \cdots p_r = q_2 \cdots q_s.$$

Applying our inductive hypothesis we conclude that $r = s$, and the primes $p_2, \ldots, p_r$ and $q_2, \ldots, q_s$ are the same up to order.

The result follows. □

1.9 A postscript

Suppose $\gcd(m, n) = 1$. Then we have seen that we can find integers x_0, y_0 such that

$$mx_0 + ny_0 = 1.$$

We can now give the general solution to this equation:

$$(x, y) = (x_0 + tn, y_0 - tm)$$

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

Certainly this is a solution. To see that it is the general solution note that

$$\begin{aligned} mx + ny = d &\implies mx + ny = mx_0 + ny_0 \\ &\implies m(x - x_0) = n(y_0 - y). \end{aligned}$$

Now n has no factor in common with m, by hypothesis. Hence all its factors divide $x - x_0$, ie

$$\begin{aligned} n \mid x - x_0 &\implies x - x_0 = tn \\ &\implies x = x_0 + tn \\ &\implies y = y_0 - tm. \end{aligned}$$

Exercise 1

In exercises 1–3 determine the gcd d of the given numbers m, n and find integers x, y such that $d = mx + my$.

* 1. $23, 39$

* 2. $87, -144$

* 3. $2317, 2009$.

** 4. Given integers $m, n > 0$ with $\gcd(m, n) = 1$ show that all integers $N \geq mn$ are expressible in the form

$$N = mx + ny$$

with $x, y \geq 0$.

** 5. Find the greatest integer n *not* expressible in the form

$$n = 17x + 23y$$

with $x, y \geq 0$.

*** 6. Which integers n are *not* expressible in the form

$$n = 17x - 23y$$

with $x, y \geq 0$?

** 7. Define the gcd

$$d = \gcd(n_1, n_2, \ldots, n_r)$$

of a finite set of integers $n_1, n_2, \ldots, n_r \in \mathbb{Z}$; and show that there exist integers $x_1, x_2, \ldots, x_r \in \mathbb{Z}$ such that

$$n_1 x_1 + n_2 x_2 + \cdots + n_r x_r = d.$$

* 8. Find $x, y, z \in \mathbb{Z}$ such that

$$24x + 30y + 45z = 1.$$

*** 9. How many ways are there of paying €10 in 1,2 and 5 cent pieces?

** 10. Show that if $m, n > 0$ then

$$\gcd(m, n) \times \operatorname{lcm}(m, n) = mn.$$

*** 11. Show that if $m, n > 0$ then

$$\gcd(m + n, mn) = \gcd(m, n).$$

** 12. Show that if $n \geq 9$ and both $n - 2$ and $n + 2$ are prime then $3 \mid n$.

*** 13. Suppose

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n$$

where $a_0, a_1, \ldots, a_n \in \mathbb{Z}$. Show that $f(n)$ cannot be a prime for all n unless $f(x)$ is constant.

*** 14. Find all integers $m, n > 1$ such that

$$m^n = n^m.$$

*** 15. If $p^e \parallel n!$ show that

$$e = [n/p] + [n/p^2] + [n/p^3] + \cdots .$$

[Note: if p is a prime we say that p^e *exactly divides* N, and we write $p \parallel N$ if $p^e \mid N$ but $p^{e+1} \nmid N$.]

*** 16. How many zeros does 1000! end with?

*** 17. Prove that $n!$ divides the product of any n successive integers.

*** 18. If F_n is the nth Fibonacci number, show that

$$\gcd(F_n, F_{n+1}) = 1$$

and

$$\gcd(F_n, F_{n+2}) = 1.$$

[Note: $F_0 = 1, F_1 = 2$ and $F_{n+2} = F_n + F_{n+1}$.]

** 19. Use the program `/usr/games/primes` on the mathematics computer system to find the next 10 primes after 1 million. [You can find how to use this program by giving the command `man primes`.]

** 20. Use the program `/usr/games/factor` on the mathematics computer system to factorise 123456789. [You can find how to use this program by giving the command `man factor`.]

** 21. Show that the product of two successive integers cannot be a perfect square.

*** 22. Can the product of three successive integers be a perfect square?

***** 23. Show that there are an infinity of integers $x, y, z > 1$ such that

$$x^x y^y = z^z.$$

Chapter 2

Euclid's Theorem

Theorem 2.1. *There are an infinity of primes.*

This is sometimes called Euclid's Second Theorem, what we have called Euclid's Lemma being known as Euclid's First Theorem.

Proof. Suppose to the contrary there are only a finite number of primes, say

$$p_1, p_2, \ldots, p_r.$$

Consider the number

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

Then N is not divisible by p_i for $i = 1, \ldots, r$, since N has remainder 1 when divided by each of these primes.

Take any prime factor q of N. (We know from the Fundamental Theorem that there is such a prime.)

Then q differs from all of the primes $p_1, \ldots, p_r$, since it divides N.

Hence our assumption that the number of primes is finite is untenable. □

2.1 Variants on Euclid's proof

Proposition 2.1. *There are an infinite number of primes of the form*

$$p = 4n - 1.$$

Proof. Suppose there are only a finite number of such primes, say

$$p_1, p_2, \ldots, p_r.$$

Consider the number

$$N = 4p_1 p_2 \cdots p_r - 1.$$

Since N is odd, it is a product of odd prime factors.

Any odd number is of the form $4n + 1$ or $4n - 1$. If all the prime factors of N were of the form $4n + 1$ their product N would be of this form. Since it is not, we conclude that N has a prime factor of the form $4n - 1$.

This must differ from $p_1, \ldots, p_r$, since none of these primes divides N.

Hence we have a further prime of the form $4n - 1$, contradicting our original assumption. □

Rather suprisingly, perhaps, we cannot show in the same way that there are an infinity of primes of the form $4n + 1$, although that is true.

2.2 The zeta function

Having established that there are an infinity of primes, the question arises: How are these primes distributed? Riemann's zeta function is the major tool in this study.

Definition 2.1. *Riemann's zeta function $\zeta(s)$ is defined by*

$$\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \cdots,$$

when this series converges.

Although Riemann's name is given to this function, it was in fact introduced by Euler. However, Euler only considered the function for real s. Riemann's contribution was to consider the function for complex s, in a revolutionary paper "On the number of primes less than a given value", published in 1859, using the theory of complex functions laid down by Cauchy some 20 years before.

Note that the terms in the series can be defined, for real and complex s, by

$$n^{-s} = e^{-s \ln n}.$$

We see from this that

$$n^{-(x+iy))} = e^{-x \ln n}\, e^{-iy \ln n},$$

and so

$$|n^{-s}| = n^{-\Re(s)},$$

since $|e^{i\theta}| = 1$ for all real θ.

A simple but useful tool allows us to determine when the series converges.

Lemma 2.1. *If $f(x)$ is a monotone function then*

$$\sum f(n) \text{ converges} \iff \int^{\infty} f(x)dx \text{ converges.}$$

The lower limits on each side so not matter; it is sufficient that $f(x)$ is defined for $x \geq X$.

One might think it should be specified that $f(x)$ is continuous. But in fact any monotone function $f(x)$ is necessarily Riemann integrable (and so Lebesgue integrable). This follows from the fact that $f(x)$ has only an enumerable set of discontinuities, so the partitions in Riemann sums can be chosen with end-points avoiding these points.

Proof. We may assume (replacing $f(x)$ by $-f(x)$ if necessary) that $f(x)$ is decreasing. We may also assume that $f(x) \to 0$ as $x \to \infty$; for we know that $f(x)$ tends to a limit ℓ (possibly $-\infty$), and if $\ell \neq 0$ then it is easy to see that both sum and integral diverge.

If $n \leq x \leq n+1$ then

$$f(n) \leq f(x) \leq f(n+1).$$

Hence

$$f(n) \leq \int_n^{n+1} f(x)dx \leq f(n+1).$$

Thus

$$f(m)+f(m+1)+\cdots+f(n-1) \geq \int_m^n f(x)dx \geq f(m+1)+f(m+2)+\cdots+f(n),$$

from which the result follows. ∎

Proposition 2.2. *The series for $\zeta(s)$ converges for $\Re(s) > 1$.*

Proof. For real $s > 1$ this follows from the previous lemma, since

$$\int x^{-s} dx = -\frac{1}{s-1} x^{-(s-1)}.$$

And it follows from this that $\sum n^{-s}$ is absolutely convergent if $\Re(s) > 1$, since $|n^{-s}| = n^{-\Re(s)}$. $\square$

2.3 Euler's Product Formula

If $a_1, a_2, \ldots$ is an infinite sequence of real of complex numbers, we say that the infinite product $a_1 a_2 \cdots$ converges to $\ell \neq 0$ if the partial products

$$A_n = a_1 a_2 \cdots a_n$$

converge to ℓ. (If $A_n \to 0$ then we say that the product *diverges* to 0.)

If the a_n are real and positive we can convert an infinite product to an infinite series by taking logarithms:

$$\prod a_n \text{ converges} \iff \sum \ln a_n \text{ converges.}$$

Because of this logarithmic connection we usually take the product in the form $\prod(1 + a_n)$. This allows us to pass to complex a_n provided $|(|a_n) < 1$, since in that case

$$\ln(1 + a_n) = a_n - \frac{1}{2}a_n^2 + \frac{1}{3}a_n^3 - \frac{1}{4}a_n^4 + \cdots.$$

Lemma 2.2. *Suppose $\sum a_n^2$ is absolutely convergent. Then*

$$\prod(1 + a_n) \text{ converges} \iff \sum a_n \text{ converges.}$$

In particular the product is convergent if the series is absolutely convergent.

Proof. Since

$$\begin{aligned}
&|\frac{1}{2}a_n^2 - \frac{1}{3}a_n^3 + \frac{1}{4}a_n^4 - \cdots| \\
&\qquad \leq \frac{1}{2}|a_n|^2 + \frac{1}{3}|a_n|^3 + \frac{1}{4}|a_n|^4 + \cdots \\
&\qquad \leq \frac{1}{2}\left(|a_n|^2 + |a_n|^3 + |a_n|^4 + \cdots\right) \\
&\qquad = \frac{1}{2}\frac{|a_n|^2}{1 - |a_n|^2} \\
&\qquad \leq |a_n|^2,
\end{aligned}$$

if $|a_n| \leq 1/2$.

It follows that

$$|\ln \prod_M^N (1 + a_n) - \sum_M^N a_n| \leq \sum_M^N |a_n|^2$$

provided $|a_n| \leq 1/2$ for $n \in [M, N]$, from which the result follows. $\square$

Theorem 2.2. *For $\Re(s) > 1$,*

$$\zeta(s) = \prod_p \left(1 - p^{-s}\right)^{-1},$$

where the infinite product extends over all prime numbers p.

Proof. The formula can be written

$$1 + 2^{-s} + 3^{-s} + 4^{-s} + \cdots =$$
$$\left(1 + 2^{-s} + 2^{-2s} + \cdots\right)\left(1 + 3^{-s} + 3^{-2s} + \cdots\right)\left(1 + 5^{-s} + 5^{-2s} + \cdots\right)\cdots.$$

If $n = 2^{e_2}3^{e^3}5^{e_5}\cdots$ then

$$n^{-s} = 2^{-e_2 s}3^{-e^3 s}5^{-e_5 s}\cdots;$$

and we see that n^{-s} on the left is matched by $2^{-e_2 s}$ from the first factor on the right, $3^{-e_3 s}$ from the second factor, and so on.

□

Theorem 2.3. *The series*

$$\sum \frac{1}{p}$$

(where p runs over the primes) diverges.

Proof. Taking $s = 1$ in the above formula, the series

$$\sum \frac{1}{n}$$

diverges. So the product

$$\prod \left(1 - \frac{1}{p}\right)^{-1}$$

also diverges.

It follows that the inverse

$$\prod \left(1 - \frac{1}{p}\right) = 0,$$

ie the partial product

$$P_n = \prod_1^n \left(1 - \frac{1}{p}\right) \to 0$$

as $n \to \infty$.

We say that the infinite product 'diverges to 0'.

Taking logarithms, it follows that

$$\sum_p \log\left(1 - \frac{1}{p}\right) = -\infty.$$

Recall that

$$\log(1 - x) = -x + x^2/2 - x^3/3 + \cdots.$$

If x is small, say $|x| < 1/2$, we can combine the second and later terms:

$$\begin{aligned} |x^2/2 - x^3/3 + \cdots| &\le x^2/2(1 + x + x^2 + \cdots) \\ &= \frac{x^2}{2(1-x)} \\ &\le x^2. \end{aligned}$$

Thus

$$\frac{1}{p} = -\log(1 - \frac{1}{p}) + a_p.$$

where $\sum a_p$ converges, since

$$|a_p| \le \frac{1}{p^2},$$

and $\sum 1/p^2$ converges with $\sum 1/n^2$.

We conclude that $\sum 1/p$ is the sum of a divergent series and a convergent series, and therefore diverges. □

Note that

$$\sum_p \frac{1}{p^r}$$

converges for $r > 1$, since

$$\sum_n \frac{1}{n^r}$$

converges (by comparison with the integral $\int 1/x^r$).

2.4 Dirichlet's Theorem

Theorem 2.4. *There are an infinity of primes in any arithmetic sequence*

$$a + dn \quad (n = 0, 1, 2, \dots)$$

with $d > 0$ *and* $\gcd(a, d) = 1$.

Exercise 2

In exercises 1–10 determine whether the given sum over $\mathbb{N}$ is convergent or not:

* 1. $\sum_n \frac{1}{n^{1/2}}$
* 2. $\sum_n \frac{1}{n^{3/2}}$
** 3. $\sum_n \frac{1}{n \ln n}$
** 4. $\sum_n \frac{1}{n \ln^2 n}$
** 5. $\sum_n \frac{\ln n}{n^2}$
* 6. $\sum_n \frac{(-1)^n}{n}$
** 7. $\sum_n \frac{(-1)^n}{n^{1/2}}$
** 8. $\sum_n \frac{\cos n}{n}$
*** 9. $\sum_n \frac{\tan n}{n}$
** 10. $\sum_n \sin n$

In exercises 11–13 determine whether the given sum over the primes is convergent or not:

** 11. $\sum_p \frac{1}{p \ln p}$
*** 12. $\sum_p \frac{(-1)^p}{p}$
*** 13. $\sum_p \frac{(-1)^p}{\sqrt{p}}$

*** 14. Determine $\zeta(2)$.
**** 15. Determine $\zeta(4)$.

Chapter 3

Fermat and Mersenne Primes

3.1 Fermat primes

Theorem 3.1. *Suppose $a, n > 1$. If*

$$a^n + 1$$

is prime then a is even and

$$n = 2^e$$

for some e.

Proof. If a is odd then $a^n + 1$ is even; and since it is ≥ 5 it is composite.

Suppose n has an odd factor r, say

$$n = rs.$$

We have

$$x^r + 1 = (x + 1)(x^{r-1} - x^{r-2} + x^{r-3} - \cdots + 1).$$

On substituting $x = a^s$,

$$a^s + 1 \mid a^n + 1,$$

and so $a^n + 1$ is composite.

Thus n has no odd factor, and so

$$a = 2^e.$$

□

Definition 3.1. *The number*

$$F(n) = 2^{2^n} + 1$$

is called a Fermat number; and if it is prime it is called a Fermat prime.

Thus

$$F(0) = 3,\ F(1) = 5,\ F(2) = 17,\ F(3) = 257,\ F(4) = 65537,\ F(5) = 4,294,967,297, \ldots$$

Fermat conjectured that the Fermat numbers are all prime. Sadly this has proved untrue.

$F(0)$ to $F(4)$ are indeed prime, but $F(5)$ is composite.

How do I know? There is a standard `Unix` program `factor` for factorizing numbers. Here is what I get:

```
tim@walton:~> /usr/games/factor 65537
65537: 65537
tim@walton:~> /usr/games/factor 4294967297
4294967297: 641 6700417
```

range:

```
tim@walton:~> /usr/games/primes 1000 1020
1009
1013
1019
```

No further Fermat primes have been found, and a heuristic argumnent suggests there probably are no more. (A *heuristic argument* is one that suggests a result is true, but does not prove it.)

The probability that

$$F(n) = 2^{2^n}$$

is prime is

$$\frac{1}{\ln(F(n)} \approx \frac{1}{2^n \ln 2}.$$

Thus the expected number of Fermat primes $F(n$ with $n \geq 5$ is

$$\frac{1}{\ln 2} \sum_{n \geq 5} \frac{1}{2^n} = \frac{1}{\ln 2} \frac{1}{16} \approx$$

So one could wager that there are no more Fermat primes after $F(4)$.

3.2 Mersenne primes

Theorem 3.2. *Suppose $a, n > 1$. If*

$$a^n - 1$$

is prime then $a = 2$ and n is prime.

Proof. We have

$$x^n - 1 = (x - 1)(x^{n-1} + x^{n-2} + \cdots + 1).$$

Thus

$$a - 1 \mid a^n - 1,$$

and so $a^n - 1$ is composite if $a > 2$.

Now suppose n is composite, say

$$n = rs,$$

with $r, s > 1$. We have

$$x^r + 1 = (x + 1)(x^{n-1} - x^{n-2} + x^{n-3} - \cdots + 1).$$

Substituting $x = a^s$,

$$a^s - 1 \mid a^n - 1,$$

and so $a^n - 1$ is composite.

Hence n is prime. □

Definition 3.2. *For each prime p the number*

$$M(p) = 2^p - 1$$

is called a Mersenne number; and if it is prime it is called a Mersenne prime.

We have

$$M(2) = 3,\ M(3) = 8,\ M(5) = 31,\ M(7) = 63,\ M(11) = 2047, \ldots$$

$$\frac{1}{\ln(2^p - 1)} \approx \frac{1}{p \ln 2}.$$

Thus the expected number of Mersenne primes is

$$\frac{1}{\ln 2} \sum \frac{1}{p},$$

where the sum runs over all primes.

But we have seen that

$$\sum \frac{1}{p}$$

is divergent. So this suggests (strongly) that the number of Mersenne primes is infinite.

We shall see later that there is a subtle test — the Lucas-Lehmer test — for the primality of the Mersenne number $M(p)$. This allows the primality of very large Mersenne numbers to be tested on the computer much more quickly than other numbers of the same size.

For this reason, the largest known prime is invariably a Mersenne prime; and the search for the next Mersenne prime is a popular pastime.

The Great Internet Mersenne Prime Search, or GIMPS (`http://www.mersenne.org/`), is a communal effort — which anyone can join — to find the next Mersenne prime. The record to date, the 48th known Mersenne prime, is

$$2^{57,885,161} - 1.$$

This was discovered in 2013, and has over 17 million digits.

We hope to join the search, and possibly win a large prize!

3.3 Perfect numbers

Definition 3.3. *We denote the sum of the divisors of $n > 0$ by $\sigma(n)$*

Note that we include 1 and n in the factors of n. Thus

$$\sigma(1) = 1,\ \sigma(2) = 3,\ \sigma(3) = 4,\ \sigma(4) = 7,\ \sigma(5) = 6,\ \sigma(6) = 12, \ldots$$

Definition 3.4. *The integer $n > 0$ is said to be perfect if it is the sum of its proper divisors, ie if*

$$\sigma(n) = 2n.$$

Thus 6 is the first perfect number.

Theorem 3.3. *If $M(p) = 2^p - 1$ is a Mersenne prime then*

$$n = 2^{p-1}(2^p - 1)$$

is perfect; and every even perfect number is of this form.

Proof. The number n above has factors

$$2^r \text{ and } 2^r M(p)$$

for $r = 0, 1, \ldots, p-1$, with sum

$$\sigma(n) = \left(1 + 2 + 2^2 + \cdots + 2^{p-1}\right)(1 + M(p)) = (2^p - 1)2^p = 2n.$$

Lemma 3.1. *The function $\sigma(n)$ is* multiplicative in the number-theoretic sense, *ie*

$$\gcd(m, n) = 1 \implies \sigma(mn) = \sigma(m)\sigma(n).$$

Proof. If $\gcd(m, n) = 1$ then the factors of mn are the numbers rs, where

$$n = 2^e m,$$

where m is odd. Then

$$\sigma(n) = (2^{e+1} - 1)\sigma(m).$$

But $\sigma(n) = 2n$. Thus

$$2^{e+1} m = (2^{e+1} - 1)\sigma(m).$$

It follows that

$$2^{e+1} - 1 \mid m,$$

say

$$m = (2^{e+1} - 1)q.$$

Then

$$\sigma(m) = 2^{e+1} q = m + q.$$

But m and q are both factors of m. It follows that they are the *only* factors of m. Hence $q = 1$ and

$$m = 2^{e+1} - 1$$

is prime. $\square$

It is not known if there are any odd perfect numbers. If there are, then the first one is $> 10^{1500}$.

Chapter 4

Modular arithmetic

4.1 The modular ring

Definition 4.1. *Suppose* $n \in \mathbb{N}$ *and* $x, y \in \mathbb{Z}$. *Then we say that* x, y are equivalent modulo n, *and we write*

$$x \equiv y \bmod n$$

if

$$n \mid x - y.$$

It is evident that equivalence modulo n is an equivalence relation, dividing $\mathbb{Z}$ into equivalence or *residue* classes.

Definition 4.2. *We denote the set of residue classes* mod n *by* $\mathbb{Z}/(n)$.

Evidently there are just n classes modulo n if $n \geq 1$;

$$\#(\mathbb{Z}/(n)) = n.$$

We denote the class containing $a \in \mathbb{Z}$ by $\bar{a}$, or just by a if this causes no ambiguity.

Proposition 4.1. *If*

$$x \equiv x', \ y \equiv y'$$

then

$$x + y \equiv x' + y', \ xy \equiv x'y'.$$

Thus we can add and multiply the residue classes mod$d$.

Corollary 4.1. *If* $n > 0$, $\mathbb{Z}/(n)$ *is a finite commutative ring (with 1).*

Example: Suppose $n = 6$. Then addition in $\mathbb{Z}/(6)$ is given by

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

while multiplication is given by

×	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

.

Theorem 4.1. *The ring $\mathbb{Z}/(n)$ is a field if and only if n is prime.*

Proof. Recall that an *integral domain* is a commutative ring A with 1 having no zero divisors, ie

$$xy = 0 \implies x = 0 \text{ or } y = 0.$$

In particular, a field is an integral domain in which every non-zero element has a multiplicative inverse.

The result follows from the following two lemmas.

Lemma 4.1. *$\mathbb{Z}/(n)$ is an integral domain if and only if n is prime.*

Proof. Suppose n is not prime, say

$$n = rs,$$

where $1 < r, s < n$. Then

$$\bar{r}\,\bar{s} = \bar{n} = 0.$$

So $\mathbb{Z}/(n)$ is not an integral domain.

Conversely, suppose n is prime; and suppose

$$\bar{r}\,\bar{s} = \overline{rs} = 0.$$

Then

$$n \mid rs \implies n \mid r \text{ or } n \mid s \implies \bar{r} = 0 \text{ or } \bar{s} = 0.$$

□

Lemma 4.2. *A finite integral domain A is a field.*

Proof. Suppose $a \in A$, $a \neq 0$. Consider the map

$$x \mapsto ax : A \to A.$$

This map is injective; for

$$ax = ay \implies a(x - y) = 0 \implies x - y = 0 \implies x = y.$$

But an injective map

$$f : X \to X$$

from a *finite* set X to itself is necessarily surjective.

In particular there is an element $x \in A$ such that

$$ax = 1,$$

ie a has an inverse. Thus A is a field. □

□

4.3 The additive group

If we 'forget' multiplication in a ring A we obtain an additive group, which we normally denote by the same symbol A. (In the language of category theory we have a 'forgetful functor' from the category of rings to the category of abelian groups.)

Proposition 4.2. *The additive group $\mathbb{Z}/(n)$ is a cyclic group of order n.*

This is obvious; the group is generated by the element 1 mod n.

Proposition 4.3. *The element a mod n is a generator of $\mathbb{Z}/(n)$ if and only if*

$$\gcd(a, n) = 1.$$

If $d > 1$ then 1 is not a multiple of a mod n, since

$$1 \equiv ra \bmod n \implies 1 = ra + sn \implies d \mid 1.$$

Conversely, if $d = 1$ then we can find $r, s \in \mathbb{Z}$ such that

$$ra + sn = 1;$$

so

$$ra \equiv 1 \bmod n,$$

Thus 1 is a multiple of a mod n, and so therefore is every element of $\mathbb{Z}/(n)$. $\square$

Note that there is only one cyclic group of order n, up to isomorphism. So any statement about the additive groups $\mathbb{Z}/(n)$ is a statement about finite cyclic groups, and vice versa. In particular, the result above is equivalent to the statement that if G is a cyclic group of order n generated by g then g^r is also a generator of G if and only if $\gcd(r, n) = 1$.

Recall that a cyclic group G of order n has just one subgroup of each order $m \mid n$ allowed by Lagrange's Theorem, and this subgroup is cyclic. In the language of modular arithmetic this becomes:

Proposition 4.4. *The additive group $\mathbb{Z}/(n)$ had just one subgroup of each order $m \mid n$. If $n = mr$ this is the subgroup*

$$\langle r \rangle = \{0, r, 2r, \dots, (m-1)r\}.$$

4.4 The multiplicative group

If A is a ring (with 1, but not necessarily commutative) then the *invertible elements* form a group; for if a, b are invertible, say

$$ar = ra = 1, \; bs = sb = 1,$$

then

$$(ab)(rs) = (rs)(ab) = 1,$$

and so ab is invertible.

We denote this group by $A^\times$.

Proposition 4.5. *The element $a \in \mathbb{Z}/(n)$ is invertible if and only if*

$$\gcd(a, n) = 1.$$

Proof. If a is invertible mod$n$, say

$$ab \equiv 1 \bmod n,$$

then

$$ab = 1 + tn,$$

and it follows that

$$\gcd(a, n) = 1.$$

Conversely, if this is so then

$$ax + ny = 1,$$

and it follows that x is the inverse of a mod n. $\square$

We see that the invertible elements in $\mathbb{Z}/(n)$ are precisely those elements that generate the additive group $\mathbb{Z}/(n)$.

Definition 4.3. *We denote the group of invertible elements in $\mathbb{Z}/(n)$ by*

$$0 \leq r < n \text{ and } \gcd(r, n) = 1.$$

This function is called *Euler's totient function.* As we shall see, it plays a very important role in elementary number theory.

Example:

$$\begin{aligned}
\phi(0) &= 0, \\
\phi(1) &= 1, \\
\phi(2) &= 1, \\
\phi(3) &= 2, \\
\phi(4) &= 2, \\
\phi(5) &= 4, \\
\phi(6) &= 2.
\end{aligned}$$

It is evident that if p is prime then

$$\phi(p) = p - 1,$$

since every number in $[0, p)$ except 0 is coprime to p.

Proposition 4.6. *The order of the multiplicative group* $(\mathbb{Z}/n)^\times$ *is* $\phi(n)$

This follows from the fact that each class can be represented by a remainder $r \in [0, n)$.

Example: Suppose $n = 10$. Then the multiplication table for the group $(\mathbb{Z}/10)^\times$ is

$$\begin{array}{c|cccc}
 & 1 & 3 & 7 & 9 \\
\hline
1 & 1 & 3 & 7 & 9 \\
3 & 3 & 9 & 1 & 7 \\
7 & 7 & 1 & 9 & 3 \\
9 & 9 & 7 & 3 & 1
\end{array}.$$

We see that this is a cyclic group of order 4, generated by 3:

$$(\mathbb{Z}/10)^\times = C_4.$$

Suppose $\gcd(a, n) = 1$. To find the inverse x of a mod n we have in effect to solve the equation

$$ax + ny = 1.$$

As we have seen, the standard way to solve this is to use the Euclidean Algorithm, in effect to determine $\gcd(a, n)$.

Example: Let us determine the inverse of 17 mod 23. Applying the Euclidean Algorithm,

$$\begin{aligned}
23 &= 17 + 6, \\
17 &= 3 \cdot 6 - 1.
\end{aligned}$$

Thus

$$\begin{aligned}
1 &= 3 \cdot 6 - 17 \\
&= 3(23 - 17) - 17 \\
&= 3 \cdot 23 - 4 \cdot 17.
\end{aligned}$$

Hence

$$17^{-1} = -4 = 19 \bmod 23.$$

Note that having found the inverse of a we can easily solve the congruence

$$ax = b \bmod n$$

In effect

$$x = a^{-1}b.$$

For example, the solution of

$$17x = 9 \bmod 23$$

Suppose $m \mid n$. Then each remainder mod n defines a remainder mod m.

For example, if $m = 3,\ n = 6$ then

$$\begin{aligned}
0 \bmod 6 &\mapsto 0 \bmod 3,\\
1 \bmod 6 &\mapsto 1 \bmod 3,\\
2 \bmod 6 &\mapsto 2 \bmod 3,\\
3 \bmod 6 &\mapsto 0 \bmod 3,\\
4 \bmod 6 &\mapsto 1 \bmod 3,\\
5 \bmod 6 &\mapsto 2 \bmod 3.
\end{aligned}$$

Proposition 4.7. *If $m \mid n$ the map*

$$r \mod n \mapsto r \mod n$$

is a ring-homomorphism

$$\mathbb{Z}/(n) \to \mathbb{Z}/(m).$$

4.6 Finite fields

We have seen that $\mathbb{Z}/(p)$ is a field if p is prime.

Finite fields are important because linear algebra extends to vector spaces over any field; and vector spaces over finite fields are central to coding theory and cryptography, as well as other branches of pure mathematics.

Definition 4.5. *The* characteristic *of a ring A is the least positive integer n such that*

$$\overbrace{1 + 1 + \cdots + 1}^{n \text{ 1's}} = 0.$$

If there is no such n then A is said to be of characteristic 0.

Thus the characteristic of A, if finite, is the order of 1 in the additive group A.

Evidently $\mathbb{Z},\ \mathbb{Q},\ \mathbb{R},\ \mathbb{C}$ are all of characteristic 0.

Proposition 4.8. *The ring $\mathbb{Z}/(n)$ is of characteristic n.*

Proposition 4.9. *The characteristic of a finite field is a prime.*

Proof. Let us write

$$n \cdot 1 \text{ for } \overbrace{1 + 1 + \cdots + 1}^{n \text{ 1's}}.$$

Suppose the order n is composite, say $n = rs$. By the distributive law,

$$n \cdot 1 = (r \cdot 1)(s \cdot 1).$$

There are no divisors of zero in a field; hence

$$r \cdot 1 = 0 \text{ of } s \cdot 1 = 0,$$

contradicting the minimality of n. ∎

The proof shows in fact that the characteristic of any field is either a prime or 0.

Proposition 4.10. *Suppose F is a finite field of characteristic p. Then F contains a subfield isomorphic to $\mathbb{Z}/(p)$.*

Proof. Consider the additive subgroup generated by 1:

$$\langle 1 \rangle = \{0, 1, 2 \cdot 1, \ldots, (p-1) \cdot 1\}.$$

It is readily verified that this set is closed under addition and multiplication;

Corollary 4.2. *There is just one field containing p elements, up to isomorphism, namely $\mathbb{Z}/(p)$.*

Theorem 4.2. *A finite field F of characteristic p contains p^n elements for some $n \geq 1$*

Proof. We can consider F as a vector space over its prime subfield P. Suppose this vector space is of dimension n. Let $e_1, \ldots, e_n$ be a basis for the space. Then each element of F is uniquely expressible in the form

$$a_1 e_1 + \cdots + a_n e_n,$$

where $a_1, \ldots, a_n \in P$. There are just p choices for each a_i. Hence the total number of choices, ie the number of elements in F, is p^n. $\square$

Theorem 4.3. *There is just one field F containing $q = p^n$ elements for each $n \geq 1$, up to isomorphism.*

Thus there are fields containing 2,3,4 and 5 elements, but not field containing 6 elements.

We are not going to prove this theorem until later.

Definition 4.6. *We denote the field containing $q = p^n$ elements by $\mathbb{F}_q$.*

The finite fields are often called *Galois fields*, after Evariste Galois who discovered them.

In Exercises 1–5 determine the additive order of the given element.

* 1. 3 mod 5

* 2. 3 mod 6

* 3. 2 mod 7

* 4. −13 mod 14

** 5. 100000 mod 123456

In Exercises 6–10 determine the multiplicative order of the given element.

* 6. 3 mod 5

* 7. 7 mod 12

** 8. 2 mod 31

** 9. −2 mod 31

*** 10. 2 mod 3^5

In Exercises 11–15 determine the multiplicative inverse of the given element.

* 11. 3 mod 5

* 12. 3 mod 13

* 13. 2 mod 111

** 14. 137 mod 253

In Exercises 16–20 determine the order of the given multiplicative group, and list its elements.

* 15. $(\mathbb{Z}/2)^\times$

* 16. $(\mathbb{Z}/6)^\times$

* 17. $(\mathbb{Z}/8)^\times$

* 18. $(\mathbb{Z}/12)^\times$

* 19. $(\mathbb{Z}/15)^\times$

* 20. Determine $\phi(45)$

* 21. Determine $\phi(3^n)$

* 22. Determine all positive integers n with $\phi(n) = n - 1$.

** 23. Determine all positive integers n with $\phi(n) = n - 2$.

** 24. What is the smallest value of $\phi(n)/n$?

** 25. Show that there is a field containing 4 elements.

** 26. Show that there is no field containing 6 elements.

Chapter 5

The Chinese Remainder Theorem

5.1 Coprime moduli

Theorem 5.1. *Suppose $m, n \in \mathbb{N}$, and*

$$\gcd(m, n) = 1.$$

Given any remainders $r \bmod m$ and $s \bmod n$ we can find N such that

$$N \equiv r \bmod m \text{ and } N \equiv s \bmod n.$$

Moreover, this solution is unique $\bmod mn$.

Proof. We use the pigeon-hole principle. Consider the mn numbers

$$0 \le N < mn.$$

For each N consider the remainders

$$r = N \bmod m, \; s = N \bmod n,$$

where r, s are chosen so that

$$0 \le r < m, \; 0 \le s < n.$$

We claim that these pairs r, s are different for different $N \in [0, mn)$. For suppose $N < N'$ have the same remainders, ie

$$N' \equiv N \bmod m \text{ and } N' \equiv N \bmod n.$$

Then

$$m \mid N' - N \text{ and } n \mid N' - N.$$

Since $\gcd(m, n) = 1$, it follows that

$$mn \mid N' - N.$$

But that is impossible, since

$$0 < N' - N < mn.$$

$\square$

Example: Let us find N such that

$$N \equiv 3 \bmod 13, \; N \equiv 7 \bmod 23.$$

One way to find N is to find a, b such that

$$a \equiv 1 \bmod m, \; a \equiv 0 \bmod n,$$
$$b \equiv 0 \bmod m, \; b \equiv 1 \bmod n.$$

For then we can take

$$23 = 2 \cdot 13 - 3,$$
$$13 = 4 \cdot 3 + 1,$$

giving

$$\begin{aligned} 1 &= 13 - 4 \cdot 3 \\ &= 13 - 4(2 \cdot 13 - 23) \\ &= 4 \cdot 23 - 7 \cdot 13. \end{aligned}$$

Thus we can take

$$a = 4 \cdot 23 = 92, \ b = -7 \cdot 13 = -91.$$

giving

$$N = 3 \cdot 92 - 7 \cdot 91 = 276 - 637 = -361.$$

Of course we can add a multiple of mn to N; so we could take

$$N = 13 \cdot 23 - 361 = 299 - 361 = -62,$$

if we want the smallest solution (by absolute value); or

$$N = 299 - 62 = 237,$$

for the smallest positive solution.

5.2 The modular ring

We can express the Chinese Remainder Theorem in more abstract language.

Theorem 5.2. *If* $\gcd(m, n) = 1$ *then the ring* $\mathbb{Z}/(mn)$ *is isomorphic to the product of the rings* $\mathbb{Z}/(m)$ *and* $\mathbb{Z}/(n)$*:*

$$\mathbb{Z}/(mn) = \mathbb{Z}/(m) \times \mathbb{Z}/(n).$$

Proof. We have seen that the maps

$$N \mapsto N \bmod m \text{ and } N \mapsto N \bmod n$$

define ring-homomorphisms

$$\mathbb{Z}/(mn) \to \mathbb{Z}/(m) \text{ and } \mathbb{Z}/(mn) \to \mathbb{Z}/(n).$$

These combine to give a ring-homomorphism

$$\mathbb{Z}/(mn) \to \mathbb{Z}/(m) \times \mathbb{Z}/(n),$$

under which

$$r \bmod mn \mapsto (r \bmod m, r \bmod n).$$

But we have seen that this map is bijective; hence it is a ring-isomorphism. □

5.3 The totient function

Proposition 5.1. *Suppose* $\gcd(m, n) = 1$. *Then*

$$\gcd(N, mn) = \gcd(N, m) \cdot \gcd(N, n).$$

Proof. Let

$$d = \gcd(N, mn).$$

Suppose

$$p^e \parallel d.$$

$$\gcd(N, mn) = 1 \iff \gcd(N, m) = 1 \textit{ and } \gcd(N, n) = 1.$$

From this we derive

Theorem 5.3. *Euler's totient function is multiplicative, ie*

$$\gcd(m, n) = 1 \implies \phi(mn) = \phi(m)\phi(n).$$

This gives a simple way of computing $\phi(n)$.

Proposition 5.2. *If*

$$n = \prod_{1 \leq i\ er} p_i^{e_i},$$

where the primes $p_1, \ldots, p_r$ are different and each $e_i/ge1$. Then

$$\phi(n) = \prod p_i^{e_i - 1}(p_i - 1).$$

Proof. Since $\phi(n)$ is multiplicative,

$$\phi(n) = \prod_i \phi(p_i^{e_i}).$$

The result now follows from

Lemma 5.1. $\phi(p^e) = p^{e-1}(p - 1)$.

Proof. The numbers $r \in [0, p^e)$ is *not* coprime to p^r if and only if it is divisible by p, ie

$$r \in \{0, p, 2p, \ldots, p^e - p\}.$$

There are

$$[p^e/p] = p^{e-1}$$

such numbers. Hence

$$\phi(p^e) = p^e - p^{e-1} = p^{e-1}(p - 1).$$

□

□

Example: Suppose $n = 1000$.

$$\begin{aligned}
\phi(1000) &= \phi(2^3 5^3) \\
&= \phi(2^3)\phi(5^3) \\
&= 2^2(2 - 1)\, 5^2(5 - 1) \\
&= 4 \cdot 1 \cdot 25 \cdot 4 \\
&= 400;
\end{aligned}$$

there are just 400 numbers coprime to 1000 between 0 and 1000.

5.4 The multiplicative group

Theorem 5.4. *If* $\gcd(m, n) = 1$ *then*

$$(\mathbb{Z}/mn)^\times = (\mathbb{Z}/m)^\times \times (\mathbb{Z}/n)^\times.$$

Proof. We have seen that the map

$$r \bmod mn \mapsto (r \bmod m, r \bmod n) : \mathbb{Z}/(mn) \to \mathbb{Z}/(m) \times \mathbb{Z}/(n)$$

maps r coprime to mn to pairs (r, s) coprime to m, n respectively. Thus the subset $(Z/mn)^\times$ maps to the product of the subsets $(\mathbb{Z}/m)^\times$ and $(\mathbb{Z}/n)^\times$, The Chinese Remainder Theorem extends to more than two moduli.

Proposition 5.3. *Suppose $n_1, n_2, \dots, n_r$ are pairwise coprime, ie*

$$i \neq j \implies \gcd(n_i, n_j) = 1;$$

and suppose we are given remainders $a_1, a_2, \dots, a_r$ moduli $n_1, n_2, \dots, n_r$, respectively. Then there exists a unique $N \bmod n_1 n_2 \cdots n_r$ such that

$$N \equiv a_1 \bmod n_1,\ N \equiv a_2 \bmod n_2, \dots, N \equiv a_r \bmod n_r.$$

Proof. This follows from the same pigeon-hole argument that we used to establish the Chinese Remainder Theorem.

Or we can prove it by induction on r; for since

$$\gcd(n_1 n_2 \cdots n_i, n_{i+1}) = 1,$$

we can add one modulus at a time,

Thus if we have found N_i such that

$$N_i \equiv a_1 \bmod n_1,\ N_i \equiv a_2 \bmod n_2, \dots, N_i \equiv a_i \bmod n_i$$

then by the Chinese Remainder Theorem we can find N_{i+1} such that

$$N_{i+1} \equiv N_i \bmod n_1 n_2 \cdots n_i \text{ and } N_{i+1} \equiv a_{i+1} \bmod n_{i+1}$$

and so

$$N_{i+1} \equiv a_1 \bmod n_1,\ N_{i+1} \equiv a_2 \bmod n_2, \dots, N_{i+1} \equiv a_{i+1} \bmod n_{i+1},$$

establishing the induction. ☐

Example: Suppose we want to solve the simultaneous congruences

$$n \equiv 4 \bmod 5,\ n \equiv 2 \bmod 7,\ n \equiv 1 \bmod 8.$$

There are two slightly different approaches to the task.

Firstly, we can start by solving the first 2 congruences. As is easily seen, the solution is

$$n \equiv 9 \bmod 35.$$

The problem is reduced to two simultaneous congruences:

$$n \equiv 9 \bmod 35,\ n \equiv 1 \bmod 8,$$

which we can solve with the help of the Euclidean Algorithm, as before.

Alternatively, we can find solutions of the three sets of simultaneous congruences

$$\begin{aligned} n_1 &\equiv 1 \bmod 5,\ n_1 \equiv 0 \bmod 7,\ n_1 \equiv 0 \bmod 8, \\ n_2 &\equiv 0 \bmod 5,\ n_2 \equiv 1 \bmod 7,\ n_2 \equiv 0 \bmod 8, \\ n_3 &\equiv 0 \bmod 5,\ n_3 \equiv 0 \bmod 7,\ n_3 \equiv 1 \bmod 8, \end{aligned}$$

ie

$$\begin{aligned} n_1 &\equiv 1 \bmod 5,\ n_1 \equiv 0 \bmod 56, \\ n_2 &\equiv 1 \bmod 7,\ n_2 \equiv 0 \bmod 40, \\ n_3 &\equiv 1 \bmod 8,\ n_3 \equiv 0 \bmod 35, \end{aligned}$$

which we can solve by our previous method. The required solution is then

$$n = 4n_1 + 2n_2 + n_3,$$

We have seen that $\phi(n)$ is multiplicative. There are several other multiplicative functions that play an important role in number theory, for example:

1. The number $d(n)$ of divisors of n, eg
$$d(2) = 1,\ d(12) = 3,\ d(32) = 5.$$

2. The sum $\sigma(n)$ of the divisors of n, eg
$$\sigma(2) = 3,\ \sigma(12) = 28,\ \sigma(32) = 63.$$

3. The Möbius function
$$\mu(n) = \begin{cases} (-1)^e & \text{if } n \text{ is square-free and has } e \text{ prime factors,} \\ 0 & \text{if } n \text{ has a square factor } n = p^2 m. \end{cases}$$

4. The function $(-1)^n$.

5. The function
$$\theta(n) = \begin{cases} 1 & \text{if } n \equiv 1 \bmod 4, \\ -1 & \text{if } n \equiv 3 \bmod 4, \\ 0 & \text{if } n \text{ is even.} \end{cases}$$

5.7 Perfect numbers

Definition 5.1. *We say that $n \in \mathbb{N}$ is* perfect *if it is the sum of all its divisors, except for n itself.*

In other words,
$$n \text{ is perfect} \iff \sigma(n) = 2n.$$

Theorem 5.5. *If $M(p) = 2^p - 1$ is prime then*
$$n = 2^{p-1} M(p)$$
is perfect. Moreover, every even *perfect number is of this form*

Remark: Euclid showed that every number of this form is perfect; Euler showed that every even perfect number is of this form.

Proof. Note that
$$\sigma(n) = n + 1 \iff n \text{ is prime}.$$
For if $n = ab$ (where $a, b > 1$) then $\sigma(n) \geq n + 1 + a$.

Also
$$\sigma(2^e) = 1 + 2 + 2^2 + \cdots + 2^e = 2^{e+1} - 1.$$

Thus if $n = 2^{p-1} M(p)$, where $P = M(p)$ is prime, then (since 2^e and $M(p)$ are coprime)
$$\begin{aligned} \sigma(n) &= \sigma(2^{p-1})\sigma(M(p) \\ &= (2^p - 1)(M(p) + 1) \\ &= (2^p - 1)(2^p) \\ &= 2n. \end{aligned}$$

Conversely, suppose n is an even perfect number. Let $n = 2^e m$, where m is odd. Then
$$\sigma(n) = \sigma(2^e)\sigma(m) = 2n,$$
ie

But x is a factor of m. So if x is not 1 or m then

$$\sigma(m) \geq m + x + 1.$$

Hence $x = 1$ or m If $x = m$ then $2^{e+1} - 1 = 1 \implies e = 0$, which is not possible since n is even.

It follows that $x = 1$, so that

$$m = 2^{e+1} - 1 = M(e+1).$$

Also

$$\sigma(m) = m + 1.$$

Thus $m = M(e+1)$ is prime (and therefore $e + 1 = p$ is prime), and

$$n = 2^{p-1} M(p),$$

as stated. $\square$

But what if n is odd? *It is not known if there are any odd perfect numbers.* This is one of the great unsolved problems of mathematics.

In Exercises 1–16 determine all solutions of the given congruence.

* 1. $3x \equiv 1 \bmod 23$

* 2. $7x \equiv 1 \bmod 47$

** 3. $5x \equiv 2 \bmod 210$

** 4. $6x \equiv 7 \bmod 25$

** 5. $8x \equiv 5 \bmod 31$

** 6. $8x \equiv 12 \bmod 32$

** 7. $12x \equiv 6 \bmod 21$

** 8. $2x \equiv 2 \bmod 16$

** 9. $20x \equiv 8 \bmod 24$

*** 10. $7x \equiv -3 \bmod 2009$

** 11. $x^2 \equiv 1 \bmod 12$

** 12. $x^2 \equiv -1 \bmod 15$

** 13. $x^2 + x + 1 \equiv 0 \bmod 3$

** 14. $x^2 - 2x + 3 \equiv 0 \bmod 5$

** 15. $x^2 - 2 \equiv 0 \bmod 7$

*** 16. $x^4 + 2x^2 + x - 2 \equiv 0 \bmod 7$

* 17. What is the order of 10 in the additive group $\mathbb{Z}/(24)$?

** 18. Determine the orders of the elements $7, 11, 21$ in the multiplicative group $(\mathbb{Z}/36)^\times$.

** 19. What is the order of the group $(\mathbb{Z}/36)^\times$?

*** 20. Is the group $(\mathbb{Z}/36)^\times$ cyclic?

$a_1, ..., a_{11}$ of numbers from $\{-1, 0, 1\}$ such that the sum

$$a_1x_1 + ... + a_{11}x_{11}$$

is divisible by 2009.

*** 23. Construct the field containing 4 elements.

**** 24. Show that there is no field containing 6 elements.

*** 25. Determine the orders of all the elements in $\mathbb{F}_{11}^{\times}$?

** 26. What is the order of the multiplicative group $\mathbb{F}_q^{\times}$?

*** 27. How many elements are there of order 4 in $\mathbb{F}_{17}^{\times}$?

*** 28. Prove that there is a multiple of 2009 which ends with the digits 000001.

Polynomial Rings

6.1 Polynomials

A polynomial of degree n over a ring A is an expression of the form

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

where $a_i \in A$ and $a_n \neq 0$.

(It is better not to think of $f(x)$ as a *function*, since a non-zero polynomial may take the value 0 for all $x \in A$, particularly if A is finite.)

We know how to add and multiply polynomials, so the polynomials over A form a ring.

Definition 6.1. *We denote the ring of polynomials over the ring A by $A[x]$.*

In practice we will be concerned almost entirely with polynomials over a field k. We will assume in the rest of the chapter that k denotes a field.

In this case we do not really distinguish between $f(x)$ and $cf(x)$, where $c \neq 0$. To this end we often restrict the discussion to *monic* polynomials, ie polynomials with leading coefficient 1:

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

6.2 Long division

Proposition 6.1. *Suppose k is a field, and suppose $f(x), g(x) \in k[x]$, with $g(x) \neq 0$. Then there exist unique polynomials $q(x), r(x) \in k[x]$ with $\deg(r(x)) < \deg(g(x))$ such that*

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

Proof. We begin by listing some obvious properties of the degree of a polynomial over a field:

Lemma 6.1.

1. *$\deg(f+g) \leq \max(\deg(f), \deg(g))$;*
2. *$\deg(fg) = \deg(f)\ \deg(g)$.*

The existence of $q(x)$ and $r(x)$ follows easily enough by induction on $\deg(f(x))$. To see that the result is unique, suppose

$$f(x) = q_1(x)g(x) + r_1(x) = q_2(x)g(x) + r_2(x)$$

Then

$$g(x)(q_1(x) - q_2(x)) = r_2(x) - r_1(x).$$

The term on the left has degree $\geq \deg(g(x))$, while that on the right has degree $< \deg(g(x))$. □

6.3 Irreducibility

Definition 6.2. *The polynomial $p(x) \in k[x]$ is said to be irreducible if it cannot be factorised into polynomials of lower degree:*

$$p(x) = g(x)h(x) \implies g(x) \text{ of } h(x) \text{ is constant.}$$

In particular, any linear polynomial (ie of degree 1) is irreducible.

$$e(x) \mid f(x),\ g(x) \implies e(x) \mid d(x).$$

Furthermore, there exist polynomials $u(x), v(x)$ such that

$$d(x) = u(x)f(x) + v(x)g(x).$$

Proof. The Euclidean Algorithm extends almost unchanged; the only difference is that $\deg(r(x))$ takes the place of $|r|$.

Thus first we divide $f(x)$ by $g(x)$:

$$f(x) = q_0(x)g(x) + r_0(x),$$

where $\deg(r_0(x)) < \deg(g(x))$.

If $r_0(x) = 0$ we are done; otherwise we divide $g(x)$ by $r_0(x)$:

$$g(x)(x) = q_1(x)r_0(x) + r_1(x),$$

where $\deg(r_1(x)) < \deg(r_0(x))$.

Since the polynomials are reducing in degree, we must reach 0 after at most $\deg(g(x))$ steps. It follows, by exactly the same argument we used with the Euclidean Algorithm in $\mathbb{Z}$, that the last non-zero remainder $r_s(x)$ is the required gcd:

$$\gcd(f(x), g(x)) = r_s(x).$$

The last part of the Proposition, the fact that $d(x)$ is a linear combination (with polynomial coefficients) of $f(x)$ and $g(x)$, follows exactly as before. □

6.5 Unique factorisation

Theorem 6.1. *A monic polynomial $f(x) \in k[x]$ can be expressed as a product of irreducible monic polynomials, and the expression is unique up to order.*

Proof. If $f(x)$ is not itself irreducible then $f(x) = g(x)h(x)$, where $g(x),\ h(x)$ are of lower degree. The result follows by induction on $\deg(f(x))$.

To prove uniqueness we establish the polynomial version of Euclid's Lemma;

Lemma 6.2. *If $p(x)$ is irreducible then*

$$p(x) \mid f(x)\, g(x) \implies p(x) \mid f(x) \text{ or } p(x) \mid g(x).$$

Proof. As with the classic Euclidean Algorithm, suppose $p(x) \nmid f(x)$. Then

$$\gcd(p(x), f(x)) = 1.$$

Hence there exist $u(x), v(x)$ such that

$$u(x)p(x) + v(x)f(x) = 1.$$

Multiplying by $g(x)$,

$$u(x)p(x)g(x) + v(x)f(x)g(x) = g(x).$$

Now $p(x)$ divides both terms on the left. Hence $p(x) \mid g(x)$, as required. □

To prove uniqueness, we argue by induction on $\deg(f(x))$. Suppose

$$f(x) = p_1(x) \cdots p_r(x) = q_1(x) \cdots q_s(x).$$

Then $p_1(x) \mid q_j(x)$, and so $p_1(x) = q_j(x)$, for some j; and the result follows on applying the inductive hypothesis to

$$f(x)/p_1(x) = p_2(x) \cdots p_r(x) = q_1(x) \cdots q_{r-1}(x)q_{r+1}(x) \cdots q_s(x).$$

Proof. Suppose $f(x)$ is coprime to $p(x)$, ie represents a non-zero element of $k[x] \bmod p(x)$. Then we can find polynomials $u(x), v(x)$ such that

$$f(x)u(x) + p(x)v(x) = 1,$$

But then

$$f(x)u(x) \equiv 1 \bmod p(x),$$

ie $fx)$ has the inverse $u(x)$ modulo $p(x)$. □

This is particularly striking if k is a prime field $\mathbb{F}_p$.

Corollary 6.1. *Suppose $f(x) \in \mathbb{F}_p[x]$ is an irreducible polynomial of degree n. Then $K = \mathbb{F}_p[x]/(f(x))$ is a finite field with p^n elements.*

Proof. This follows from the fact that the residues modulo $f(x)$ are represented by the p^n polynomials

$$a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} \qquad (0 \le a_0, a_1, \ldots, a_{n-1} < p).$$

□

Example: Let us look at the first irreducible polynomials in $\mathbb{F}_2[x]$.

Every linear polynomial $x - c$ in $k[x]$ is irreducible, by definition. Thus there are two irreducible polynomials of degree 1 in $\mathbb{F}_2[x]$: x and $x + 1$.

If one of the four polynomials of degree 2 is not irreducible then it must be one of the 3 products of x and $x + 1$,

$$x^2, x(x+1) = x^2 + x, (x+1)^2 = x^2 + 1.$$

This leave one irredicible polynomial of degree 2: $x^2 + x + 1$.

Turning to the eight polynomials of degree 3, there are four linear products:

$$x^3,\ x^2(x+1) = x^3 + x,\ x(x+1)^2 = x^3 + x,\ (x+1)^3 = x^3 + x^2 + x + 1.$$

There are two other 'composite' polynomials:

$$x(x^2 + x + 1) = x^3 + x^2 + x + 1,\ (x+1)(x^2 + x + 1) = x^3 + 1.$$

We are left with two irreducibles:

$$x^3 + x^2 + 1,\ x^3 + x + 1.$$

Each polynomial of degree d in $F_2[x]$ can be represented by d digits. Thus the irreducible polynomials listed above can be written:

$$10,\ 11,\ 111,\ 1101,\ 1011, \ldots.$$

These compare with the familar prime numbers, in binary form:

$$10,\ 11,\ 101,\ 111,\ 1001, \ldots.$$

The field $\mathbb{F}_2[x]/(x^2 + x + 1)$ has 4 elements, represented by the residues $0, 1, x, x + 1$. The addition and multiplication tables for this field of order 4 are

$+$	0	1	x	$x+1$
0	0	1	x	$x+1$
1	1	0	$x+1$	x
x	x	$x+1$	0	1
$x+1$	$x+1$	x	1	0

$\times$	0	1	x	$x+1$
0	0	0	0	0
1	0	1	x	$x+1$
x	0	x	$x+1$	1
$x+1$	0	$x+1$	1	x

Proof.

Lemma 6.3. *Each polynomial $f(x) \in \mathbb{Q}[x]$ can be expressed in the form*

$$f(x) = qF(x)$$

where $q \in \mathbb{Q}$, $F(x) \in \mathbb{Z}[x]$ and the coefficients of $F(x)$ are coprime; moreover, this expression is unique up to $\pm$.

Proof. It is evident that $f(x)$ can be brought to this form, by multiplying by the lcm of the coefficients and then taking out the gcd of the resulting integer coefficients.

If there were two such expressions, then multiplying across we would have

$$n_1F_1(x) = n_2F_2(x).$$

The gcd of the coefficients on the left is $|n_1|$, while the gcd of those on the right is $|n_2|$. Thus $n_1 = \pm n_2$, and the result follows. □

Lemma 6.4. *Suppose*

$$u(x) = v(x)w(x),$$

where $u(x), v(x), w(x) \in \mathbb{Z}[x]$. If the coefficients of $v(x)$ are coprime, and those of $w(x)$ are also coprime, then the same is true of $u(x)$.

Proof. Suppose to the contrary that the prime p divides all the coefficients of $f(x)$. Let

$$v(x) = b_rx^r + \cdots + b_0,\ w(x) = c_sx^s + \cdots + c_0,\ u(x) = a_{r+s}x^{r+s} + \cdots + a_0.$$

By hypothesis, p does not divide all the b_i, or all the c_j. Suppose

$$p \mid b_r, b_{r-1}, \ldots, b_{i+1} \text{ but } p \nmid b_i,$$

and similarly

$$p \mid c_s, c_{s-1}, \ldots, c_{j+1} \text{ but } p \nmid c_j,$$

Then

$$p \nmid a_{i+j} = b_{i+j}c_0 + b_{i+j-1}c_1 + \cdots + b_ic_j + b_{i-1}c_{j+1} + \cdots + b_0c_{i+j},$$

for p divides every term in the sum except b_ic_j, which it does not divide since

$$p \mid b_ic_j \implies p \mid b_i \text{ or } p \mid c_j.$$

So p does not divide all the coefficients of $u(x)$, contrary to hypothesis. □

Writing $f(x), g(x), h(x)$ in the form of the first Lemma,

$$q_1F(x) = (q_2G(x))(q_3H(x)),$$

where the coefficients of each of $F(x), G(x), H(x)$ are coprime integers. Thus

$$F(x) = (q_2q_3/q_1)G(x)H(x).$$

Since the coefficients of both $F(x)$ and $G(x)H(x)$ are coprime, by the second Lemma they are equal up to sign, and the result follows. □

6.8 Euclidean domains, PIDs and UFDs

Definition 6.3. *An integral domain A is said to be a euclidean domain if there exists a function $N : A \to \mathbb{N}$ such that $N(a) = 0 \iff a = 0$, and given $a, b \in A$ with $b \neq 0$ there exists $q, r \in A$ with*

2. $a \in A, b \in I \implies ab \in I$,

Example: The whole ring A is an ideal in A, and so is the set $\{0\}$.

If $a \in A$ then $(a) = \{ax : x \in A\}$ is an ideal. An ideal of this form is said to be principal.

If $a, b \in A$ then

$$b \mid a \iff (a) \subset (b).$$

Also

$$(a) = (b) \iff b = eb,$$

where e is a unit.

Definition 6.6. *An integral domain A is said to be a* principal ideal domain *(PID) if every ideal $I \subset A$ is principal: $I = (a)$ for some $a \in A$.*

Proposition 6.5. *A euclidean domain is a principal ideal domain.*

Proof. Suppose I is an ideal in the euclidean domain A. If $I \neq (0)$ let $d \in I$ be a non-zero element with minimal $N(d)$. Suppose $a \in I$. Then $d \mid a$, for else

$$a = qd + r,$$

with $N(r) < N(d)$; and then $r \in I$ contradicts the definition of d. □

Definition 6.7. *An element p in an integral domain A is said to be* primitive *if $p \mid ab \implies p \mid a$ or $p \mid b$.*

Proposition 6.6. *A primitive element p cannot be factored; if $p = ab$ then either a or b*

Proof. Since $p \mid p = ab$, $p \mid a$ or $p \mid b$. Suppose $p \mid a$, say $a = pc$. Then $p = pcb \implies bc = 1$, so that b is a unit. □

Definition 6.8. *A* unique factorisation domain *(UFD) is an integral domain A with the property that every non-zero element $a \in A$ is expressible in the form*

$$a = ep_1p_2 \ldots p_r,$$

where e is a unit and $p_1, p_2, \ldots, p_r$ are primitive elements.

We allow $a = e$ with $r = 0$. Also, we note that we can omit e if $r \geq 1$ since ep is primitive if p is primitive.

Theorem 6.3. *A principal ideal domain is a unique factorisation domain:*

$$\textit{PID} \implies \textit{UFD}.$$

Proof. Suppose A is a PID; and suppose $a \in A$, $a \neq 0$. We may assume that a is not a unit, since the result holds trivially (with no primitive elements) in that case.

We must show that a cannot be factorised into an arbitrarily large number of non-units. Suppose that is false.

Then in particular $x = y_0z_0$, where y_0, z_0 are non-units. One of y_0, z_0, say y_0, can be factorised into an arbitrarily large number of non-units. In particular $y_0 = y_1z_1$, where y_1, z_1 are non-units. One of y_1, z_1, say y_1, can be factorised into an arbitrarily large number of non-units. In particular $y_1 = y_2z_2$, where y_2, z_2 are non-units.

Continuing in this way, we obtain an infinite sequence

$$y_1, y_2, y_3, \ldots,$$

such that $y_{i+1} \mid y_i$ for all i. Thus

$$(y_1) \subset (y_2) \subset (y_3) \subset \cdots$$

Let

$$I = (y_1) \cup (y_2) \cup (y_3) \cup \cdots.$$

It is readily verified that I is an ideal. Since A is a PID, it follows that $I = (d)$ for some $d \in A$. Thus $d \in (y_n)$ for some n. But $y_{n+1} \in (d)$. It follows

*** 5. Determine the irreducible polynomials of degree 3 over $\mathbb{F}_3$.

*** 6. How many irreducible polynomials are there of degree 4 over $\mathbb{F}_3$?

** 7. Determine the irreducible polynomials of degree 2 over $\mathbb{F}_5$.

** 8. Determine the irreducible polynomials of degree 2 over $\mathbb{F}_7$.

** 9. Show that an irreducible polynomial over $\mathbb{R}$ is of degree 1 or 2.

** 10. Determine the irreducible polynomials over $\mathbb{C}$.

In exercises 11–20 determine if the given polynomial is irreducible over $\mathbb{Q}$.

** 11. $x^2 + x + 1$

** 12. $x^3 + 2x + 1$

*** 13. $x^4 + 1$

*** 14. $x^4 + 2$

*** 15. $x^4 + 4$

*** 16. $x^4 + 4x^3 + 1$

** 17. Determine the irreducible polynomials of degree 2 over $\mathbb{F}_7$.

Finite fields

7.1 The order of a finite field

Definition 7.1. *The* characterisitic *of a ring A is the additive order of 1, ie the smallest integer $n > 1$ such that*

$$n \cdot 1 = \underbrace{1 + 1 + \cdots + 1}_{n \text{ terms}} = 0,$$

if there is such an integer, or ∞ if there is not.

Examples: $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ all have infinite characteristic.
$\mathbb{F}_p = \mathbb{Z}/(p)$ has characteristic p.

Proposition 7.1. *The characteristic of an integral domain A is either a prime p, or else ∞.*

In particular, a finite field has prime characteristic.

Proof. Suppose A has characteristic $n = ab$ where $a, b > 1$. By the distributive law,

$$\underbrace{1 + \cdots + 1}_{n \text{ terms}} = (\underbrace{1 + \cdots + 1}_{a \text{ terms}})(\underbrace{1 + \cdots + 1}_{b \text{ terms}}).$$

Hence

$$\underbrace{1 + \cdots + 1}_{a \text{ terms}} = 0 \text{ or } \underbrace{1 + \cdots + 1}_{b \text{ terms}} = 0,$$

contrary to the minimal property of the characteristic. □

Proposition 7.2. *Suppose the finite field F has characteristic p. Then F contains p^n elements, for some n.*

Proof. The elements $\{0, 1, 2, \ldots, p - 1\}$ form a subfield of F isomorphic to $\mathbb{F}_p$. We can consider F as a vector space over this subfield. Let $e_1, e_2, \ldots, e_n$ be a basis for this vector space. Then the elements of F are

$$x_1 e_1 + x_2 e_2 + \cdots + x_n e_n \qquad (0 \le x_1, x_2, \ldots, x_n < p).$$

Thus the order of F is p^n. □

7.2 On cyclic groups

Let us recall some results from elementary group theory.

Proposition 7.3. *The element g^i in the cyclic group C_n has order $n/\gcd(n, i)$.*

Proof. This follows from

$$(g^i)^e = 1 \iff n \mid ie \iff \frac{n}{\gcd(n, i)} \mid e.$$

□

Corollary 7.1. *C_n contains $\phi(n)$ generators, namely the elements g^i with $0 \le i < n$ for which $\gcd(n, i) = 1$.*

Proposition 7.4. *The cyclic group $C_n = \langle g \rangle$ has just one subgroup of each order $d \mid n$, namely the cyclic subgroup $C_d = \langle g^{n/d} \rangle$.*

$$\mu(n) = \begin{cases} 0 & \text{if } n \text{ has a square factor} \\ (-1)^r & \text{if } n \text{ is square-free and has } r \text{ prime factors} \end{cases}$$

Thus

$$\mu(1) = 1,\ \mu(2) = -1,\ \mu(3) = -1,\ \mu(4) = 0,\ \mu(5) = -1,$$
$$\mu(6) = 1,\ \mu(7) = -1,\ \mu(8) = 0,\ \mu(9) = 0,\ \mu(10) = 1.$$

Theorem 7.1. *Given an arithmetic function $f(n)$, suppose*

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

Then

$$f(n) = \sum_{d|n} \mu(n/d)g(n).$$

Proof. Given arithmetic functions $u(n), v(n)$ let us defined the arithmetic function $u \circ v$ by

$$(u \circ v)(n) = \sum_{d|n} u(d)v(n/d) = \sum_{n=xy} u(x)v(y).$$

(Compare the convolution operation in analysis.) This operation is commutative and associative, ie $v \circ u = u \circ v$ and $(u \circ v) \circ w = u \circ (v \circ w)$. (The latter follows from

$$((u \circ v) \circ w)(n) = \sum_{n=xyz} u(x)v(y)w(z).)$$

Lemma 7.1. *We have*

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \textit{if } n = 1 \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. Suppose $n = p_1^{e_1} \cdots p_n^{e_n}$. Then it is clear that only the factors of $p_1 \cdots p_r$ will contribute to the sum, so we may assume that $n = p_1 \cdots p_r$.

But in this case the terms in the sum correspond to the terms in the expansion of

$$\underbrace{(1-1)(1-1)\cdots(1-1)}_{r \text{ products}}$$

giving 0 unless $r = 0$, ie $n = 1$. □

Let us define $\delta(n)$, $\epsilon(n)$ by

$$\delta(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{otherwise,} \end{cases}$$
$$\epsilon(n) = 1 \text{ for all } n$$

It is easy to see that

$$\delta \circ f = f$$

for all arithmetic functions f. Also the lemma above can be written as

$$\mu \circ \epsilon = \delta,$$

while the result we are trying to prove is

$$g = \epsilon \circ f \implies f = \mu \circ g.$$

This follows since

$$x^{p^n-1} = 1,$$

ie

$$U(x) = x^{p^n-1} - 1 = 0.$$

Since this polynomial has degree $p^n - 1$, and we have $p^n - 1$ roots, it factorizes completely into linear terms:

$$U(x) = \prod_{a \in F^\times} (x - a).$$

Now suppose $d \mid p^n - 1$. Since

$$f(x) = x^d - 1 \mid U(x)$$

it follows that $x^d - 1$ factorizes completely into linear terms, say

$$f(x) = \prod_{0 \le i < d} (x - a_i).$$

Lemma 7.2. *Suppose there are $\sigma(d)$ elements of order d in $F^\times$. Then*

$$\sum_{e \mid d} \sigma(e) = d.$$

Proof. Any element of order $e \mid d$ must satisfy the equation $f(x) = 0$; and conversely any root of the equation must be of order $e \mid d$. The result follows on adding the elements of each order. □

Lemma 7.3. *We have*

$$\sum_{e \mid d} \phi(e) = d.$$

Proof. Since the function $\phi(d)$ is multiplicative, so (it is easy to see) is $\sum_{e \mid d} \phi(d)$. Hence it is only necessary to prove the result for $d = p^n$, ie to show that

$$\phi(p^d) + \phi(p^{d-1}) + \cdots + \phi(1) = p^d,$$

which follows at once from the fact that $\phi(p^n) = p^n - p^{n-1}$. □

From the two Lemmas, on applying Möbius inversion,

$$\sigma(d) = \sum_{e \mid d} e = \phi(d).$$

In particular,

$$\sigma(p^n - 1) = \phi(p^n - 1) \ge 1,$$

from which the theorem follows, since any element of this order will generate $F^\times$. □

Remarks:

1. It is not necessary to invoke Möbius inversion to deduce from the two Lemmas that $\sigma(d) = \phi(d)$, since it follows by simple induction that if the result holds for $e < d$ then it holds for d.

2. For a slight variant on this proof, suppose $a \in F^\times$ has order d. Then a satisfies the equation $f(x) = x^d - 1 = 0$, as do the d elements $a^i (0 \le i < d)$. Moreover any element of order d satisfies this equation. It follows that the elements of order d are all in the cyclic subgroup $C_d = \langle a \rangle$.

Since the order of an element divides the order of the group, which is 6 in this case, it follows that 3 has order 6 mod 7, and so is a primitive root.

If g generates the cyclic group G then so does g^{-1}. Hence

$$3^{-1} \equiv 5 \bmod 7$$

is also a primitive root mod 7.

Proposition 7.5. *There are $\phi(p-1)$ primitive roots* mod p. *If π is one primitive root then the others are π^i where $0 \le i < p-1$ and* $\gcd(p-1, i) = 1$.

This follows from Proposition 7.3 above.

Examples: Suppose $p = 11$. Then $(\mathbb{Z}/11)^\times$ has order 10, so its elements have orders 1,2,5 or 10. Now

$$2^5 = 32 \equiv -1 \bmod 11.$$

So 2 must be a primitive root mod 11.

There are

$$\phi(10) = 4$$

primitive roots mod 11, namely

$$2, 2^3, 2^7, 2^9 \bmod 11,$$

ie

$$2, 8, 7, 6.$$

Suppose $p = 23$. Then $(\mathbb{Z}/23)^\times$ has order 22, so its elements have orders 1,2,11 or 28.

Note that since $a^{22} = 1$ for all $a \in (\mathbb{Z}/29)^\times$, it follows that $a^{11} = \pm 1$.

Working always modulo 23,

$$2^5 = 32 \equiv 9 \implies 2^{10} \equiv 81 \equiv 12 \implies 2^{11} \equiv 24 \equiv 1.$$

So 2 has order 11. Also

$$3^2 \equiv 2^5 \implies 3^{10} \equiv 2^{25} \equiv 2^3 \implies 3^{11} \equiv 3 \cdot 8 \equiv 1.$$

So 3 also has order 11. But

$$5^2 \equiv 2 \implies 5^{10} \equiv 2^5 \equiv 9 \implies 5^{11} \equiv 45 \equiv -1.$$

Since $5^2 \equiv 2 \implies 5^4 \equiv 2^2 = 4$, we conclude that 5 is a primitive root modulo 23.

7.5 Uniqueness

Theorem 7.3. *Two fields F, F' of the same order p^n are necessarily isomorphic.*

Proof. If $a \in F^\times$ then $a^{p^n-1} = 1$, ie a is a root of the polynomial

$$U(x) = x^{p^n-1} - 1.$$

Hence

$$U(x) = \prod_{a \in F^\times} (x - a),$$

since the number $p^n - 1$ of elements is equal to the degree of $U(x)$.

Now suppose $U(x)$ factorises over $\mathbb{F}_p$ into irreducible polynomials

$$U(x) = f_1(x) \cdots f_r(x).$$

of $f_1(x)$ in F'. Then we claim that the map $\Theta : F \to F'$ given by

$$\pi^r \mapsto \pi'^r \qquad (0 \le r < p^n - 1)$$

(together with $0 \mapsto 0$) is a homomorphism.

It is easy to see that $\Theta(xy) = \Theta(x)\Theta(y)$. It remains to show that $\Theta(x + y) = \Theta(x) + \Theta(y)$. Suppose $x = \pi^a$, $y = \pi^b$, $x + y = \pi^c$. Then π satisfies the equation

$$f(x) = x^a + x^b - x^x.$$

It follows that

$$f_1(x) \mid f(x).$$

On passing to F',

$$f(\pi') = 0 \implies \pi'^a + \pi'^b = \pi'^c,$$

as required.

Finally, a homomorphism $\Theta : F \to F'$ from one field to another is necessarily injective. For if $x \neq 0$ then x has an inverse y, and then

$$\Theta(x) = 0 \implies \Theta(1) = \Theta(xy) = \Theta(x)\Theta(y) = 0,$$

contrary to fact that $\Theta(1) = 1$. (We are using the fact that Θ is a homomorphism of additive groups, so that $\ker \Theta = 0$ implies that Θ is injective.) Since F and F' contain the same number of elements, we conclude that Θ is bijective, and so an isomorphims. □

7.6 Existence

Theorem 7.4. *There exists a field F of every prime power p^n.*

Proof. We know that if $f(x) \in \mathbb{F}_p[x]$ is of degree d, then $\mathbb{F}_p[x]/(f(x))$ is a field of order p^n. Thus the result will follow if we can show that there exist irreducible polynomials $f(x) \in \mathbb{F}_p[x]$ of all degrees $n \geq 1$.

There are p^n monic polynomials of degree n in $\mathbb{F}_p[x]$. Let us associate to each such polynomial the term x^n. Then all these terms add up to the generating function

$$\sum_{n \in \mathbb{N}} p^n x^n = \frac{1}{1 - px}.$$

Now consider the factorisation of each polynomial

$$f(x) = f_1(x)^{e_1} \cdots f_r(x)^{e_r}$$

into irreducible polynomials. If the degree of $f_i(x)$ is d_i this product corresponds to the power

$$x^{d_1 e_1 + \cdots + d_r e_r}.$$

Putting all these terms together, we obtain a product formula analagous to Euler's formula. Suppose there are $\sigma(n)$ irreducible polynomials of degree n. Let $d(f)$ denote the degree of the polynomial $f(x)$. Then

$$\begin{aligned}
\frac{1}{1 - px} &= \prod_{\text{irreducible } f(x)} \left(1 + x^{d(f)} + x^{2d(f)} + \cdots\right) \\
&= \prod_{\text{irreducible } f(x)} \frac{1}{1 - x^{d(f)}} \\
&= \prod_{d \in \mathbb{N}} (1 - d^n)^{-\sigma(d)}.
\end{aligned}$$

As we have seen, we can pass from infinite products to infinite series by taking logarithms. When dealing with infinite products of functions it is usually easier to use logarithmic differentiation:

Applying Möbius inversion,

$$n\sigma(n) = \sum_{d|n} \mu(n/d)p^d.$$

The leading term p^n (arising when $d = 1$) will dominate the remaining terms. For these will consist of terms $\pm p^e$ for various different $e < n$. Thus their absolute sum is

$$\begin{aligned} &\leq \sum_{e \leq n-1} p^e \\ &= \frac{p^n - 1}{p - 1} \\ &< p^n. \end{aligned}$$

It follows that $\sigma(n) > 0$. ie there exists at least one irreducible polynomial of degree n. $\square$

Corollary 7.2. *The number of irreducible polynomials of degree n over $\mathbb{F}_p$ is*

$$\frac{1}{n} \sum_{d|n} \mu(n/d)p^d.$$

Examples: The number of polynomials of degree 3 over $\mathbb{F}_2$ is

$$\frac{1}{3}\left(\mu(1)2^3 + \mu(3)2\right) = \frac{2^3 - 2}{3} = 2,$$

namely the polynmials $x^3 + x^2 + 1,\ x^3 + x + 1$.

The number of polynomials of degree 4 over $\mathbb{F}_2$ is

$$\frac{1}{4}\left(\mu(1)2^4 + \mu(3)2^2 + \mu(1)2\right) = \frac{2^4 - 2^2}{4} = 3.$$

(Recall that $\mu(4) = 0$, since 4 has a square factor.)

The number of polynomials of degree 10 over $\mathbb{F}_2$ is

$$\frac{1}{10}\left(2^{10} - 2^5 - 2^2 + 2\right) = \frac{990}{10} = 99$$

The number of polynomials of degree 4 over $\mathbb{F}_3$ is

$$\frac{1}{4}\left(3^4 - 3^2\right) = \frac{72}{8} = 9.$$

*** 5. Determine the irreducible polynomials of degree 3 over $\mathbb{F}_3$.

*** 6. How many irreducible polynomials are there of degree 4 over $\mathbb{F}_3$?

** 7. Determine the irreducible polynomials of degree 2 over $\mathbb{F}_5$.

** 8. Determine the irreducible polynomials of degree 2 over $\mathbb{F}_7$.

** 9. Show that an irreducible polynomial over $\mathbb{R}$ is of degree 1 or 2.

** 10. Determine the irreducible polynomials over $\mathbb{C}$.

In exercises 11–20 determine if the given polynomial is irreducible over $\mathbb{Q}$.

** 11. $x^2 + x + 1$

** 12. $x^3 + 2x + 1$

*** 13. $x^4 + 1$

*** 14. $x^4 + 2$

*** 15. $x^4 + 4$

*** 16. $x^4 + 4x^3 + 1$

** 17. Determine the irreducible polynomials of degree 2 over $\mathbb{F}_7$.

Fermat's Little Theorem

8.1 Lagrange's Theorem

Let us recall (without proof) this basic result of group theory: *If G is a finite group of order n then*

$$g^n = 1$$

for all $g \in G$.

If G is commutative (as all the groups we consider will be) there is a simple way of proving this: Let

$$G = \{g_1, \ldots, g_n\}.$$

Then

$$\{gg_1, gg_2, \ldots, gg_n\}$$

are the same elements, in a different order (unless $g = 1$). Multiplying these elements together:

$$(gg_1)(gg_2)\cdots(gg_n) = g_1 g_2 \cdots g_n,$$

ie

$$g^n (g_1 g_2 \cdots g_n) = (g_1 g_2 \cdots g_n).$$

Multiplying by $(g_1 g_2 \cdots g_n)^{-1}$,

$$g^n = 1.$$

8.2 Euler's Theorem

Theorem 8.1 (Euler's Theorem)**.** *For all x coprime to n,*

$$x^{\phi(n)} \equiv 1 \bmod n.$$

Proof. The group $(\mathbb{Z}/n)^\times$ has order $\phi(n)$. The result follows on applying Lagrange's Theorem. □

8.3 Fermat's Little Theorem

As a particular case of Euler's Theorem, since $\phi(p) = p - 1$ if p is prime, we have

Theorem 8.2 (Fermat's Little Theorem)**.** *If p is prime then*

$$x^{p-1} \equiv 1 \bmod p$$

for all x coprime to p.

The title 'Fermat's Little Theorem' is sometimes given to the following variant.

Corollary 8.1. *If p is prime then*

$$x^p \equiv x \bmod p$$

for all x.

Proof. If $p \nmid x$ the result follows on multiplying the congruence in the Theo-

Unfortunately, it turns out that some composite numbers can satisfy Fermat's test for all x.

Definition 8.1. *We say that $n \in \mathbb{N}$ is a* Carmichael number *if n is composite but*

$$x^n \equiv x \bmod n \text{ for all } x.$$

Example: The smallest Carmichael number is

$$561 = 3 \cdot 11 \cdot 17.$$

To see that 561 is a Carmichael number, note that $3 - 1 = 2, 11 - 1 = 10$ and $17 - 1 = 16$ all divide $561 - 1 = 560$.

Suppose first that x is coprime to 561. By Fermat's Little Theorem,

$$x^2 \equiv 1 \bmod 3 \implies x^{560} \equiv 1 \bmod 3$$

Similarly,

$$\begin{aligned} x^{10} \equiv 1 \bmod 11 &\implies x^{560} \equiv 1 \bmod 11, \\ x^{16} \equiv 1 \bmod 17 &\implies x^{560} \equiv 1 \bmod 17. \end{aligned}$$

Putting these together, we deduce that

$$x^{560} \equiv 1 \bmod 3 \cdot 11 \cdot 17 = 561 \implies x^{561} \equiv x \bmod 561.$$

But what if x is not coprime to 561, say $17 \mid x$ but $3, 11 \nmid x$? Then $x = 17y$, where $\gcd(y, 33) = 1$.

The congruence is trivially satisfied mod 17:

$$(17y)^{561} \equiv 17y \bmod 17.$$

So we only have to show that

$$(17y)^{561} \equiv 17y \bmod 33,$$

Now $\phi(33) = 2 \cdot 10 = 20$. Since 17 and y are coprime to 33, it follows by Euler's Theorem that

$$17^{20} \equiv 1 \bmod 33 \text{ and } y^{20} \equiv 1 \bmod 33.$$

Hence

$$\begin{aligned} (17y)^{20} \equiv 1 \bmod 33 &\implies (17y)^{560} \equiv 1 \bmod 33 \\ &\implies (17y)^{561} \equiv 17y \bmod 33. \end{aligned}$$

The other cases where x is divisible by one or more of $3, 11, 17$ can be dealt with similarly.

We shall prove the following result later. The argument is similar to that above, but requires one more ingredient, which we shall meet in the next Chapter.

Proposition 8.1. *The number n is a Carmichael number if and only if it is square-free, and*

$$n = p_1 p_2 \cdots p_r$$

where $r \geq 2$ and

$$p_i - 1 \mid n - 1$$

for $i = 1, 2, \ldots, r$.

There are in fact an infinity of Carmichael numbers — this was only proved about 20 years ago — although they are sparsely distributed. (There are about $N^{1/3}$ Carmichael numbers $\leq N$.)

Note that if a number fails Fermat's test then it is certainly composite. The converse is not true, as we have seen; a number may pass the test but not be prime.

However, Fermat's test does provide a reasonable *probabilistic* algorithm, for determining "beyond reasonable doubt" if a large number n is prime: Choose a random number $x_1 \in [2, n-1]$, and see if

$$\left(x^{2^{e-1}m}\right)^2 \equiv 1 \bmod p.$$

It follows that

$$x^{2^{e-1}m} \equiv \pm 1 \bmod p;$$

for $\mathbb{Z}/(p)$ is a field; so if $x \in \mathbb{Z}/(p)$ then

$$x^2 = 1 \implies (x-1)(x+1) = 1 \implies x = \pm 1$$

Now suppose

$$x^{2^{e-1}m} \equiv 1 \bmod p.$$

Then we can repeat the argument, if $e > 1$, to see that

$$x^{2^{e-2}m} \equiv \pm 1 \bmod p.$$

Continuing in this way, we see that either

$$x^{2^i m} \equiv -1 \bmod p$$

for some $i \in [0, e-1]$. or else

$$x^m \equiv 1 \bmod p.$$

That is the Miller-Rabin test. It turns out that if a number n passes the test for all x coprime to n then it must be prime; there is no analogue of Carmichael numbers.

But we shall need the results of the next chapter to establish this ....

8.6 The AKS algorithm

The Miller-Rabin test (like the Fermat test) is *probabilistic*. It will only determine *up to a given probability* if a number is prime. Just over 10 years ago, three Indian mathematicians — Agrawal, Kayal and Saxena — found a deterministic polynomial-time primality algorithm.

This algorithm is based on a simple extension of Fermat's Little Theorem to polynomias.

Theorem 8.3. *The integer $n \geq 2$ is prime if and only if*

$$(x+a)^n \equiv x^n + a \bmod n$$

for all a.

Remark: Suppose $f(x) = \sum a_i x^i,\ g(x) = \sum b_i x^i \in \mathbb{Z}[x]$. We say that $f(x) \equiv g(x) \bmod n$ if $a_i \equiv b_i \bmod n$ for all i.

Proof.

Lemma 8.1. *If p is prime then*

$$p \mid \binom{i}{p}$$

for $i \neq 0, p$.

Proof. We have

$$\binom{i}{p} = \frac{p(p-1)\cdots(p-i+1)}{i(i-1)\cdots 1}.$$

The only term divisible by p is the first term in the numerator. □

It follows from this lemma that the relation in the theorem holds if n is prime.

Suppose n is not prime, say $p^i \parallel n$ where p is prime. Then

$$p^{i-1} \parallel \binom{n}{p}.$$

For

** 6. 7

** 7. 11

** 8. 13

** 9. 19

** 10. 29

*** 11. Show that if p is a prime then there are $\phi(d)$ elements of order d in the group $(\mathbb{Z}/p)^\times$.

*** 12. Show that the group $(\mathbb{Z}/n)^\times$ is cyclic if and only if $n = p^e$ or $2p^e$, where p is prime.

*** 13. How many elements of each order are there in $(\mathbb{Z}/32)^\times$?

**** 14. What is the order of 7 mod 2^e for each e?

** 15.

** 16.

** 17.

** 18.

*** 19. Show that if p and q are primes and $q \mid (a^p - 1)$ then either $q \mid (a-1)$ or $p \mid (q-1)$.

** 20.

Quadratic Residues

9.1 Introduction

Definition 9.1. *We say that $a \in \mathbb{Z}$ is a* quadratic residue mod n *if there exists $b \in \mathbb{Z}$ such that*

$$a \equiv b^2 \bmod n.$$

If there is no such b we say that a is a quadratic non-residue mod n.

Example: Suppose $n = 10$.

We can determine the quadratic residues mod n by computing b^2 mod n for $0 \le b < n$. In fact, since

$$(-b)^2 \equiv b^2 \bmod n,$$

we need only consider $0 \le b \le [n/2]$.

Thus the quadratic residues mod 10 are $0, 1, 4, 9, 6, 5$; while $3, 7, 8$ are quadratic non-residues mod 10.

Proposition 9.1. *If a, b are quadratic residues* mod n *then so is ab.*

Proof. Suppose

$$a \equiv r^2, \ b \equiv s^2 \bmod p.$$

Then

$$ab \equiv (rs)^2 \bmod p.$$

$\square$

9.2 Prime moduli

Proposition 9.2. *Suppose p is an odd prime. Then the quadratic residues coprime to p form a subgroup of $(\mathbb{Z}/p)^\times$ of index 2.*

Proof. Let Q denote the set of quadratic residues in $(\mathbb{Z}/p)^\times$. If $\theta : (\mathbb{Z}/p)^\times \to (\mathbb{Z}/p)^\times$ denotes the homomorphism under which

$$r \mapsto r^2 \bmod p$$

then

$$\ker\theta = \{\pm 1\}, \ \operatorname{im}\theta = Q.$$

By the first isomorphism theorem of group theory,

$$|ker\theta| \cdot |\operatorname{im}\theta| = |(\mathbb{Z}/p)^\times|.$$

Thus Q is a subgroup of index 2:

$$|Q| = \frac{p-1}{2}.$$

$\square$

Corollary 9.1. *Suppose p is an odd prime; and suppose a, b are coprime to p. Then*

1. *$1/a$ is a quadratic residue if and only if a is a quadratic residue.*
2. *If both of a, b, or neither, are quadratic residues, then ab is a quadratic residue;*
3. *If one of a, b is a quadratic residue and the other is a quadratic non-residue then ab is a quadratic non-residue.*

9.4 Euler's criterion

Proposition 9.4. *Suppose p is an odd prime. Then*

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \bmod p.$$

Proof. The result is obvious if $p \mid a$.

Suppose $p \nmid a$. Then

$$\left(a^{(p-1)/2}\right)^2 = a^{p-1} \equiv 1 \bmod p,$$

by Fermat's Little Theorem. It follows that

$$\left(\frac{a}{p}\right) \equiv \pm 1 \bmod p.$$

Suppose a is a quadratic residue, say $a \equiv r^2 \bmod p$. Then

$$a^{(p-1)/2} \equiv b^{p-1} \equiv 1 \bmod p$$

by Fermat's Little Theorem.

These provide all the roots of the polynomial

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

Hence

$$a^{(p-1)/2} \equiv -1 \bmod p$$

if a is a quadratic non-residue.

□

9.5 Gauss's Lemma

Suppose p is an odd prime. We usually take $r \in [0, p-1]$ as representatives of the residue-classes mod p But it is sometimes more convenient to take $r \in [-(p-1)/2, (p-1)/2]$, ie $\{-p/2 < r < p/2\}/$

Let P denote the strictly positive residues in this set, and N the strictly negative residues:

$$P = \{1, 2, \ldots, (p-1)/2\},\ N = -P = \{-1, -2, \ldots, -(p-1)/2\}.$$

Thus the full set of representatives is $N \cup \{0\} \cup P$.

Now suppose $a \in (\mathbb{Z}/p)^\times$. Consider the residues

$$aP = \{a, 2a, \ldots, \frac{p-1}{2}a\}.$$

Each of these can be written as $\pm s$ for some $s \in P$, say

$$ar = \epsilon(r)\pi(r),$$

where $\epsilon(r) = \pm 1$. It is easy to see that the map

$$\pi : P \to P$$

is injective; for

$$\begin{aligned}\pi(r) = \pi(r') &\implies ar \equiv \pm ar' \bmod p\\ &\implies r \equiv \pm r' \bmod p\\ &\implies r \equiv r' \bmod p,\end{aligned}$$

since s and s' are both positive.

Thus π is a permutation of P (by the pigeon-hole principle, if you like). It follows that as r runs over the elements of P so does $\pi(r)$.

Thus if we multiply together the congruences

1. Note that we could equally well choose the residues in $[1, p-1]$, and define t to be the number of times the residue appears in the second half $(p+1)/2, (p-1)$.

2. The map $a \mapsto (-1)^t$ is an example of the *transfer homomorphism* in group theory. Suppose H is an abelian subgroup of finite index r in the group G. We know that G is partitioned into H-cosets:

$$G = g_1 H \cup \cdots \cup g_r H.$$

If now $g \in G$ then

$$gg_i = g_j h_i$$

for $i \in [1, r]$. Now it is easy to see — the argument is similar to the one we gave above — that the product $h = h_1 \cdots h_r$ is independent of the choice of coset representatives $g_1, \ldots, g_r$, and the map

$$\tau : G \to S$$

is a homomorphism, known as the transfer homomorphism from G to S.

If G is abelian — which it is in all the cases we are interested in — we can simply multiply together all the equations $gg_i = g_j h_i$, to get

$$\tau(g) = g^r.$$

9.6 Computation of $\left(\dfrac{-1}{p}\right)$

Proposition 9.5. *If p is an odd prime then*

$$\left(\frac{-1}{p}\right) = \begin{cases} 1 \text{ if } p \equiv 1 \bmod 4, \\ -1 \text{ if } p \equiv -1 \bmod 4. \end{cases}$$

Proof. The result follows at once from Euler's Criterion

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \bmod p.$$

But it is instructive to deduce it by Gauss's Lemma.

We have to consider the residues

$$-1, -2, \ldots, -(p-1)/2 \bmod p.$$

All these are in the range $N = [-(p-1)/2, (p-1)/2]$. It follows that $t = (p-1)/2$; all the remainders are negative.

Hence

$$\begin{aligned} \left(\frac{-1}{p}\right) &= (-1)^{(p-1)/2} \\ &= \begin{cases} 1 \text{ if } p \equiv 1 \bmod 4, \\ -1 \text{ if } p \equiv -1 \bmod 4. \end{cases} \end{aligned}$$

□

Example: According to this,

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

$$2, 4, 6, \ldots, (p-1) \bmod p.$$

We have to determine the number t of these residues in the first half of $[1, p-1]$, and the number in the second. We can describe these two ranges as $\{0 < r < p/2\}$ and $\{p/2 < r < p\}$. Since

$$p/2 < 2x < p \iff p/4 < x < p/2$$

it follows that

$$t = \lfloor p/2 \rfloor - \lfloor p/4 \rfloor.$$

Suppose

$$p = 8n + r,$$

where $r = 1, 3, 5, 7$. Then

$$\lfloor p/2 \rfloor = 4n + \lfloor r/2 \rfloor, \ \lfloor p/4 \rfloor = 2n + \lfloor r/4 \rfloor.$$

Thus

$$t \equiv \lfloor r/2 \rfloor + \lfloor r/4 \rfloor \bmod 2.$$

The result follows easily from the fact that

$$\lfloor r/2 \rfloor = \begin{cases} 0 & \text{for } r = 1 \\ 1 & \text{for } r = 3 \\ 2 & \text{for } r = 5 \\ 3 & \text{for } r = 7, \end{cases}$$

while

$$\lfloor r/4 \rfloor = \begin{cases} 0 & \text{for } r = 1, 3 \\ 1 & \text{for } r = 5, 7 \end{cases}.$$

□

Example: Since $71 \equiv -1 \bmod 8$,

$$\left(\frac{2}{71}\right) = 1,$$

Can you find the solutions of

$$x^2 \equiv 2 \bmod 71?$$

Again Since $19 \equiv 3 \bmod 8$,

$$\left(\frac{2}{19}\right) = -1.$$

So by Euler's criterion,

$$2^9 \equiv -1 \bmod 19.$$

Checking,

$$2^4 \equiv 3 \implies 2^8 \equiv 9 \implies 2^9 \equiv 18 \bmod 19.$$

9.8 Composite moduli

Proposition 9.7. *Suppose m, n are coprime; and suppose a is coprime to m and n. Then a is a quadratic residue modulo mn if and only if it is a quadratic residue modulo m and modulo n.*

which

$$t \mapsto t^2 \bmod p^e$$

then

$$\ker \theta = \{\pm\}.$$

Proof. Suppose

$$a^2 - 1 = (a-1)(a+1) \equiv 0 \bmod p^e.$$

Then

$$p \mid a-1 \text{ and } p \mid a+1 \implies p \mid 2a \implies p \mid a,$$

which we have excluded. If $p \mid a+1$ then $p^e \mid a-1$; and if $p \mid a-1$ then $p^e \mid a+1$. Thus

$$a \equiv \pm 1 \bmod p^e.$$

□

It follows that the quadratic residues modulo p^e coprime to p form a subgroup of index 2 in $(\mathbb{Z}/p^e)^\times$, ie just half the elements of $(\mathbb{Z}/p^e)^\times$ are quadratic residues modulo p^e. Since just half are also quadratic residues modulo p, the result follows. □

Remark: For an alternative proof, we can argue by induction of e. Suppose a is a quadratic residue modp^e, say

$$a \equiv r^2 \bmod p^e,$$

ie

$$a = r^2 + tp^e.$$

Set

$$s = r + xp^e.$$

Then

$$\begin{aligned} s^2 &= r^2 + 2xp^e + x^2p^{2e} \\ &\equiv r^2 + 2xp^e \bmod p^{e+1} \\ &\equiv a + (t+2x)p^e \bmod p^{e+1} \\ &\equiv ap^e \bmod p^{e+1} \end{aligned}$$

if

$$t + 2x \equiv 0 \bmod p,$$

ie

$$x = -t/2 \bmod p,$$

using the fact that 2 is invertible modulo an odd prime p.

Corollary 9.2. *The number of quadratic residues in* $(\mathbb{Z}/p^e)^\times$ *is*

$$\frac{\phi(p^e)}{2} = \frac{(p-1)p^{e-1}}{2}.$$

The argument above extends to moduli 2^e with a slight modification.

Proposition 9.9. *Suppose p is an odd prime; and suppose $a \in \mathbb{Z}$ is coprime to p. Then a is a quadratic residue modulo p^e (where $e \geq 1$) if and only if it is quadratic residue modulo p.*

Proof. The argument we gave above for quadratic residues modulo p still applies here.

Lemma 9.2. *If* $\theta : (\mathbb{Z}/p^e)^\times \to (\mathbb{Z}/p^e)^timesimes$ *is the homomorphism under* ie

$$a = r^2 + tp^e.$$

Set

$$s = r + xp^e.$$

Then

$$\begin{aligned} s^2 &= r^2 + 2xp^e + x^2p^{2e} \\ &\equiv r^2 + 2xp^e \bmod p^{e+1} \\ &\equiv a + (t + 2x)p^e \bmod p^{e+1} \\ &\equiv a \bmod p^{e+1} \end{aligned}$$

if

$$t + 2x \equiv 0 \bmod p,$$

ie

$$x = -t/2 \bmod p,$$

using the fact that 2 is invertible modulo an odd prime p.

Corollary 9.3. *The number of quadratic residues in* $(\mathbb{Z}/p^e)^\times$ *is*

$$\frac{\phi(p^e)}{2} = \frac{(p-1)p^{e-1}}{2}.$$

The argument above extends to moduli 2^e with a slight modification.

Proposition 9.10. *Suppose a is an odd integer. Then a is a quadratic residue modulo 2^e (where $e \geq 3$) if and only if $a \equiv 1 \bmod 8$*

Proof. It is readily verified that 1 is the only odd quadratic residue modulo 8; 3,5 and 7 are quadratic non-residues.

We show by induction on e that if a is an odd quadratic residue modulo 2^e then it is a quadratic residue modulo 2^{e+1}. For suppose

$$a \equiv r^2 \bmod 2^e,$$

say

$$a = r^2 + t2^e.$$

Let

$$s = r + t2^{e-1}.$$

Then

$$\begin{aligned} s^2 &\equiv r^2 + t2^e \bmod 2^{e+1} \\ &= a. \end{aligned}$$

□

Corollary 9.4. *The number of quadratic residues in* $(\mathbb{Z}/2^e)^\times$ *(where $e \geq 3$) is*

$$\frac{\phi(2^e)}{4} = 2^{e-3}.$$

Remarks:

1. It is easy to see that p^f (where $f < e$) is a quadratic residue modulo p^e

** 4. $\left(\frac{5}{5}\right)$

** 5. $\left(\frac{5}{7}\right)$

In exercises 6-15, determine if the given congruence has a solution, and if it does find the smallest solution $x \geq 0$.

** 6. $x^2 \equiv 5 \bmod 10$

** 7. $x^2 \equiv 5 \bmod 11$

** 8. $x^2 \equiv 5 \bmod 12$

** 9. $x^2 \equiv 4 \bmod 15$

** 10. $x^2 \equiv -1 \bmod 105$

** 11. $x^2 + 3x + 1 \equiv 0 \bmod 13$

*** 12. $x^2 + 3x + 1 \equiv 0 \bmod 13$

*** 13. $x^2 \equiv 2 \bmod 27$

*** 14. $x^2 + 2 \equiv 0 \bmod 81$

*** 15. $x^2 \equiv 4 \bmod 25$

*** 16. Show that if p is a prime satisfying $p \equiv 1 \bmod 4$ then $x = ((p-1)/2)!$ satisfies

$$x^2 + 1 \equiv 0 \bmod p.$$

Quadratic Reciprocity

10.1 Gauss' Law of Quadratic Reciprocity

This has been described as 'the most beautiful result in Number Theory'.

Theorem 10.1. *Suppose p, q are distinct odd primes. Then*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}} = \begin{cases} -1 & \text{if } p \equiv q \equiv 3 \bmod 4 \\ 1 & \text{otherwise.} \end{cases}$$

More than 200 proofs of this have been given. Gauss himself gave 11.

We give a short proof of the Theorem below. It is due to Rousseau, and is fairly recent (1989), although it is said to be based on Gauss' 5th proof. It is subtle, but requires nothing we have not met.

10.2 Wilson's Theorem

We start with a preliminary result which is not really necessary, but which simplifies the formulae in the proof.

Proposition 10.1. *If p is an odd prime then*

$$(p-1)! \equiv -1 \bmod p.$$

Proof. Consider the numbers $1, 2, \ldots, p-1$. Each number x has a reciprocal $x^{-1} \bmod p$ in this set. The number x is equal to its reciprocal if and only if

$$x^2 \equiv 1 \implies x \equiv \pm 1 \bmod p.$$

It follows that the remaining $p-3$ numbers divide into pairs, each with product $1 \bmod p$. Hence the product of all $p-1$ numbers is

$$1 \cdot -1 = -1 \bmod p.$$

$\square$

We shall find our formulae are simplified if we set

$$P = (p-1)/2, \; Q = (q-1)/2.$$

Corollary 10.1. $(P!)^2 \equiv (-1)^{P+1} \bmod p$.

Proof. This follows from Wilson's Theorem on replacing the numbers $\{P+1, \ldots, p-1\}$ by $\{-1, -2, ..., -P \bmod p\}$. $\square$

Recall the definition of the quotient-group G/H, where H is a normal subgroup of G. (We will only be interested in abelian groups, in which case every subgroup is normal.) The elements of G/H are the cosets of H in G. If we write $x' \sim x$ to mean that x', x are in the same H-coset, ie $x' = xh$ for some $h \in H$, then the basic step in defining the product operation on G/H is to show that

$$x' \sim x, \; y' \sim y \implies x'y' \sim xy.$$

It follows from this that if we take representatives $x_1, \ldots, x_r$ of all the cosets of H then the coset containing the product $x_1 \cdots x_r$ is independent of the choice of representatives:

$$x_i' \sim x_i \text{ for } 1 \le i \le r \implies x_i' \cdots x_r' \sim x_1 \cdots x_r,$$

$$\{(x,y) : x \in \{1,\dots,p-1\},\ y \in \{1,\dots,q-1\}\}.$$

We are going to consider the quotient of this group by the subgroup

$$\{\pm 1\} = C_2.$$

In other words, we are going to divide the group into pairings $\{(x,y),(-x,-y)\}$. The group has order $(p-1)(q-1) = 4PQ$, so there are $2PQ$ pairings.

We are going to choose one representative from each pairing, in two different ways. In each case we will form the product of these representatives. by the argument above, the two products will differ by a factor ± 1.

For our first division, let us take the first half of $(\mathbb{Z}/p)^\times$, and the whole of $(\mathbb{Z}/q)^\times$. In other words, we take the representatives

$$\{(x,y) :\ 1 \le x \le P,\ 1 \le y \le q-1\}.$$

We want to compute the product of these elements.

The x-components are $1,2,\dots,P$, repeated $q-1$ times. Their product is

$$(P!)^{q-1} = ((P!)^2)^Q \equiv (-1)^{(P+1)Q} \bmod p,$$

by the Corollary to Wilson's Theorem.

The y-components are $1,2,\dots,q-1$, repeated P times. By Wilson's Theorem, their product is

$$(-1)^P \bmod q.$$

Thus the product of the representatives is

$$((-1)^{(P+1)Q} \bmod p,\ (-1)^P \bmod q).$$

We could equally well choose representatives by taking the whole of $(\mathbb{Z}/p)^\times$ and the first half of $(\mathbb{Z}/q)^\times$. The product of these representatives would be

$$((-1)^Q \bmod p,\ (-1)^{P(Q+1)} \bmod q).$$

However, what we need is a third way of choosing representatives, by choosing the first half of $(\mathbb{Z}/pq)^\times$. By this we mean the pairs ($n \mod p$, $n \mod q$), where n runs through the numbers $1,\dots,(pq-1)/2$ not divisible by p or q, ie the set of numbers $A \setminus B$, where

$$A = \{1,2,\dots,p-1,p+1,p_2,\dots,2p-1,\dots,Qp+1,\dots,Qp+P\},$$

while B denotes the numbers in this set divisible by q, ie

$$B = \{q, 2q, \dots, Pq\}.$$

Again, we compute the product ($X \bmod p$, $Y \bmod q$) of these representatives. The first component $X \mod p$ is

$$((p-1)!)^Q \cdot P!/q^P \cdot P! = ((p-1)!)^Q/q^P \equiv (-1)^Q/q^P \bmod p.$$

But by Eisenstein's criterion,

$$q^P = \left(\frac{q}{p}\right) \bmod p.$$

Thus

$$X = (-1)^Q \left(\frac{q}{p}\right) \bmod p.$$

Similarly, the second component $Y \mod q$ is

$$Y = (-1)^P \left(\frac{p}{q}\right) \bmod q.$$

Comparing the products of the two choices of representatives,

$$((-1)^{(P+1)Q} \bmod p,\ (-1)^P \bmod q) = \pm((-1)^Q \left(\frac{q}{p}\right) \bmod p, (-1)^P \left(\frac{p}{q}\right) \bmod q).$$

** 4. $\left(\dfrac{36}{61}\right)$

** 5. $\left(\dfrac{2009}{2011}\right)$

In exercises 6-15, determine if the given congruence has a solution, and if it does find the smallest solution $x \geq 0$.

** 6. $x^2 \equiv 10 \bmod 36$

** 7. $x^2 + 12 \equiv 0 \bmod 75$

*** 8. $x^2 \equiv 8 \bmod 2009$

*** 9. $x^2 \equiv 56 \bmod 2317$

*** 10. $x^2 + 2x + 17 \equiv 0 \bmod 35$

*** 11. $x^2 + 3x + 1 \equiv 0 \bmod 13$

** 12. $x^3 \equiv -1 \bmod 105$

*** 13. $x^7 \equiv 3 \bmod 17$

*** 14. $x^3 + 2 \equiv 0 \bmod 27$

*** 15. $x^5 + 3x + 1 \equiv 0 \bmod 25$

**** 16. If $n > 0$ is an odd number, and $n = p_1 \dots p_r$, we define the Jacobi symbol $\left(\dfrac{a}{n}\right)$ by

$$\left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right) \cdots \left(\frac{a}{p_r}\right).$$

Show that if $m, n > 0$ are both odd then

$$\left(\frac{m}{n}\right)\left(\frac{n}{m}\right) = \begin{cases} -1 \text{ if } m \equiv n \equiv -1 \bmod 4, \\ 1 \text{ otherwise }. \end{cases}$$

$\left(2009\right)$

** 21. $\left(\frac{2009}{2317}\right)$

**** 22. Is there a power 7^n which ends with the digits 000011? If so, what is the smallest such n?

**** 23. Is there a power of 2009 which ends with the digits 2317?

**** 24. Is there a power of 2319 which ends with the digits 2009?

*** 25. Determine $\left(\frac{3}{p}\right)$ for an odd prime p without using Quadratic Reciprocity.

Gaussian Integers

11.1 Gaussian Numbers

Definition 11.1. *A* gaussian number *is a number of the form*

$$z = x + iy \quad (x, y \in \mathbb{Q}).$$

If $x, y \in \mathbb{Z}$ we say that z is a gaussian integer*.*

Proposition 11.1. *The gaussian numbers form a field.*

The gaussian integers form a commutative ring.

Proof. The only part that is not, perhaps, obvious is that the inverse of a gaussian number $z = x + iy$ is a gaussian number. In fact

$$\begin{aligned}\frac{1}{z} &= \frac{1}{x+iy} \\ &= \frac{x-iy}{(x+iy)(x-iy)} \\ &= \frac{x}{x^2+y^2} - i\frac{y}{x^2+y^2}.\end{aligned}$$

□

We denote the gaussian numbers by $\mathbb{Q}(i)$, and the gaussian integers by $\mathbb{Z}[i]$ or Γ. (We will be mainly interested in this ring.)

11.2 Conjugates and norms

Definition 11.2. *The* conjugate *of the gaussian number*

$$z = x + iy \in \mathbb{Q}(i)$$

is

$$\bar{z} = x - iy.$$

Proposition 11.2. *The map*

$$z \mapsto \bar{z} : \mathbb{Q}(i) \to \mathbb{Q}(i)$$

is an automorphism of $\mathbb{Q}(i)$. In fact it is the only automorphism apart from the trivial map $z \mapsto z$.

Proof. It is evident that $z \mapsto \bar{z}$ preserves addition. To see that it preserves multiplication, note that

$$(x + iy)(u + iv) = (xu - yv) + i(xv + yu) \mapsto (xu - yv) - i(xv + yu),$$

while

$$(x - iy)(u - iv) = (xu - yv) - i(xv + yu).$$

Suppose θ is an automorphism of $\mathbb{Q}(i)$. By definition,

$$\theta(0) = 0,\ \theta(1) = 1.$$

Hence

$$\theta(n) = 1 + \cdots + 1 = n$$

for $n \in \mathbb{N}$. It follows easily that $\theta(n) = n$ for $n \in \mathbb{Z}$, and that if $q = n/d \in \mathbb{Q}$ then

$$\begin{aligned}\mathcal{N}(zw) &= (zw)(\overline{zw}) \\ &= zw\bar{z}\bar{w} \\ &= (z\bar{z})(w\bar{w}) \\ &= \mathcal{N}(z)\mathcal{N}(w).\end{aligned}$$

□

11.3 Units

Recall that an element ϵ of a ring A is said to be a *unit* if it is invertible, ie if there exists an element $\eta \in A$ such that

$$\epsilon\eta = 1 = \eta\epsilon.$$

The units in A form a group $A^\times$.

Evidently $\mathbb{Z}^\times = \{\pm 1\}$.

Proposition 11.4. *The units in Γ are: $\pm 1, \pm i$*

Proof. Evidently $\pm 1, \pm i$ are units.

Lemma 11.1. *If $\epsilon \in \Gamma$ then*

$$\epsilon \textit{ is a unit} \iff \mathcal{N}(\epsilon) = 1.$$

Proof. Suppose ϵ is a unit, say

$$\epsilon\eta = 1.$$

Then

$$\begin{aligned}\epsilon\eta = 1 &\implies \mathcal{N}(\epsilon)\mathcal{N}(\eta) = \mathbb{N}(1) = 1 \\ &\implies \mathcal{N}(\epsilon) = \mathcal{N}(\eta) = 1.\end{aligned}$$

□

Suppose $\epsilon = m + in \in \Gamma$ is a unit. Then

$$\mathcal{N}(\epsilon) = m^2 + n^2 = 1.$$

Evidently the only solutions to this are

$$(m, n) = (\pm 1, 0) \text{ or } (0, \pm 1),$$

giving $\pm 1, \pm i$. □

11.4 Division in Γ

Proposition 11.5. *Suppose $z, w \in \Gamma$, with $w \neq 0$. Then we can find $q, r \in \Gamma$ such that*

$$z = qw + r,$$

with

$$\mathcal{N}(r) < \mathcal{N}(w).$$

Proof. Suppose

$$\frac{z}{w} = x + iy,$$

where $x, y \in \mathbb{Q}$.

Let $m, n \in \mathbb{Z}$ be the nearest integers to x, y, respectively. Then

$$|x - m| \leq \frac{1}{2}, \ |y - n| \leq \frac{1}{2}.$$

Set

$$q = m + in.$$

Then

$$\frac{z}{w} - q = (x - m) + i(y - n).$$

$$uz + vw = \delta.$$

Proof. We follow the Euclidean Algorithm as in $\mathbb{Z}$, except that we use $\mathcal{N}(z)$ in place of $|n|$.

We start by dividing z by w:

$$z = q_0 w + r_0, \quad \mathcal{N}(r_0) < \mathcal{N}(w).$$

If $r_0 = 0$, we are done. Otherwise we divide w by r_0:

$$w = q_1 r_0 + r_1, \quad \mathcal{N}(r_1) < \mathcal{N}(r_0).$$

If $r_1 = 0$, we are done. Otherwise we continue in this way. Since

$$\mathcal{N}(w) > \mathcal{N}(r_0) > \mathcal{N}(r_1) > \cdots,$$

and the norms are all positive integers, the algorithm must end, say

$$r_i = q_i r_{i-1}, \; r_{i+1} = 0.$$

Setting

$$\delta = r_i,$$

we see successively that

$$\delta \mid r_{i-1}, r_{i-2}, \ldots, r_0, w, z.$$

Conversely, if $\delta' \mid z, w$ then

$$\delta' \mid z, w, r_0, r_1, \ldots, r_i = \delta.$$

The last part of the Proposition follows as in the classic Euclidean Algorithm; we see successively that $r_1, r_2, \ldots, r_i = \delta$ are each expressible as linear combinations of z, w with coefficients in Γ. □

11.6 Unique factorisation

If A is an integral domain, we say that $a \in A$ is a *prime element* if

$$a = bc \implies b \text{ is a unit, or } c \text{ is a unit.}$$

(We often just say "a is prime" if that cannot cause confusion.) We say that two prime elementsπ, π' are *equivalent*, and we write $\pi \sim \pi'$, if

$$\pi' = \epsilon \pi$$

for some unit ϵ.

Definition 11.4. *We say that an integral domain A is a* Unique Factorisation Domain *(UFD) if each non-zero element $a \in A$ is expressible in the form*

$$a = \epsilon p_1 \cdots p_r,$$

where ϵ is a unit, and $p_1, \ldots, p_r$ are prime elements, and if moreover this expression is unique up to order and multiplication by units, ie if

$$a = \epsilon' p'_1 \ldots p'_s$$

then $r = s$, and after re-ordering if necessary,

$$p'_i \sim p_i.$$

If $r \geq 1$ we could of course combine ϵ with one of the prime elements, and write

$$a = p_1 \cdots p_r.$$

Multiplying by w,

$$u\pi w + vzw = w.$$

Since π divides both terms on the left,

$$\pi \mid w.$$

□

Now the proof is as before. Again, we argue by induction on $\mathcal{N}(z)$. Suppose

$$z = \epsilon p_1 \cdots p_r = \epsilon' p'_1 \ldots p'_s.$$

Then

$$\pi_1 \mid \pi'_i$$

for some i. Hence

$$\pi'_i \sim \pi.$$

Now we can divide both sides by π_1 and apply the inductive hypothesis. □

Definition 11.5. *If A is a unique factorisation domain we use the term* prime *for a prime element, with the understanding that equivalent prime elements define the same prime.*

More precisely perhaps, a prime is a set $\{\epsilon\pi : \epsilon \in A^\times\}$ of equivalent prime elements.

11.7 Gaussian primes

Having established unique factorisation in Γ, we must identify the primes.

Proposition 11.7. *Each prime π in Γ divides just one rational prime p.*

Proof. Let us factorise $\mathcal{N}(\pi)$ in $\mathbb{N}$:

$$\mathcal{N}(\pi) = \pi\bar{\pi} = p_1 \ldots p_r.$$

On factorising both sides in Γ, it follows that

$$\pi \mid p_i$$

for some i.

Now suppose π divides two primes p, q. Since p, q are coprime, we can find $u, v \in \mathbb{Z}$ such that

$$up + vq = 1.$$

But now

$$\pi \mid p, q \implies \pi \mid 1,$$

which is absurd. □

Proposition 11.8. *Each rational prime p splits into at most 2 primes in Γ.*

Proof. Suppose

$$p = \pi_1 \cdots \pi_r.$$

Then

$$\mathcal{N}(p) = p^2 = \mathcal{N}(\pi_1) \cdots \mathcal{N}(\pi_r).$$

Since $\mathcal{N}(\pi_i) > 1$, it follows that

$$r \leq 2.$$

□

Proposition 11.9. *If the rational prime p splits in Γ, say*

But this is impossible, since

$$a^2 \equiv 0 \text{ or } 1 \bmod 4.$$

$\square$

Proposition 11.11. *If* $p \equiv 1 \bmod 4$ *(where* p *is a rational prime) then* p *splits in* Γ *into two distinct but conjugate primes:*

$$p = \pi\bar{\pi}.$$

Proof. This is more subtle. We know that

$$\left(\frac{-1}{p}\right) = 1.$$

Thus there exists an r such that

$$r^2 \equiv -1 \bmod p,$$

where we may suppose that $0 < r < p$. Then

$$r^2 + 1 \equiv 0 \bmod p$$

ie

$$p \mid r^2 + 1 = (r + i)(r - i).$$

If p does not split in Γ then

$$p \mid r + i \text{ or } p \mid r - i.$$

But either implies that

$$p \mid 1,$$

which is absurd.

Thus

$$p = \pi\sigma,$$

where π, σ are primes. But then

$$\mathcal{N}(\pi) = \pi\bar{\pi} = p,$$

ie p is the product of two conjugate primes in Γ.

Finally,

$$\pi \not\sim \bar{\pi}.$$

For

$$\bar{\pi} = \epsilon\pi \implies p = \mathcal{N}(\pi) = \pi\bar{\pi} = \epsilon\pi^2.$$

But if $\pi = m + in$ this implies that

$$m^2 + n^2 = \epsilon(m^2 - n^2 + 2imn).$$

The coefficient of i on the right must vanish. If $\epsilon = \pm 1$ this gives $mn = 0$, which is absurd. If $\epsilon = \pm i$ it gives

$$m^2 - n^2 = 0 \implies m = \pm n \implies p = 2m^2 \implies p = 2.$$

$\square$

The rational prime 2 has a special property in Γ.

Proposition 11.12. *The rational prime 2* ramifies *in* Γ, *ie it splits into 2* and suppose $p \mid n$, where $p \equiv 3 \bmod 4$. Then

$$p \mid x + iy \text{ or } p \mid x - iy.$$

In either case

$$p \mid x \text{ and } p \mid y.$$

But $p^2 \mid n$ and we can divide the equation by p^2:

$$n/p^2 = (x/p)^2 + (y/p)^2.$$

But now the result for n follows from that for n/p^2.

Now suppose that n has this form, say

$$n = 2^e p_1^{e_1} \cdots p_r^{e_r} q_1^{2f_1} \ldots q_s^{2f_s},$$

where $p_1, \ldots, p_r$ are primes $\equiv 1 \mod 4$ and $q_1, \ldots, q_s$ are primes $\equiv 3 \mod 4$.

Each rational prime p_i splits into conjugate primes, say

$$p = \pi_i \bar{\pi}_i.$$

Let

$$\theta = m + in = (1+i)^e \pi_1^{e_1} \cdots \pi_r^{e_r} q_1^{f_1} \cdots q_s^{f_s}.$$

Then

$$\begin{aligned}
\mathcal{N}(\theta) &= m^2 + n^2 \\
&= \mathcal{N}(1+i)^e \mathcal{N}(1+i)^e \mathcal{N}(\pi_1)^{e_1} \cdots \mathcal{N}(\pi_r)^{e_r} \mathcal{N}(q_1)^{f_1} \cdots \mathcal{N}(q_s)^{f_s} \\
&= 2^e p_1^{e_1} \cdots p_r^{e_r} q_1^{2f_1} \ldots q_s^{2f_s} \\
&= n.
\end{aligned}$$

□

Example: Since

$$2317 = 7 \cdot 331,$$

7 occurs just once in 2317. So 2317 is not the sum of two squares.

But

$$2009 = 7 \cdot 7 \cdot 41.$$

Here 7 occurs twice, while $41 \equiv 1 \bmod 4$. Hence 2009 *is* the sum of two squares.

Our argument shows that if

$$2009 = m^2 + n^2$$

then

$$7 \mid m, n.$$

If we set

$$m = 7a, \ n = 7b,$$

then

$$41 = a^2 + b^2.$$

Now it is easy to see that $a, b = 5, 7$ (if we restrict to positive solutions), ie

$$2009 = 35^2 + 40^2.$$

The argument also gives the *number* of ways of expressing a number as the sum of two squares.

Proposition 11.14. *Suppose*

$$n = 2^e p_1^{e_1} \cdots p_r^{e_r} q_1^{2f_1} \ldots q_s^{2f_s},$$

where $p_1, \ldots, p_r$ are primes $\equiv 1 \mod 4$ and $q_1, \ldots, q_s$ are primes $\equiv 3 \mod 4$. Then n can be expressed as primes.

** 6. $3+5i$

** 7. $5+3i$

*** 8. $23+17i$

** 9. $11+2i$

** 10. $29-i$

In exercises 11-15, either express the given number as a sum of two squares, or else show that this is not possibles.

** 11. 233

** 12. 317

** 13. 613

** 14. 1009

** 15. 2010

*** 16. Find a formula expressing

$$(x^2+y^2+z^2+t^2)(X^2+Y^2+Z^2+T^2)$$

as a sum of 4 squares.

*** 17. Show that every prime p can be expressed as a sum of 4 squares.

** 18. Deduce from the last 2 exercises that every $n \in \mathbb{N}$ can be expressed as a sum of 4 squares.

** 19. Show that if $n \equiv 7 \bmod 8$ then n cannot be expressed as a sum of 3 squares.

*** 20. Show that if $n = 4^e(8m+7)$ then n cannot be expressed as a sum of 3 squares.

*** 23. Show that if the prime $p = m^2 + n^2$ and $p \equiv \pm 1 \bmod 10$ then

$$5 \mid xy.$$

*** 24. Find the smallest $n \in \mathbb{N}$ such that $n, n+1, n+2$ are each a sum of 2 squares, but none is a perfect square.

**** 25. Show that there are arbitrarily long gaps between successive integers expressible as a sum of 2 squares.

Algebraic numbers and algebraic integers

12.1 Algebraic numbers

Definition 12.1. *A number $\alpha \in \mathbb{C}$ is said to be* algebraic *if it satisfies a polynomial equation*

$$f(x) = x^n + a_1 x^{n-1} + \cdots + a_n = 0$$

with rational coefficients $a_i \in \mathbb{Q}$.

For example, $\sqrt{2}$ and $i/2$ are algebraic.

A complex number is said to be *transcendental* if it is not algebraic. Both e and π are transcendental. It is in general extremely difficult to prove a number transcendental, and there are many open problems in this area, eg it is not known if π^e is transcendental.

Theorem 12.1. *The algebraic numbers form a field $\bar{\mathbb{Q}} \subset \mathbb{C}$.*

Proof. If α satisfies the equation $f(x) = 0$ then $-\alpha$ satisfies $f(-x) = 0$, while $1/\alpha$ satisfies $x^n f(1/x) = 0$ (where n is the degree of $f(x)$). It follows that $-\alpha$ and $1/\alpha$ are both algebraic. Thus it is sufficient to show that if α, β are algebraic then so are $\alpha + \beta, \alpha\beta$.

Lemma 12.1. *Suppose $V \subset \mathbb{C}$ is a finite-dimensional vector space over $\mathbb{Q}$, with $V \neq 0$; and suppose $x \in \mathbb{C}$. If*

$$xV \subset V$$

then $x \in \bar{\mathbb{Q}}$.

Proof. Let $e_1, \ldots, e_n$ be a basis for V. Suppose

$$\begin{aligned} xe_1 &= a_{11}e_1 + \cdots a_{1n}e_n \\ xe_2 &= a_{21}e_1 + \cdots a_{2n}e_n \\ &\cdots \\ xe_n &= a_{n1}e_1 + \cdots a_{nn}e_n. \end{aligned}$$

Then

$$\det(xI - A) = 0,$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

This is a polynomial equation with coefficients in $\mathbb{Q}$. Hence $x \in \bar{\mathbb{Q}}$. □

Consider the vector space

$$V = \langle \alpha^i \beta^j : 0 \leq i < m,\ 0 \leq j < n \rangle$$

over $\mathbb{Q}$ spanned by the mn elements $\alpha^i \beta^j$. Evidently

$$\alpha V \subset V,\ \beta V \subset V.$$

Thus

$$(\alpha + \beta)V \subset V,\ (\alpha\beta)V \subset V.$$

Hence $\alpha + \beta$ and $\alpha\beta$ are algebraic. □

Lemma 12.2. *Suppose $S \subset \mathbb{C}$ is a finitely-generated abelian group, with $S \neq 0$; and suppose $x \in \mathbb{C}$. If*

$$xS \subset S$$

then $x \in \bar{\mathbb{Z}}$.

Proof. Let $s_1, \dots, s_n$ generate S. Suppose

$$\begin{aligned} xs_1 &= a_{11}s_1 + \cdots a_{1n}s_n \\ xs_2 &= a_{21}s_1 + \cdots a_{2n}s_n \\ &\cdots \\ xs_n &= a_{n1}s_1 + \cdots a_{nn}s_n. \end{aligned}$$

Then

$$\det(xI - A) = 0.$$

This is a monic equation with coefficients in $\mathbb{Z}$. Hence $x \in \bar{\mathbb{Z}}$. □

Consider the abelian group

$$S = \langle \alpha^i \beta^j : 0 \leq i < m,\ 0 \leq j < n \rangle$$

generated by the mn elements $\alpha^i \beta^j$. Evidently

$$\alpha S \subset S,\ \beta S \subset S.$$

Thus

$$(\alpha + \beta)S \subset S,\ (\alpha\beta)S \subset S.$$

Hence $\alpha + \beta$ and $\alpha\beta$ are algebraic integers. □

Proposition 12.1. *A rational number $c \in \mathbb{Q}$ is an algebraic integer if and only if it is a rational integer:*

$$\bar{\mathbb{Z}} \cap \mathbb{Q} = \mathbb{Z}.$$

Proof. Suppose $c = m/n$, where $\gcd(m, n) = 1$; and suppose c satisfies the equation

$$x^d + a_1 x^{d-1} + \cdots + a_d = 0 \quad (a_i \in \mathbb{Z}).$$

Then

$$m^d + a_1 m^{d-1} n + \cdots + a_d n^d = 0.$$

Since n divides every term after the first, it follows that $n \mid m^d$. But that is incompatible with $\gcd(m, n) = 1$, unless $n = 1$, ie $c \in \mathbb{Z}$. □

12.3 Number fields and number rings

Suppose $F \subset \mathbb{C}$ is a field. Then $1 \in F$, by definition, and so

$$\mathbb{Q} \subset F \subset \mathbb{C}.$$

We can consider F as a vector space over $\mathbb{Q}$.

Definition 12.3. *An* algebraic number field *(or simply* number field *is a subfield $F \subset \mathbb{C}$ which is a finite-dimensional vector space over $\mathbb{Q}$. The* degree *of F is the dimension of this vector space:*

$$\deg F = \dim_{\mathbb{Q}} F.$$

Proposition 12.2. *The elements of a number field F are algebraic numbers:*

$$\mathbb{Q} \subset F \subset \bar{\mathbb{Q}}.$$

is a gaussian number. We have to show that z is an algebraic integer if and only if $x, y \in \mathbb{Z}$.

If $m, n \in \mathbb{Z}$ then $m + in \in \bar{\mathbb{Z}}$, since $m, n, i \in \bar{\mathbb{Z}}$ and $\bar{\mathbb{Z}}$ is a ring.

Conversely, suppose

$$z = x + iy \in \bar{\mathbb{Z}}.$$

Then

$$\bar{z} = x - iy \in \bar{\mathbb{Z}}$$

since z and $\bar{z}$ satisfy the same polynomials over $\mathbb{Q}$. Hence

$$z + \bar{z} = 2x \in \bar{\mathbb{Z}} \cap \mathbb{Q} = \mathbb{Z}.$$

Similarly

$$-iz = y - ix \in \bar{\mathbb{Z}} \implies 2y \in \mathbb{Z}.$$

Thus

$$z = \frac{m + in}{2},$$

with $m, n \in \mathbb{Z}$.

But now

$$\mathcal{N}(z) = z\bar{z} \in \bar{\mathbb{Z}} \cap \mathbb{Q} = \mathbb{Z},$$

ie

$$x^2 + y^2 = \frac{m^2 + n^2}{4} \in \mathbb{Z},$$

ie

$$m^2 + n^2 \equiv 0 \bmod 4.$$

But $m^2, n^2 \equiv 0$ or $1 \bmod 4$. So

$$\begin{aligned} m^2 + n^2 \equiv 0 \bmod 4 &\implies 2 \mid m, n \\ &\implies z \in \Gamma. \end{aligned}$$

$\square$

Example: $\sqrt{2}$ is an algebraic integer, since it satisfies the equation

$$x^2 - 2 = 0.$$

But $\sqrt{2}/2$ is not an algebraic integer. For if it were,

$$(\sqrt{2}/2)^2 = 1/2$$

would be an algbraic integer (since $\bar{\mathbb{Z}}$ is a ring), which we have just seen is not so.

Algebraic number theory is the study of number rings. The first question one might ask is whether a given number ring is a Unique Factorisation Domain.

We have seen that the number rings Z and Γ are. But in general number rings are not UFDs.

The foundation of algebraic number theory was Dedekind's amazing discovery that unique factorisation could be recovered if one added what Dedekind called 'ideal numbers', and what are today called 'ideals'.

However, we are not going into that theory. We shall only be looking at a small number of quadratic number rings which are UFDs.

12.4 Integral closure

Recall that any integral domain A can be extended to its *field of fractions*, which we shall denote by $Q(A)$, since we follows exactly the same process as in creating the field of rational numbers $\mathbb{Q}$ from the ring of integers $\mathbb{Z}$. We define $Q(A)$ to be the quotient set X/E, where X is the set of pairs (n, d),

Quadratic fields and quadratic number rings

12.1 Quadratic number fields

Definition 12.1. *A quadratic number field is a number field of degree 2.*

The integer $d \in \mathbb{Z}$ is said to be *square-free* if it has no square factor, ie

$$a^2 \mid d \implies a = \pm 1.$$

Thus the square-free integers are

$$\pm 1, \pm 2, \pm 3, \pm 5, \ldots.$$

Proposition 12.1. *Suppose $d \neq 1$ is square-free. Then the numbers*

$$x + y\sqrt{d} \quad (x, y \in \mathbb{Q})$$

form a quadratic number field $\mathbb{Q}(\sqrt{d}$.

Moreover, every quadratic number field is of this form; and different square-free integers $d, d' \neq 1$ give rise to different quadratic number fields.

Proof. Recall the classic proof that $\sqrt{d}$ is irrational;

$$\sqrt{d} = \frac{m}{n} \implies n^2 d = m^2,$$

and if any prime factor $p \mid d$ divides the left hand side to an odd power, and the right to an even power.

It is trivial to see that the numbers $x + y\sqrt{d}$ form a commutative ring, while

$$\begin{aligned} \frac{1}{x + y\sqrt{d}} &= \frac{x - y\sqrt{d}}{(x - y\sqrt{d})(x + y\sqrt{d})} \\ &= \frac{x - y\sqrt{d}}{x^2 - dy^2}, \end{aligned}$$

where $x^2 - dy^2 \neq 0$ since $\sqrt{d} \notin \mathbb{Q}$.

It follows that these numbers form a field; and the degree of the field is 2 since $1, \sqrt{d}$ form a basis for the vector space.

Conversely, suppose F is a quadratic number field. Let $1, \theta$ be a basis for the vector space. Then $1, \theta, \theta^2$ are linearly independent, ie θ satisfies a quadratic equation

$$a\theta^2 + b\theta + c = 0 \quad (a, b, c \in \mathbb{Q}).$$

Since F is of degree 2, $a \neq 0$, and we can take $a = 1$. Thus

$$\theta = \frac{-b \pm \sqrt{D}}{2},$$

with $D = b^2 - 4c$.

Now

$$D = a^2 d,$$

where d is a square-free integer (with $a \in \mathbb{Q}$). It follows easily that

$$F = \mathbb{Q}(\sqrt{d}).$$

$$\bar{z} = x - y\sqrt{d}$$

If $d < 0$ then this coincides with the complex conjugate; but if $d > 0$ then both z and $\bar{z}$ are real; and

$$z = \bar{z} \iff z \in \mathbb{Q}.$$

Proposition 12.2. *The map*

$$z \mapsto \bar{z} : \mathbb{Q}(\sqrt{d}) \to \mathbb{Q}(\sqrt{d})$$

is an automorphism of $\mathbb{Q}(\sqrt{d})$. *In fact it is the only such automorphism apart from the trivial map* $z \mapsto z$.

The proof is identical to that we gave for gaussian numbers.

Definition 12.3. *The* norm *of* $z = x + y\sqrt{d} \in \mathbb{Q}(\sqrt{d})$ *is*

$$\mathcal{N}(z) = z\bar{z} = x^2 - dy^2.$$

Proposition 12.3.

1. $\mathcal{N}(z) \in \mathbb{Q}$;
2. $\mathcal{N}(z) = 0 \iff z = 0$;
3. $\mathcal{N}(zw) = \mathcal{N}(z)\mathcal{N}(w)$;
4. *If* $a \in \mathbb{Q}$ *then* $\mathcal{N}(a) = a^2$;

Again, the proof is identical to that we gave for the corresponding result for gaussian numbers.

12.3 Quadratic number rings

We want to determine the number ring

$$A = \mathbb{Q}(\sqrt{d}) \cap \bar{\mathbb{Z}}$$

associated to the number field $\mathbb{Q}(\sqrt{d})$, ie we want to find which numbers $x + y\sqrt{d}$ are algebraic integers.

Theorem 12.1. *Suppose*

$$z = x + y\sqrt{d} \in \mathbb{Q}(\sqrt{d}).$$

Then

1. *If* $d \not\equiv 1 \bmod 4$
$$z \in \bar{\mathbb{Z}} \iff z = m + n\sqrt{d},$$
where $m, n \in \mathbb{Z}$.
2. *If* $d \equiv 1 \bmod 4$ *then*
$$z \in \bar{\mathbb{Z}} \iff z = \frac{m + n\sqrt{d}}{2},$$
where $m, n \in \mathbb{Z}$ *and* $m \equiv n \bmod 2$.

Proof. If

$$z = x + y\sqrt{d} \in \bar{\mathbb{Z}}$$

then

$$\bar{z} = x \in y\sqrt{d} \in \bar{\mathbb{Z}}$$

since z and $\bar{z}$ satisfy the same polynomials over $\mathbb{Q}$. Hence

$$z + \bar{z} = 2x \in \bar{\mathbb{Z}} \cap \mathbb{Q} = \mathbb{Z}$$

the other hand if m, n are both odd then

$$m^2 \equiv n^2 \equiv 1 \bmod 4.$$

It follows that

$$d \equiv 1 \bmod 4.$$

In other words, if $d \not\equiv 1 \bmod 4$ then m, n are even, and so

$$z = a + b\sqrt{d},$$

with $a, b \in \mathbb{Z}$.

On the other hand, if $d \equiv 1 \bmod 4$ then m, n are both even or both odd.

It only remains to show that if $d \equiv 1 \bmod 4$ and m, n are both odd then

$$z = \frac{m + n\sqrt{d}}{2} \in \bar{\mathbb{Z}},$$

It is sufficient to show that

$$\theta = \frac{1 + \sqrt{d}}{2} \in \bar{\mathbb{Z}},$$

since

$$z = (a + b\sqrt{d}) + \theta,$$

where

$$a = (m-1)/2,\ b = (n-1)/2 \in \mathbb{Z}.$$

But

$$(\theta - 1/2)^2 = d/4,$$

ie

$$\theta^2 - \theta + (1-d)/4.$$

But $(1-d)/4 \in \mathbb{Z}$ if $d \equiv 1 \bmod 4$. Hence

$$\theta \in \bar{\mathbb{Z}}.$$

$\square$

12.4 Units I: Imaginary quadratic fields

Suppose F is a number field, with associated number ring A (the algebraic integers in F). By 'abuse of language', as the French say, we shall speak of the units of F when we are really referring to the units in A.

Proposition 12.4. *Suppose $z \in \mathbb{Q}(\sqrt{d})$ is an algebraic integer. Then*

$$z \text{ is a unit} \iff \mathcal{N}(z) = \pm 1.$$

Proof. Suppose z is a unit, say

$$zw = 1,$$

where w is also an integer. Then

$$\mathcal{N}(zw) = \mathcal{N}(z)\mathcal{N}(w) = \mathcal{N}(1) = 1^2 = 1.$$

Since $\mathcal{N}(z), \mathcal{N}(w) \in \mathbb{Z}$ it follows that

$$\mathcal{N}(z) = \mathcal{N}(w) = \pm 1.$$

If $d \equiv 1 \bmod 4$ then

$$\epsilon = \frac{m + n\sqrt{d}}{2},$$

where $m, n \in \mathbb{Z}$ with $m \equiv n \bmod 2$. In this case,

$$\mathcal{N}(\epsilon) = \frac{m^2 - dn^2}{4} = 1,$$

ie

$$m^2 - dn^2 = 4.$$

If $d \leq -7$ then this implies that $m = \pm 1,\ n = 0$. This only leaves the case $d = -3$, where

$$m^2 + 3n^2 = 4.$$

This has 6 solutions: $m = \pm 2,\ n = 0$, giving $\epsilon = \pm 1$; and $m = \pm 1,\ n = \pm 1$, giving $\epsilon = \pm\omega, \pm\omega^2$. □

Units in real quadratic fields (where $d > 0$) have a very different character, requiring a completely new idea from the theory of *diophantine approximation*; we leave this to another Chapter.

*** 6. Show that the real number ring $\mathbb{Z}[\sqrt{2}]$ is a Unique Factorisation Domain, and determine the primes in this ring.

In exercises 7-10, determine the prime factorisation of the given number in the ring $\mathbb{Z}[\sqrt{2}]$.

*** 7. 2

*** 8. 7

*** 9. $2+\sqrt{2}$

*** 10. $3+\sqrt{3}$

*** 11. Show that the ring $\mathbb{Z}[\sqrt{5}]$ is not a Unique Factorisation Domain. [Note: this is not the number ring associated to the field $\mathbb{Q}(\sqrt{5})$.]

*** 12. Show that the imaginary number ring $\mathbb{Z}[\omega]$ (where $\omega^3 = 1,\ \omega \neq 1$) is a Unique Factorisation Domain, and determine the primes in this ring.

In exercises 13-15, determine the prime factorisation of the given number in the ring $\mathbb{Z}[\omega]$.

*** 13. $1-\omega$

*** 14. $2+\omega$

*** 15. $2-\omega$

*** 16. Show that the imaginary number ring $\mathbb{Z}[\sqrt{-5}]$ is not a Unique Factorisation Domain, by considering the factorisations of the number 6 in this ring, or in any other way.

**** 17. Determine if the imaginary number ring $\mathbb{Z}[\sqrt{-6}]$ is a Unique Factorisation Domain.

**** 18. Determine if the imaginary number ring $\mathbb{Z}[\sqrt{-7}]$ is a Unique Factorisation Domain.

**** 19. Show that the real number ring $\mathbb{Z}[\sqrt{6}]$ is a Unique Factorisation Domain.

**** 20. Show that the real number ring $\mathbb{Z}[\sqrt{7}]$ is a Unique Factorisation Domain.

Pell's Equation

14.1 Kronecker's Theorem

Diophantine approximation concerns the approximation of real numbers by rationals. Kronecker's Theorem is a major result in this subject, and a very nice application of the Pigeon Hole Principle.

Theorem 14.1. *Suppose $\theta \in \mathbb{R}$; and suppose $N \in \mathbb{N}$, $N \neq 0$. Then there exists $m, n \in \mathbb{Z}$ with $0 < n \leq N$ such that*

$$|n\theta - m| < \frac{1}{N}.$$

Proof. If $x \in \mathbb{R}$ we write $\{x\}$ for the fractional part of x, so that

$$x = [x] + \{x\}.$$

Consider then $N + 1$ fractional parts

$$0, \{\theta\}, \{2\theta\}, \ldots \{N\theta\};$$

and consider the partition of $[0, 1)$ into N equal parts;

$$[0, 1/N), [1/N, 2/N), \ldots, [(N-1)/N, 1).$$

By the pigeon-hole principal, two of the fractional parts must lie in the same partition, say

$$\{i\theta\}, \{j\theta\} \in [t/N, (t+1)/N],$$

where $0 \leq i < j < N$. Setting

$$[i\theta] = r, \ [j\theta] = s,$$

we can write this as

$$i\theta - r, \ j\theta - s \in [t/N, (t+1)/N).$$

Hence

$$|(j\theta - s) - (i\theta - r)| < 1/N,$$

ie

$$|n\theta - m| < 1/N,$$

where $n = j - i$, $m = r - s$ with $0 < n \leq N$. □

Corollary 14.1. *If $\theta \in \mathbb{R}$ is irrational then there are an infinity of rational numbers m/n such that*

$$|\theta - \frac{m}{n}| < \frac{1}{n^2}.$$

Proof. By the Theorem,

$$\begin{aligned} |\theta - \frac{m}{n}| &< \frac{1}{nN} \\ &\leq \frac{1}{n^2}. \end{aligned}$$

□

which in turn gives

$$\mathcal{N}(z) = 1,$$

where

$$z = x + ay\sqrt{d'}.$$

Thus z is a unit in the quadratic number field $\mathbb{Q}(\sqrt{d'}$.

Let us denote the group of units in this number field by U. Every unit $\epsilon \in U$ is not necessarily of this form. Firstly the coefficient of $\sqrt{d'}$ must be divisible by a; and secondly, if $d' \equiv 1 \bmod 4$ then we are omitting the units of the form $(m + n\sqrt{d'})/2$.

But it is not difficult to see that these units form a subgroup $U' \subset U$ of finite index in U. It follows that U' is infinite if and only if U is infinite.

However, we shall not pursue this line of enquiry, since it is just as easy to work with these numbers in the form

$$z = x + y\sqrt{d}.$$

In particular, if

$$z = m + n\sqrt{d}, \ w = M + N\sqrt{d}$$

then

$$zw = (mM + dnN) + (mN + nM)\sqrt{d};$$

and on taking norms (ie multiplying each side by its conjugate),

$$(m^2 - dn^2)(M^2 - dN^2) = (mM + dnN)^2 - d(mN + nM)^2$$

Similarly,

$$\begin{aligned}\frac{z}{w} &= \frac{(m + n\sqrt{d})(M - N\sqrt{d})}{M^2 - dN^2} \\ &= \frac{(mM + dnN) - (mN - nM)\sqrt{d}}{M^2 - dN^2}.\end{aligned}$$

On taking norms,

$$\frac{m^2 - dn^2}{M^2 - dN^2} = u^2 - dv^2,$$

where

$$u = \frac{mM + dnN}{M^2 - dN^2}, \ \frac{mN - nM}{M^2 - dN^2}.$$

Now to the proof.

Proof. By the Corollary to Kronecker's Theorem there exist an infinity of $m, n \in \mathbb{Z}$ such that

$$|\sqrt{d} - \frac{m}{n}| < \frac{1}{n^2}.$$

Since

$$\sqrt{d} + \frac{m}{n} = 2\sqrt{d} - (\sqrt{d} - \frac{m}{n})$$

it follows that

$$|\sqrt{d} + \frac{m}{n}| < 2\sqrt{d} + 1.$$

Hence

$$\begin{aligned}|d - \frac{m^2}{n^2}| &= |\sqrt{d} - \frac{m}{n}| \cdot |\sqrt{d} + \frac{m}{n}| \\ &< \frac{2\sqrt{d} + 1}{n^2}.\end{aligned}$$

Thus

$$|m^2 - dn^2| < 2\sqrt{d} + 1.$$

$$\begin{aligned} mM - dnN \equiv m^2 - dn^2 &= t \bmod T \\ &\equiv 0 \bmod T \end{aligned}$$

(since $t = \pm T$); and similarly

$$\begin{aligned} mN - nM &\equiv mn - nm \bmod T \\ &\equiv 0 \bmod T. \end{aligned}$$

Thus

$$T \mid mM - dnN,\ mN - nM$$

and so

$$u, v \in \mathbb{Z}.$$

□

14.3 Units II: Real quadratic fields

Theorem 14.3. *Suppose $d > 1$ is square-free. Then there exists a unique unit $\epsilon > 1$ in $\mathbb{Q}(\sqrt{d})$ such that the units in this field are*

$$\pm\epsilon^n$$

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

Proof. We know that the equation

$$x^2 - dy^2 = 1$$

has an infinity of solutions. In particular it has a solution $(x, y) \neq (\pm 1, 0)$.

Let

$$\eta = x + y\sqrt{d}.$$

Then

$$\mathcal{N}(\eta) = 1;$$

so η is a unit $\neq \pm 1$.

We may suppose that $\eta > 1$; for of the 4 units $\pm\eta, \pm\eta^{-1}$ just one appears in each of the intervals $(-\infty, -1),\ (-1, 0),\ (0, 1),\ (1, \infty)$.

Lemma 14.1. *There are only a finite number of units in $(1, C)$, for any $C > 1$.*

Proof. Suppose

$$\epsilon = \frac{m + n\sqrt{d}}{2} \in (1, C)$$

is a unit. Then

$$\bar{\epsilon} = \frac{m - n\sqrt{d}}{2} = \pm\epsilon^{-1}.$$

Thus

$$-1 \leq \frac{m - n\sqrt{d}}{2} \leq 1.$$

Hence

$$0 < m < C + 1.$$

Since

$$m^2 - dn^2 = \pm 4$$

it follows that

$$n^2 < m^2 + 4 < (C + 1)^2 + 4.$$

□

We have seen that there is a unit $\eta > 1$. Since there are only a finite number of units in $(1, \eta]$ there is a least such unit ϵ.

Now suppose $\eta > 1$ is a unit. Since $\epsilon > 1$,

$$\epsilon^n \to \infty \text{ as } n \to \infty.$$

$\mathbb{Q}(\sqrt{5})$ and the golden ratio

15.1 The field $\mathbb{Q}(\sqrt{5})$

Recall that the quadratic field

$$\mathbb{Q}(\sqrt{5}) = \{x + y\sqrt{5} : x, y \in \mathbb{Q}\}.$$

Recall too that the conjugate and norm of

$$z = x + y\sqrt{5}$$

are

$$\bar{z} = x - y\sqrt{5},\ \mathcal{N}(z) = z\bar{z} = x^2 - 5y^2.$$

We will be particularly interested in one element of this field.

Definition 15.1. *The* golden ratio *is the number*

$$\phi = \frac{1+\sqrt{5}}{2}.$$

The Greek letter ϕ (phi) is used for this number after the ancient Greek sculptor Phidias, who is said to have used the ratio in his work.

Leonardo da Vinci explicitly used ϕ in analysing the human figure.

Evidently

$$\mathbb{Q}(\sqrt{5}) = \mathbb{Q}(\phi),$$

ie each element of the field can be written

$$z = x + y\phi \quad (x, y \in \mathbb{Q}).$$

The following results are immediate:

Proposition 15.1.

1. $\bar{\phi} = \frac{1-\sqrt{5}}{2}$;
2. $\phi + \bar{\phi} = 1,\ \phi\bar{\phi} = -1$;
3. $\mathcal{N}(x + y\phi) = x^2 + xy - y^2$;
4. $\phi, \bar{\phi}$ *are the roots of the equation*

$$x^2 - x - 1 = 0.$$

15.2 The number ring $\mathbb{Z}[\phi]$

As we saw in the last Chapter, since $5 \equiv 1 \bmod 4$ the associated number ring

$$\mathbb{Z}(\mathbb{Q}(\sqrt{5})) = \mathbb{Q}(\sqrt{5}) \cap \bar{\mathbb{Z}}$$

consists of the numbers

$$\frac{m + n\sqrt{5}}{2},$$

where $m \equiv n \bmod 2$, ie m, n are both even or both odd. And we saw that this is equivalent to

Proposition 15.2. *The number ring associated to the quadratic field* $\mathbb{Q}(\sqrt{5})$ *is*

$$\mathbb{Z}[\phi] = \{m + n\phi : m, n \in \mathbb{Z}\}.$$

15.3 Unique Factorisation

Then

$$\frac{z}{w} - q = (x - m) + (y - n)\phi.$$

Hence

$$\mathcal{N}(\frac{z}{w} - q) = (x - m)^2 + (x - m)(y - n) - (y - n)^2.$$

It follows that

$$-\frac{1}{2} < \mathcal{N}(\frac{z}{w} - q) < \frac{1}{2},$$

and so

$$|\mathcal{N}(\frac{z}{w} - q)| \leq \frac{1}{2} < 1,$$

ie

$$|\mathcal{N}(z - qw)| < |\mathcal{N}(w)|.$$

□

This allows us to apply the euclidean algorithm in $\mathbb{Z}[\phi]$, and establish

Lemma 15.2. *Any two numbers $z, w \in \mathbb{Z}[\phi]$ have a greatest common divisor δ such that*

$$\delta \mid z, w$$

and

$$\delta' \mid z, w \implies \delta' \mid \delta.$$

Also, δ is uniquely defined up to multiplication by a unit.

Moreover, there exists $u, v \in \mathbb{Z}[\phi]$ such that

$$uz + vw = \delta.$$

From this we deduce that irreducibles in $\mathbb{Z}[\phi]$ are primes.

Lemma 15.3. *If $\pi \in \mathbb{Z}[\phi]$ is irreducible and $z, w \in \mathbb{Z}[phi]$ then*

$$\pi \mid zw \implies \pi \mid z \text{ or } \pi \mid w.$$

Now Euclid's Lemma , and Unique Prime Factorisation, follow in the familiar way. □

15.4 The units in $\mathbb{Z}[\phi]$

Theorem 15.2. *The units in $\mathbb{Z}[\phi]$ are the numbers*

$$\pm\phi^n \quad (n \in \mathbb{Z}).$$

Proof. We saw in the last Chapter that any real quadratic field contains units $\neq \pm 1$, and that the units form the group

$$\{\pm\epsilon^n : n \in \mathbb{Z}\},$$

where ϵ is the smallest unit > 1.

Thus the theorem will follow if we establish that ϕ is the smallest unit > 1 in $\mathbb{Z}[\phi]$.

Suppose $\eta \in \mathbb{Z}[\phi]$ is a unit with

$$1 < \eta = m + n\phi \leq \phi.$$

Then

$$\mathcal{N}(\eta) = \eta\bar{\eta} = \pm 1,$$

But
$$-1 + \phi < 1.$$
Hence
$$m \geq 0,$$
and so
$$\eta \geq \epsilon.$$
$\square$

15.5 The primes in $\mathbb{Z}[\phi]$

Theorem 15.3. *Suppose $p \in \mathbb{N}$ is a rational prime.*

1. *If $p \equiv \pm 1 \bmod 5$ then p splits into conjugate primes in $\mathbb{Z}[\phi]$:*
$$p = \pm \pi \bar{\pi}.$$

2. *If $p \equiv \pm 2 \bmod 5$ then p remains prime in $\mathbb{Z}[\phi]$.*

Proof. Suppose p splits, say
$$p = \pi \pi'.$$
Then
$$\mathcal{N}(p) = p^2 = \mathcal{N}(\pi)\mathcal{N}(\pi').$$
Hence
$$\mathcal{N}(\pi) = \mathcal{N}(\pi') = \pm p.$$

Suppose
$$\pi = m + n\phi.$$
Then
$$\mathcal{N}(\pi) = m^2 - mn - n^2 = \pm p,$$
and in either case
$$m^2 - mn - n^2 \equiv 0 \bmod p.$$

If $p = 2$ then m and n must both be even. (For if one or both of m, n are odd then so is $m^2 - mn - n^2$.) Thus
$$2 \mid \pi,$$
which is impossible.

Now suppose p is odd, Multiplying by 4,
$$(2m - n)^2 - 5n^2 \equiv 0 \bmod p.$$

But
$$n \equiv 0 \bmod p \implies m \equiv 0 \bmod p \implies p \mid \pi,$$
which is impossible. Hence $n \not\equiv 0 \bmod p$, and so
$$r^2 \equiv 5 \bmod p,$$
where
$$r \equiv (2m - n)/n \bmod p.$$

Thus
$$\left(\frac{5}{p}\right) = 1.$$

$$p \mid n - \sqrt{5} \text{ or } p \mid n + \sqrt{5},$$

both of which imply that $p \mid 1$, which is absurd.

We conclude that

$$p \equiv \pm 1 \bmod 5 \implies p \text{ splits in } \mathbb{Z}[\phi].$$

Finally we have seen in this case that if $\pi \mid p$ then

$$\mathcal{N}(\pi) = \pm p \implies p = \pm \pi \bar{\pi}.$$

□

15.6 Fibonacci numbers

Recall that the Fibonacci sequence consists of the numbers

$$0, 1, 1, 2, 3, 5, 8, 13, \ldots$$

defined by the *linear recurrence relation*

$$F_{n+1} = F_n + F_{n-1},$$

with initial values

$$F_0 = 0, \; F_1 = 1.$$

There is a standard way of solving a general linear recurrence relation

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \cdots + a_d x_{n-d}.$$

Let the roots of the *associated polynomial*

$$p(t) = t^d - c_1 t^{d-1} - c_2 t^{d-2} + \cdots + c_d.$$

be $\lambda_1, \ldots, \lambda_d$.

If these roots are distinct then the general solution of the recurrence relation is

$$x_n = C_1 \lambda_1^n + C_2 \lambda_2^n + \cdots + C_d \lambda_d^n.$$

The coefficients $C_1, \ldots, C_d$ are determined by d 'initial conditions', eg by specifying $x_0, \ldots, x_{d-1}$.

If there are multiple roots, eg if λ occurs e times then the term $C\lambda^n$ must be replaced by $\lambda^n p(\lambda)$, where p is a polynomial of degree e.

But these details need not concern us, since we are only interested in the Fibonacci sequence, with associated polynomial

$$t^2 - t - 1.$$

This has roots $\phi, \bar{\phi}$. Accordingly,

$$F_n = A\phi^n + B\bar{\phi}^n.$$

Substituting for $F_0 = 0, \; F_1 = 1$, we get

$$A + B = 0, \; A\phi + B\bar{\phi} = 1.$$

Thus

$$B = -A, \; A(\phi - \bar{\phi}) = 1.$$

Since

$$\phi - \bar{\phi} = \frac{1 + \sqrt{5}}{2} - \frac{1 - \sqrt{5}}{2} = \sqrt{5},$$

this gives

when $p \equiv 3 \bmod 4$. In the next Chapter we shall give a stronger version which works for all primes.

Proposition 15.4. *Suppose the prime* $p \equiv 3 \bmod 4$. *Then*

$$P = 2^p - 1$$

is prime if and only if

$$\phi^{2^p} \equiv -1 \bmod P.$$

Proof. Suppose first that P is a prime.

Since $p \equiv 3 \bmod 4$ and $2^4 \equiv 1 \bmod 5$,

$$\begin{aligned} 2^p &\equiv 2^3 \bmod 5 \\ &\equiv 3 \bmod 5. \end{aligned}$$

Hence

$$P = 2^p - 1 \equiv 2 \bmod 5.$$

Now

$$\begin{aligned} \phi^P &= \left(\frac{1+\sqrt{5}}{2}\right)^P \\ &\equiv \frac{1^P + (\sqrt{5})^P}{2^P} \bmod P, \end{aligned}$$

since P divides all the binomial coefficients except the first and last. Thus

$$\phi^P \equiv \frac{1 + 5^{(P-1)/2}\sqrt{5}}{2} \bmod P,$$

since $2^P \equiv 2 \bmod P$ by Fermat's Little Theorem.

But

$$5^{(P-1)/2} \equiv \left(\frac{5}{P}\right),$$

by Euler's criterion. Hence by Gauss' Quadratic Reciprocity Law,

$$\begin{aligned} \left(\frac{5}{P}\right) &= \left(\frac{P}{5}\right) \\ &= -1, \end{aligned}$$

since $P \equiv 2 \bmod 5$. Thus

$$5^{(P-1)/2} \equiv -1 \bmod P,$$

and so

$$\phi^P \equiv \frac{1 - \sqrt{5}}{2} \bmod P.$$

But

$$\begin{aligned} \frac{1-\sqrt{5}}{2} &= \bar{\phi} \\ &= -\phi^{-1}. \end{aligned}$$

It follows that

$$\phi^{P+1} \equiv -1 \bmod P,$$

ie

$$\phi^{2^p} \equiv -1 \bmod P.$$

and so, by the argument above, the order of $\phi \bmod Q$ is 2^{p+1}.

We want to apply Fermat's Little Theorem, but we need to be careful since we are working in $\mathbb{Z}[\phi]$ rather than $\mathbb{Z}$.

Lemma 15.4 (Fermat's Little Theorem, extended). *If the rational prime Q does not split in $\mathbb{Z}[\phi]$ then*

$$z^{Q^2-1} \equiv 1 \bmod Q$$

for all $z \in \mathbb{Z}[\phi]$ with $z \not\equiv 0 \bmod Q$.

Proof. The quotient-ring $A = \mathbb{Z}[\phi] \bmod Q$ is a field, by exactly the same argument that $\mathbb{Z} \bmod p$ is a field if p is a prime. For if $z \in A$, $z \neq 0$ then the map

$$w \mapsto zw : A \to A$$

is injective, and so surjective (since A is finite). Hence there is an element z' such that $zz' = 1$, ie z is invertible in A.

Also, A contains just Q^2 elements, represented by

$$m + n\sqrt{5} \quad (0 \leq m, n < Q).$$

Thus the group

$$A^\times = A \setminus 0$$

has order $Q^2 - 1$, and the result follows from Lagrange's Theorem. □

In particular, it follows from this Lemma that

$$\phi^{Q^2-1} \equiv 1 \bmod Q,$$

ie the order of $\phi \bmod Q$ divides $Q^2 - 1$. But we know that the order of $\phi \bmod Q$ is 2^{p+1}. Hence

$$2^{p+1} \mid Q^2 - 1 = (Q-1)(Q+1).$$

But

$$\gcd(Q-1, Q+1) = 2.$$

It follows that either

$$2 \parallel Q-1,\ 2^p \mid Q+1 \text{ or } 2 \parallel Q+1,\ 2^p \mid Q-1.$$

Since $Q \leq P = 2^p - 1$, the only possibility is

$$2^p \mid Q+1,$$

ie $Q = P$, and so P is prime. □

This result can be expressed in a different form, more suitable for computation.

Note that

$$\phi^{2^p} \equiv -1 \bmod P$$

can be re-written as

$$\phi^{2^{p-1}} + \phi^{2^{-(p-1)}} \equiv 0 \bmod P.$$

Let

$$t_i = \phi^{2^i} + \phi^{2^{-i}}$$

Then

$$\begin{aligned} t_i^2 &= \phi^{2^{i+1}} + 2 + \phi^{2^{-(i+1)}} \\ &= t_{i+1} + 2, \end{aligned}$$

16.1 The field $\mathbb{Q}(\sqrt{3})$

We have

$$\mathbb{Q}(\sqrt{3}) = \{x + y\sqrt{3} : x, y \in \mathbb{Q}\}.$$

The conjugate and norm of

$$z = x + y\sqrt{3}$$

are

$$\bar{z} = x - y\sqrt{3},\ \mathcal{N}(z) = z\bar{z} = x^2 - 3y^2.$$

16.2 The ring $\mathbb{Z}[\sqrt{3}]$

Since $3 \not\equiv 1 \bmod 4$,

$$\mathbb{Z}(\mathbb{Q}(\sqrt{3})) = \mathbb{Q}(\sqrt{3}) \cap \bar{\mathbb{Z}} = \{m + n\sqrt{3} : m, n \in \mathbb{Z}\} = \mathbb{Z}[\sqrt{3}].$$

16.3 The units in $\mathbb{Z}[\sqrt{3}]$

Evidently

$$\epsilon = 2 + \sqrt{3}$$

is a unit, since

$$\mathcal{N}(\epsilon) = 2^2 - 3 \cdot 1^2 = 1,$$

Theorem 16.1. *The units in $\mathbb{Z}[\phi]$ are the numbers*

$$\pm\epsilon^n \quad (n \in \mathbb{Z}),$$

where

$$\epsilon = 2 + \sqrt{3}.$$

Proof. We have to show that ϵ is the smallest unit > 1.

Suppose $\eta = m + n\sqrt{3}$ is a unit satisfying

$$1 < \eta \le \epsilon.$$

Since $\mathcal{N}(\eta) = \eta\bar{\eta} = \pm 1$,

$$\bar{\eta} = m - n\sqrt{3} = \pm\eta^{-1} \in (-1, 1).$$

Hence

$$\eta - \bar{\eta} = 2n\sqrt{3} \in (0, 1 + \epsilon),$$

ie

$$0 < n < (3 + \sqrt{3})/2\sqrt{3} < 2.$$

Thus

$$n = 1.$$

But now

$$\mathcal{N}(\eta) = \pm 1 \implies m^2 - 3 = \pm 1$$
$$\implies m = \pm 2$$

and choose the nearest integers m, n to x, y, so that

$$|x-m|, |y-m| \leq \frac{1}{2}.$$

Then we set

$$q = m + n\sqrt{3},$$

so that

$$\frac{z}{w} - q = (x-m) + (y-n)\sqrt{3},$$

and

$$\mathcal{N}(\frac{z-qw}{w}) = (x-m)^2 - 3(y-n)^2.$$

Now

$$-\frac{3}{4} \leq \mathcal{N}(\frac{z-qw}{w}) \leq \frac{1}{4}.$$

In particular,

$$|\mathcal{N}(\frac{z-qw}{w})| < 1,$$

ie

$$|\mathcal{N}(z-qw)| < |\mathcal{N}(w)|.$$

This allows the Euclidean Algorithm to be used in $\mathbb{Z}[\sqrt{3}]$, and as a consequence Eulid's Lemma holds, and unique factorisation follows. □

16.5 The primes in $\mathbb{Z}[\sqrt{3}]$

Theorem 16.3. *Suppose $p \in \mathbb{N}$ is a rational prime. Then*

1. *If $p = 2$ or 3 then p ramifies in $\mathbb{Z}[\sqrt{3}]$;*
2. *If $p \equiv \pm 1 \bmod 12$ then p splits into conjugate primes in $\mathbb{Z}[\sqrt{3}]$,*

$$p = \pm\pi\bar{\pi};$$

3. *If $p \equiv \pm 5 \bmod 12$ then p remains prime in $\mathbb{Z}[\sqrt{3}]$.*

Proof. To see that 2 ramifies, note that

$$(1+\sqrt{3})^2 = 2\epsilon,$$

where $epsilon = 2 + \sqrt{3}$ is a unit. It is evident that $3 = \sqrt{3}^2$ ramifies.

Suppose $p \neq 2, 3$.

If p splits, say

$$p = \pi\pi',$$

then

$$\mathcal{N}(p) = p^2 = \mathcal{N}(\pi)\mathcal{N}(\pi').$$

Hence

$$\mathcal{N}(\pi) = \mathcal{N}(\pi') = \pm p.$$

Thus if $\pi = m + n\sqrt{3}$ then

$$m^2 - 3n^2 = \pm p.$$

In particular,

$$m^2 - 3n^2 \equiv 0 \bmod p.$$

Now

$$n \equiv 0 \bmod p \implies m \equiv 0 \bmod p \implies p \mid \pi$$

in which case we can find a such that

$$a^2 \equiv 3 \bmod p,$$

ie

$$p \mid (a^2 - 3) = (a - \sqrt{3})(a + \sqrt{3}).$$

If now p does *not* split then this implies that

$$p \mid a - \sqrt{3} \text{ or } p \mid a + \sqrt{3}.$$

But both these imply that $p \mid 1$, which is absurd. □

16.6 The Lucas-Lehmer test for Mersenne primality

Theorem 16.4. *If p is prime then*

$$P = 2^p - 1$$

is prime if and only if

$$\epsilon^{2^{p-1}} \equiv -1 \bmod P,$$

where

$$\epsilon = 2 + \sqrt{3}.$$

Proof. Suppose P is prime. Then

$$\epsilon^P \equiv 2^P + (\sqrt{3})^P \bmod P,$$

since

$$P \mid \binom{r}{P}$$

for $r \neq 0, P$.

But

$$2^P \equiv 2 \bmod P$$

by Fermat's Little Theorem, while

$$(\sqrt{3})^{P-1} = 3^{\frac{P-1}{2}} \equiv \left(\frac{3}{P}\right) \bmod P$$

by Euler's criterion. Thus

$$\epsilon^P \equiv 2 + \left(\frac{3}{P}\right)\sqrt{3}.$$

Now

$$2^p \equiv (-1)^p \equiv -1 \bmod 3 \implies P \equiv 1 \bmod 3,$$

while

$$4 \mid 2^p \implies P \equiv -1 \bmod 4.$$

So by Gauss' Reciprocity,

$$\begin{aligned}\left(\frac{3}{P}\right) &= -\left(\frac{P}{3}\right)\\ &= -\left(\frac{1}{3}\right)\end{aligned}$$

But now

$$(1-\sqrt{3})(1+\sqrt{3}) = -2,$$

and so

$$1-\sqrt{3} = -2(1+\sqrt{3})^{-1}.$$

Thus

$$(1+\sqrt{3})^{P+1} \equiv -2 \bmod P,$$

ie

$$(1+\sqrt{3})^{2^p} \equiv -2 \bmod P,$$

ie

$$(2\epsilon)^{2^{p-1}} \equiv -2 \bmod P.$$

To deal with the powers of 2, note that by Euler's criterion

$$2^{(P-1)/2} \equiv \left(\frac{2}{P}\right) \bmod P.$$

Recall that

$$\left(\frac{2}{P}\right) = \begin{cases} 1 \text{ if } P \equiv \pm 1 \bmod 8, \\ -1 \text{ if } P \equiv \pm 1 \bmod 8. \end{cases}$$

In this case,

$$P = 2^p - 1 \equiv -1 \bmod 8.$$

Thus

$$2^{(P-1)/2} \equiv 1 \bmod P,$$

and so

$$2^{(P+1)/2} \equiv 2 \bmod P,$$

ie

$$2^{2^{p-1}} \equiv 2 \bmod P.$$

So our previous result simplifies to

$$\epsilon^{2^{p-1}} \equiv -1 \bmod P.$$

This was on the assumption that P is prime. Suppose now that P is not prime, but that the above result holds.

Then P has a prime factor $Q \leq \sqrt{P}$. Also

$$\epsilon^{2^{p-1}} \equiv -1 \bmod Q.$$

It follows that the order of $\epsilon \bmod Q$ is 2^p.

But consider the quotient-ring

$$A = \mathbb{Z}[\sqrt{3}]/(Q).$$

This ring contains just Q^2 elements, represented by

$$m + n\sqrt{5} \quad (0 \leq m, n < Q).$$

Then

$$\begin{aligned} s_i^2 &= \epsilon^{2^{i+1}} + 2 + \epsilon^{2^{-(i+1)}} \\ &= s_{i+1} + 2, \end{aligned}$$

ie

$$s_{i+1} = s_i^2 - 2.$$

Since

$$s_0 = \epsilon + \epsilon^{-1} = 4$$

it follows that $s_i \in \mathbb{N}$ for all i, with the sequence starting $4, 14, 194, \ldots$.

Now we can re-state our result.

Corollary 16.1. *Let the integer sequence s_i be defined recursively by*

$$s_{i+1} = s_i^2 - 2, \; s_0 = 4.$$

Then

$$P = 2^p - 1 \text{ is prime} \iff P \mid s_{p-2}.$$

17.1 Finite continued fractions

Definition 17.1. *A* finite continued fraction *is an expression of the form*

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}},$$

where $a_i \in \mathbb{Z}$ with $a_1, \ldots, a_n \geq 1$. We denote this fraction by

$$[a_0, a_1, \ldots, a_n].$$

Example: The continued fraction

$$[2, 1, 3, 2] = 2 + \cfrac{1}{1 + \cfrac{1}{3 + \cfrac{1}{2}}}$$

represents the rational number

$$\begin{aligned} 2 + \cfrac{1}{1 + \cfrac{2}{7}} &= 2 + \frac{7}{9} \\ &= \frac{25}{9}. \end{aligned}$$

Conversely, suppose we start with a rational number, say

$$\frac{57}{33}.$$

To convert this to a continued fraction:

$$\frac{57}{33} = 1 + \frac{14}{33}.$$

Now invert the remainder:

$$\frac{33}{14} = 2 + \frac{5}{14}.$$

Again:

$$\frac{14}{5} = 2 + \frac{4}{5},$$

and again:

$$\frac{5}{4} = 1 + \frac{1}{4},$$

and finally:

$$\frac{4}{1} = 4.$$

Thus

$$\frac{57}{33} = [1, 2, 2, 1, 4]$$

below, using induction on the *length* n of the continued fraction.

We start with the continued fraction

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

setting

$$p = a_0, \ q = 1,$$

Now suppose that we have defined p, q for continued fractions of length $< n$; and suppose that under this definition

$$\alpha_1 = [a_1, a_2, \ldots, a_n] = \frac{p'}{q'}.$$

Then

$$\begin{aligned} \alpha &= a_0 + \frac{1}{\alpha_1} \\ &= a_0 + \frac{q'}{p'} \\ &= \frac{a_0 p' + q'}{p'}. \end{aligned}$$

So we set

$$p = a_0 p' + q', \ q = p'$$

as the definition of p, q for a continued fraction of length n. We set this out formally in

Definition 17.2. *The 'canonical representation' of a continued fraction*

$$[a_0, a_1, a_2, \ldots, a_n] = \frac{p}{q}$$

is defined by induction on n, setting

$$p = a_0 p' + q', \ q = p',$$

where

$$[a_1, a_2, \ldots, a_n] = \frac{p'}{q'}$$

is the canonical representation for a continued fraction of length $n - 1$. The induction is started by setting

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

Henceforth if we write

$$[a_0, a_1, a_2, \ldots, a_n] = \frac{p}{q},$$

then p, q will refer to the canonical representation defined above.

17.3 Successive approximants

Definition 17.3. *If*

$$\alpha = [a_0, a_1, \ldots, a_n]$$

then we call

$$[a_0, a_1, \ldots, a_i] = \frac{p_i}{q_i}$$

the ith convergent *or* approximant *to α (for $0 \le i \le n$).*

Example: Continuing the previous example, the successive approximants to convergents are very good approximants to the given number.

Theorem 17.1. *If*

$$\alpha = [a_0, a_1, \dots, a_n]$$

then

$$\begin{aligned} p_i &= a_i p_{i-1} + p_{i-2}, \\ q_i &= a_i q_{i-1} + q_{i-2}, \end{aligned}$$

for $i = 2, 3, \dots, n$.

Proof. We argue by induction on n.

The result follows by induction for $i \neq n$, since the convergents involved are — or can be regarded as — convergents to

$$[a_0, a_1, \dots, a_{n-1}],$$

covered by our inductive hypothesis.

It remains to prove the result for $i = n$. In this case, by the inductive definition of p, q,

$$\begin{aligned} p_n &= a_0 p'_{n-1} + q'_{n-1}, \\ p_{n-1} &= a_0 p'_{n-2} + q'_{n-2}, \\ p_{n-2} &= a_0 p'_{n-3} + q'_{n-3}. \end{aligned}$$

But now by our inductive hypothesis,

$$p'_{n-1} = a_n p'_{n-2} + p'_{n-3}, q'_{n-1} \qquad = a_n q'_{n-2} + q'_{n-3},$$

since

$$a'_{n-1} = a_n,$$

ie the $(n-1)$th entry in α' is the nth entry in α.

Hence

$$\begin{aligned} p_n &= a_0 p'_{n-1} + q'_{n-1}, \\ &= a_0(a_n p'_{n-2} + p'_{n-3}) + (a_n q'_{n-2} + q'_{n-3}), \\ &= a_n(a_0 p'_{n-2} + q'_{n-2}) + (a_0 p'_{n-3} + q'_{n-3}), \\ &= a_n p_{n-1} + p_{n-2}; \end{aligned}$$

with the second result

$$q_n = a_n q_{n-1} + q_{n-2}$$

following in exactly the same way. □

We can regard this as a recursive definition of $\frac{p_i}{q_i}$, starting with

$$\frac{p_0}{q_0} = \frac{a_0}{1}, \frac{p_1}{q_1} = \frac{a_0 a_1 + 1}{a_1},$$

and defining

$$\frac{p_2}{q_2}, \frac{p_3}{q_3}, \frac{p_4}{q_4}, \dots$$

successively.

Actually, we can go back two futher steps.

Proposition 17.1. *If we set*

$$\begin{aligned} p_{-2} = 1, \; q_{-2} = 0, \\ p_{-1} = 0, \; q_{-1} = 1, \end{aligned}$$

then

$$\frac{37}{33} = [1, 2, 2, 1, 4] = [1, 2, 2, 1, 3, 1].$$

So there are at least 2 ways of expressing x as a continued fraction.

Proposition 17.3. *A rational number $x \in \mathbb{Q}$ has just two representations as a continued fraction: one with $n = 0$ or $n > 1$, $a_n > 1$, and the other with $n > 0$ and $a_n = 1$.*

Proof. It is sufficient to show that x has just one representation of the first kind. Suppose

$$x = [a_0, a_1, \ldots, a_m] = [b_0, b_1, \ldots, b_n],$$

We may assume that $m \le n$.

We argue by induction on n. The result is trivial if $m = n = 0$.

Lemma 17.1. *If $n > 0$ and $a_n > 1$ then*

$$a_0 < [a_0, a_1, a_2, \ldots, a_n] < a_0 + 1.$$

Proof. We argue, as usual, by induction on n. This tells us that

$$[a_1, a_2, \ldots, a_n] > 1,$$

from which the result follows, since

$$[a_0, a_1, a_2, \ldots, a_n] = a_0 + \frac{1}{[a_1, a_2, \ldots, a_n]}.$$

∎

It follows that

$$[x] = a_0 = b_0.$$

Thus

$$x - a_0 = \frac{1}{[a_1, a_2, \ldots, a_m]} = \frac{1}{[b_1, b_2, \ldots, b_n]} \implies [a_1, a_2, \ldots, a_m] = [b_1, b_2, \ldots, b_n],$$

from which the result follows by induction. ∎

We will take the first form for the continued fraction of a rational number as standard, ie we shall assume that the last entry $a_n > 1$ unless the contrary is stated.

17.5 A fundamental identity

Theorem 17.2. *Successive convergents $p_i/q_i, p_{i+1}/q_{i+1}$ to the continued fraction $[a_0, a_1, \ldots, a_n]$ satisfy the identity*

$$p_i q_{i+1} - q_i p_{i+1} = (-1)^{i+1}.$$

Proof. We argue by induction on i, using the relations

$$\begin{aligned} p_i &= a_i p_{i-1} + p_{i-2}, \\ q_i &= a_i q_{i-1} + q_{i-2}. \end{aligned}$$

Thus

$$\begin{aligned} p_i q_{i+1} - q_i p_{i+1} &= p_i(a_{i+1} q_i + q_{i-1}) = q_i(a_{i+1} p_i + p_{i-1}) \\ &= p_i q_{i-1} - q_i p_{i-1} \\ &= -(p_{i-1} q_i - q_{i-1} p_i) \\ &= -(-1)^i \\ &= (-1)^{i+1}. \end{aligned}$$

It follows that p_{i+2}/q_{i+2} is closer than p_i/q_i to p_{i+1}/q_{i+1}. Hence

$$\frac{p_i}{q_i} < \frac{p_{i+2}}{q_{i+2}} < \frac{p_{i+1}}{q_{i+1}}.$$

So the even convergents are increasing; and similarly the odd convergents are decreasing.

Also, any even convergent is less than any odd convergent; for if i is even and j is odd then

$$\frac{p_i}{q_i} < \frac{p_{i+j-1}}{q_{i+j-1}} < \frac{p_{i+j}}{q_{i+j}} < \frac{p_j}{q_j}.$$

And since x is equal to the last convergent, it must be sandwiched between the even and odd convergents. ☐

17.6 Infinite continued fractions

So far we have been considering continued fraction expansions of *rational* numbers. But the concept extends to any *real* number $\alpha \in \mathbb{R}$.

Suppose α is irrational. We set

$$a_0 = [\alpha],$$

and let

$$\alpha_1 = \frac{1}{\alpha - a_0}.$$

Then we define $a_1, a_2, \dots,$ successively, setting

$$\begin{aligned} a_1 &= [\alpha_1], \\ \alpha_2 &= \frac{1}{\alpha_1 - a_1}, \\ a_2 &= [\alpha_2], \\ \alpha_3 &= \frac{1}{\alpha_2 - a_2}, \end{aligned}$$

and so on.

Proposition 17.5. *Suppose*

$$a_0, a_1, a_2, \dots \in \mathbb{Z} \text{ with } a_1, a_2, \dots > 0.$$

Let

$$[a_0, a_1, \dots, a_i] = \frac{p_i}{q_i}.$$

Then the sequence of convergents converges:

$$\frac{p_i}{q_i} \to x \text{ as } i \to \infty.$$

Proof. It follows from the finite case that the even convergents are increasing, and the odd convergents are decreasing, with the former bounded by the latter, and conversely:

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \frac{p_4}{q_4} < \cdots < \frac{p_5}{q_5} < \frac{p_3}{q_3} < \frac{p_1}{q_1}.$$

It follows that the even convergents must converge, to α say, and the odd convergents must also converge, to β say.

But if i is even,

$$\frac{p_i}{q_i} - \frac{p_{i+1}}{q_{i+1}} = \frac{1}{}.$$

Then

$$\alpha < \beta \textit{ if } n \textit{ is even,}$$
$$\alpha > \beta \textit{ if } n \textit{ is odd.}$$

Proof. This follows easily from the fact that even convergents are increasing, odd convergents decreasing. ☐

Now let a_0 be the largest first entry among rational $x < \alpha$; let a_1 be the least second entry among those rationals with a_0 as first entry; let a_2 be the largest third entry among those rationals with a_0, a_1 as first two entries; and so on. Then it is a simple exercise to show that

$$\alpha = [a_0, a_1, a_2, dots].$$

(Note that if the a_n (with given $a_0, \dots, a_{n-1}$) at the $(n+1)$th stage were unbounded then it would follow that α is rational, since

$$[a_0, \dots, a_{n-1}, x] \to [a_0, \dots, a_{n-1}]$$

if $x \to \infty$.)

☐

17.7 Diophantine approximation

Theorem 17.3. *If p_n/q_n is a convergent to $\alpha = [a_0, a_1, a_2, \dots]$ then*

$$|\alpha - \frac{p_n}{q_n}| \le \frac{1}{q_n^2}.$$

Proof. Recall that α lies between successive convergents p_n/q_n, p_{n+1}/q_{n+1}. Hence

$$\begin{aligned} |\alpha - \frac{p_n}{q_n}| &\le |\frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n}| \\ &= \frac{1}{q_n q_{n+1}} \\ &\le \frac{1}{q_n^2}. \end{aligned}$$

☐

Remarks:

1. There is in fact inequality in the theorem except in the very special case where α is rational, p_n/q_n is the last but one convergent, and $a_{n+1} = 1$; for except in this case $q_n < q_{n+1}$.

2. Since
$$\frac{1}{q_n q_{n+1}} = \frac{1}{q_n(a_n q_n + q_{n-1})} \le \frac{1}{a_n q_n^2},$$
if $a_n > 1$ then
$$|\alpha - \frac{p_n}{q_n}| \le \frac{1}{2q_n^2}.$$
In particular, if α is irrational then there are an infinity of convergents satisfying
$$|\alpha - \frac{p}{q}| \le \frac{1}{2q^2}$$
unless $a_n = 1$ for all $n \ge N$.

In this case

Solving for x,

$$\begin{aligned} x &= -\frac{q_n\alpha - p_n}{q_{n-1}\alpha - p_{n-1}} \\ &= -\frac{\alpha - p_n/q_n}{\alpha - p_{n-1}/q_{n-1}} \end{aligned}$$

We want to ensure that $x > 0$. This will be the case if

$$(\alpha - \frac{p_n}{q_n}) \text{ and } (\alpha - \frac{p_{n-1}}{q_{n-1}})$$

are of opposite sign, ie α lies between the two convergents.

At first this seems a matter of good or bad luck. But recall that there are two ways of representing p/q as a continued fraction, one of even length and one odd. (One has last entry $a_n > 1$, and the other has last entry 1.)

We can at least ensure in this way that α lies on the same side of p_n/q_n as p_{n-1}/q_{n-1}, since even convergents are $<$ odd convergents; so if $\alpha > p/q$ then we choose n to be even, while if $\alpha < p/q$ we choose n to be odd.

This ensures that $x > 0$. Now we must show that $x \geq 1$; for then if

$$x = [b_0, b_1, b_2, \ldots]$$

we have

$$\alpha = [a_0, \ldots, a_n, b_0, b_1, b_2, \ldots],$$

and

$$\frac{p}{q} = [a_0, \ldots, a_n]$$

is a convergent to α, as required.

But now

$$|\alpha - \frac{p_n}{q_n}| \leq \frac{1}{2q_n^2};$$

and since

$$|\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}}| = \frac{1}{q_n q_{n-1}}$$

it follows that

$$\begin{aligned} |\alpha - \frac{p_{n-1}}{q_{n-1}}| &\geq |\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}}| - |\alpha - \frac{p_n}{q_n}| \\ &\geq \frac{1}{q_n q_{n-1}} - \frac{1}{2q_n^2} \\ &\geq \frac{1}{q_n^2} - \frac{1}{2q_n^2} \\ &= \frac{1}{2q_n^2}, \end{aligned}$$

and so

$$|x| = \frac{|\alpha - p_n/q_n|}{|\alpha - p_{n-1}/q_{n-1}|} \geq 1.$$

$\square$

17.8 Quadratic surds and periodic continued fractions

Recall that a quadratic surd is an irrational number of the form

$$\alpha = x + y\sqrt{d},$$

where $x, y \in \mathbb{Q}$, and $d > 1$ is square-free. In other words, The converse is more difficult. Suppose

$$\alpha = [a_0, a_1, \dots]$$

satisfies the quadratic equation

$$F(x) \equiv Ax^2 + 2Bx + C = 0 \quad (A, B, C \in \mathbb{Z}).$$

Let

$$\alpha_n = [a_n, a_{n+1}, \dots].$$

We have to show that

$$\alpha_{m+n} = \alpha_n$$

for some $m, n \in \mathbb{N},\ m > 0$.

We shall do this by showing that α_n satisfies a quadratic equation with bounded coefficients.

Writing θ for a_{n+1}, for simplicity,

$$\begin{aligned} \alpha &= [a_0, \dots, a_n, \theta] \\ &= \frac{\theta p_n + p_{n-1}}{\theta q_n + q_{n-1}}. \end{aligned}$$

Thus

$$A(\theta p_n + p_{n-1})^2 + 2B(\theta p_n + p_{n-1})(\theta q_n + q_{n-1}) + C(\theta q_n + q_{n-1})^2 = 0,$$

ie

$$A'\theta^2 + 2B'\theta + C',$$

where

$$\begin{aligned} A' &= Ap_n^2 + 2Bp_n q_n + Cq_n^2, \\ B' &= Ap_n p_{n-1} + 2B(p_n q_{n-1} + p_{n-1} q_n) + Cq_n q_{n-1}, \\ C' &= Ap_{n-1}^2 + 2Bp_{n-1} q_{n-1} + Cq_{n-1}^2. \end{aligned}$$

Now

$$A' = q_n^2 F(p_n/q_n).$$

Since $F(\alpha) = 0$ and p_n/q_n is close to α, $F(p_n/q_n)$ is small.

More precisely, since

$$|\alpha - \frac{p_n}{q_n}| \le \frac{1}{q_n^2},$$

it follows by the Mean Value Theorem that

$$\begin{aligned} F(p_n/q_n) &= -(F(\alpha) - F(p_n/q_n)) \\ &= -F'(t)(\alpha - p_n/q_n), \end{aligned}$$

where $t \in [\alpha, \alpha + p_n/q_n]$.

Thus if we set

$$M = \max_{t \in [\alpha-1, \alpha+1]} |F'(t)|$$

then

$$|F(p_n/q_n)| \le \frac{M}{q_n^2}$$

and so

$$|A'| \le M.$$

Similarly

Thus A', B', C' are bounded for all n. We conclude that one (at least) of these equations occurs infinitely often; and so one of the α_n occurs infinitely often, ie α is periodic. $\square$

Example: Let us determine the continued fraction for $\sqrt{3}$. We have

$$\begin{aligned}
\sqrt{3} &= 1 + (\sqrt{3} - 1), \\
\frac{1}{\sqrt{3}-1} &= \frac{\sqrt{3}+1}{2} = 1 + \frac{\sqrt{3}-1}{2}, \\
\frac{2}{\sqrt{3}-1} &= \sqrt{3} + 1 = 2 + (\sqrt{3} - 1), \\
\frac{1}{\sqrt{3}-1} &= 1 + \frac{\sqrt{3}-1}{2}, \\
&\dots
\end{aligned}$$

Thus

$$\sqrt{3} = [1, \overline{1, 2}],$$

where we have overlined the periodic part.

*** 8. $\sqrt{11}$

*** 9. $\frac{\sqrt{3}+1}{2}$

*** 10. $7\sqrt{3}$

*** 11. Suppose the quadratic surd

$$\alpha = [a_0, a_1, \ldots]$$

satisfies the equation

$$Ax^2 + 2Bx + c = 0.$$

where $A, B, C \in \mathbb{Z}$ with $\gcd(A, B, C) = 1$. If the corresponding equation for

$$\alpha_n = [a_n, a_{n+1}, \ldots]$$

is

$$A_n x^2 + 2B_n x + c_n = 0$$

show that

$$B^2 - AC = B_n^2 - A_n C_n.$$

*** 12. Find the first 5 convergents to π.

***** 13. Show that

$$e = [2, 1, 2, 1, 1, 4, 1, 1, 6, \ldots].$$

squares

A.1 Sum of two squares

Theorem A.1. *The positive integer n is expressible as a sum of two squares,*

$$n = a^2 + b^2 \quad (a, b \in \mathbb{Z}),$$

if and only if every prime $p \equiv 3 \bmod 4$ divides n to an even power.

Proof.

Lemma A.1. *If m, n are each expressible as the sum of two squares then so is mn.*

Proof. If

$$m = a^2 + b^2,\ n = x^2 + y^2$$

then

$$mn = (ax + by)^2 + (ay - bx)^2.$$

□

Remark: The formula can be derived from the norms of complex numbers, taking

$$z = a + ib,\ w = x + iy,$$

and using the fact that

$$|zw| = |z||w|.$$

Lemma A.2. $2n$ *is a sum of two squares if and only n is a sum of two squares.*

Proof. If

$$2n = x^2 + y^2$$

then either x, y are both even, or both are odd. Thus $x \pm y$ are both even, and

$$n = \left(\frac{x+y}{2}\right)^2 + \left(\frac{x-y}{2}\right)^2.$$

Conversely,

$$n = x^2 + y^2 \implies 2n = (x + y)^2 + (x - y)^2.$$

□

Corollary A.1. *If $n = 2^e m$, where m is odd, that n is a sum of two squares if and only if that is true of m.*

Lemma A.3. *Every prime $p \equiv 1 \bmod 4$ is expressible as the sum of two squares.*

Proof. Since $\left(\dfrac{-1}{p}\right) = 1$,

$$-1 \equiv r^2 \bmod p \implies p \mid r^2 + 1.$$

Let the smallest sum of two squares divisible by p be

$$pd = a^2 + b^2.$$

If $d = 1$ we are done. Suppose $d > 1$. Let q be a prime diviisor of d. We can find x, y coprime to q such that

$$ax + by \equiv 0 \bmod q.$$

(We can regard this as a linear equation over the field $\mathbb{F}_q = \mathbb{Z}/(p)$.) We may assume that $|x|, |y| < q/2$, so that $x^2 + y^2 < q^2/2$.

Now

This argument also shows that

$$a^2 + b^2 \equiv 0 \bmod p \iff p \mid a, b.$$

□

Corollary A.2. *If the prime* $p \equiv 3 \bmod 4$ *divides* $n = a^2 + b^2$ *then* p *divides* n *to an even power.*

Proof. $p^2 \mid n$ since $b \mid a, b$. But now we can apply the same argument to

$$n/p^2 = (a/p)^2 + (b/p)^2;$$

and repeating this as often as necessary we conclude that p divides n to an even power. □

□

A.2 Sum of three squares

Theorem A.2. *The number* $n \in \mathbb{N}$ *is expressible as the sum of three squares if and only if it is not of the form*

$$n = 4^e(8m + 7).$$

The proof of the "if" part of this theorem would take us far beyond the reach of the course. It depends on the study of quadratic forms in 3 variables over $\mathbb{Z}$. But it is easy to prove the "only if" part.

Proposition A.1. *A postive integer of the form* $n = 4^e(8m + 7)$ *cannot be expressed as the sum of three squares.*

Proof. The result follows from the following two Lemmas.

Lemma A.5. *A number* $n \equiv 7 \bmod 8$ *is not expressible as the sum of three squares.*

Proof. The quadratic residues modulo 8 are 0,1,4. It is not possible to express 7 as the sum of three numbers, each equal to 0,1 or 4. □

Lemma A.6. $4n$ *is a sum of three squares if and only* n *is a sum of three squares.*

Proof. If $n = a^2 + b^2 + c^2$ then $4n = (2a)^2 + (2b)^2 + (2c)^2$.

Conversely, if $4n = a^2 + b^2 + c^2$ then by the argument in the proof of the previous Lemma a, b, c are all even, say $a = 2A, b = 2B, c = 2C$; and then $n = A^2 + B^2 + C^2$. □

□

A.3 Sum of four squares

Theorem A.3. *Every* $n \in \mathbb{N}$ *is the sum of four squares:*

$$n = a^2 + b^2 + c^2 + d^2 \quad (a, b, c, d \in \mathbb{Z}).$$

Proof. The basic idea is exactly the same as our proof that a prime $p \equiv 1 \bmod 4$ is the sum of two squares.

Lemma A.7. *If* m, n *are each expressible as the sum of four squares then so is* mn.

Proof. If

then

$$t = rs \equiv y^2 s \equiv (xy)^2 + y^2.$$

□

Applying this to $-1 \bmod p$ gives

Corollary A.3. *If p is an odd prime then every $n \in \mathbb{Z}$ is expressible as the sum of three squares modulo p, at least one of which is coprime to p:*

$$n \equiv a^2 + b^2 + c^2 \bmod p.$$

Suppose p is an odd prime.

Let the smallest sum of four squares divisible by p be

$$pd = a^2 + b^2 + c^2 + d^2$$

If $d = 1$ we are done. Suppose $d > 1$.

Let q be a prime divisor of d. If we set

$$L_1 = ax - by - cz - dt,\ L_2 = ay + bx + ct - dy,\ L_3 = az - bt + cx + dy,\ L_4 = at + bz - cy + dx$$

then

$$pd(x^2 + y^2 + z^2 + t^2) = L_1^2 + L_2^2 + L_3^2 + L_4^2.$$

Consider the 4 linear equivalences

$$L_i(x, y.z.t) \equiv 0 \bmod q \quad (i = 1, 2, 3, 4)$$

We can regard these as 4 linear equations over $\mathbb{F}_q = \mathbb{Z}/(q)$. Recall that $m < n$ simultaneous linear equations in n unknowns always have a non-trivial solution. It follows that we can find x, y, z, t, not all divisible by q, such that the last 3 equivalences hold:

$$L_2 \equiv 0,\ L_3 \equiv 0,\ L_4 \equiv 0 \bmod q;$$

and we may assume that $|x|, |x|, |x|, |x| < q/2$, so that $x^2 + y^2 + z^2 + t^2 < q^2$.

But now, since $q \mid pd$, it follows that

$$L_1 \equiv 0 \bmod q$$

also. Let

$$A = \frac{L_1}{q},\ B = \frac{L_2}{q},\ C = \frac{L_3}{q},\ D = \frac{L_4}{q}.$$

Then

$$pd' = A^2 + B^2 + C^2 + D^2,$$

where

$$d' = d\frac{x^2 + y^2 + z^2 + t^2}{q^2} < d,$$

contradicting the minimality of d. Hence $d = 1$ and

$$p = a^2 + b^2 + c^2 + d^2.$$

□

Theorem B.1. *Every finite abelian group A can be expressed as a direct sum of cyclic groups of prime-power order:*

$$A = \mathbb{Z}/(p_1^{e_1}) \oplus \cdots \oplus \mathbb{Z}/(p_r^{e_r}).$$

Moreover the powers $p_1^{e_1}, \ldots, p_r^{e_r}$ are uniquely determined by A.

Note that the primes $p_1, \ldots, p_r$ are not necessarily distinct.

We prove the result in two parts. First we divide A into its primary components A_p. Then we show that each of these components is expressible as a direct sum of cyclic groups of prime-power order.

B.2 Primary decomposition

Proposition B.1. *Suppose A is a finite abelian group. For each prime p, the elements of order p^n in A for some $n \in N$ form a subgroup*

$$A_p = \{a \in A : p^n a = 0 \textit{ for some } n \in \mathbb{N}\}.$$

Proof. Suppose $a, b \in A_p$. Then

$$p^m a = 0, \ p^n b = 0,$$

for some m, n. Hence

$$p^{m+n}(a + b) = 0,$$

and so $a + b \in A_p$. □

Definition B.1. *We call the A_p the* primary components *or* p-component *of A.*

Proposition B.2. *A finite abelian group A is the direct sum of its primary components A_p:*

$$F = \oplus_p A_p.$$

Proof. Suppose $a \in A$ By Lagrange's Theorem, $na = 0$ for some $n > 0$ Let

$$n = p_1^{e_1} \cdots p_r^{e_r};$$

and set

$$m_i = n/e_i^{p_i}.$$

Then $\gcd(m_1, \ldots, m_r) = 1$, and so we can find $n_1, \ldots, n_r$ such that

$$m_1 n_1 + \cdots + m_r n_r = 1.$$

Thus

$$a = a_1 + \cdots + a_r,$$

where

$$a_i = m_i n_i a.$$

But

$$p_i^{e_i} a_i = (p_i^{e_i} m_i) n_i a = n n_i a = 0$$

(since $na = 0$). Hence

$$a_i \in A_{p_i}.$$

Thus A is the sum of the subgroups A_p.

To see that this sum is direct, suppose

$$a_1 + \cdots + a_r = 0,$$

where $a_i \in A_{p_i}$, with distinct primes $p_1, \ldots, p_r$. Suppose

$$p_i^{e_i} a_i = 0.$$

Let

$$m_i = p_1^{e_1} \cdots p_{i-1}^{e_{i-1}} p_{i+1}^{e_{i+1}} \cdots p_r^{e_r}.$$

Then

$$pA = \{pa : a \in A\}.$$

For pA is stricty smaller than A, since

$$pA = A \implies p^n A = A,$$

while we know from Lagrange's Theorem that $p^n A = 0$.

Suppose

$$pA = \langle pa_1 \rangle \oplus \cdots \oplus \langle pa_r \rangle.$$

Then the sum

$$\langle a_1 \rangle + \cdots + \langle a_r \rangle = B,$$

say, is direct. For suppose

$$n_1 a_1 + \cdots + n_r a_r = 0.$$

If $p \mid n_1, \ldots, n_r$, say $n_i = pm_i$, then we can write the relation in the form

$$m_1(pa_1) + \cdots + m_r(pa_r) = 0,$$

whence $m_i pa_i = n_i a_i = 0$ for all i.

On the other hand, if p does not divide all the n_i then

$$n_1(pa_1) + \cdots + n_r(pa_r) = 0,$$

and so $pn_i a_i = 0$ for all i. But if $p \nmid n_i$ this implies that $pa_i = 0$. (For the order of a_i is a power of p, say p^e; while $p^e \mid n_i p$ implies that $e \leq 1$.) But this contradicts our choice of pa_i as a generator of a direct summand of pA. Thus the subgroup $B \subset A$ is expressed as a direct sum

$$B = \langle a_1 \rangle \oplus \cdots \oplus \langle a_r \rangle.$$

Let

$$K = \{a \in A : pa = 0\}.$$

Then

$$A = B + K.$$

For suppose $a \in A$. Then $pa \in pA$, and so

$$pa = n_1(pa_1) + \cdots + n_r(pa_r)$$

for some $n_1, \ldots, n_r \in \mathbb{Z}$. Thus

$$p(a - n_1 a_1 - \cdots - n_r a_r) = 0,$$

and so

$$a - n_1 a_1 - \cdots - n_r a_r = k \in K.$$

Hence

$$a = (n_1 a_1 + \cdots + n_r a_r) + k \in B + K.$$

If $B = A$ then all is done. If not, then $K \not\subset B$, and so we can find $k_1 \in K, k_1 \notin B$. Now the sum

$$B_1 = B + \langle k_1 \rangle$$

is direct. For $\langle k_1 \rangle$ is a cyclic group of order p, and so has no proper subgroups. Thus

$$B \cap \langle k_1 \rangle = \{0\},$$

and so

$$B_1 = B \oplus \langle k_1 \rangle$$

If now $B_1 = A$ we are done. If not we can repeat the construction, by choosing $k_2 \in K, k_2 \notin B_1$. As before, this gives us a direct sum

Choose an exponent e coprime to $\phi(n)$, and let $\alpha : \mathbb{Z}/(n) \to \mathbb{Z}/(n)$ be the map

$$\alpha : x \mapsto x^e.$$

Then we can determine f such that

$$ef \equiv 1 \bmod \phi(n),$$

eg by using the Euclidean algorithm. Let $\beta : \mathbb{Z}/(n) \to \mathbb{Z}/(n)$ be the map

$$\alpha : x \mapsto x^f.$$

Then if x is coprime to n

$$x^{ef} \equiv x \bmod n,$$

ie

$$\beta(\alpha(x)) = x;$$

β is the *inverse* of α, at least for x not divisible by p or q.

C.2 Encryption

Let us choose very large primes p, q, say with about 150 digits, or about 500 bits, each.

This will not take long, using either the Miller-Rabin or the AKS test. If we take an odd integer u with about 150 digits at random, and then test $u, u+1, u+2, \ldots$ for primality we can be be reasonably sure that we will meet a prime in about $\ln u \approx 15 \ln 10$ steps, by the Prime Number Theorem. (Of course we can reduce the number of tests by omitting even numbers, and perhaps numbers divisible by small primes, so the number might be reduced to a dozen or so.)

Next we choose $e \in (1, \phi(n))$ at random. We *publish* the numbers n and e — RSA is a *public key encryption* system, and these are our public keys.

Now if someone wants to send us a secret message they encode it using our public keys. We have computed the secret key f, and thus can decode the message.

We are betting that nobody can determine the factors p and q by factorising n, or determine f in some other way. In effect, we are relying on the belief that *factorisation cannot be computed in polynomial time.* More precisely, there is no algorithm that can factorise any number n in less that $P(\ln n)$ steps, where $P(x)$ is some fixed polynomial.

For example, dividing by all numbers up to $\sqrt{n}$ is an *exponential time algorithm* since

$$\sqrt{x} = e^{\ln x/2}.$$

Remarks:

1. If we want 1000-bit security, we would probably choose n to have 1024 bits, to simplify computation.

2. Note that $x^e \bmod n$ can be computed in polynomial time (probably in quadratic time) by repeatedly squaring x, always working modulo n.

3. There is an extremely small probability that some block x of the message will be divisible by p or q, and will therefore be "corrupted". However, we can ignore this possibility on the grounds that is far more likely to be corrupted in other ways.

tations involve smaller numbers, so can be carried out in less time.

'Arithmetic on elliptic curves' is probably the most active area of research in number theory today, and was the basic tool in Wiles' proof of Fermat's Last Theorem. Elliptic curves give rise to zeta functions like Riemann's, with Euler-like factorisation into terms corresponding to primes.

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}} = \begin{cases} -1 & \textit{if } p \equiv q \equiv 3 \bmod 4 \\ 1 & \textit{otherwise.} \end{cases}$$

Proof. Let

$$S = \{1, 2, \dots, \frac{p-1}{2}\},\ T = \{1, 2, \dots, \frac{q-1}{2}\}.$$

We shall choose remainders mod$p$ from the set

$$\{-\frac{p}{2} < i < \frac{p}{2}\} = -S \cup \{0\} \cup S,$$

and remainders mod$q$ from the set

$$\{-\frac{q}{2} < i < \frac{q}{2}\} = -T \cup \{0\} \cup T.$$

By Gauss' Lemma,

$$\left(\frac{q}{p}\right) = (-1)^{\mu},\ \left(\frac{p}{q}\right) = (-1)^{\nu}.$$

Writing $\#X$ for the number of elements in the set X,

$$\mu = \#\{i \in S : qi \bmod p \in -S\},\ \nu = \#\{i \in T : pi \bmod q \in -T\}.$$

By '$qi \bmod p \in -S$' we mean that there exists a j (necessarily unique) such that

$$qi - pj \in -S.$$

But now we observe that, in this last formula,

$$0 < i < \frac{p}{2} \implies 0 < j < \frac{q}{2}.$$

The basic idea of the proof is to associate to each such contribution to μ the 'point' $(i, j) \in S \times T$. Thus

$$\mu = \#\{(i, j) \in S \times T : -\frac{p}{2} < qi - pj < 0\};$$

and similarly

$$\nu = \#\{(i, j) \in S \times T : 0 < qi - pj < \frac{q}{2}\},$$

where we have reversed the order of the inequality on the right so that both formulae are expressed in terms of $(qi - pj)$.

Let us write $[R]$ for the number of integer points in the region $R \subset \mathbb{R}^2$. Then

$$\mu = [R_1],\ \nu = [R_2],$$

where

$$R_1 = \{(x, y) \in R : -\frac{p}{2} < qx - py < 0\},\ R_2 = \{(x, y) \in R : 0 < qx - py < \frac{q}{2}\},$$

and R denotes the rectangle

$$R = \{(x, y) : 0 < x < \frac{p}{2},\ 0 < y < \frac{p}{2}\}.$$

The line

$$qx - py = 0$$

is a diagonal of the rectangle R, and R_1, R_2 are strips above and below the diagonal (Fig D).

This leaves two triangular regions in R,

$$R_3 = \{(x, y) \in R : qx - py < -\frac{p}{2}\},\ R_4 = \{(x, y) \in R : qx - py > \frac{q}{2}\}.$$

Now we take P to be the centre of this rectangle, ie

$$P = (\frac{p+1}{2}, \frac{q+1}{2}).$$

The reflection is then given by

$$(x, y) \mapsto (X, Y) = (p + 1 - x, q + 1 - y).$$

It is clear that reflection in P will send the integer points of R into themselves. But it is not clear that it will send the integer points in R_3 into those in R_4, and vice versa. To see that, let us shrink these triangles as we shrank the rectangle. If $x, y \in \mathbb{Z}$ then

$$qx - py < -\frac{p}{2} \implies qx - py \leq -\frac{p+1}{2};$$

and similarly

$$qx - py > \frac{q}{2} \implies qx - py \geq \frac{q+1}{2}.$$

Now reflection in P *does* send the two lines

$$qx - py = -\frac{p+1}{2}, \; qx - py = \frac{q+1}{2}$$

into each other; for

$$qX - pY = q(p + 1 - x) - p(q + 1 - y) = (q - p) - (qx - py),$$

and so

$$qx - py = -\frac{p+1}{2} \iff qX - pY = (q - p) + \frac{p+1}{2} = \frac{q+1}{2}.$$

We conclude that

$$[R_3] = [R_4].$$

Hence

$$[R] = [R_1] + [R_2] + [R_3] + [R_4] \equiv \mu + \nu \bmod 2,$$

and so

$$\mu + \nu \equiv [R] = \frac{p-1}{2}\frac{q-1}{2}.$$

□

Example: Take $p = 37, \; q = 47$. Then

$$\begin{aligned}
\left(\frac{37}{47}\right) &= \left(\frac{47}{37}\right) \text{ since } 37 \equiv 1 \bmod 4 \\
&= \left(\frac{10}{37}\right) \\
&= \left(\frac{2}{37}\right)\left(\frac{5}{37}\right) \\
&= -\left(\frac{5}{37}\right) \text{ since } 37 \equiv -3 \bmod 8 \\
&= -\left(\frac{37}{5}\right) \text{ since } 5 \equiv 1 \bmod 4 \\
&= -\left(\frac{2}{5}\right) \\
&= -(-1) = 1.
\end{aligned}$$

Thus 37 *is* a quadratic residue mod47.