Contents

Theory of Numbers
Math 356, Rutgers
Claire Burrin

These are notes prepared for the course *Theory of Numbers*, taught at Rutgers in Spring 2019. We follow the organization of Davenport's *Higher Arithmetic* (Cambridge University Press, 8th Ed.) with some additional supplements. In particular, the material on Farey fractions is based on Rademacher's *Lectures on Elementary Number Theory*.

> *[...] the subject matter is so attractive that only extravagant incompetence could make it dull.*
>
> – G.H. Hardy, E.M. Wright, *Introduction to the Theory of Numbers*

CHAPTER 1

Numbers and primes

1.1. The elementary operations

Arithmetic (from the Greek, *arithmos*, number, *tike*, art) is the study of numbers and their properties under the elementary operations of adding, subtracting, multiplying, and dividing.

The set of (natural) numbers[1] is

$$\mathbf{N} = \{0, 1, 2, ...\}.$$

This set is infinite countable, and totally ordered (i.e. for any two elements $a, b \in \mathbf{N}$, either $a \leq b$ or $b \leq a$). We take for granted the following 'laws of arithmetic': for any $a, b, c \in \mathbf{N}$,

(1) addition and multiplication are commutative: $a + b = b + a$, $ab = ba$,
(2) addition and multiplication are associative: $a + (b + c) = (a + b) + c$, $a(bc) = (ab)c$,
(3) multiplication is distributive with respect to addition: $a(b + c) = ab + ac$.

If $a, b \in \mathbf{N}$, then $a - b \in \mathbf{N}$ if and only if $a - b \geq 0$. That is, $\mathbf{N}$ is not closed under subtraction ($1 - 2 = -1 \notin \mathbf{N}$). For this, we have the ring of integers

$$\mathbf{Z} = \{..., -2, -1, 0, 1, 2, ...\}.$$

Similarly, $\frac{a}{b} \in \mathbf{N}$ if and only if a is a multiple of b. For this, we have the field of fractions

$$\mathbf{Q} = \{\tfrac{a}{b} : a \in \mathbf{Z}, b \in \mathbf{N}\}.$$

We won't introduce here the algebraic terminology of rings and fields. Just think that a ring is a set closed under addition, multiplication, and subtraction, such that each element has an additive inverse $a + (-a) = 0$, while a field is a set closed under all four operations, and such that each element has an additive and a multiplicative inverse $a \cdot a^{-1} = 1$. Other examples of fields are $\mathbf{R}$ and $\mathbf{C}$. These are successively larger than $\mathbf{Q}$ as $\mathbf{R}$ is also closed under taking the square root of positive integers, and $\mathbf{C}$ is closed under taking all square roots (since $\sqrt{-1} \in \mathbf{C}$).

1.2. Divisibility

We write $a|b$ to say a divides b.

Example 1. $1, 2, 3, 4, 6, 12, 18, 36|36$, $5 \not| 36$.

[1] Here, we will work in the French tradition of considering 0 as a number.

Note that if $a, b \in \mathbf{N}$ and $a|b$, there exists $n \in \mathbf{N}$ such that $b = an$. If $a, b \in \mathbf{Z}$ and $a|b$, there exists $n \in \mathbf{Z}$ such that $b = an$.

Proposition 2. *Take note of the following basic properties of division.*

(1) *If $a|b$, $b \neq 0$, then $a \leq b$*
(2) *If $a|b$ and $b|c$ then $a|c$*
(3) *If $a|b$ and $a|c$, then for any $u, v \in \mathbf{Z}$, $a|ub + vc$.*

PROOF. (1) Since $b = an$ for $n \geq 1$, $b \geq a$.
(2) If $b = an$, $c = bm$, then $c = (an)m = a(nm)$, i.e. $a|c$.
(3) If $b = an$, $c = am$, then for any $u, v \in \mathbf{Z}$,

$$ub + vc = u(an) + v(am) = a(un + vm).$$

□

If $b \geq a$ but $a \not| b$, then the division of b by a leaves a remainder/rest r, which is smaller than a. This remainder is unique:

THEOREM 1 (Euclid's division algorithm). *Let $a, b \in \mathbf{N}$, with $b \geq a$. Then there exists exactly two numbers $q, r \in \mathbf{N}$ such that $b = aq + r$ and $0 \leq r < a$.*

PROOF. Let $r \in \mathbf{N}$ be the *smallest* number of the form $b - aq$. Suppose for contradiction that $r \geq a$, then

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

contradicts the minimality of r. □

1.3. Even and odd

Definition 3. *We say that n is even if $2|n$ and odd otherwise.*

In particular, any even number is of the form $2k$ (for some $k \in \mathbf{N}$) and any odd number is of the form $2k + 1$. E.g. $3 = 2 + 1$, $17 = 2 \cdot 8 + 1$, $14 = 2 \cdot 7$. Of course, k is then itself either even or odd... Here are some warm-up exercises.

Proposition 4. *Any odd number is either of the form $4k + 1$ or $4k - 1$.*

PROOF. Let $n = 2k + 1$. Then k is either odd, $k = 2j - 1$, or even $k = 2j$. Hence either $n = 4j - 1$ or $n = 4j + 1$. □

Proposition 5. *If $n \in \mathbf{N}$ is odd, then n^2 is odd.*

PROOF. Since n is odd, $n = 2k + 1$ for some $k \in \mathbf{N}$. Then

$$n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1.$$

□

Proposition 6. *The sum of the first n odd numbers is $= n^2$.*

PROOF. The statement is formalized by the equation

$$\sum_{k=0}^{n-1}(2k+1) = n^2$$

and easily proven by induction. First, it clearly holds for $n = 1$. Assume it is true for n. Then

$$\sum_{k=0}^{n}(2k+1) = \sum_{k=0}^{n-1}(2k+1) + 2n + 1 = n^2 + 2n + 1 = (n+1)^2.$$

□

1.4. Prime numbers

Any $n \in \mathbf{N}$ has at least two divisors: 1 and itself. We refer to those as the trivial divisors of n.

Definition 7. *If $n \geq 2$ is only divisible by 1 and n, it is* **prime**. *Otherwise we say that n is* **composite**.

In particular, if n is composite, there exist $1 < u, v < n$ such that $n = uv$.

THEOREM 2. *Any $n \geq 2$ is either prime or factorizes as a product of primes, i.e.*

$$n = p_1 \cdots p_k.$$

PROOF. Suppose that the statement is true up to $n-1$. Suppose further that n is composite. There exist $1 < u, v < n$ such that $n = uv$. By our induction hypothesis, both u and v factor into a product of primes: $u = p_1 \cdots p_j$, $v = q_1 \cdots q_l$. Hence $n = uv = p_1 \cdots q_l$ also factorizes as a product of primes. □

Example 8. $1000 = 2^3 \cdot 5^3$, $999 = 3 \cdot 37$.

The prime factorization of any number is unique (up to rearrangements); this is called the **fundamental theorem of arithmetic**. We will prove this soon.

Example 9. *Here are all the primes below 100:*

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97$$

How do we check that a given number is prime ? Is there a smarter way to proceed than just by brute force, i.e. checking that no smaller number is a divisor ? For the moment, we only record the following observation.

THEOREM 3. *If n has no prime divisors $\leq \sqrt{n}$ then n is prime.*

PROOF. Suppose for contradiction that n is composite. In particular, n has at least two distinct prime factors p, q and there is a $m \geq 1$ such that $n = pqm$. By assumption, $p, q > \sqrt{n}$, hence $n > nm \geq n$, which is absurd. □

Example 10. *Since $2, 3, 5 \not| 37$, it is prime.*

Another question is whether we can determine efficiently the prime factorization of a given large number. Here we present a trick of Fermat for factorizing using the 'difference of squares.' Suppose you are given $n \in \mathbf{N}$. Pick the smallest number a such that $a^2 > n$. At some point in the sequence

$$a^2 - n, (a+1)^2 - n, (a+2)^2 - n, \dots$$

there is a perfect square b^2, i.e. for some $k \in \mathbf{N}$, $(a+k)^2 - n = b^2$, or equivalently

$$n = (a + k - b)(a + k + b).$$

Let's now see a numerical example. Let $n = 10'001$. It is easy to see that $100^2 = 10'000 < n < 101^2$. Hence $a = 101$. Now we can quickly compute

$$101^2 - n = 200, 102^2 - n = 403, \dots, 105^2 - n = 1024 = 32^2$$

leading to the factorization $10'001 = (105 - 32)(105 + 32) = 73 \cdot 137$. These factors are small enough that we can very quickly check that 73 and 137 are both primes.

1.5. The set of primes

THEOREM 4 (Euclid). *There are infinitely many primes.*

PROOF. Suppose for contradiction that there are only finitely many primes: $2, 3, 5, ..., p$. Consider

$$N = 2 \cdot 3 \cdots p + 1.$$

By assumption, N is not prime since $N > p$. Since it is odd, it must have an odd prime factor $q|N$. Then q is an element of our finite list of primes $2, 3, \dots, p$. In particular, $q|(2 \cdot 3 \cdots p)$, hence

$$q|(N - 2 \cdot 3 \cdots p) = 1,$$

which is absurd. □

Now that we know that the set of primes is infinite, we may want to know how it is distributed in $\mathbf{N}$. This is not well understood, and relates to the famous (yet unproven) Riemann hypothesis. Let us explore a first question in this direction. Apart from 2, all prime numbers are odd. Now, we have seen that an odd number is either of the form $4k + 1$ or of the form $4k - 1$. Are there infinitely many primes of both forms ?

THEOREM 5. *There are infinitely many primes of the form* $4k - 1$.

PROOF. We mimic Euclid's argument: suppose that there are only finitely many primes of the form $4k - 1$: $3, 7, 11, ..., p$, and consider

$$N = 4(3 \cdot 7 \cdots p) - 1.$$

By assumption N is not prime and odd. We show that not all prime divisors of N can be of the form $4k + 1$. In fact, since

$$(4k + 1)(4l + 1) = 4(4kl + k + l) + 1,$$

if all prime divisors of N were of the form $4k + 1$, then N would also be of the form $4k + 1$. Hence N has a prime divisor q of the form $4k - 1$. Hence $q|(3 \cdot 7 \cdots p)$ also and $q|(3 \cdot 7 \cdots p - N) = 1$, which is absurd. □

Remark that the same argument can not be used to prove that there are infinitely many primes of the form $4k+1$ (why ?). This is nonetheless also true, as we will see later. Actually, a much stronger result is known; this is Dirichlet's theorem on arithmetic progressions.

THEOREM 6 (Dirichlet). *Any arithmetic progressions contains infinitely many primes. In other words, there are infinitely many primes of the form $ax+b$.*

The proof of Dirichlet's theorem is unfortunately outside of the scope of this course, but should be studied by anyone who wants to delve deeper in number theory !

1.6. Diophantine equations

If we know that every arithmetic progressions contains infinitely many primes, we can think of a more general question: given a polynomial $f(x) = a_n x^n + \cdots + a_1 x + a_0$ with integer coefficients $a_0, a_1, \ldots, a_n \in \mathbf{Z}$, are there infinitely many primes in its image $f(\mathbf{Z})$? Well, currently, as soon as $n > 1$, nothing is known ! A reasonable guess is that things would be easier if we have instead a polynomial with several variables, $f(x_1, x_2, \ldots, x_n)$...

A **Diophantine equation** is a polynomial equation with at least two unknowns. For example:

Example 11.
(1) *$ax+by=c$. In general, a polynomial equation where all terms have degree 1 is called a* **linear Diophantine equation**.
(2) *$x^2+y^2=z^2$. The integer solutions to this equation are called Pythagorean triples and have been completely described by Euclid.*
(3) *$x^n+y^n=z^n$, for $n \geq 3$.* **Fermat's last theorem** *(proven by Wiles in 1994) states that these have no integer solutions.*
(4) *$x^2-ny^2=1$, $(n \neq k^2)$, is called* **Pell's equation**.

Cracking a nut with a sledgehammer:

THEOREM 7. *$\sqrt[n]{2}$ is irrational for $n \geq 3$.*

PROOF. Suppose for contradiction that $\sqrt[n]{2} = \frac{a}{b}$. This is equivalent to $2b^n = b^n + b^n = a^n$, which contradicts Fermat's last theorem. □

Definition 12. *We say that $a, b \in \mathbf{N}$ are* **coprime** *if they have no common prime factor.*

Example 13. *$15 = 3 \cdot 5$ and $4 = 2 \cdot 2$ are coprime.*

THEOREM 8. *If $a, b \in \mathbf{N}$ are coprime, then the linear Diophantine equation*

$$ax + by = c$$

has integer solutions.

PROOF. Let $d \geq 1$ be the smallest positive number of the form $ax+by$:

$$d = ax_0 + by_0.$$

We will show that $d|a$ and $d|b$. Since a and b are coprime, we must have $d = 1$. Then $a(x_0c) + b(y_0c) = c$ is a solution. To show that $d|a$, we divide a by d: there are two numbers q, r such that $a = qd + r$ and $0 \leq r < d$. Then

$$r = a - qd = a - q(ax_0 + by_0) = a(1 - qx_0) + b(-qy_0)$$

is either 0 or contradicts the minimality of d. This proves that $d|a$. We can make the exact same argument to prove $d|b$. □

A moment of consideration should convince you that there are solvable equations $ax + by = c$ where a and b are not coprime; for example, $2x + 4y = 6$. You will also notice in this example that all coefficients have 2 as a common divisor.

1.7. gcd and lcm

Definition 14 (gcd)**.** *Let $a, b \in \mathbf{Z}$. The* **greatest common divisor** *of a and b is the positive integer d such that*

1. *$d|a$ and $d|b$ and*
2. *for any other common divisor c of a and b, $c \leq d$. The shorthand notation for the gcd is $d = (a, b)$.*

THEOREM 9. *The linear Diophantine equation $ax + by = c$ has integer solutions if and only if $(a, b)|c$.*

PROOF. Choose d as previously; let it be the smallest positive number of the form $ax + by$. Then we have seen that $d|a$ and $d|b$. Moreover, for any m such that $m|a, b$, then $m|ax + by = d$. Hence $d = (a, b)$.

If $d|c$, there is a k such that $c = dk$, and $a(xk) + b(yk) = dk = c$ is an example of solution. Conversely, if $ax + by = c$, we have that $d|c$ since $d|a$ and $d|b$. □

A **very** useful representation of the gcd for practical manipulations is given by the following corollary.

Corollary 15 (Bézout's Lemma)**.** *For any $a, b \in \mathbf{Z}$, there exist $u, v \in \mathbf{Z}$ such that $au + bv = (a, b)$.*

Exercise 16. *Use Bézout's lemma to show that if $c|a$ and $c|b$, then $c|(a, b)$.*

To compute the gcd of two numbers one can rely on their unique prime factorization or on the Euclidean division algorithm.

Example 17. $(36, 45) = (2^2 \cdot 3^2, 3^2 \cdot 5) = 3^2 = 9$, $(422, 7'491) = (2 \cdot 211, 3 \cdot 11 \cdot 227) = 1$.

THEOREM 10. *The gcd (a, b) is the last nonzero remainder when running the division algorithm.*

PROOF. We may assume that $a \geq b$. by the division algorithm, $a = bq + c$ with $0 \leq c < b$. If $d|a$ and $d|b$, then $d|r = a - bq$. Conversely if $d|r$ and $d|b$ then $d|a$. Hence $(a, b) = (b, c)$. Repeating this process a finite number of time, we have

$$(a, b) = (b, c) = \cdots = (l, m) = (m, n)$$

with $l = q'm + n$, $0 \leq n < m$, and $m = q''n$. Hence $(a, b) = (m, n) = n$. □

Example 18. $\underline{7491} = 17 \cdot \underline{422} + 317 \to \underline{422} = \underline{317} + 105 \to \underline{317} = 3 \cdot \underline{105} + 2 \to \underline{105} = 52 \cdot \underline{2} + \mathbf{1}$.

Definition 19. *The* **least common multiplier** $\{a, b\}$ *is the number such that*

(1) $a|\{a, b\}$ *and* $b|\{a, b\}$*, and*
(2) *for any other common multiple* $a|c$*,* $b|c$*,* $\{a, b\} \leq c$*.*

THEOREM 11. $ab = \{a, b\}(a, b)$.

PROOF. Set $x = \frac{ab}{(a,b)}$. Note that $a|a\frac{b}{(a,b)}$ since $(a, b)|b$, and similarly $b|\frac{a}{(a,b)}b$; so, $a, b|x$. Let c such that $a|c$ and $b|c$, and represent (a, b) as $(a, b) = au + bv$. Then

$$\frac{c}{x} = c\frac{(a, b)}{ab} = c\frac{au + bv}{ab} = \frac{c}{b}u + \frac{c}{a}v$$

and since $a, b|c$, this is an integer, call it n. We thus conclude that $c = nx$, i.e. $x|c$. This goes to show that $x = \{a, b\}$. □

Exercise 20. *Show that if* $a|c$ *and* $b|c$*, then* $\{a, b\}|c$*.*

1.8. The fundamental theorem of arithmetic

THEOREM 12 (Euclid's lemma). *If* a, b *are coprime and* $a|bc$*, then* $a|c$*.*

PROOF. By Bézout's lemma, there exist $u, v \in \mathbf{Z}$ such that $au+bv = 1$. Multiplying both sides by c yields

$$auc + bcv = auc + anv = a(uc + nv) = c.$$

□

Euclid's lemma is actually equivalent to the fundamental theorem of arithmetic. In fact, we will prove the latter using Euclid's lemma in the next section, but we can also prove Euclid's lemma with the fundamental theorem of arithmetic:

PROOF OF EUCLID'S LEMMA USING THE FUND. THM. OF ARITHMETIC. Consider the prime factorizations $b = p_1 \cdots p_k$, $c = q_1 \cdots q_l$. Since a and b have no common prime factors, none of $p_1, \ldots, p_k$ appear in the prime factorization of a. Hence $a|c$. □

Two neat consequences of Euclid's lemma.

THEOREM 13. *Let* (x_0, y_0) *be a solution of the linear Diophantine equation* $ax+by = 1$ *with* a, b *coprime. Then the set of all solutions to* $ax + by = 1$ *is*

$$\{(x, y) : x = x_0 - bn,\ y = y_0 + an,\ n \in \mathbf{Z}\}.$$

PROOF. If (x_0, y_0) is a solution, then

$$a(x_0 - bn) + b(y_0 + an) = ax_0 + by_0 = 1$$

for each $n \in \mathbf{Z}$. Conversely, suppose that $ax + by = 1$. Then subtracting this equation to $ax_0 + by_0 = 1$, we have

$$a(x_0 - x) + b(y_0 - y) = 1 - 1 = 0,$$

hence $a|b(y-y_0)$ and by Euclid's lemma, $a|(y-y_0)$. Say $y-y_0=an$. Then

$$a(x_0-x)-ban=0$$

and $x=x_0-bn$. □

THEOREM 14. *A number $n\geq 2$ is prime if and only if $n|ab$ implies that $n|a$ or $n|b$.*

PROOF. Suppose that $n=p$ is prime. Then the claim follows directly by Euclid's lemma. Conversely, assume for contradiction that n is composite, i.e. $n=uv$ for $1<u,v<n$. Then by assumption, either $n|u$ or $n|v$. But since $u,v<n$, this is absurd. □

THEOREM 15 (Fund. theorem of arithmetic). *Each $n\geq 2$ is either a prime or can be uniquely (up to rearrangement of the factors) factorized as a product of primes.*

PROOF. Suppose that $n\geq 2$ has two prime factorizations: $n=p_1\cdots p_k$ and $n=q_1\cdots q_l$. Then $p_1|q_q\cdots q_l$. By Euclid's lemma, $p_1|q_i$ for some i. In fact, up to rearranging the terms, we can suppose that $p_1|q_1$. Since q_1 is prime, we conclude that $p_1=q_1$ and our original equality reduces to $p_2\cdots p_k=q_1\cdots q_l$. Repeating this argument a finite number of times, we can conclude that $k=l$ and primes on both sides agree. □

Observe that the argument above motivates the convention of excluding 1 from the list of prime numbers. We note the following application of the fundamental theorem of arithmetic, due to Euclid.

THEOREM 16. *Let p be prime, then $\sqrt{p}$ is irrational.*

PROOF. Suppose for contradiction that $\sqrt{p}=\frac{a}{b}$. Then $pb^2=a^2$. Consider the prime decomposition $b=p_1\cdots p_k$. Then $b^2=p_1^2\cdots p_k^2$ has twice as many prime factors, and in particular, has an *even* number of prime factors. Hence if we look at the equation $pb^2=a^2$, the left hand side has an odd number of prime factors, while the right hand side has an even number of prime factors, which is absurd. □

Exercise 21. *Generalize the proof above to show that if n is not a perfect square, then $\sqrt{n}$ is irrational.*

CHAPTER 2

Congruences

From the fundamental theorem of arithmetic, we see that the most natural (and interesting) properties of numbers are multiplicative. In this way, we understand small numbers very well, yet $\mathbf{N}$ is infinite. In fact, given very large numbers a and b

(1) How to check that $n|m$?
(2) How to compute the prime factorization of m ?
(3) How to check whether n is prime ?

The latter two questions are difficult computationally. The first one is more accessible for the reason that the Euclidean division algorithm is very fast and efficient. Of course $a|b$ if and only if the rest after division is $r = 0$. Recall that the rest must be an element of the finite set $\{0, 1, \ldots, n-1\}$. In this chapter, by studying systematically the finite set of remainders, or **residues** of division by n, we will get some more insight on these three essentially questions about the structures of numbers.

Fix $n \geq 1$. We know that each $a \in \mathbf{Z}$ can be written (in a unique way) as $a = qn+r$, where $q \in \mathbf{Z}$, and $r \in \{0, \ldots, n-1\}$. In other words, when divided by n, any number has its remainder (or **residue**) in the (finite) set $\{0, 1, \ldots, n-1\}$. We call two numbers a, b that have the same residue **congruent**. We call n the **modulus**.

For illustration, order all integers as follows

				$\ldots$	$-n-1$
$-n$	$-n+1$	$-n+2$	$\ldots$	-2	-2
0	1	2	$\ldots$	$n-2$	$n-1$
n	$n+1$	$n+2$	$\ldots$	$2n-2$	$2n-1$
$2n$	$2n+1$	$2n+2$	$\ldots$	$3n-2$	$3n-1$
$3n$	$\ldots$				

then all members of a column are congruent.

2.1. Gauss' notation for congruences

Definition 22. *Let $n \geq 1$. We say that a is* **congruent to** b **modulo** n*, written $a \equiv b$ (mod n), or even shorter, $a \equiv b$ (n), if a and b differ by a multiple of n.*

In other words, $a \equiv b$ (n) if $n|(b-a)$. Notice that $n|(b-a)$ if and only if $n|(a-b)$: the equation $b - a = kn$ is equivalent to $a - b = (-k)n$.

Example 23. $63 \equiv 0$ (3), $64 \equiv 1$ (3), $7 \equiv -1$ (8), $5^2 \equiv -1$ (13).

Exercise 24. *Show that if $a, b \in \mathbf{N}$ have the same parity (i.e. are either both odd, or both even), then $a \equiv b$ (2).*

THEOREM 17. *Being congruent to modulus n is an* **equivalence relation**.

PROOF. Recall that $\sim$ is an equivalence relation on a set S if it is reflexive ($a \sim a$ for all $a \in S$), symmetric ($a \sim b$ if and only if $b \sim a$ for all $a, b \in S$), and transitive (if $a \sim b$ and $b \sim c$ then $a \sim c$). We only check that being congruent is transitive, and leave the rest to the reader. Suppose $a \equiv b\ (n)$ and $b \equiv c\ (n)$. By definition, there exist $k, l \in \mathbf{Z}$ such that $kn = (a-b)$ and $ln = (b-c)$. Hence $a-c = a-b+b-c = kn+ln = (k+l)n$. Thus $n|a-c$. □

As a result, congruences to same modulus behave (*mostly*) like equations.

THEOREM 18. *Fix $n \geq 1$. Suppose that $a \equiv b\ (n)$ and $c \equiv d\ (n)$. Then*

(1) *$a + c \equiv b + d\ (n)$ and $a - c \equiv b - d\ (n)$,*
(2) *$am \equiv bm\ (n)$ for all $m \in \mathbf{Z}$,*
(3) *$ac \equiv bd\ (n)$.*

PROOF. By assumption $nk = a - b$ and $nl = c - d$, for some $k, l \in \mathbf{Z}$. Then $a + c - (b + d) = (a - b) + (c - d) = (k + l)n$, which proves the first assertion in (1). The second assertion is proven similarly. For (2), $n|(b - a)$ and hence $n|m(b - a)$ for all $m \in \mathbf{Z}$. For (3), we note that by by (2), $ac \equiv bc\ (n)$ and also $bc \equiv bd\ (n)$. Hence by transitivity, $ac \equiv bd\ (n)$. □

2.2. Applications

Gauss' congruence notation is useful in devising **obstructions** on a number being of a certain form. For example,

Proposition 25. *If $n \equiv 2\ (4)$ or $n \equiv 3\ (4)$ then n is not a perfect square.*

PROOF. Let $n = a^2$. If $a = 2k$, then $n = 4k^2 \equiv 0\ (4)$. Id $a = 2k + 1$, then $n = 4(k^2 + k) + 1 \equiv 1\ (4)$. □

Proposition 26. *If $n \geq 3$ is odd, then $n^2 + 1$ is not a prime.*

PROOF. n is odd, hence $n \equiv 1\ (2)$, hence $n^2 \equiv 1\ (2)$ and $n^2 + 1 \equiv 0\ (2)$. □

Another application: **divisibility tests.** For example,

Proposition 27. *Let $n \in \mathbf{N}$. If the sum of the digits of n is divisible by 3, then n is divisible by 3.*

PROOF. Consider the decimal expansion of n; $n = a_0 + a_1 \cdot 10 + a_2 \cdot 100 + \ldots a_k \cdot 10^k$, or,

$$\begin{aligned} n &= a_0 + a_1 \cdot (3^2 + 1) + a_2(3^4 + 1) + \ldots a_k \cdot (3^{2k} + 1) \\ &\equiv a_0 + a_1 + a_2 + \cdots + a_k\ (3). \end{aligned}$$

If the sum of the digits $a_0 + a_1 + \ldots a_k \equiv 0\ (3)$, then $n \equiv 0\ (3)$. □

Remark 28. *The same argument works with 9 instead of 3.*

Proposition 29. *If the last two digits of n are divisible by 4, then n is divisible by 4.*

PROOF. Consider the decimal expansion of n;

$$\begin{aligned} n &= a_0 + a_1 \cdot 10 + a_2 \cdot 10^2 + \dots a_k \cdot 10^k \\ &= a_0 + a_1 \cdot (2 \cdot 5) + a_2 \cdot (2^2 \cdot 5^2) + \dots a_k \cdot (2^k \cdot 5^k) \\ &= a_0 + a_1 \cdot 10 + a_2 \cdot (4 \cdot 5^2) + \dots a_k \cdot (4 \cdot 2^{k-2} \cdot 5^k) \\ &\equiv a_0 + a_1 \cdot 10 \mod 4. \end{aligned}$$

□

Exercise 30. *Show that if the alternating sum of the digits of n is a multiple of 11, then n is a multiple of 11.*

2.3. Arithmetic properties

One can generalize Theorem 18 to show that given $a_i \equiv b_i\ (n)$, $i = 1, \dots, k$, then

$$a_1 + \dots + a_k \equiv c_1 + \dots + c_k\ (n), \quad a_1 \cdots a_k \equiv c_1 \cdots c_k\ (n).$$

However the **cancellation law** ($ab = ac \implies b = c$) does not always hold in congruence equations.

Example 31. $12 \equiv 24\ (10)$ *and* $2 \cdot 6 = 12$, $7 \cdot 6 = 42$, *but* $2 \not\equiv 7\ (10)$.

Proposition 32. *If $ab \equiv ac\ (n)$ and $(a, n) = 1$ then $b \equiv c\ (n)$.*

PROOF. We will see that this is really only a reformulation of Euclid's lemma. Indeed, $ab \equiv ac\ (n) \iff n | a(b - c)$ and since $(a, n) = 1$, $n | (b - c)$ and this is equivalent to $b \equiv c\ (n)$. □

And in general,

Proposition 33. *If $ab \equiv ac\ (n)$ then $b \equiv c\ (\frac{n}{(a,n)})$.*

PROOF. Let $d := (a, n)$ and $n = kd$, $a = ld$. Then $(l, k) = 1$. (Why is this true ?) With these new variables, $ab \equiv ac\ (n)$ becomes $ldb \equiv ldc\ (kd)$. Observe that this is equivalent to $lb \equiv lc\ (k)$. Since $(k, l) = 1$, the statement follows from Proposition 32. □

We will denote by $\mathbf{Z}_n$ the set of distinct congruences $x \equiv 0, 1, \dots, n - 1\ (n)$. (This notation, as the discussion about to follows, will be familiar to anyone who took abstract algebra.) We now have seen that $\mathbf{Z}_n$ is closed under addition and multiplication. It is easy to see that $a + x \equiv 0\ (n)$ has a solution – take $x \equiv -a\ (n)$) – and that all solutions are congruent to it mod n. What about $ax \equiv 1\ (n)$? If x is a solution, then it is necessarily congruent to one of $1, 2, \dots, n - 1$. So we only have finitely many options to check.

Example 34. *Consider the following two examples.*

1. $3x \equiv 1\ (5)$. *Then* $3 \cdot 2 = 6 \equiv 1\ (5)$ *is a solution.*
2. $3x \equiv 1\ (6)$. *Here there are no solutions. Indeed,* $3 \not\equiv 1\ (6)$, $3 \cdot 2 \equiv 0\ (6)$, $3 \cdot 3 \equiv 3\ (6)$, $3 \cdot 4 = 12 \equiv 0\ (6)$, $3 \cdot 5 \equiv 3\ (6)$.

Proposition 35. *$ax \equiv 1\ (n)$ has a (unique) solution if and only if $(a, n) = 1$.*

PROOF. Suppose that x is a solution to $ax \equiv 1\ (n)$, and suppose for contradiction that $d := (a, n) > 1$. Then since $d|n$ and $n|(ax - 1)$, $ax \equiv 1\ (d)$. Since $d|a$ also holds, $ax \equiv 0\ (d)$, and we are left with $0 \equiv 1\ (d)$, which is absurd.

Conversely, suppose that $(a, n) = 1$. By Corollary 15 (Bézout' Lemma), there exist $u, v \in \mathbf{Z}$ such that $au + nv = 1$. Then $au \equiv 1\ (n)$ and u is a solution to $ax \equiv 1\ (n)$. Suppose u' were another solution. That is, $au' \equiv 1\ (n)$. Then $au \equiv au'\ (n)$. Since $(a, n) = 1$, we may apply cancellation, and obtain that $u \equiv u'\ (n)$. This proves that the solution is unique. □

Exercise 36. *Show that more generally, $ax \equiv b\ (n)$ has a solution if and only if $(a, n)|b$.*

The only $a \in \mathbf{Z}$ having a multiplicative inverse, i.e. for which $ax = 1$ has a solution, are $a = \pm 1$. And in particular, these are self-reciprocal (their own inverses). Proposition 35 above tells us that in $\mathbf{Z}_n$ there are

$$\varphi(n) = \#\{1 \leq a < n : (a, n) = 1\}$$

elements that admit a multiplicative inverse. (The function $\varphi(n)$ is called Euler's totient function, and we will meet it again...)

Example 37. *Check that $\varphi(4) = 2$, $\varphi(17) = 16$, and that in general $\varphi(p) = p - 1$.*

Question: How many elements in $\mathbf{Z}_n$ are self-reciprocal ? For the moment, we will answer this for $n = p$ prime. That is, how many $a \in \mathbf{Z}_p$ satisfy $a^2 \equiv 1\ (p)$? Observe

$$a^2 \equiv 1\ (p) \iff a^2 - 1 = (a - 1)(a + 1) \equiv 0\ (p) \implies a \equiv 1\ (p) \text{ and } a \equiv -1\ (p),$$

hence there are only two self-reciprocal elements mod p: 1 and $p - 1$. Using this, we can present Gauss' proof of **Wilson's theorem**:

THEOREM 19 (Wilson's theorem). *p is a prime if and only if $(p - 1)! \equiv -1\ (p)$.*

PROOF. If $p = 2$, this is clear. Suppose that $p > 2$. Suppose that p is prime. By the discussion above, each of $1, 2, \ldots, p-1$ has a (unique) multiplicative inverse among $1, ..., p - 1$. Moreover, in this set, only 1 and $p - 1$ are self-reciprocal. Consider the product $2 \cdot 3 \cdots (p - 2)$. This is a product of $p - 2 - 1 = p - 3$, i,e, an even number of factors. If we pair each factor with its reciprocal, we obtain that

$$(p - 2)! = 2 \cdot 3 \cdots (p - 2) \equiv 1\ (p).$$

And so multiplying both sides by $p - 1$: $(p - 1)! \equiv p - 1 \equiv -1\ (p)$.

Conversely, assume for contradiction that p has a non-trivial divisor, $d|p$. Then $(p-1)! \equiv -1\ (d)$, but since $d < p$, $(p-1)! \equiv 0\ (d)$. We are left with $0 \equiv -1\ (d)$ which is absurd for $d \geq 2$. □

We have a new **primality test** ! ...but considering it requires computing a factorial, this is rather only of theoretical value. In the next section, we see another similar primality test – via Fermat's little theorem – that as practical significance.

2.4. Fermat and Euler theorems

THEOREM 20 (Fermat's little theorem (1640)). *Let p be a prime, and $(a,p)=1$. Then*

$$a^{p-1} \equiv 1 \ (p).$$

PROOF. We present a proof of Ivory (1806): $a, 2a, 3a, \ldots, (p-1)a$ are congruent (in some order) to $1, 2, \ldots, p-1$. In fact this is a one-to-one correspondence: if $ja \equiv ka \ (p)$ then since $(a,p)=1$, $j \equiv k \ (p)$, but $1 \leq j,k \leq p-1$, hence $j=k$.) Then

$$a(2a)(3a)\cdots((p-1)a) \equiv (p-1)! \ (p)$$

and

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

Thus $(p-1)!a^{p-1} \equiv (p-1)! \ (p)$ and since $(p,(p-1)!)$, we may divide $(p-1)!$ from both sides. □

Remark 38. *For applications (as we will see an example below), it is sometimes useful to keep in mind that $a^{p-1} \equiv 1 \ (p) \iff a^p \equiv a \ (p)$.*

We next discuss two applications of Fermat's little theorem.

(1) Primality testing. We can rule out that a large number is prime using Fermat's Little Theorem (how?). Conversely, a number n for which $(a,n)=1$ and $a^{n-1} \equiv 1 \ (n)$ has a good chance to be prime, but this is not sufficient. An example is $561 = 3 \cdot 11 \cdot 17$. We test with Fermat using repeated squaring to reduce the power to compute. First note $560 = 2 \cdot 280 = 2 \cdot 2 \cdot 140 = 2^3 \cdot 70 = 2^4 \cdot 35$. Good enough, 2^{35} is computable. In fact,

$$2^{35} \equiv 263 \ (561).$$

Squaring both sides,

$$2^{70} \equiv 166 \ (561),$$

and repeating the process enough many times: $2^{140} \equiv 67 \ (561)$, $2^{280} \equiv 1 \ (561)$, and from here it follows that

$$2^{560} \equiv 1 \ (561).$$

Hence 561 is a 'false positive'. We call such numbers **pseudo-primes** (or Carmichael numbers) – 561 is the smallest pseudo-prime.

(2) Solving linear congruence equations. Earlier, we have seen that the equation $ax \equiv b \ (p)$, for p prime, has a solution. Fermat's Little Theorem gives us a simple way of finding this solution: multiply both sides of the congruence equation by a^{p-2}, then using that $a^{p-1} \equiv 1$,

$$ax \equiv b \ (p) \implies x \equiv a^{p-2}b \ (p).$$

Example 39. *Let's solve $5x \equiv 2 \mod 17$. Applying Fermat, $x \equiv 5^{15}2 \mod 17$ is a solution. We now want to find the remainder of $5^{15}2$ when divided by 17:*

$$5^{15} = 25^7 \cdot 5 \equiv 8^7 \cdot 5 = 64^3 \cdot 40 \equiv (-4)^3 \cdot 6 = -16 \cdot 24 \equiv 1 \cdot 7 = 7 \mod 17$$

hence $x \equiv 5^{15} \cdot 2 \equiv 14 \mod 17$.

From this, can we infer the case of general modulus n ? This is Euler's generalization of Fermat's theorem:

THEOREM 21 (Euler's theorem, 1760). *If $(x, n) = 1$, then*

$$x^{\varphi(n)} \equiv 1 \ (n),$$

where $\varphi(n) = \#\{0 \leq m < n : (m, n) = 1\}$ is Euler's totient function.

Remark 40. *We have already met Euler's totient function once. For now, let's simply observe that if p is prime, $\varphi(p) = p - 1$. So this indeed generalizes Fermat's theorem.*

PROOF. We apply the same argument we applied to prove Fermat's Little Theorem. We first list all $1 \leq a < n$ satisfying $(a, n) = 1$ as follows:

$$a_1, a_2, \ldots, a_k$$

where $k = \varphi(n)$, since this is by definition the number of positive integers below n that are coprime to it. We now consider

$$a_1 \cdot a, a_2 \cdot a, \ldots, a_k \cdot a.$$

These are congruent (in some order) to $a_1, a_2, \ldots, a_k$ such that

$$(a_1 \cdots a_k)a^k = (a_1 \cdot a)(a_2 \cdot a) \cdots (a_k \cdot a) \equiv a_1 \cdots a_k \mod n.$$

To cancel $(a_1 \cdots a_k)$ on both sides, we only need to show that $(a_1 \cdots a_k, n) = 1$ and this follows by repeated applications of Euclid's lemma. □

The latter point in the proof is sufficiently important and general to stand as a Lemma on its own.

Lemma 41. *Let $a_1, \ldots, a_k \in \mathbf{Z}$ such that $(a_i, n) = 1$ for each $i = 1, \ldots, k$. Then $(a_1 \cdots a_k, n) = 1$.*

PROOF. Exercise. □

2.5. Chinese remainder theorem

As we have just seen with the Fermat–Euler theorem, solving an equation of the form $ax \equiv b \pmod n$ reduces to finding the smallest residue x such that $x \equiv a^{\varphi(n)-1}b \pmod n$, and an efficient way of doing division by hand is using the method of repeated squaring. As is clear from Example 112, the smaller the modulus is, the quickest the operation. Let's see a numerical example with a larger modulus: What is the remainder of $(102^{73} + 55)^{37}$ divided by 111 ? In other words, we want the smallest positive residue r such that $r \equiv (102^{73} + 55)^{37} \pmod{111}$. An efficient way of treating such a problem is to break it into several congruence equations of smaller *prime* modulus. Here, $111 = 3 \cdot 37$. Observe that if $r \equiv (102^{73} + 55)^{37} \pmod{111}$, then

$$r \equiv (102^{73} + 55)^{37} \pmod 3$$
$$r \equiv (102^{73} + 55)^{37} \pmod{37}$$

and vice versa. Let us formalize this observation:

Proposition 42. *Fix $n \geq 2$ and consider its prime factorization $n = p_1^{k_1} \cdots p_l^{k_l}$, where $p_1, \ldots, p_l$ are distinct primes. Then X is a solution of $ax \equiv b \ (mod\ n)$ if and only if X is a solution of*

$$ax \equiv b \ (mod\ p_1^{k_1})$$
$$\vdots$$
$$ax \equiv b \ (mod\ p_l^{k_l})$$

PROOF. Exercise.. □

Coming back to our numerical example, we note that

$$102 \equiv 0 \ (3), \quad 55 \equiv 1 \ (3)$$

hence $(102^{73} + 55)^{37} \equiv 1 \ (3)$ and by Fermat's Little Theorem,

$$102^{73} = (102^{36})^2 \cdot 102 \equiv 102 \equiv -9 \ (37)$$

hence $102^{73} + 55 \equiv 9 \ (37)$ and applying Fermat's Little Theorem once more,

$$(102^{73} + 55)^{37} \equiv 9 \ (37)$$

In summary, we have the system of equation

$$r \equiv 1 \ (\text{mod } 3)$$
$$r \equiv 9 \ (\text{mod } 37)$$

THEOREM 22 (Chinese remainder theorem). *Let $n_1, n_2, \ldots, n_k$ be pairwise coprime. Then the system of congruence equations*

$$x \equiv a_1 \ (n_1)$$
$$x \equiv a_2 \ (n_2)$$
$$\vdots$$
$$x \equiv a_k \ (n_k)$$

admits a unique solution $X \ (mod\ n_1 \cdot n_2 \cdots n_k)$.

PROOF. The proof is by induction over k. Suppose $k = 2$. Then we have a system of two congruence equations

$$x \equiv a \ (m)$$
$$x \equiv b \ (n)$$

Since $(m, n) = 1$, there exist $u, v \in \mathbf{Z}$ such that $mu + nv = 1$ (Bézout's Corollary 15). Note that $mu \equiv 1 \ (\text{mod } n)$ and $nv \equiv 1 \ (\text{mod } m)$, hence $x \equiv anv \ (\text{mod } m)$ and $x \equiv bmu \ (\text{mod } n)$. Thus, $X = anv + bmu$ is a solution. Suppose there were another solution X'. Then $X \equiv X' \ (\text{mod } m)$ and $X \equiv X' \ (\text{mod } n)$. Hence there exists $j \in \mathbf{Z}$ such that $X - X' = jn$ and $m|jn$. Then since $(m, n) = 1$, by Euclid's Lemma, we conclude that $m|j$, and hence $X \equiv X' \ (\text{mod } mn)$.

Suppose now the claim holds for systems of $< k$ equations, and consider the system above with k equations. By assumption, the first $k - 1$ congruence equations have a

simultaneous solution $a \pmod{n_1 \cdots n_{k-1}}$. Set $m := n_1 \cdots n_{k-1}$, $n := n_k$, $b := a_k$. Our system therefore reduces to

$$x \equiv a \ (m)$$
$$x \equiv b \ (n)$$

If we can show that $(m, n) = 1$, then we're done. Hence, suppose that p is a common prime factor to m and n. Since $p|m = n_1 \cdots n_{k-1}$, by Euclid's lemma, $p|n_i$ for some i. Hence p is a common prime factor to n_i and $n = n_k$. This is absurd since $(n_i, n_k) = 1$. Hence $(m, n) = 1$. □

Let's conclude with a numerical application: what is the remainder of $(102^{73}+55)^{37}$ divided by 111 ? That is, we want to find $x \equiv (102^{73} + 55)^{37}$ (111). First note the prime factorization $111 = 3 \cdot 37$. Applying Fermat,

$$(102^{73} + 55)^{37} \equiv 102^{73} + 55 \equiv 102^{73} + 18 \ (37)$$

where

$$102^{73} = 102^{2 \cdot 37 - 1} \equiv 102^{2-1} \equiv 28 \ (37),$$

hence

$$(102^{73} + 55)^{37} \equiv 28 + 18 \equiv 9 \ (37).$$

The same process mod 3 yields

$$(102^{73} + 55)^{37} \equiv 2 \ (3).$$

Hence the remainder is the solution to the system of equation

$$x \equiv 9 \ (37)$$
$$x \equiv 2 \ (3).$$

Exercise 43. *Show using the Chinese remainder theorem that* $x \equiv 7$ (111).

2.6. Euler's totient function

Recall Euler's totient function

$$\varphi(n) = \#\{1 \leq a < n : (a, n) = 1\}.$$

This is our first example of a **multiplicative function**.

THEOREM 23. *If* $(m, n) = 1$, *then* $\varphi(mn) = \varphi(m)\varphi(n)$.

PROOF. Consider the sets

$$A = \{1 \leq a < mn : (a, mn) = 1\},$$
$$B = \{1 \leq a < m : (a, m) = 1\}, \quad C = \{1 \leq a < n : (a, n) = 1\}.$$

Note that $|\varphi(mn) = |A|$, $\varphi(m) = |B|$, $\varphi(n) = |C|$. Define

$$F : A \to B \times C$$

by

$$F(a) = (a \pmod m, a \pmod n).$$

Our claim is that this function is a *bijection* (or, one-to-one correspondence). But first we need to make sure that this map is well defined: if $1 \leq a < mn$ satisfies $(a, mn) = 1$, then $a' = a \pmod{m}$ satisfies $(a', m) = 1$, and $b = a \pmod{n}$ satisfies $(b, n) = 1$. Let us check the latter claim. First, suppose that $a'' \equiv 0 \pmod{n}$ then $n|a$ and this contradicts $(a, mn) = 1$. Second, let d be a common divisor to b and n, then since $n|(b - a)$, d is also a common divisor to a and mn. Since $(a, mn) = 1$, the only possible common divisor is $d = 1$, hence $(b, n) = 1$.

By the Chinese Remainder Theorem, for every pair $(a \pmod{m}, b \pmod{n}) \in B \times C$ with $(a, m) = (b, n) = 1$ there exists a unique $1 \leq x < mn$ such that

$$x \equiv a \pmod{m}$$
$$x \equiv b \pmod{n}.$$

Since $(a, m) = (a, n) = 1$, it follows that $(x, m) = (x, n) = 1$, and hence $(x, mn) = 1$. This shows that F is a one-to-one correspondence, and therefore its domain and codomain must have the same cardinality: $\varphi(mn) = \varphi(m)\varphi(n)$. □

Hence, to compute $\varphi(n)$, we need

(1) to know the prime factorization $n = p_1^{k_1} \cdots p_l^{k_l}$ of n, and
(2) to know the value of $\varphi(p^k)$ for $k \in \mathbf{N}$.

Proposition 44. *Let p be a prime. For $k \in \mathbf{N}$, $\varphi(p^k) = p^k - p^{k-1}$.*

Proof. We count $1 \leq a \leq p^k$ such that $(a, p^k) \neq 1$: these are $a = m \cdot p$, for $m = 1, \ldots, p^{k-1}$, hence $\varphi(p^k) = p^k - p^{k-1}$. □

Combining these results,

$$\varphi(n) = \varphi(p_1^{k_1} \cdots p_l^{k_l}) = p_1^{k_1}(1 - p_1^{-1}) \cdots p_l^{k_l}(1 - p_l^{-1}) = n \prod_{p|n} (1 - p^{-1}),$$

where the product is taken over all prime divisors of n. This is **Euler's formula**. Finally, we take note of the beautiful identity noted by Gauss: for every n,

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

2.7. Higher order congruences

In this section

$$f(x) = a_k x^k + a_{k-1} x^{k-1} + \cdots + a_2 x^2 + a_1 x + a_0$$

is a polynomial of degree k (in particular, $a_k \neq 0$) and with integer coefficients $a_0, a_1, \ldots, a_k \in \mathbf{Z}$. We say call $f(x) \equiv 0 \pmod{n}$ a congruence equation of order k. So far, we have studied congruence equations of order 1, also called linear congruence equations. Just as was the case then, solving higher order congruence equations boils down to solving system of equations to prime moduli; this is the cumulative content of the next two propositions.

Proposition 45. *Fix $n \geq 2$ and consider its prime factorization $n = p_1^{k_1} \cdots p_l^{k_l}$, where $p_1, \ldots, p_l$ are distinct primes. Then X is a solution to $f(x) \equiv 0 \pmod{n}$ if and only if X is a solution to*

$$f(x) \equiv 0 \ (mod\ p_1^{k_1})$$
$$\vdots$$
$$f(x) \equiv 0 \ (mod\ p_l^{k_l})$$

PROOF. Adapt the proof of Proposition 42. □

Proposition 46. *Let $k \geq 1$, and let p be a prime. Then,*

1. *if X is a solution to $f(x) \equiv 0 \pmod{p^k}$*
2. *if Y is a solution to $f'(X)x + \frac{f(X)}{p^k} \equiv 0 \pmod{p}$,*

then $X + p^k Y$ is a solution to $f(x) \equiv 0 \pmod{p^{k+1}}$.

PROOF. By Taylor expansion,

$$f(X + p^k Y) = f(X) + p^k Y f'(X) + m \cdot p^{k+1}$$

for some integer m. By assumption $f'(X) \cdot Y = -\frac{f(X)}{p^k} + m' \cdot p$ for some integer m'. Plugging this in the formula above:

$$f(X + p^k Y) = (m' + m)p^{k+1} \equiv 0 \pmod{p^{k+1}}.$$

□

We will study in depth quadratic congruence equations next chapter.

Exercise 47. *Prove that if $(4a, n) = 1$, solving $ax^2 + bx + c \equiv 0 \pmod{n}$ can be reduced to solving a congruence equation of the form $y^2 \equiv q \pmod{n}$. (Hint: multiply $ax^2 + bx + c$ by $4a$ and complete the square.)*

CHAPTER 3

Quadratic reciprocity

In this chapter, p will always denote an odd prime. The set of non-zero residues mod p is $\mathbf{Z}_p^* = \{1, 2, \ldots, p-1\}$. This set has the algebraic structure of a **group**.

Definition 48. *A set G equipped with an operation $\circ : G \times G \to G$, $(g, h) \mapsto g \circ h$ is called a group if*

(1) *(G is closed under $\circ$) for every $g, h \in G$, $g \circ h \in G$,*
(2) *(existence of identity) there exists (a unique) $e \in G$ such that $g \circ e = e \circ g = g$,*
(3) *(existence of inverse) for each $g \in G$, there exists (a unique) $\overline{g} \in G$ such that $g\overline{g} = \overline{g}g = e$.*

Proposition 49. *$\mathbf{Z}_p^*$ is a group with respect to multiplication mod p:*

$$(k, l) \in \mathbf{Z}_p^* \times \mathbf{Z}_p^* \mapsto kl \ (mod\ p) \in \mathbf{Z}_p^*.$$

Proof. We check that (1)-(3) hold. First we show (1); that $kl \pmod p \in \{1, \ldots, p-1\}$. The remainder of kl divided by p is in $\{0, \ldots, p-1\}$. If $kl \equiv 0\ (p)$ then either $k \equiv 0\ (p)$ or $l \equiv 0\ p$, which are impossible since $k, l \in \mathbf{Z}_p^*$. We now show (2); $k \cdot 1 = 1 \cdot k = k$ for each $k \in \mathbf{Z}_p^*$, and we claim that this holds only with 1. Indeed, using cancellation, $ak \equiv k\ (p)$ reduces to $a \equiv 1\ (p)$. Finally, consider the equation $kx \equiv 1\ (p)$. This equation has a unique solution (cf. Proposition 35). This proves (3). □

3.1. Quadratic residues

Let $p \geq 3$.

Definition 50. *An integer a is called a quadratic residue mod p* **(QR)** *if*

$$x^2 \equiv a\ (mod\ p)$$

has a solution, and it is called a quadratic nonresidue **(QNR)** *otherwise.*

Example 51. *What is the set of QR mod 5 ? First we note that the answers depends only on a (mod p). Then*

$$1^2 \equiv 1\ (5), \quad 2^2 \equiv 4\ (5), \quad 3^2 \equiv 4\ (5), \quad 4^2 \equiv 1\ (5)$$

that is, 1 and 4 are QR mod 5, 2 and 3 are QNR mod 5.

In this chapter, we will seek to answer the following questions.

(1) Given p, determine the set of QR mod p
(2) Given a and p, determine quickly whether a is a QR mod p.
(3) Given a, determine the set of odd primes p such that a is a QR mod p

For a number theorist, task (3) is particularly alluring because it would provide new information about primes. Is it at all feasible? Clearly $a = 1$ is a QR for all p.

THEOREM 24. *$a = -1$ is a QR (mod p) if and only if $p \equiv 1$ (mod 4).*

PROOF. Suppose that $x^2 \equiv -1 \pmod{p}$ has a solution. (In particular, this implies that $x \not\equiv 0 \pmod{p}$.) Assume for contradiction that $p = 4k + 3$. Then

$$x^{p-1} = x^{4k+2} = x^{2(2k+1)} \equiv (-1)^{2k+1} = -1 \pmod{p},$$

and this contradicts Fermat's theorem. This shows that our assumption is absurd, and since p is odd, we must therefore have $p \equiv 1 \pmod{4}$.

For the converse, we present an argument of Lagrange. Let $p = 4k+1$. By Wilson's theorem, $(p-1)! = (4k)! \equiv -1\ (p)$. On the other hand, $4k \equiv -1 \pmod{p}$, and so, $4k - 1 \equiv -2 \pmod{p}$, $4k - 2 \equiv -3 \pmod{p}$, etc. Thus

$$\begin{aligned}(4k)! = 1 \cdot 2 \cdots (2k) \cdot (2k+1) \cdots (4k-1) \cdot (4k) &\equiv 1 \cdot 2 \cdots (2k) \cdot (-2k) \cdots (-2) \cdot (-1) \\ &= (-1)^{2k}((2k)!)^2 = ((2k)!)^2.\end{aligned}$$

□

In the coming section, we deduce from this that there are infinitely many primes of the form $4k + 1$ (recall the discussion in Section 1.5).

Exercise 52. *Why can't the proof given to show that there are infinitely many primes of the form $4k - 1$ be also used to show that there are infinitely many primes of the form $4k + 1$?*

3.2. Primes of the form $4k + 1$

THEOREM 25. *There are infinitely many primes of the form $4k + 1$.*

PROOF. Suppose for contradiction there are only finitely many such primes:

$$5, 9, 13, \ldots, p.$$

Let

$$N = 4(5 \cdot 9 \cdots p)^2 + 1.$$

Thus $N \equiv 1 \pmod{4}$, and since $N > p$, it is by assumption not a prime. In particular, it must have an odd prime factor $q | N = (2 \cdot 5 \cdots p)^2 + 1$. Then

$$(2 \cdot 5 \cdots p)^2 \equiv -1 \pmod{q}.$$

Thus by Theorem 24, $q \equiv 1 \pmod{4}$. Therefore q is in our list of primes of the form $4k + 1$, hence $q | (5 \cdot 9 \cdots p)$. But then

$$q | (N - 4(5 \cdot 9 \cdots p)^2) = 1,$$

which is absurd. □

Let's see the first few examples of primes of the form $4k+1$ and $4k-1$.

$p \equiv 1\ (4)$	$p \equiv -1\ (4)$
$5 = 2^2 + 1$	7
$13 = 3^2 + 2^2$	11
$17 = 4^2 + 1$	19
$29 = 5^2 + 2^2$	23
$37 = 6^2 + 1$	31
$\vdots$	$\vdots$

Exercise 53. *Show that no prime of the form $p \equiv -1$ (mod 4) can be written as a sum of two squares.*

Lemma 54 (Fibonacci)**.**

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

PROOF. Direct computation. □

Exercise 55. *Check Euler's identity*

$$(a^2+b^2+c^2+d^2)(A^2+B^2+C^2+D^2) = (aA+bB+cC+dD)^2 + (aB-bA-cD+dC)^2 + (aC+bD-cA-dB)^2 + (aD-bC+cB-dA)^2.$$

With this identity, one can Lagrange showed that to show that every number can be written as the sum of four squares, it suffices to show that every prime can be written as the sum of four squares. This is Lagrange's theorem.

THEOREM 26 (Fermat). *Every prime of the form $p \equiv 1$ (mod 4) can be written as a sum of two squares, in a unique way.*

PROOF. Since $p \equiv 1 \pmod 4$, -1 is a quadratic residue, and $p|(x^2+1)$. Choose m to be the smallest positive integer such that mp can be represented, as the sum of two squares, i.e.

$$mp = x^2 + y^2$$

for some x, y. Such m exists since for some n, $pn = x^2 + 1$.

For contradiction, we will assume that $m \geq 2$. Then we may choose $-\frac{1}{2}m \leq a, b \leq \frac{1}{2}m$ such that $a \equiv x,\ b \equiv y\ (m)$. Then $a^2+b^2 \equiv x^2+y^20 \pmod m$; in particular, there exists $k \in \mathbf{Z}$ such that

$$km = a^2 + b^2.$$

Multiplying the above two equations, and applying Fibonacci's identity,

$$km^2p = (a^2+b^2)(x^2+y^2) = (ax+by)^2 + (ay-bx)^2. \tag{3.1}$$

The right hand side is divisible by m^2: recall that $a \equiv x, b \equiv y\ (m)$;

$$ax + by \equiv a^2 + b^2 \equiv 0\ (m)$$

and

$$ay - bx \equiv ab - ba \equiv 0\ (m).$$

Hence

$$kp = \left(\frac{ax+by}{m}\right)^2 + \left(\frac{ay-bx}{m}\right)^2.$$

Now recall that

$$km = a^2 + b^2 \leq \frac{m^2}{4} + \frac{m^2}{4}$$

and hence $k \leq \frac{m}{2} < m$, contradicting the minimality of m.

We now prove the uniqueness of the representation. Let $p = x^2 + y^2 = a^2 + b^2$. Since p is prime, $(x, y) = (a, b) = 1$.

Mod p, $x^2 \equiv -y^2$. Then

$$p^2 = (x^2 + y^2)(a^2 + b^2) = (xa + yb)^2 + (xb - ay)^2. \tag{3.2}$$

Since $p \equiv 1 \pmod 4$, -1 is a QR mod p. Let h be a solution to $z^2 \equiv -1 \pmod p$. Then $x^2 \equiv -y^2 \equiv (hy)^2 \pmod p$ and $a^2 \equiv (hb)^2 \pmod p$ have solutions $x = hy$, $a = hb$, and

$$xa + yb = h^2yb + yb \equiv -yb + yb = 0 \pmod p, \qquad xb - ya = hyb - hyb = 0 \pmod p.$$

In particular, both terms on the RHS of (3.2) are divisible by p^2. We thus have

$$1 = \left(\frac{xa+yb}{p}\right)^2 + \left(\frac{xb-ya}{p}\right)^2.$$

There are only two ways to represent 1 as a sum of two squares, either (a) $1 = 0^2 + 1^2$ or (b) $1 = 1^2 + 0^2$.

Let's consider case (a) first. Then $xa = -yb$. Since $(x, y) = 1$, $x|b$, and since $(a, b) = 1$, $b|x$. Thus $b = x$. Similarly, one shows that $a = y$. Case (b) is left as an exercise. □

More generally,

THEOREM 27. *A positive integer $n \geq 1$ is a sum of two squares if and only if it is of the form $n = d^2 \cdot 2^l \cdot p_1 \cdots p_k$, where $d \geq 1$, $l \in \{0, 1\}$ and $p_1, \ldots, p_k$ are distinct odd primes of the form $4k + 1$.*

3.3. Legendre symbol

Recall our list of tasks

(1) Given p, determine (efficiently) the set of QR mod p
(2) Given p, determine (efficiently) whether a is a QR mod p
(3) Given a, determine (completely) the set of prime moduli p such that a is a QR mod p

In this section, we address (1) and (2).

Proposition 56. *There are exactly $\frac{p-1}{2}$ QR (mod p), and $\frac{p-1}{2}$ QNR (mod p).*

PROOF. We have to count the number of distinct residues $a^2 \pmod p$ where a ranges over $1, \ldots, p-1$. Since

$$a^2 \equiv (p-a)^2 \pmod p$$

there are at most $\frac{p-1}{2}$ distinct residues. We now show that there are exactly that many QRs. Let $a, b \in \{1, \ldots, \frac{p-1}{2}\}$, and suppose that $a^2 \equiv b^2 \pmod{p}$. Then

$$a^2 - b^2 \equiv 0 \ (p) \iff (a-b)(a+b) \equiv 0 \ (p) \implies a \equiv \pm b \ (p).$$

Since a and b are both in $\{1, \ldots, \frac{p-1}{2}\}$, $a = b$. □

Example 57. *We determine the set of QR mod 17. There will be 8, and since $a^2 \equiv (17-a)^2$ (mod 17), we only need to compute the first 8 squares to find the QR. These are*

$$1, \ 2^2 = 4, \ 3^2 = 9, \ 4^2 = 16, \ 5^2 \equiv 8, \ 6^2 \equiv 2, \ 7^2 \equiv 15, \ 8^2 \equiv 13.$$

Lemma 58. *We note the following multiplicativity behavior.*

(1) *If a and b are QR mod p, then ab is a QR mod p,*
(2) *If a is a QR and b is a QNR, then ab is a QNR,*
(3) *If a and b are QNRs, then ab is a QR.*

Proof. The first statement is easy: if $x^2 \equiv a \ (p)$ and $y^2 \equiv b \ (p)$ then $(xy)^2 \equiv ab \ (p)$.

For the next two statements, we list all residues mod p in function of whether they are QRs,

$$r_1, \ldots, r_k,$$

or QNRs,

$$n_1, \ldots, n_k,$$

where $k = \frac{p-1}{2}$. Let a be a QR mod p. Then

$$ar_1, \ldots, ar_k, an_1, \ldots, an_k$$

are all distinct. By (1), the first half, namely $ar_1, \ldots, ar_k$ are QRs, and this implies, by Proposition 56, that the second half are QNR's. In particular, each product of a QR and a QNR is a QNR. This proves (2). The same argument can be used to prove (3) (exercise). □

Thus QR behave like $+1$, QNR like -1; this motivates the following definition.

Definition 59 (Legendre symbol)**.** *For $a \in \mathbf{Z}$, p an odd prime, the Legendre symbol is*

$$\left(\frac{a}{p}\right) = \begin{cases} +1 & a \text{ is a QR mod } p \\ -1 & a \text{ is a QNR mod } p \end{cases}$$

and $(\frac{a}{p}) = 0$ if $a \equiv 0 \ (p)$.

Proposition 60. *We record the following desired properties.*

(1) $\left(\frac{a}{p}\right)$ *depends only on a (p),*
(2) $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$,
(3) *(Euler's criterium)* $\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \ (p)$.

PROOF. For (1): Suppose that $b \equiv a\ (p)$. Then if a is a QR, there exists $X \in \mathbf{Z}_p^*$ such that $X^2 \equiv a \equiv b\ (p)$, hence b is also a QR. Similarly, if a is a QNR, then b is a QNR. And if $a \equiv 0\ (p)$, then $b \equiv 0\ (p)$.

The second point is simply a reformulation of Lemma 58.

To prove Euler's criterium, suppose first that a is a QR: $x^2 \equiv a \pmod p$. By Fermat's Little Theorem,

$$a^{\frac{p-1}{2}} \equiv x^{p-1} \equiv 1 \pmod p$$

Suppose instead that a is a QNR. For each $b \in \{1, \dots, p-1\}$, we know that $bx \equiv a \pmod p$ has a unique solution $\overline{b}$, and $\overline{b} \neq b$ since otherwise a would be a QR. Therefore, we can pair each residue below with its multiplicative inverse:

$$(p-1)! = 1 \cdot 2 \cdots (p-1) \equiv a \cdots a = a^{\frac{p-1}{2}} \pmod p$$

and Wilson's theorem states that $(p-1)! \equiv -1 \pmod p$. □

Example 61.

$$\left(\frac{72}{97}\right) = \left(\frac{9}{97}\right)\left(\frac{8}{97}\right) = \left(\frac{3}{97}\right)^2\left(\frac{2}{97}\right)^3 = \left(\frac{2}{97}\right) \equiv 2^{48} \equiv 1\ (\textit{mod}\ 97)$$

3.4. Gauss's Lemma

Example 62. *By Euler's criterion,*

$$\left(\frac{-1}{p}\right) \equiv (-1)^{\frac{p-1}{2}}\ (\textit{mod}\ p).$$

This is $+1$ if $\frac{p-1}{2} = 2k$ for some $k \in \mathbf{Z}$, and -1 if $\frac{p-1}{2} = 2k+1$ for some $k \in \mathbf{Z}$. In other words, -1 is a QR exactly when p is of the form $4k+1$.

Proposition 63. *2 is a quadratic residue mod p if and only if $p \equiv \pm 1$ (mod 8).*

PROOF. By Euler's criterium, $(\frac{2}{p}) \equiv 2^{\frac{p-1}{2}} \pmod p$. Observe

$$\begin{aligned} 2^{\frac{p-1}{2}}(\tfrac{p-1}{2})! &= 2^{\frac{p-1}{2}} \cdot \left[1 \cdot 2 \cdots (\tfrac{p-1}{2} - 1)(\tfrac{p-1}{2})\right] \\ &= 2 \cdot 4 \cdots (p-3)(p-1), \end{aligned}$$

and notice that $p-1 \equiv -1\ (p)$, $p-3 \equiv -3\ (p)$, etc: regrouping even and odd numbers, we conclude that

$$2^{\frac{p-1}{2}}(\tfrac{p-1}{2})! \equiv (-1)^{\nu}(\tfrac{p-1}{2})! \pmod p$$

and hence

$$2^{\frac{p-1}{2}} \equiv (-1)^{\nu} \pmod p$$

for

$$\nu = \#\{k \in \{1, \dots, \tfrac{p-1}{2}\} : \tfrac{p-1}{2} < 2k \leq p-1\} = \#\{k \in \{1, \dots, \tfrac{p-1}{2}\} : \tfrac{p-1}{4} < k \leq \tfrac{p-1}{2}\}.$$

Let $p = 4j + r$, where $r \in \{1, 3\}$. Then

$$\nu = \#\{j + \tfrac{r-1}{4} < k \leq 2j + \tfrac{r-1}{2}\}.$$

Now if $r = 1$, $\nu = j$, and if $r = 3$, $\nu = \#\{j < k \leq 2j + 1\} = j + 1$.

If j is even, $p \equiv r$ (8), and ν is even if and only if $r = 1$. If j is odd, then $p \equiv 4 + r$ (8). and ν is even if and only if $r = 3$. This establishes the claim. □

Example 64.

$$\left(\frac{72}{97}\right) = \left(\frac{9}{97}\right)\left(\frac{8}{97}\right) = \left(\frac{3}{97}\right)^2\left(\frac{2}{97}\right)^3 = \left(\frac{2}{97}\right) = 1$$

since $97 \equiv 1$ *(mod 8).*

Lemma 65 (Gauss's lemma)**.** *Let* $a \in \mathbf{Z}_p^*$. *Then*

$$\left(\frac{a}{p}\right) = (-1)^\nu,$$

where ν *is the number of residues* ka *(mod* p*), for* $k = 1, \ldots, \frac{p-1}{2}$ *that lie in the interval* $(\frac{p}{2}, p)$.

Proof. First recall that the $a, 2a, \ldots, \frac{p-1}{2}a$ have distinct residues mod p. Define $a_1, \ldots, a_{\frac{p-1}{2}}$ such that $ka \equiv a_k \pmod p$ and $-\frac{p}{2} < a_k < \frac{p}{2}$. Let

$$\nu = \#\{k \in \{1, \ldots, \tfrac{p-1}{2}\} : -\tfrac{p}{2} < a_k <\}.$$

Then

$$a^{\frac{p-1}{2}}(\tfrac{p-1}{2})! = a(2a)\cdots(\tfrac{p-1}{2}a) \equiv (-1)^\nu(\tfrac{p-1}{2})! \pmod p.$$

Dividing both sides by $(\frac{p-1}{2})!$ together with Euclid's criterion concludes. □

The application of Gauss's lemma to the following proof is perhaps more instructive.

Proposition 66. *3 is a QR mod* p *if and only if* $p \equiv \pm 1$ *(mod 12).*

Proof. Let $p = 12k + r$. We need to count the number n of integers b in the interval

$$\frac{p-1}{2} < 3b < p \iff \frac{p-1}{6} < b < \frac{p}{3} \iff 2k + \frac{r-1}{6} < b < 4k + \frac{r}{3}.$$

Since we only care whether n is even or odd, we may remove $2k$ and $4k$ from both sides:

$$\frac{r-1}{6} < b < \frac{r}{3}.$$

If $r = \pm 1$, $n = 0$. If $r = 5, 7$, $n = 1$. □

3.5. Quadratic reciprocity law

Proposition 67. *If* p, q *are primes such that* $p \equiv \pm q$ *(mod 4a) then*

$$\left(\frac{a}{p}\right) = \left(\frac{a}{q}\right).$$

PROOF. Let $p = q + 4ar$, and $a \in \mathbf{Z}_p^*$. Let n be the number of integers b in the interval

$$\frac{p-1}{2a} < b < \frac{p}{a} \iff 2r + \frac{q-1}{2a} < b < 4r + \frac{q}{a}.$$

Let n' be the number of integers b in the interval

$$\frac{q-1}{2a} < b < \frac{q}{a}.$$

Then n and n' have the same parity (namely, n is even if and only if n' is even). This, together with Gauss' lemma, implies that $(\frac{a}{p}) = (\frac{a}{q})$. The same argument applies when $p \equiv -q \pmod{4a}$. □

THEOREM 28 (Quadratic reciprocity law). *If p and q are odd primes, then*

$$\left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}} \left(\frac{q}{p}\right) = \begin{cases} -(\frac{q}{p}) & p \equiv q \equiv 3\ (4) \\ +(\frac{q}{p}) & \textit{otherwise.} \end{cases}$$

PROOF. Assume that $p \equiv q \pmod 4$. Let $p = q + 4a$. Then

$$\left(\frac{p}{q}\right) = \left(\frac{q+4a}{q}\right) = \left(\frac{4a}{q}\right) = \left(\frac{2}{q}\right)^2 \left(\frac{a}{q}\right) = \left(\frac{a}{q}\right),$$

$$\left(\frac{q}{p}\right) = \left(\frac{p-4a}{p}\right) = \left(\frac{-4a}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{2}{p}\right)^2 \left(\frac{a}{p}\right) = (-1)^{\frac{p-1}{2}} \left(\frac{a}{p}\right),$$

Since $p \equiv q\ (4)$, by Proposition 67,

$$\left(\frac{a}{q}\right) = \left(\frac{a}{p}\right),$$

and as such, if $p \equiv q\ (4)$,

$$\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}} \left(\frac{p}{q}\right).$$

Otherwise, $p \equiv -q \pmod 4$. Let $p + q = 4a$. Then

$$\left(\frac{p}{q}\right) = \left(\frac{4a-q}{q}\right) = \left(\frac{4a}{q}\right) = \left(\frac{a}{q}\right),$$

$$\left(\frac{q}{p}\right) = \left(\frac{4a-p}{p}\right) = \left(\frac{4a}{p}\right) = \left(\frac{a}{p}\right),$$

and again by Proposition 67, $(\frac{p}{q}) = (\frac{q}{p})$. □

3.6. Applications

In the applications below, we use that for two odd primes $(\frac{p}{q})(\frac{q}{p}) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$ and $(\frac{-1}{p}) = 1$ if and only if $p \equiv 1 \pmod 4$, and $(\frac{2}{p}) = 1$ if and only if $p \equiv \pm 1 \pmod 8$.

Exercise 68. *Show that for all odd primes p, q $(\frac{p}{q}) = (\frac{q}{p})$ except when $p \equiv q \equiv 3$ (mod 4), in which case, $(\frac{p}{q}) = -(\frac{q}{p})$*

Example 69.

$$\left(\frac{164}{257}\right) = \left(\frac{2^2 41}{257}\right) = \left(\frac{41}{257}\right) = \left(\frac{257}{41}\right) = \left(\frac{11}{41}\right) = \left(\frac{41}{11}\right) = \left(\frac{8}{11}\right) = \left(\frac{2}{11}\right) = -1.$$

Example 70.

$$\left(\frac{34}{97}\right) = \left(\frac{2}{97}\right)\left(\frac{17}{97}\right) = \left(\frac{17}{97}\right) = \left(\frac{97}{17}\right) = \left(\frac{12}{17}\right) = \left(\frac{2}{17}\right)^2\left(\frac{3}{17}\right) = \left(\frac{3}{17}\right) = -1.$$

Example 71. *We determine the set of odd primes p for which $\left(\frac{5}{p}\right) = 1$. By the Quadratic Reciprocity Law, $(\frac{5}{p}) = (\frac{p}{5})$, which depends only on p (mod 5). We quickly check that*

$$\left(\frac{1}{5}\right) = \left(\frac{4}{5}\right) = +1, \quad \left(\frac{2}{5}\right) = \left(\frac{3}{5}\right) = -1.$$

Hence $(\frac{5}{p}) = 1$ if and only if $p \equiv \pm 1$ (mod 5) and since p is odd, this is really $p \equiv \pm 1$ (mod 10).

Exercise 72. *Let p be an odd prime. Explain why $(\frac{7}{p}) = (-1)^{\frac{p-1}{2}}(\frac{p}{7})$. The first factor on the right hand side depends on p (mod 4) and the second on p (mod 7). Hence the right hand side depends on p (mod 28). Determine the set of odd primes p such that $(\frac{7}{p}) = 1$.*

Proposition 73. *There are infinitely many primes of the form $5k \pm 1$.*

Proof. To complete We give a "Euclidean proof" of this statement. Assume for contradiction that there are only finitely many primes, $p_1 < \ldots, < p_n$ of the form $5k \pm 1$, and set

$$N = (p_1 \cdots p_n)^2 - 5.$$

Then $N > p_n$ and $N \equiv 1 \pmod 5$. Then N is not prime, and factors as a product of primes. Let q be an odd prime divisor of N. Then $(p_1 \cdots p_n)^2 \equiv 5 \pmod q$. By the example above, $q \equiv \pm 1 \pmod 5$. Hence $q|(p_1 \cdots p_n)^2 - N = 5$, which implies that $q = 5$, which is absurd. □

CHAPTER 4

Farey fractions

The field of fractions is

$$\mathbf{Q} = \{\tfrac{a}{b} : a \in \mathbf{Z}, b \in \mathbf{N}^*\}$$

and addition and multiplication are given by, respectively

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}, \qquad \frac{a}{c}\frac{b}{d} = \frac{ac}{bd}.$$

Every fraction $\frac{a}{b} \in \mathbf{Q}$ has an additive and a multiplicative inverse:

$$\frac{a}{b} + \frac{-a}{b} = 0, \qquad \frac{a}{b} \cdot \frac{b}{a} = 1.$$

Exercise 74. *Prove the following statements.*

(1) $\mathbf{Q}$ *is a field*
(2) $\mathbf{Q}$ *is an infinite countable set*
(3) $\mathbf{Q}$ *is a totally ordered set with respect to the following order:*

$$\frac{a}{b} < \frac{c}{d} \iff ad < bc.$$

(4) *Show that this order would not be well-defined if* $bd < 0$. *(This is why we impose that the denominator of each fraction in* $\mathbf{Q}$ *be positive!)*

4.1. Farey sequence

Two observations:

(1) Each real number x can be written as the sum $x = \lfloor x \rfloor + \{x\}$, where $\lfloor x \rfloor$ (**the integral part of** x) is the largest integer $\leq x$, and $\{x\} := x - \lfloor x \rfloor$ (**the fractional part of** x). Note that $\{x\} \in [0, 1)$. In particular each fraction in $\mathbf{Q}$ is the sum of an integer and a fraction in $[0, 1)$.
(2) Fractions are not unique in the following sense:

$$\frac{3}{4} = \frac{15}{20} = \frac{12}{16}.$$

A fraction $\frac{a}{b}$ is called **reduced** if $(a, b) = 1$.

Exercise 75. *Show that each fraction has a unique reduced form.*

Definition 76. *The Farey sequence* $\mathcal{F}_n$ *of order* n *is the ordered set of all fractions* $0 \leq \frac{a}{b} \leq 1$ *in* $\mathbf{Q}$ *such that* $0 < b \leq n$.

Example 77.

$$\mathcal{F}_1 = \{0, 1\}, \quad \mathcal{F}_2 = \{0, \tfrac{1}{2}, 1\}, \quad \mathcal{F}_3 = \{0, \tfrac{1}{3}, \tfrac{1}{2}, \tfrac{2}{3}, 1\}, \quad \mathcal{F}_4 = \{0, \tfrac{1}{4}, \tfrac{1}{3}, \tfrac{1}{2}, \tfrac{2}{3}, \tfrac{3}{4}, 1\}, \quad \dots$$

Proposition 78. *For each $n \in \mathbf{N}^*$,*

$$\mathcal{F}_n = \mathcal{F}_{n-1} \cup \{\tfrac{a}{n} : (a, n) = 1\}.$$

PROOF. By definition,

$$\mathcal{F}_n = \{\tfrac{a}{b} \in [0, 1] : 0 < b \leq n\} = \{\tfrac{a}{b} \in [0, 1] : 0 < b \leq n - 1\} \cup \{\tfrac{a}{n} : (a, n) = 1\}.$$

□

Corollary 79.

$$|\mathcal{F}_n| = 1 + \sum_{k=1}^{n} \varphi(k).$$

PROOF. By recursion,

$$|\mathcal{F}_n| = |\mathcal{F}_{n-1}| + \varphi(n) = |\mathcal{F}_1| + \sum_{k=2}^{n} \varphi(k) = 1 + \sum_{k=1}^{n} \varphi(k)$$

for the convention $\varphi(1) = 1$. □

Proposition 80. *If $\frac{a}{b} < \frac{c}{d}$ are fractions in $\mathbf{Q}$, then we call $\frac{a+c}{b+d}$* **the mediant***, and*

$$\frac{a}{b} < \frac{a + c}{b + d} < \frac{c}{d}.$$

PROOF.

$$\frac{a + c}{b + d} < \frac{c}{d} \iff (a + c)d < (b + d)c \iff ad < bc$$

and this is true by assumption. The other inequality is proven the same way. □

Definition 81. *Two consecutive fractions $\frac{a}{b} < \frac{c}{d}$ in a Farey sequence are called Farey neighbors.*

Example 82. *$\frac{1}{4}$ and $\frac{1}{3}$ are Farey neighbors in $\mathcal{F}_4$, 0, $\frac{1}{3}$ are Farey neighbors in $\mathcal{F}_3$ but not in $\mathcal{F}_4$.*

THEOREM 29. *If $\frac{a}{b} < \frac{c}{d}$ are Farey neighbors, then $cb - ad = 1$.*

PROOF. We prove the statement by induction over n, the order of the Farey sequence. For $n = 1$, $\frac{0}{1} < \frac{1}{1}$ are the only Farey neighbors, and $1 - 0 = 1$. Suppose the assertion holds for all Farey neighbors in $\mathcal{F}_n$. Let $\frac{x}{y}$ be a reduced fraction that is *not* in $\mathcal{F}_n$. Then one can find two Farey neighbors in $\mathcal{F}_n$ such that

$$\frac{a}{b} < \frac{x}{y} < \frac{c}{d}.$$

Let

$$\alpha = xb - ay, \qquad \beta = cy - dx.$$

By the inequalities above, $\alpha, \beta \geq 1$. We now solve this linear system of equations.

$$x = x \iff \alpha d + ady = cby - \beta b \iff y = \alpha d + \beta b,$$

where we used that $cb - ad = 1$ (our induction hypothesis). Similarly, we show that

$$x = \beta a + \alpha c.$$

Hence every reduced fraction between $\frac{a}{b}$ and $\frac{c}{d}$ can be written as

$$\frac{x}{y} = \frac{\beta a + \alpha c}{\alpha d + \beta b}.$$

The next reduced fraction to appear between $\frac{a}{b}$ and $\frac{c}{d}$ in the Farey sequence is, by defintion, the one with smallest denominator, and this happens for $\alpha = \beta = 1$, i.e.

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

the mediant of $\frac{a}{b}$, $\frac{c}{d}$. Since all other $\frac{x}{y}$ have higher denominator, they will only appear in later Farey sequences, and our new Farey neighbors are $\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}$. We now easily check that they satisfy

$$(a+c)b - a(b+d) = cb - ad = 1, \quad c(b+d) - d(a+c) = cb - ad = 1.$$

□

Corollary 83. *The fractions that belong to $\mathcal{F}_{n+1}$ but not $\mathcal{F}_n$ are mediants of fractions in $\mathcal{F}_n$.*

Exercise 84. *Show that if $|cb - ad| = 1$, then $\frac{a}{b}$, $\frac{c}{d}$ must be Farey neighbors in some Farey sequence.*

4.2. Geometry of Farey fractions

Definition 85 (Ford circle)**.** *The Ford circle $\mathcal{C}_{\frac{a}{b}}$ associated to the reduced fraction $\frac{a}{b}$ is the circle with center at*

$$\left(\frac{a}{b}, \frac{1}{2b^2}\right)$$

and radius $1/2b^2$.

Hence each Ford circle $\mathcal{C}_{\frac{a}{b}}$ "lies on" the x-axis, and is tangent to it exactly at $x = \frac{a}{b}$. Below we see the Ford circles for all Farey fractions in $\mathcal{F}_5$.

THEOREM 30. *The interiors of two distinct Ford circles are disjoint. Moreover, two distinct Ford circles $\mathcal{C}_{\frac{a}{b}}$ and $\mathcal{C}_{\frac{c}{d}}$ are tangent if and only if $\frac{a}{b}$ and $\frac{c}{d}$ are Farey neighbors (in some Farey sequence).*

PROOF. Let D denote the distance between the centers of $\mathcal{C}_{\frac{a}{b}}$ and $\mathcal{C}_{\frac{c}{d}}$, that is

$$D^2 = \left(\frac{a}{b} - \frac{c}{d}\right)^2 + \left(\frac{1}{2b^2} - \frac{1}{2d^2}\right)^2.$$

Let R denote the sum of the radii of $\mathcal{C}_{\frac{a}{b}}$ and $\mathcal{C}_{\frac{c}{d}}$, that is

$$R = \frac{1}{2b^2} + \frac{1}{2d^2}.$$

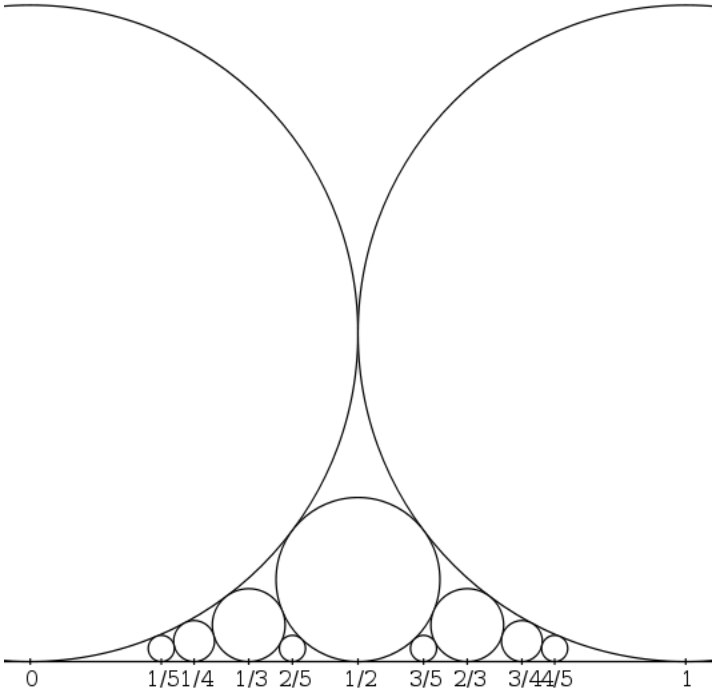

FIGURE 1. From Wikimedia Commons

The interiors of the two circles intersect if $D < R$, they are disjoint if $D \geq R$. This is easily checked:

$$D^2 - R^2 = \left(\frac{a}{b} - \frac{c}{d}\right)^2 - \frac{1}{b^2d^2} = \frac{(ad-bc)^2 - 1}{b^2d^2}.$$

Since $\mathcal{C}_{\frac{a}{b}} \neq \mathcal{C}_{\frac{c}{d}}$, $\frac{a}{b} \neq \frac{c}{d}$, hence $ad - bc \neq 0$. Therefore $(ad-bc)^2 - 1 \geq 1 - 1 = 0$. Moreover, $D = R$ if and only if the two circles are tangent if and only if $|ad - bc| = 1$ if and only if they are Farey neighbors in some Farey sequence (see Exercise 84). □

Exercise 86. *Let $\mathcal{C}_{\frac{a}{b}}$, $\mathcal{C}_{\frac{c}{d}}$ be two tangent Ford circle. Then their point of tangency is given by*

$$\left(\frac{ab+cd}{b^2+d^2}, \frac{1}{b^2+d^2}\right).$$

4.3. Some simple proofs using fractions

Proposition 87. *Let $a, b \in \mathbf{N}^*$ such that $(a, b) = 1$. There exist $u, v \in \mathbf{Z}$ such that $au - bv = 1$.*

PROOF. Suppose that $a < b$. Then $\frac{a}{b}$ is a Farey fraction in $\mathcal{F}_b$. Let $\frac{v}{u} < \frac{a}{b}$ be its Farey neighbor, then by Theorem 29, $au - bv = 1$. □

Proposition 88 (Gauss)**.**

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

PROOF. Consider the set $S = \{\frac{1}{n}, \ldots, \frac{n-1}{n}, \frac{n}{n}\}$. Among the n fractions of this set $\varphi(n)$ are reduced. Suppose $\frac{k}{n}$ is not reduced. Then there is a divisor $d|n$ such that $\frac{k}{n} = \frac{k'}{d}$ and $(k', d) = 1$. There are $\varphi(d)$ fractions in S that, in reduced form, have denominator k'. Hence

$$n = |S| = \varphi(n) + \sum_{\substack{d|n \\ d \neq n}} \varphi(d).$$

□

CHAPTER 5

Continued fractions

5.1. Rationals and finite continued fractions

We expand $\frac{67}{24}$ in a continued fraction. First note $\frac{67}{24} \approx 2.79$. Hence $\lfloor \frac{67}{24} \rfloor = 2$ and $\{\frac{67}{24}\} = \frac{67}{24} - 2$. We can express the fractional part more precisely using Euclid's division algorithm:

$$\frac{67}{24} = \frac{2 \cdot 24 + 19}{24} = 2 + \frac{19}{24} = 2 + \frac{1}{\frac{24}{19}} = 2 + \frac{1}{\frac{1 \cdot 19 + 5}{19}} = 2 + \frac{1}{1 + \frac{1}{\frac{19}{5}}} = 2 + \frac{1}{1 + \frac{1}{3 + \frac{1}{1 + \frac{1}{4}}}}.$$

The right hand side is called a continued fraction expansion, its terms $2, 1, 3, 1, 4$ are called its **partial quotients**. These are precisely the quotients appearing in the Euclidean division algorithm. In fact, compare the latter continued fraction expansion to

$$\begin{aligned} 67 &= \mathbf{2} \cdot 24 + 19 \\ 24 &= \mathbf{1} \cdot 19 + 5 \\ 19 &= \mathbf{3} \cdot 5 + 4 \\ 5 &= \mathbf{1} \cdot 4 + 1 \\ 4 &= \mathbf{4} \cdot 1. \end{aligned}$$

We call $\frac{67}{24}, \frac{24}{19}, \frac{19}{5}, \frac{5}{4}$ the **complete quotients**. However, the partial quotients determine the continued fraction expansion, and this motivates the following notation

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \frac{1}{a_n}}}} = [a_0, a_1, \ldots, a_n].$$

Here are a few additional facts about the partial quotients we can deduce from Theorem 1 and Theorem 10 that

(1) $a_0 \in \mathbf{N}$, $a_1, \ldots, a_n \in \mathbf{N}^*$.
(2) If we expand using the division algorithm as above $\frac{a}{b} = [a_0, a_1, \ldots, a_n]$ then $a_n > (a, b) \geq 1$, i.e. $a_n > 1$.
(3) However note that

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

Example 89.

$$\frac{67}{24} = 2 + \frac{1}{1 + \frac{1}{3 + \frac{1}{1 + \frac{1}{4}}}} = 2 + \frac{1}{1 + \frac{1}{3 + \frac{1}{1 + \frac{1}{3 + \frac{1}{1}}}}}, \qquad \frac{1}{2} = 0 + \frac{1}{2} = 0 + \frac{1}{1 + \frac{1}{1}}$$

THEOREM 31. *A rational $\frac{a}{b} \in \mathbf{Q}$ has a unique continued fraction expansion whose last partial quotient is > 1.*

PROOF. Suppose that $\frac{a}{b}$ had two continued fraction expansions:

$$[a_0, a_1, \ldots, a_m] = [b_0, b_1, \ldots, b_n].$$

Taking the integral part of $\frac{a}{b}$, we see that $\lfloor \frac{a}{b} \rfloor = a_0 = b_0$. Hence canceling a_0 out on both sides, we are left with

$$[a_1, \ldots, a_m] = [b_1, \ldots, b_n].$$

Repeating this argument, we see that $m = n$, and $a_i = b_i$ for $i = 0, \ldots, n$. □

Our **convention** will be to use the continued fraction whose last partial quotient is > 1.

Examples 90. *Some examples:*

(1) $\frac{17}{11} = [1, 1, 1, 5]$
(2) $\frac{11}{31} = [0, 2, 1, 4, 2]$
(3) $\sqrt{2} = 1 + (\sqrt{2} - 1) = 1 + \frac{1}{1+\sqrt{2}} = [1, 2, 2, \ldots]$. *Here there is obviously a convergence issue we need to take care of, but we will see that this infinite continued fraction expansions does indeed represent $\sqrt{2}$.*

5.2. Convergents

Consider a continued fraction, finite or infinite, $[a_0, a_1, ...]$.

Definition 91. *Its k^{th} convergent is $[a_0, \ldots, a_k]$.*

Example 92. *We already computed that*

$$\frac{67}{24} = [2, 1, 3, 1, 4].$$

Its convergents are

$$[2] = 2$$
$$[2, 1] = 2 + \frac{1}{1} = 3$$
$$[2, 1, 3] = 2 + \frac{1}{1 + \frac{1}{3}} = 2 + \frac{3}{4} = \frac{11}{4}$$
$$[2, 1, 3, 1] = 2 + \frac{1}{1 + \frac{1}{3+1}} = 2 + \frac{4}{5} = \frac{14}{5}$$

and, finally, $[2,1,3,1,4] = \frac{67}{24}$. *So the successive convergents are*

$$2,\ 3, \frac{11}{4},\ \frac{14}{5},\ \frac{67}{24}.$$

And if we order them by size,

$$2 < \frac{11}{4} < \frac{67}{24} < \frac{14}{5} < 3.$$

In fact, each further convergent is closer to $\frac{67}{24}$, *with even convergents on its left, and odd convergents on the right. We will later see that this phenomenon is always true.*

Following the definition, the first few convergents of $[a_0, a_1, a_2, a_3, \dots]$ are given by the formulas

$$\begin{aligned}
[a_0] &= a_0 \\
[a_0, a_1] &= a_0 + \frac{1}{a_1} = \frac{a_0a_1 + 1}{a_1} \\
[a_0, a_1, a_2] &= a_0 + \frac{1}{a_1 + \frac{1}{a_2}} = a_0 + \frac{1}{[a_1, a_2]} = a_0 + \frac{a_2}{a_1a_2 + 1} = \frac{a_0a_1a_2 + a_0 + a_2}{a_1a_2 + 1} \\
[a_0, a_1, a_2, a_3] &= a_0 + \frac{a_2a_3 + 1}{a_1a_2a_3 + a_1 + a_3} = \frac{a_0a_1a_2a_3 + a_0a_1 + a_0a_3 + a_2a_3 + 1}{a_1a_2a_3 + a_1 + a_3}
\end{aligned}$$

etc. Looking at these (and further) computations, you can observe that there is a discernable, recursive pattern in the determination of the convergents. Of course, this only works in the above fractions are already in reduced forms. We settle these two points at once with the following lemma.

Lemma 93. *Given a continued fraction (finite or infinite)* $[a_0, a_1, \dots]$, *define* $p_{-1} := 1, q_{-1} := 0$, $p_0 := a_0$, $q_0 := 1$, *and*

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

for $k \geq 1$. *Then* $p_kq_{k-1} - p_{k-1}q_k = (-1)^{k+1}$.

Proof. We can immediately check that $p_0q_{-1} - p_{-1}q_0 = -1$. Suppose that the relation holds for k. Then

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

□

Corollary 94. *For each* $k \geq 0$, $(p_k, q_k) = 1$.

Proof. If d is a positive divisor of both p_k and q_k, then by the previous lemma, it also divides $(-1)^{k+1}$. Hence $d = 1$. □

Corollary 95. *Given a continued fraction (finite or infinite) $[a_0, a_1, \dots]$, its convergents are*

$$[a_0, \dots, a_k] = \frac{p_k}{q_k}$$

with p_k, q_k defined as in Lemma 93.

PROOF. We again proceed by induction. Suppose this holds for k^{th} convergents. Formally,

$$\begin{aligned}[a_0, \dots, a_{k-1}, a_k] = [a_0, \dots, a_{k-1} + \tfrac{1}{a_k}] &= \frac{(a_{k-1} + \frac{1}{a_k})p_{k-2} + p_{k-3}}{(a_{k-1} + \frac{1}{a_k})q_{k-2} + q_{k-3}} \\ &= \frac{a_k a_{k-1} p_{k-2} + a_k p_{k-3} + p_{k-2}}{a_k a_{k-1} q_{k-2} + a_k q_{k-3} + q_{k-2}} \\ &= \frac{a_k p_{k-1} + p_{k-2}}{a_k q_{k-1} + q_{k-2}}.\end{aligned}$$

□

Exercise 96. *Compute the convergents of $\frac{67}{24}$ using the recursion formulas given by Lemma 93.*

Corollary 97. *Given a (finite or infinite) continued fraction $[a_0, a_1, \dots]$ where $a_k \geq 1$ for $k \geq 1$, then the convergent's denominators (q_k) form a strictly increasing sequence from $k \geq 1$ on, and*

$$q_k \geq k$$

for each $k \geq 0$.

PROOF. We immediately note that $q_0 = 1$, $q_1 = a_1 \geq 1$, $q_2 = a_2 a_1 + 1 \geq 2$. Suppose that the statements hold for all denominators of convergents up to k, with $k \geq 3$. Then

$$q_k = a_k q_{k-1} + q_{k-2} \geq q_{k-1} + q_{k-2} \geq q_{k-1} + 1 > q_{k-1}$$

and

$$q_k \geq q_{k-1} + q_{k-2} \geq k - 1 + k - 2 \geq k.$$

□

5.3. Solutions to $ax + by = 1$

Recall that by Bézout's lemma, for any two coprime numbers a, b, there exist $x, y \in \mathbf{Z}$ such that

$$ax + by = 1.$$

We saw two proofs of this result. The first one used a minimality argument that did not provide us with a construction of the solutions. The second one used Farey fractions and showed that the solution can be constructed with Farey neighbors. Here, we use continued fractions to construct solutions even more quickly : Since a, b are coprime, $\frac{a}{b}$ is in reduced form, and let $[a_0, \dots, a_n]$ be its continued fraction expansion. The last convergent is

$$\frac{p_n}{q_n} = \frac{a}{b}$$

and since $aq_{n-1} - bp_{n-1} = (-1)^{n+1}$, the second to last convergent provides an explicit solution to the equation $ax + by = 1$!

Example 98. *If $a = 67$ and $b = 24$, then the second to last convergent is $\frac{p_3}{q_3} = \frac{14}{5}$ and $aq_3 - bp_3 = -1$, hence*

$$67(-5) + 24(14) = 1.$$

We can also use this to compute the multiplicative inverse of an integer $a \pmod n$ (where $(a, n) = 1$). In fact, if $(a, n) = 1$, then there exist $x, y \in \mathbf{Z}$ such that $ax+by = 1$. If $\frac{a}{n} = [a_0, \ldots, a_k]$, then $aq_{k-1} - np_{k-1} = (-1)^{k+1}$, and thus

$$a((-1)^{k-1}q_{k-1}) \equiv 1 \pmod n.$$

Example 99. *To find the multiplicative inverse of 13 (mod 76), we solve the congruence equation $13x \equiv 1$ (mod 76). We first compute the continued fraction expansion*

$$\frac{13}{76} = [0, 5, 1, 5, 2].$$

Its convergents are

$$0, \frac{1}{5}, \frac{1}{6}, \frac{6}{35}, \frac{13}{76}.$$

Then

$$13(-35) \equiv 13(41) \equiv 1 \ (mod\ 76).$$

5.4. Irrationals and infinite continued fractions

Recall that we saw

$$\sqrt{2} = 1 + \frac{1}{1+\sqrt{2}} = 1 + \frac{1}{2 + \frac{1}{1+\sqrt{2}}} = 1 + \frac{1}{2 + \cfrac{1}{2 + \cfrac{1}{\ddots\, 2 + \frac{1}{1+\sqrt{2}}}}}} = [1, 2, 2, \ldots, 2, 1 + \sqrt{2}],$$

but does the infinite continued fraction $[1, \overline{2}]$ eventually converge to $\sqrt{2}$? Let $\frac{p_n}{q_n}$ denote the last rational convergent in the finite expansion above, i.e. $\frac{p_n}{q_n} = [1, 2, 2, \ldots, 2]$. Then

$$\begin{aligned}
\left|\sqrt{2} - \frac{p_n}{q_n}\right| &= \left|\frac{(1+\sqrt{2})p_n + p_{n-1}}{(1+\sqrt{2})q_n + q_{n-1}} - \frac{p_n}{q_n}\right| \\
&= \left|\frac{p_{n-1}q_n - p_nq_{n-1}}{((1+\sqrt{2})q_n + q_{n-1})q_n}\right| \\
&= \frac{1}{((1+\sqrt{2})q_n + q_{n-1})q_n} \\
&\leq \frac{1}{n^2}.
\end{aligned}$$

From this, we conclude that

$$\sqrt{2} = \lim_{n\to\infty} \frac{p_n}{q_n},$$

where $\frac{p_n}{q_n}$ are the convergents of the infinite continued fraction $[1, \overline{2}]$.

More generally, let α be an irrational number. We can decompose α into its integral and fractional parts, $\alpha = \lfloor \alpha \rfloor + \{\alpha\}$. Let

$$a_0 := \lfloor \alpha \rfloor, \qquad \alpha_1 := \frac{1}{\{\alpha\}} = \frac{1}{\alpha - a_0}.$$

Note that $a_0 \in \mathbf{Z}$, $\alpha_1 > 1$, and $\alpha_1 \notin \mathbf{Q}$ (prove these facts!). Then

$$\alpha = a_0 + \frac{1}{\alpha_1}.$$

Now let

$$a_1 := \lfloor \alpha_1 \rfloor, \qquad \alpha_2 := \frac{1}{\{\alpha_1\}} = \frac{1}{\alpha_1 - a_1}.$$

Here, $a_1 \in \mathbf{N}^*$, $\alpha_2 > 1$, and $\alpha_2 \notin \mathbf{Q}$. Then

$$\alpha = a_0 + \frac{1}{a_1 + \frac{1}{\alpha_2}}.$$

Iterating this process over n steps yields

$$\alpha = [a_0, a_1, \ldots, a_{n-1}, \alpha_n],$$

where

- $a_0 \in \mathbf{Z}$
- $a_k \in \mathbf{N}^*$ for $k = 1, \ldots, n-1$,
- $\alpha_n > 1$ and $\alpha_n \notin \mathbf{Q}$,
- the first n convergents $\frac{p_k}{q_k} = [a_0, \ldots, a_k] \in \mathbf{Q}$,
- $$[a_0, \ldots, \alpha_n] = \frac{\alpha_n p_{n-1} + p_{n-2}}{\alpha_n q_{n-1} + q_{n-2}}.$$

Exercise 100. *Prove these statements, and conclude that*

$$\alpha = \lim_{n \to \infty} \frac{p_n}{q_n}.$$

THEOREM 32. *Let $(\frac{p_n}{q_n})$ be the sequence of convergents of a (finite or infinite) continued fraction $x = [a_0, a_1, \ldots]$. Then all even convergents lie in increasing order to the left of the final value $\frac{a}{b}$ and all odd convergents lie in decreasing order to its right,*

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \cdots < x < \cdots < \frac{p_3}{q_3} < \frac{p_1}{q_1}.$$

Moreover, each convergent is closer to the final value x than the previous one, that is

$$\left| \frac{p_n}{q_n} - x \right| < \left| \frac{p_{n-1}}{q_{n-1}} - x \right|.$$

PROOF. Consider

$$\frac{p_{n+2}}{q_{n+2}} - \frac{p_n}{q_n} = \frac{a_{n+2}p_{n+1} + p_n}{a_{n+2}q_{n+1} + q_n} - \frac{p_n}{q_n} = \frac{a_{n+1}(p_{n+1}q_n - p_n q_{n+1})}{q_{n+2}q_n} = \frac{a_{n+2}(-1)^{n+1}}{q_{n+2}q_n}.$$

The RHS is positive if n is odd, and negative if n is even. This shows that even convergents form an increasing sequence, and odd convergents, a decreasing one. Write

$$x = [a_0, \dots] = [a_0, \dots, a_n, y] \qquad \text{with } y = [a_{n+1}, \dots].$$

Here note that $y > 1$. Then

$$x = \frac{yp_n + p_{n-1}}{yq_n + q_{n-1}}$$

which is equivalent to

$$y(xq_n - p_n) = p_{n-1} - xq_{n-1}.$$

Dividing both sides by yq_n,

$$x - \frac{p_n}{q_n} = \frac{p_{n-1} - xq_{n-1}}{yq_n} = \frac{q_{n-1}}{yq_n}\left(\frac{p_{n-1}}{q_{n-1}} - x\right).$$

Since

$$\frac{q_{n-1}}{yq_n} < \frac{q_{n-1}}{q_n} < 1,$$

taking absolute values on both sides yields the last assertion. □

THEOREM 33. *Let $[a_0, a_1, \dots]$ be an infinite continued fraction, where $a_0 \in \mathbf{Z}$ and $a_i \in \mathbf{N}^*$ for all $i \geq 1$. Then*

$$\lim_{n\to\infty} \frac{p_n}{q_n} \quad \textit{exists}$$

and is irrational.

PROOF. Note that for each convergent $\frac{p_n}{q_n}$, $n \geq 0$,

$$\frac{p_0}{q_0} \leq \frac{p_n}{q_n} \leq \frac{p_1}{q_1}.$$

Moreover, the subsequence

$$\left(\frac{p_n}{q_n}\right)_{n \text{ even}}$$

is increasing and bounded by above, while the subsequence

$$\left(\frac{p_n}{q_n}\right)_{n \text{ odd}}$$

is decreasing and bounded below. Hence both subsequences have a limit, and since

$$\left|\frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n}\right| = \frac{1}{q_n q_{n+1}} \leq \frac{1}{n(n+1)} \leq \frac{1}{n^2} \to 0$$

as $n \to \infty$, these two limits coincide. □

5.5. Quadratic irrationals

Quadratic irrationals are irrational solutions of a quadratic polynomial

$$ax^2 + bx + c = 0,$$

with $a, b, c \in \mathbf{Z}$. Let $d = b^2 - 4ac$ denote the discriminant of this equation, recall that for $d > 0$, the equation has two solutions, explicitly given by

$$\alpha = \frac{-b + \sqrt{d}}{2a}, \qquad \overline{\alpha} = \frac{-b - \sqrt{d}}{2a},$$

where $\overline{\alpha}$ is called the conjugate of α.

Exercise 101. *$\sqrt{2}, 1 + \sqrt{2}$ are quadratic irrationals. Find the polynomials for which they are roots.*

We already have seen that $\sqrt{2} = [1, \overline{2}]$. Or in other words,

$$\sqrt{2} = 1 + \frac{1}{[\overline{2}]}$$

and this is equivalent to $\sqrt{2} + 1 = [\overline{2}]$. Such a continued fraction is called **purely periodic**, while the continued fraction expansion of $\sqrt{2}$ is **eventually periodic**.

Examples 102. (1)

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

(2)

$$\alpha = [\overline{1, 2}] = 1 + \frac{1}{2 + \frac{1}{\alpha}} = 1 + \frac{\alpha}{2\alpha + 1} = \frac{3\alpha + 1}{2\alpha + 1}$$

and this is equivalent to

$$2\alpha^2 - 2\alpha - 1 = 0.$$

Solving this equation, we find the positive solution

$$\alpha = \frac{1}{2} + \frac{\sqrt{3}}{2}.$$

We will see a special case of the following theorem of Lagrange: every quadratic irrational has an eventually periodic continued fraction expansion.

Definition 103. *We call a quadratic irrational α is reduced if $\alpha > 1$ and $-1 < \overline{\alpha} < 0$.*

Examples 104. (1) *$\alpha = \frac{1}{2} + \frac{\sqrt{3}}{2}$ is a reduced quadratic irrational.*
(2) *$\alpha = \sqrt{n}$ is a quadratic irrational ($\alpha^2 - n = 0$) but not reduced: $\overline{\alpha} = -\sqrt{n} \leq -1$.*
(3) *However $\alpha = \sqrt{n} + \lfloor \sqrt{n} \rfloor$ is a reduced quadratic irrational. In fact, $\alpha \geq 1 + 1 = 2$, and*

$$\overline{\alpha} = -\sqrt{n} + \lfloor \sqrt{n} \rfloor = -\{\sqrt{n}\} \in (-1, 0).$$

THEOREM 34 (Galois). *A quadratic irrational α is reduced if and only if α has a purely periodic continued fraction expansion $\alpha = [\overline{a_0, a_1, \ldots, a_n}]$.*

PROOF. Suppose that α is reduced. Let $\alpha = \frac{-b+\sqrt{\Delta}}{2a}$. Observe that since α is reduced,

$$0 < a < \sqrt{\Delta}, \qquad -\sqrt{\Delta} < b < 0.$$

(Exercise.) More generally, this shows that there are only finitely many reduced quadratic irrationals to discriminant Δ.

Consider α_1 in

$$\alpha = a_0 + \frac{1}{\alpha_1}.$$

Since $\alpha > 1$, $a_0 \geq 1$, and $\alpha_1 > 1$ is also a quadratic irrational to discriminant d. Indeed,

$$\alpha_1 = \frac{1}{\alpha - a_0} = \frac{2a}{-(b+2aa_0)+\sqrt{\Delta}} = \frac{2a(b+2aa_0+\sqrt{\Delta})}{\Delta-(b+2aa_0)^2}.$$

Set

$$b_1 = -b - 2aa_0,$$

and

$$a_1 = \frac{\Delta-(b+2aa_0)^2}{4a} = -c - a_0 - aa_0^2 \neq 0.$$

Then α_1 is written

$$\alpha_1 = \frac{-b_1+\sqrt{\Delta}}{2a_1}$$

and its conjugate is

$$\overline{\alpha_1} = \frac{-b_1-\sqrt{\Delta}}{2a_1} = \frac{1}{\overline{\alpha}-a_0}.$$

In particular, since $\overline{\alpha} \in (-1,0)$, $\overline{\alpha_1} \in (-1,0)$. Hence, every new quadratic irrational arising in the continued fraction expansion of α is itself a reduced quadratic irrational to discriminant d. Since there are only finitely many reduced quadratic irrationals to discriminant d, the continued fraction expansion must be eventually periodic. That is,

$$\alpha = [a_0, \ldots, a_n, \alpha_{n+1}]$$

and there is some $m \leq n$ for which $\alpha_{n+1} = \alpha_m$. Since $\alpha_{n+1} = \alpha_m$, their conjugates are also equal: $\overline{\alpha_{n+1}} = \overline{\alpha_m}$. Set

$$\beta_n = -\frac{1}{\overline{\alpha_n}} > 1.$$

Then

$$\overline{\alpha_n} = a_n + \frac{1}{\overline{\alpha_{n+1}}} \iff \beta_{n+1} = a_n + \frac{1}{\beta_n}.$$

We conclude that $a_n = \lfloor \beta_{n+1} \rfloor$. In particular, if $\beta_{n+1} = \beta_m$ then $a_{n+1} = a_m$ and $\beta_n = \beta_{m-1}$. This in turns implies that $\alpha_n = \alpha_{m-1}$. This shows that

$$\alpha = [a_0, \ldots, a_{n-m}, \alpha] = [\overline{a_0, \ldots, a_{n-m}}].$$

Suppose now that α has the completely periodic continued fraction expansion

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots \cfrac{1}{a_n + \frac{1}{\alpha}}}}$$

Then

$$\alpha = \frac{\alpha p_n + p_{n-1}}{\alpha q_n + q_{n-1}} \implies q_n\alpha^2 + (q_{n-1} - p_n)\alpha - p_{n-1} = 0.$$

The discriminant of this equation is

$$d = (q_{n-1} - p_n)^2 + 4q_n p_{n-1} = (p_n + q_{n-1})^2 + 4(-1)^n.$$

In particular, $d > 0$ and is not a perfect square. Hence α is a quadratic irrational. We still have to check that it is reduced. Since the expansion is purely periodic, $a_0, \ldots, a_n \geq 1$, and since $a_0 = \lfloor \alpha \rfloor$, we have that $\alpha > 1$. Then we note that

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

The second equation establishes that α and its conjugate have opposite signs. From the first one, we get

$$\overline{\alpha} = \frac{p_n}{q_n} - \alpha - \frac{q_{n-1}}{q_n} > \frac{p_n}{q_n} - \alpha - 1$$

and this converges to -1 as $n \to \infty$. □

We do not know anything about the structure of continued fraction expansions for other algebraic irrational than the quadratic ones.

Example 105. *We compute the value of* $\alpha = [\overline{1}]$. *We first note the recurrence relation*

$$\alpha = 1 + \frac{1}{\alpha}.$$

This is equivalent to the quadratic equation $\alpha^2 - \alpha - 1 = 0$. *The positive solution is*

$$\alpha = \frac{1 + \sqrt{5}}{2},$$

the golden mean.

5.6. Pell's equation

Let $N \geq 1$. Then

$$x^2 - Ny^2 = 1$$

is called a Pell's equation. We wish to find integer solutions $x, y \in \mathbf{Z}$ to this equation.

If N is a perfect square, meaning that $N = k^2$ for some $k \in \mathbf{N}$, then

$$x^2 - Ny^2 = (x - ky)(x + ky) = 1$$

and this has integer solutions if and only if

$$x - ky = x + ky = \pm 1.$$

But this only holds for $y = 0$, $x = \pm 1$.

From here on, we will assume that N is not a perfect square. We nonetheless have the factorization

$$x^2 - Ny^2 = (x - \sqrt{N}y)(x + \sqrt{N}y) = 1.$$

The continued fraction expansion of $\sqrt{N}$ is of the form

$$\sqrt{N} = [a_0, \overline{a_1, \ldots, a_k, 2a_0}].$$

(Exercise.) Let

$$\sqrt{N} = [a_0, a_1, \ldots, a_k, \alpha_{k+1}] = \frac{\alpha_{k+1}p_k + p_{k-1}}{\alpha_{k+1}q_k + q_{k-1}}. \tag{5.1}$$

Note that here

$$\alpha_{k+1} = [\overline{2a_0, a_1, \ldots, a_k}] = a_0 + \sqrt{N}.$$

Plugging this in (5.1), we obtain

$$\sqrt{N}(a_0q_k + q_{k-1}) + Nq_k = a_0p_k + p_{k-1} + \sqrt{N}p_k.$$

and this implies

$$p_{k-1} = Nq_k - a_0p_k, \tag{5.2}$$

$$q_{k-1} = p_k - a_0q_k. \tag{5.3}$$

Exercise 106. *Prove that if $a + b\sqrt{N} = c + d\sqrt{N}$. then $a = c$ and $b = d$.*

Recall that the k-th $(k-1)$-th convergents are related by

$$p_kq_{k-1} - p_{k-1}q_k = (-1)^{k-1}.$$

Plugging (5.2) in :

$$p_k(p_k - a_0q_k) - (Nq_k - a_0p_k)q_k = p_k^2 - Nq_k^2 = (-1)^{k-1}.$$

Hence if k **is odd**, then

$$p_k^2 - Nq_k^2 = 1.$$

If k **is even** instead, we have $p_k^2 - Nq_k^2 = -1$. Squaring both sides of the equation yields

$$(p_k^2 + Nq_k^2)^2 - N(2p_kq_k)^2 = 1.$$

Example 107. *Let $N = 21$. Then*

$$\sqrt{21} = [4, \overline{1, 1, 2, 1, 1, 8}].$$

The first few convergents are

$$4, 5, \frac{9}{2}, \frac{23}{5}, \frac{32}{7}, \frac{55}{12} = \frac{p_5}{q_5}.$$

Hence

$$p_5^2 - 21q_5^2 = 55^2 - 21 \cdot 12^2 = 1.$$

Example 108. *Let $N = 29$. Then*

$$\sqrt{29} = [5, \overline{2, 1, 1, 2, 10}].$$

The convergents are

$$5, \frac{11}{2}, \frac{16}{3}, \frac{27}{5}, \frac{70}{13} = \frac{p_4}{q_4}.$$

Hence

$$(p_4^2 + 29q_4^2)^2 - 29(2p_4q_4)^2 = 9801^2 - 29(1820)^2 = 1.$$

The parity of "k" in the continued fraction expansion of $\sqrt{N}$ is rather mysterious. In fact, we do not have a characterization of N for which k is even (or odd).

The equation

$$x^2 - Ny^2 = -1$$

is called the **negative Pell's equation**. Note that if N is such that k is even, then the negative Pell's equation is soluble. Otherwise this is not necessarily the case.

Example 109. *We show that $x^2 - 21y^2 = -1$ is not soluble. In fact, if it were, then $x^2 \equiv -1$ (mod 21) would be soluble and in particular $x^2 \equiv -1$ (mod 3) would be soluble. But this is not the case.*

The next fact are not harder to prove but we will omit its proof.

Fact 110. *All solutions to the Pell equations $x^2 - Ny^2 = \pm 1$ arise from the convergents of $\sqrt{N}$. In particular,*

$$x^2 - Ny^2 = -1 \text{ is soluble } \iff k \text{ is even.}$$

Moreover, (p_k, q_k) (rec. $(p_k^2 + Nq_k^2, 2p_kq_k)$) is the smallest solution to $x^2 - Ny^2 = 1$, and every other solutions is given by

$$(x_n, y_n) = (p_{k+n(k+1)}, q_{k+n(k+1)})$$

[TO FIGURE OUT]

5.7. On the shape of $\sqrt{N}$

Proposition 111. *Let $\alpha = [\overline{a_0, \ldots, a_n}]$ and $\beta = [\overline{a_n, \ldots, a_0}]$. Then*

$$-\frac{1}{\beta} = \overline{\alpha}.$$

Example 112. *Consider $\alpha = [\overline{1, 2}]$. Then*

$$\alpha = [1, 2, \alpha] = \frac{3\alpha + 1}{2\alpha + 1}$$

and this is equivalent to

$$2\alpha^2 - 2\alpha - 1 = 0,$$

which has solution

$$\alpha = \frac{1 + \sqrt{3}}{2}.$$

Similarly $\beta = [\overline{2, 1}]$ leads to the quadratic equation

$$\beta^2 - 2\beta - 2 = 0,$$

which has solution

$$\beta = 1 + \sqrt{3}.$$

PROOF. Let

$$\alpha = [a_0, \dots, a_n, \alpha] = \frac{\alpha p_n + p_{n-1}}{\alpha q_n + q_{n-1}}$$

which yields the quadratic equation

$$q_n \alpha^2 + (q_{n-1} - p_n)\alpha - p_{n-1} = 0. \tag{5.4}$$

Similarly

$$\beta = [a_n, \dots, a_0, \beta] = \frac{\beta p'_n + p'_{n-1}}{\beta q'_n + q'_{n-1}}$$

yields

$$q'_n \beta^2 + (q'_{n-1} - p'_n)\beta - p'_{n-1} = 0.$$

Given Lemma 113 below, the LHS here is equal to

$$p_{n-1}\beta^2 + (q_{n-1} - p_n)\beta - q_n = 0. \tag{5.5}$$

Comparing (5.4) and (5.5), we remark that there is a symmetry in the coefficients of the two quadratic equations. In fact, if we divide both sides of (5.5) by $-\frac{1}{\beta^2}$, we obtain

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

which is equivalent to

$$q_n \left(-\frac{1}{\beta}\right)^2 - (q_{n-1} - p_n)\left(-\frac{1}{\beta}\right) - p_{n-1} = 0.$$

Hence the equation $q_n x^2 + (q_{n-1} - p_n)x - p_{n-1} = 0$ has two solutions: α and $-1/\beta$. Since $\alpha \neq -1/\beta$, we must have that $-1/\beta = \overline{\alpha}$. □

Lemma 113. *Let*

$$\frac{p_n}{q_n} = [a_0, \dots, a_n], \quad \frac{p'_n}{q'_n} = [a_n, \dots, a_0].$$

Then

$$\frac{p_n}{p_{n-1}} = \frac{p'_n}{q'_n}, \quad \frac{q_n}{q_{n-1}} = \frac{p'_{n-1}}{q'_{n-1}}.$$

PROOF. Recall the recurrence relation $p_n = a_n p_{n-1} + p_{n-2}$. Applying it repeatedly, we can write

$$\frac{p_n}{p_{n-1}} = a_n + \frac{1}{\frac{p_{n-1}}{p_{n-2}}} = a_n + \frac{1}{a_{n-1} + \frac{1}{\frac{p_{n-2}}{p_{n-3}}}} = \dots = a_n + \frac{1}{a_{n-1} + \frac{1}{\ddots + \frac{p_0}{p_{-1}}}} = \frac{p'_n}{q'_n}.$$

Similarly,

$$\frac{q_n}{q_{n-1}} = a_n + \frac{1}{\frac{q_{n-1}}{q_{n-2}}} = \dots = a_n + \frac{1}{a_{n-1} + \frac{1}{\ddots + \frac{q_1}{q_0}}} = [a_n, \dots, a_1] = \frac{p'_{n-1}}{q'_{n-1}}.$$

□

Let $\sqrt{N} = [a_0, \overline{a_1, \ldots, a_k, 2a_0}]$. Note that

$$\alpha = \sqrt{N} + a_0 = [\overline{2a_0, a_1, \ldots, a_k}]$$

is purely periodic. Hence by Proposition 111,

$$\overline{\alpha} = \frac{-1}{[\overline{a_k, \ldots, a_1, 2a_0}]}.$$

On the other hand $\overline{\alpha} = a_0 - \sqrt{N}$. If we compare both expressions,

$$[\overline{a_k, \ldots, a_1, 2a_0}] = \frac{1}{\sqrt{N} - a_0} = \frac{1}{[0, \overline{a_1, \ldots, a_k, 2a_0}]} = [\overline{a_1, \ldots, a_k, 2a_0}].$$

Since continued fraction expansions are unique, we conclude that

$$a_k = a_1, \ a_{k-1} = a_2, \ a_{k-2} = a_3, \ \ldots.$$

This proves that if k **is even**, then

$$\sqrt{N} = [a_0, \overline{a_1, \ldots, a_m, a_m, \ldots, a_1, 2a_0}],$$

and if k **is odd**, then

$$\sqrt{N} = [a_0, \overline{a_1, \ldots, a_m, \ldots, a_1, 2a_0}]$$

where the single term a_m is called the **central term**.

Examples 114.

$$\sqrt{7} = [2, \overline{1, 1, 1, 4}]$$
$$\sqrt{14} = [3, \overline{1, 2, 1, 6}]$$
$$\sqrt{21} = [4, \overline{1, 1, 2, 1, 1, 8}]$$
$$\sqrt{29} = [5, \overline{2, 1, 1, 2, 10}]$$

5.8. Back to sums of squares

Suppose that the continued fraction expansion of $\sqrt{N}$ has no central term. In other words, it is of the form

$$[a_0, \overline{a_1, \ldots, a_m, a_m, \ldots, a_1, 2a_0}].$$

Let us write it formally as a finite continued fraction,

$$\sqrt{N} = [a_0, a_1, \ldots, a_m, \alpha],$$

where α is necessarily irrational and has continued fraction expansion

$$\alpha = [\overline{a_m, \ldots, a_1, 2a_0, a_1, \ldots, a_m}].$$

Observe the symmetry: not only is α purely periodic, but it is also a palindrome with a central term. Gathering what we have learned so far, α must be a reducible quadratic irrational, hence is of the form

$$\alpha = \frac{a + \sqrt{N}}{b}$$

for some integers $a, b \in \mathbf{Z}$, $b \neq 0$, and its algebraic conjugate, then

$$\overline{\alpha} = \frac{a - \sqrt{N}}{b},$$

must satisfy $\overline{\alpha} = -\frac{1}{\alpha}$

Exercise 115. *Why does it need to be of the form* $\frac{a+\sqrt{N}}{b}$*?*

Hence

$$\overline{\alpha} = \frac{a + \sqrt{N}}{b}\frac{a - \sqrt{N}}{b} = \frac{a^2 - N}{b^2} = -1,$$

which is equivalent to

$$N = a^2 + b^2,$$

that is N can be written as the sum of two (explicitly computable) squares. Of course, this construction is conditional on the continued fraction expansion of $\sqrt{N}$ having no central part. Recall from the previous two sections that

$$\sqrt{N} \text{ has no central part} \iff x^2 - Ny^2 = -1 \text{ is soluble.}$$

Hence we have just proven that

THEOREM 35. *If $x^2 - Ny^2 = -1$ is soluble, then N can be written as a sum of (explicitly computable) squares, $N = a^2 + b^2$.*

We know that every prime $p \equiv 1 \pmod 4$ can be represented as a sum of two squares. Thanks to the following proof of Legendre, we now have an explicit way of constructing this representation. (Skip the proof to see numerical examples.)

THEOREM 36 (Legendre). *If $p \equiv 1 \pmod 4$, then $x^2 - py^2 = -1$ is soluble.*

PROOF. Let (a, b) be the smallest solution to Pell's equation $x^2 - py^2 = +1$. Then

$$a^2 - pb^2 \equiv a^2 - b^2 \equiv 1 \pmod 4. \tag{5.6}$$

Recall that a square mod 4 can only satisfy $a^2 \equiv 0, 1 \pmod 4$ (depending on whether a is even or odd). Hence (5.6) implies that a must be odd and b must be even. Write $b = 2k$. Then

$$pk^2 = \frac{(a-1)}{2}\frac{(a+1)}{2}.$$

We leave the following general fact as exercise: Let $n, a, b \in \mathbf{N}$. If $n^2 = ab$ and $(a, b) = 1$, then $a = u^2$ and $b = v^2$ for some $u, v \in \mathbf{Z}$.

Since

$$\frac{a+1}{2} - \frac{a-1}{2} = 1, \tag{5.7}$$

the two terms are coprime, and hence by the general fact cited above, one of two things is true:

(1) either $\frac{a-1}{2} = pu^2$ and $\frac{a+1}{2} = v^2$,
(2) or $\frac{a-1}{2} = u^2$ and $\frac{a+1}{2} = pv^2$.

Let's consider case (1) first. Equation (5.7) implies that

$$\frac{a+1}{2} - \frac{a-1}{2} = v^2 - pu^2 = 1.$$

That is, (v, u) is a solution to Pell's equation. Note that $a^2 = pb^2 + 1 > 1$ and hence

$$v = \sqrt{\frac{a+1}{2}} < \sqrt{\frac{2a}{2}} = \sqrt{a} < a.$$

This is absurd since a is the minimal x-solution to Pell's equation $x^2 - py^2 = 1$. Hence case (1) can not occur.

Instead, in case (2),

$$\frac{a+1}{2} - \frac{a-1}{2} = pv^2 - u^2 = 1,$$

and this is equivalent to $u^2 - pv^2 = -1$. We have thus found a solution to the negative Pell equation. □

Example 116. *Take $p = 677$. This is very easy to represent as a sum of two squares if you observe that $26^2 = 676$, but let's pretend we didn't notice this. The continued fraction expansion of $\sqrt{677}$ is $[26, \overline{52}]$. Hence the purely periodic palindrome part is here simply $\alpha = [\overline{52}]$, which we need to write down as a quadratic irrational. The quadratic equation α satisfies is*

$$\alpha = 52 + \frac{1}{[\overline{52}]} = 52 + \frac{1}{\alpha} \iff \alpha^2 - 52\alpha - 1 = 0.$$

Computing the positive solution to $x^2 - 52x - 1 = 0$, we find that

$$\alpha = 26 + \sqrt{677}.$$

Hence $677 = 26^2 + 1^2$.

Example 117. *Take this time $p = 1009$. This prime is $\equiv 1$ (mod 4); how can we write it as a sum of two squares? We proceed exactly as above:*

$$\sqrt{1009} = [31, \overline{1, 3, 3, 1, 62}],$$

and we find that

$$\alpha = [\overline{3, 1, 62, 1, 3}] = \frac{28 + \sqrt{1009}}{15}.$$

Hence

$$1009 = 28^2 + 15^2.$$

CHAPTER 6

Diophantine approximation

THEOREM 37. *The set* $\mathbf{Q}$ *is* ε*-dense in* $\mathbf{R}$*. That is, for any* $\alpha \in \mathbf{R}$ *and* $\varepsilon > 0$*, there exists* $\frac{a}{b} \in \mathbf{Q}$ *such that*

$$\left|\alpha - \frac{a}{b}\right| < \varepsilon.$$

PROOF. Pick b large enough such that $\frac{1}{b} < \varepsilon$. There exists $a \in \mathbf{Z}$ such that

$$\frac{a}{b} < \alpha < \frac{a+1}{b}.$$

Then

$$\left|\alpha - \frac{a}{b}\right| < \frac{1}{b} < \varepsilon.$$

□

You can see this quite easily knowing continued fractions: let $\frac{p_n}{q_n}$ be the convergents of α, in particular,

$$\alpha = \lim_{n\to\infty} \frac{p_n}{q_n}.$$

By the definition of a limit, for any $\varepsilon > 0$, there exists $N > 0$ such that for all $n \geq N$,

$$\left|\alpha - \frac{p_n}{q_n}\right| < \varepsilon.$$

Diophantine approximation studies the approximation of irrationals by rationals. For instance, given α, what is the best rational approximation of α in $\mathcal{F}_n$ and how good is it ?

6.1. Dirichlet's theorem and best approximants

We denote by $\mathbf{R} \setminus \mathbf{Q}$ the set of all irrational numbers in $\mathbf{R}$. Recall Dirichlet's box principle: given N boxes and $N+1$ objects that are randomly distributed among the boxes, there is one box that will contain at least 2 objects.

THEOREM 38 (Dirichlet's approximation theorem). *Let* $\alpha \in \mathbf{R} \setminus \mathbf{Q}$*,* $N \in \mathbf{N}^*$*. Then there exists* $\frac{a}{b} \in \mathcal{F}_N$ *such that*

$$\left|\alpha - \frac{a}{b}\right| < \frac{1}{bN}.$$

Remark 118. *If* $\frac{a}{b}, \frac{c}{d} \in \mathcal{F}_N$ *are distinct Farey fractions, then*

$$\left|\frac{a}{b} - \frac{c}{d}\right| = \frac{|ad - bc|}{bd} \geq \frac{1}{bd} \geq \frac{1}{bN}.$$

PROOF. Consider the fractional parts

$$0 < \{\alpha\}, \{2\alpha\}, \ldots, \{N\alpha\} < 1,$$

and the intervals

$$[0, \tfrac{1}{N}), [\tfrac{1}{N}, \tfrac{2}{N}), \ldots, [\tfrac{N-1}{N}, 1).$$

One of two things can happen: either one of the fractional parts above, say $\{m\alpha\}$, is in the first subinterval $[0, \frac{1}{N})$, or $[0, \frac{1}{N})$ contains none of the N fractional parts above.

In the first case, we have

$$\{m\alpha\} = m\alpha - \lfloor m\alpha \rfloor < \frac{1}{N}$$

and dividing both sides of this equation: $\alpha - \frac{\lfloor m\alpha \rfloor}{m} < \frac{1}{mN}$. In this case, $b = m$ and $a = \lfloor m\alpha \rfloor$ yields the inequality we wanted.

If instead, $[0, \frac{1}{N})$ contains none of $\{\alpha\}, \{2\alpha\}, \ldots, \{m\alpha\}$, then by Dirichlet's box principle, one of the other subintervals, say $[\frac{k}{N}, \frac{k+1}{N})$, contains $\{m\alpha\}$, $\{n\alpha\}$ with $m \neq n$. Then

$$|\{n\alpha\} - \{m\alpha\}| = |(n-m)\alpha - (\lfloor n\alpha \rfloor - \lfloor m\alpha \rfloor)| < \frac{1}{N}.$$

In this case, take $b = n - m$, $a = \lfloor n\alpha \rfloor - \lfloor m\alpha \rfloor$. □

Corollary 119. *Let $\alpha \in \mathbf{R} \setminus \mathbf{Q}$. Then there exists an infinite sequence $\frac{a_n}{b_n} \in \mathbf{Q}$ of rational approximants such that*

$$\left| \alpha - \frac{a_n}{b_n} \right| < \frac{1}{b_n^2}.$$

PROOF. For each $N \geq 1$, Dirichlet's approximation theorem states that there exists $\frac{a}{b} \in \mathbf{Q}$ such that

$$\left| \alpha - \frac{a}{b} \right| < \frac{1}{bN} \leq \frac{1}{b^2}.$$

□

In the rest of this section, we will examine how the most effective rational approximations are provided by convergents to α. First recall that

Lemma 120.

$$\left| \alpha - \frac{p_n}{q_n} \right| < \frac{1}{q_n q_{n+1}}$$

PROOF. Consider the following formal finite continued fraction expansion for α:

$$\alpha = [a_0, \ldots, a_n, \alpha_{n+1}] = \frac{\alpha_{n+1} p_n + p_{n-1}}{\alpha_{n+1} q_n + q_{n-1}}.$$

Then

$$\begin{aligned} \left| \alpha - \frac{p_n}{q_n} \right| &= \left| \frac{\alpha_{n+1} p_n + p_{n-1}}{\alpha_{n+1} q_n + q_{n-1}} - \frac{p_n}{q_n} \right| = \frac{|p_{n-1} q_n - p_n q_{n-1}|}{(\alpha_{n+1} q_n + q_{n-1}) q_n} \\ &= \frac{1}{(\alpha_{n+1} q_n + q_{n-1}) q_n} < \frac{1}{q_{n+1} q_n} \end{aligned}$$

since $\alpha_{n+1} = a_{n+1} + \{\alpha_{n+1}\} > a_{n+1}$ and $a_{n+1} q_n + q_{n-1} = q_{n+1}$. □

We can use this bound to give effective proofs of the two results we have seen so far.

Proof that Q is ε-dense in R. Choose n large enough such that $\frac{1}{n^2} < \varepsilon$. Using that $q_n \geq n$, we see that

$$\left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q_n q_{n+1}} \leq \frac{1}{n(n+1)} < \frac{1}{n^2} < \varepsilon.$$

□

Proof of Dirichlet's approximation theorem. Choose n such that $q_n \leq N < q_{n+1}$. Then

$$\left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q_n q_{n+1}} < \frac{1}{q_n N}.$$

□

Example 121. *We wish to find a rational number $\frac{a}{b}$ such that $|\pi - \frac{a}{b}| < 0.01$. Consider the first two convergents of π, $\frac{p_1}{q_1} = \frac{22}{7}$, $\frac{p_2}{q_2} = \frac{333}{106}$. Then*

$$\left|\pi - \frac{22}{7}\right| < \frac{1}{742} < 0.01.$$

In fact, the convergents of $\alpha \in \mathbf{R} \setminus \mathbf{Q}$ provide the best possible approximants to α:

THEOREM 39 (Best approximation theorem). *Let $n \geq 2$, $\alpha \in \mathbf{R} \setminus \mathbf{Q}$. Let $(\frac{p_k}{q_k})$ be the convergents to α. Then for any $\frac{a}{b} \neq \frac{p_n}{q_n}$ with $b \leq q_n$,*

$$\left|\alpha - \frac{p_n}{q_n}\right| < \left|\alpha - \frac{a}{b}\right|.$$

PROOF. Suppose first that $0 < b < q_n$. We can find two integers u, v such that

$$\begin{aligned} a &= p_n u + p_{n-1} v \\ b &= q_n u + q_{n-1} v \end{aligned} \tag{6.1}$$

Then

$$b\alpha - a = u(q_n\alpha - p_n) + v(q_{n-1}\alpha - p_{n-1}).$$

Since convergents approximate α from below and above alternatively (see Theorem 32), we know that

$$q_n\alpha - p_n = -(q_{n-1}\alpha - p_{n-1}).$$

On the other hand, it follows from (6.1) and the assumption that $0 < b < q_n$ that if $u > 0$, then $v < 0$ and if $u < 0$, then $v > 0$. Note that if $u = 0$ then $\frac{a}{b} = \frac{p_{n-1}}{q_{n-1}}$ and the inequality to prove holds by Theorem 32. We will suppose henceforth that $u \neq 0$. We conclude that $u(q_n\alpha - p_n)$ and $v(q_{n-1}\alpha - p_{n-1})$ have the *same sign*, therefore:

$$|b\alpha - a| = |u||q_n\alpha - p_n| + |v||q_{n-1}\alpha - p_{n-1}| > |u||q_n\alpha - p_n| \geq |q_n\alpha - p_n|.$$

To conclude, we use that $0 < b < q_n$ and divide both sides of the latter inequality by b:

$$\left|\alpha - \frac{a}{b}\right| < \frac{1}{b}|q_n\alpha - p_n| < \frac{1}{q_n}|q_n\alpha - p_n| = \left|\alpha - \frac{p_n}{q_n}\right|.$$

The remaining case to prove, namely $b = q_n$ is proved using the triangle inequality. Indeed, observe that

$$\left|\alpha - \frac{a}{b}\right| \geq \left|\frac{a}{b} - \frac{p_n}{q_n}\right| - \left|\alpha - \frac{p_n}{q_n}\right| = \frac{|a - p_n|}{q_n} - \left|\alpha - \frac{p_n}{q_n}\right| \geq \frac{1}{q_n} - \left|\alpha - \frac{p_n}{q_n}\right| > \frac{1}{q_n} - \frac{1}{q_n q_{n+1}}.$$

Since $n \geq 2$, $q_{n+1} > q_1 > q_0 = 1$, hence

$$\frac{1}{q_n} - \frac{1}{q_n q_{n+1}} > \frac{1}{q_n} - \frac{1}{2q_n} = \frac{1}{2q_n} > \frac{1}{q_n q_{n+1}} > \left|\alpha - \frac{p_n}{q_n}\right|.$$

□

Example 122. *The first few convergents of* π *are* $3, \frac{22}{7}, \frac{333}{106}, \frac{355}{113}, \frac{103993}{33102}$. *They have, respectively, the same 0,2,4,6,9 digits as* π.

6.2. Hurwitz's theorem

Proposition 123. *Let* $\varphi = \frac{1+\sqrt{5}}{2}$, *and* $0 < c < 1$. *There exist only finitely many* $\frac{a}{b} \in \mathbf{Q}$ *such that*

$$\left|\varphi - \frac{a}{b}\right| < \frac{c}{\sqrt{5}b^2}.$$

Proof. Suppose that $\frac{a}{b}$ satisfies the inequality above. We will show that there is only a finite number of choices for the integer parameter b. Recall that

$$x^2 - x - 1 = (x - \varphi)(x - \overline{\varphi}).$$

Plugging in $\frac{a}{b}$ and taking the absolute value gives us the relation

$$\left|\frac{a^2}{b^2} - \frac{a}{b} - 1\right| = \left|\frac{a}{b} - \varphi\right| \left|\frac{a}{b} - \overline{\varphi}\right| < \frac{c}{\sqrt{5}b^2} \left|\frac{a}{b} - \overline{\varphi}\right|.$$

Multiplying both sides by b^2 and applying the triangle inequality,

$$|a^2 - ab - b^2| < \frac{c}{\sqrt{5}} \left(\left|\frac{a}{b} - \varphi\right| + \underbrace{|\varphi - \overline{\varphi}|}_{=\sqrt{5}}\right) < \frac{c}{\sqrt{5}} \left(\frac{c}{\sqrt{5}b^2} + \sqrt{5}\right)$$

The LHS is a non-zero integer, hence implies the inequality

$$1 < \frac{c^2}{5b^2} + c \iff b^2 < \frac{c^2}{5(1-c)}.$$

Since c is fixed, there are only finitely many integers b that satisfy this inequality. □

Theorem 40 (Hurwitz, 1891). *For any* $\alpha \in \mathbf{R} \setminus \mathbf{Q}$, *there exists an infinite sequence* $\frac{a_n}{b_n} \in \mathbf{Q}$ *such that*

$$\left|\alpha - \frac{a_n}{b_n}\right| < \frac{1}{\sqrt{5}b_n^2}.$$

One can prove this theorem using that among a triplet of successive convergents $(\frac{p_n}{q_n}, \frac{p_{n+1}}{q_{n+1}}, \frac{p_{n+2}}{q_{n+2}})$, at least one satisfies the inequality above. Instead, we will give a *geometric* proof.

PROOF. Fix $\alpha \in \mathbf{R} \setminus \mathbf{Q}$. Let L_α be the vertical line in the plane passing through $(\alpha, 0)$. That is, $L = \{(\alpha, t) : t \in \mathbf{R}\}$. Then L intersects infinitely many of the triangular interstices in the Farey–Ford packing, see Figure 1 in Section 4.2.

Each triangular region is completely determined by its three vertices, and these are the tangency points of neighboring Ford circles. The coordinates of these tangency points are rational (see Exercise 86). For instance, the tangency point of the Ford circles $\mathcal{C}_{\frac{a}{b}}$ and $\mathcal{C}_{\frac{c}{d}}$ is

$$\left(\frac{ab+cd}{b^2+d^2}, \frac{1}{b^2+d^2} \right). \tag{6.2}$$

In particular, if T is a triangular interstice through which L_α passes, then L_α must pass through the *interior* of T, since the x-coordinates of the vertices are rational, but $\alpha \in \mathbf{R} \setminus \mathbf{Q}$.

Let T be the triangular region between the Ford circles $C_1 = \mathcal{C}_{\frac{a}{b}}$, $C_2 = \mathcal{C}_{\frac{c}{d}}$ and $C_3 = \mathcal{C}_{\frac{a+c}{b+d}}$. Its vertices are then

$$A = C_1 \cap C_2 = (a_1, a_2) = \left(\frac{ab+cd}{b^2+d^2}, \frac{1}{b^2+d^2} \right) \tag{6.3}$$

$$B = C_1 \cap C_3 = (b_1, b_2) = \left(\frac{ab+(a+c)(b+d)}{b^2+(b+d)^2}, \frac{1}{b^2+(b+d)^2} \right) \tag{6.4}$$

$$C = C_2 \cap C_3 = (c_1, c_2) = \text{exercise}. \tag{6.5}$$

WLOG, we assume that C_2 is the largest circle, or in other word, that

$$d < b.$$

Using the computations above, one may check that $c_1 > b_1$, $c_1 > a_1$, and

$$b_1 - a_1 = \frac{s^2 - s - 1}{d^2(s^2+1)(2s^2+2s+1)} = \frac{(s-\varphi)(s-\overline{\varphi})}{d^2(s^2+1)(2s^2+2s+1)},$$

where $s := \frac{b}{d}$, and $\varphi = \frac{1+\sqrt{5}}{2}$ is the golden ratio. Since by assumption $s = \frac{b}{d} > 1$, we easily see that

$$b_1 - a_1 = (s - \varphi) \cdot (\text{something positive}).$$

We will consider the following two cases separately:

(1) $b_1 < a_1 \iff s < \varphi \iff b_1 < \alpha < c_1$
(2) $b_1 > a_1 \iff s > \varphi \iff a_1 < \alpha < c_1$

In Case (1), we will approximate α by $\frac{a+c}{b+d}$. Since by assumption, C_2 is larger than C_1, we have that $c_2 > b_2$ and hence that

$$\frac{a+c}{b+d} - b_1 > c_1 - \frac{a+c}{b+d}.$$

(A drawing helps figuring this out.) Then

$$\left| \alpha - \frac{a+c}{b+d} \right| < \left| b_1 - \frac{a+c}{b+d} \right| = \frac{1}{(b+d)^2} \cdot \frac{s(s+1)}{s^2+(s+1)^2} < \frac{1}{\sqrt{5}(b+d)^2}.$$

We leave it to the reader to check the steps of this computation.

In Case (2), the shape of T (draw it!) indicates that the circle C_2 must be rather large, and hence that α must be closed to the fraction $\frac{c}{d}$. Then

$$\frac{c}{d} - \alpha < \frac{c}{d} - a_1 = \frac{c}{d} - \frac{ab+cd}{b^2+d^2} = \frac{s}{d^2(s^2+1)} < \frac{1}{\sqrt{5}d^2}.$$

In fact, for the last inequality, we need to show that $s^2 - \sqrt{5}s + 1 > 0$. Factoring it,

$$s^2 - \sqrt{5}s + 1 = (s - \varphi)(s + \overline{\varphi}).$$

The second factor is always positive, and by assumption (Case (2)), the first factor is as well.

Since L_α intersects infinitely many triangles such as T, and we have just seen that for each such triangle, we can find a fraction $\frac{k}{l}$ such that $\left|\alpha - \frac{k}{l}\right| < \frac{1}{\sqrt{5}l^2}$, the statement is proved. □

CHAPTER 7

Applications and Outlook

7.1. Billards

In our model, our billiard table is a square, with unit area, our ball is a point, and there is no friction. In particular, once the ball rolls, it never slows down. We start with an initial position on the table and an initial direction in which we shoot. We will call this data the initial vector. Once shot, the ball will eventually hit an edge of the square table, at which point it bounces back with angle of reflection equal to angle of incidence. (For simplicity, we ignore here the case where one shoots directly in one of the four corners.) And so on, and so on, *ad nauseam*.

We call the trajectory **periodic** if after some (finite) time, it starts repeating itself exactly. If it never does, we say that the trajectory is **chaotic**.

THEOREM 41. *If the line passing through the initial vector has rational slope, then the trajectory is periodic. If instead the slope is irrational, then the trajectory is chaotic.*

PROOF. "Unfold" the billiard: namely, instead of bouncing back against the "walls" of the table, develop the ball's trajectory into a straight line in the plane, and consider the plane as a union of identical copies of the unit square glued side by side.

The ball's trajectory now has a line equation

$$L: \quad y = ax + b.$$

We may assume that $b < 0, a > 0$. Under this identification, record x_n for which

$$2n + 1 = ax_n + b,$$

for each $n \geq 0$. Set $b_n = \{x_n\} = \{\frac{2n+1-b}{a}\}$ for each $n \geq 0$. The sequence b_n parametrizes exactly the consecutive places of the table's edge $[(0,1),(1,1)]$ where the ball bounces.

Suppose that the slope of L is rational, i.e. $a = \frac{a_1}{a_2} \in \mathbf{Q}$. Then b_n is *periodic*. Indeed, $b_{a_1} = \{\frac{a_2(1-b)}{a_1}\} = b_0$, $b_{a_1+1} = b_1$, etc. Dynamically, this means that the ball keeps bouncing against the top edge at exactly the same points over and over again. The trajectory must then be periodic. Conversely, if (b_n) is periodic, say $b_m = b_n$, then for some $l \in \mathbf{Z}$,

$$\frac{2m+1-b}{a} = \frac{2n+1-b}{a} + l \iff a = \frac{2(m-n)}{l} \in \mathbf{Q}.$$

□

To approximate a chaotic trajectory, approximate the "irrational angle"!

7.2. Public key cryptography – RSA

We describe here the RSA algorithm, named after Rivest, Shamir, Adlemann (1977). This algorithm concerns one-way secure transmission of messages, and uses a **public key**, openly available, and a secret, **private key**.

For the public key, take two primes p, q, large ($\sim 2^{512}$) and chosen at "random"; the security of the algorithm relies of $n = p \cdot q$ being hard to factor. Then choose a number e such that $(e, \varphi(n)) = 1$. Recall that $\varphi(n) = \varphi(pq) = (p-1)(q-1)$ and note that the computation of $\varphi(n)$ relies on the factorization of n. Then find d such that $de \equiv 1 \pmod{\varphi(n)}$.

Definition 124. *The data (n, e) is called the public key. The data (n, d) is called the private key.*

Remark 125.
(1) *Hint: e stands for "encryption", d for "decryption".*
(2) *Even though (n, e) is public, one needs to know the factorization of n to compute d. Hence, (n, d) is indeed "private".*
(3) *The arithmetic condition $(e, \varphi(n)) = 1$ ensures that the equation $de \equiv 1$ (mod $\varphi(n)$) has a* unique *solution.*

To understand how this works, let's say that Bob wants to send Alice the message HELLO. Bob first encodes his message according to some agreed upon protocol. For example, using the ASCII standard, where $H = 072$, $E = 101$, $L = 108$, $O = 111$, so that his message is

$$a = 07210110811.$$

The message needs to be shorter that the public key's modulus n. Hence Bob might need to chop up a in smaller blocks $a_1, a_2, \ldots, a_k$, each $< n$. For additional security, Bob might also use some prescribed permutation on the digits of each block to make sure his message is not too easy to decode. Henceforth, Bob's message is encoded as $a < n$. To encrypt his message, Bob computes

$$b \equiv a^e \pmod{n}$$

using Alice's public key, and then sends Alice the encrypted message b.

Using her private key, Alice will decrypt b by computing

$$x \equiv b^d \pmod{n}.$$

In fact, $de = 1 + k\varphi(n)$ for some integer k and

$$x \equiv b^d \equiv (a^e)^d \equiv a \cdot (a^{\varphi(n)})^k \equiv a \pmod{n}$$

by Euler's theorem. Since $a < n$, $x = a$.

The security of the algorithm therefore relies essentially on the following two problems being *hard*: factoring large numbers and taking large modular roots.

7.3. The distribution of primes

The larger a number n is, the more possible divisors it can have. We thus expect that primes become sparser as we consider sets of large numbers. Fix $x > 0$, and let

$$\pi(x) = \#\{p \leq x : p \text{ is prime}\}$$

count the number of primes below x. E.g. $\pi(10) = 4$, $\pi(1000) = 168$, etc.

THEOREM 42 (Prime Number Theorem (PNT) (1896)). *As $x \to \infty$,*

$$\pi(x) \sim \frac{x}{\log(x)}. \tag{7.1}$$

The asymptotic notation above means the following. Given two functions $f(x)$, $g(x)$, we say that $f(x) \sim g(x)$ if

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1.$$

In words, $f(x)$ and $g(x)$ have the same order of growth. The asymptotic (7.1) was conjectured by Gauss, Legendre, Riemann,... but finally proved, independently, by Hadamard and de la Vallée-Poussin in 1896. The proof uses complex analysis, so instead we will see Chebyshev's proof (1850) of the following weaker statement.

THEOREM 43 (Weak PNT (1850)). *For x sufficiently large,*

$$0.66\frac{x}{\log x} \;<\; \pi(x) \;<\; 1.7\frac{x}{\log x}.$$

Remark 126. *Actually, with some more careful analysis, Chebyshev proved the above with* 0.89 *instead of* 0.66 *and* 1.11 *instead of 1.7, thus proving the PNT correct with a relative error of 11%.*

PROOF. The proof argues by using divisibility properties of binomial coefficients

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

By the fundamental theorem of arithmetic,

$$\binom{n}{k} = \prod_{p\le n} p^m,$$

where p^m is the largest power of the prime p dividing $\binom{n}{k}$.

For the lower bound, we rely on the following lemma

Lemma 127. *Let p be a prime, and p^m the largest power of p that divides $\binom{n}{k}$. Then $p^m \le n$.*

Then

$$\binom{n}{k} \le \prod_{p\le n} n = n^{\pi(n)}.$$

On the other hand, by the binomial theorem,

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

hence $2^n \le (n+1)n^{\pi(n)}$. Since log is monotonic increasing, we have

$$n\log(2) \le \log(n+1) + \pi(n)\log(n).$$

Here, $\log 2 \approx 0.69$, and we can show that for $n > 200$,

$$\pi(n) \geq \log 2 \frac{n}{\log n} - \frac{\log(n+1)}{\log n} > \frac{2}{3}\frac{n}{\log n}.$$

For the upper bound, we will use that $\pi(x) < 1.7\frac{x}{\log x}$ is true for $x < 1200$ and proceed by induction to prove that for $n > 1200$, both $\pi(2n)$ and $\pi(2n+1)$ have the correct upper bound. Again by the binomial theorem,

$$\binom{2n}{n} < \sum_{k=0}^{2n} \binom{2n}{k} = (1+1)^{2n} = 4^n.$$

Each prime $n < p \leq 2n$ appears in the numerator of

$$\binom{2n}{n} = \frac{(n+1)\cdot(n+2)\cdots 2n}{1\cdot 2\cdots n}$$

and never appears in the denominator. Hence each prime $n < p \leq 2n$ divides $\binom{2n}{n}$, and therefore

$$\prod_{n<p\leq 2n} p \mid \binom{2n}{n}.$$

Then

$$n^{\pi(2n)-\pi(n)} \leq \prod_{n<p\leq 2n} p \leq \binom{2n}{n} < 4^n$$

and $(\pi(2n) - \pi(n)) \log n < \log 4n \approx 1.39n$. By our induction hypothesis,

$$\pi(2n) < 1.39\frac{n}{\log n} + \pi(n) < 3.09\frac{n}{\log n}$$

and, as above, we can show for $n > 1200$ that

$$\pi(2n) < 3.09\frac{n}{\log n} < 1.7\frac{2n}{\log(2n)}.$$

For $\pi(2n+1)$, we note that $\pi(2n+1) \leq \pi(2n) + 1$, and similarly,

$$\pi(2n+1) < 3.09\frac{n}{\log n} + 1 < 1.7\frac{2n+1}{\log(2n+1)}$$

for $n > 1200$. □

The PNT suggest that we can use $\frac{x}{\log x}$ to approximate $\pi(x)$. This tells us that the probability for a large number n to be prime is roughly $\frac{1}{\log(n)}$.

Exercise 128. *Explain why the probability that a large number with $2n$ digits is prime is $1/2$ the probability that a large number with n digits is prime.*

In fact, there are better approximations of $\pi(x)$. Observe first for any constant c, $\pi(x) \sim \frac{x}{\log x - c}$ as $x \to \infty$. In turns out that taking $c = 1$ yields a consistently better

approximation of $\pi(x)$ than $c = 0$. For illustration, here are a (very) few values.

x	$\pi(x)$	$\frac{x}{\log x}$	$\frac{x}{\log x - 1}$	$R(x)$
10^3	168	145	169	168
10^5	9′592	8′686	9′512	9′587
10^8	5′761′455	5′428′681	5′740′304	5′761′552

The function on the very right, which gives the better approximation, was introduced by Riemann and can be expressed as

$$R(x) = 1 + \sum_{n\geq 1} \frac{1}{n\zeta(n+1)} \frac{(\log x)^n}{n!},$$

where ζ is the **Riemann ζ-function**

$$\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s} = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \dots.$$

Here $s \in \mathbf{C}$ is a parameter controlling the convergence of the infinite series. If s is a positive integer ≥ 2, then clearly $\zeta(s)$ converges, while if $s = 1$, this is the harmonic series, which diverges. More generally,

$$|\zeta(s)| \leq \sum_{n\geq 1} \frac{1}{|n^s|} = \sum_{n\geq 1} \frac{1}{n^\sigma} < \infty$$

if and only if $\sigma > 1$, where $n^s = n^{\sigma+it} = n^\sigma e^{it\log n}$. Riemann could prove that $\zeta(s)$ has an analytic continuation to all of $s \in \mathbf{C}$, which we also denote $\zeta(s)$. To understand why it is not so surprising to see the Riemann ζ-function appear in relation to primes, one needs only to look at Euler's proof of the infinitude of primes.

Euler's proof that there are infinitely many primes. Let us consider the following formal manipulation using the fundamental theorem of arithmetic:

$$\sum_{n\geq 1} \frac{1}{n} = \sum_{n\geq 1} \frac{1}{p_1^{k_1}\cdots p_l^{k_l}} = \prod_p \sum_{k\geq 0} \frac{1}{p^k} = \prod_p \frac{1}{1-\frac{1}{p}}.$$

If there are only finitely many primes, then this product is necessarily finite, and hence the harmonic series converges, which is absurd. □

Remark 129. *The product expansion above, called an* **Euler product***, also holds for the ζ-function, namely*

$$\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s} = \prod_p (1 - p^{-s})^{-1},$$

whenever $\sigma > 1$. In particular, since ζ can be expressed as an infinite product, the function has no zeros in the part of the plane where $\sigma > 1$.

Instead of $\zeta(s)$, we can consider the **completed ζ-function**

$$\xi(s) = \begin{cases} \frac{1}{2}s(s-1)\pi^{s/2}\Gamma(\frac{s}{2})\zeta(s) & \text{if } \sigma > 0 \\ \xi(1-s) & \text{if } \sigma \leq 0 \end{cases}$$

where the Γ-function is defined by

$$\Gamma(s) = \int_0^\infty e^{-t}t^{s-1}dt$$

for $\sigma > 0$.

Exercise 130. *Show that the Γ-function extends the classical notion of factorial:*

$$\Gamma(n) = (n-1)!.$$

This is a nicer function: the **Riemann ξ-function** is *entire*, which means that it is complex-analytic (and in particular well-defined) for all $s \in \mathbf{C}$. Moreover, it has not only no zeros for $\sigma > 1$, but also by symmetry for $\sigma < 0$. The vertical strip $0 < \sigma < 1$ of the complex plane where ξ can have zeros is called the **critical strip**. In the critical strip, the zeros of ξ are exactly the zeros of ζ.

Although Riemann couldn't prove the PNT, he stated the following **exact formula**

$$\pi(x) = R(x) - \sum_\rho R(x^\rho),$$

where the sum is taken over the set of all zeros of the Riemann ξ-function. This means that the fluctuations of $\pi(x)$ depend on the location of the zeros of ξ. The **Riemann Hypothesis** states that all of them lie exactly on the vertical line passing through $\sigma = 1/2$. In turns, this implies that primes have the nicest distribution one can hope for.

7.4. Gaussian integers

A **Gaussian integer** is a complex number of the form $a+ib$ where $a, b \in \mathbf{Z}$. (Recall here that i is the imaginary unit, namely it is the solution to the equation $x^2 + 1 = 0$, and so $i^2 = -1$.) The set of all Gaussian integers,

$$\mathbf{Z}[i] = \{u = a + ib : a, b \in \mathbf{Z}\},$$

can be thought of as a square grid in the complex plane $\mathbf{C}$. To understand Gaussian integers, let's compare them to integers in $\mathbf{Z}$. An integer a in $\mathbf{Z}$ has "size" $|a|$, this is the natural number that measures it's difference/distance to 0. Similarly, the **norm** of a Gaussian integer,

$$N(a + ib) = a^2 + b^2$$

is the natural number that measures the distance of $a+ib$ to the origin of the complex plane. Also similarly to the absolute value, the norm is multiplicative, i.e. for any complex numbers $u, v \in \mathbf{C}$,

$$N(uv) = N(u)N(v).$$

Exercise 131. *Use polar coordinates to prove the multiplicativity of the norm. Show also that $N(z) \geq 0$ and that $N(z) = 0$ if and only if $z = 0$.*

In $\mathbf{Z}$, we write $a|b$ to say that there exists an integer $m \in \mathbf{Z}$ such that $b = ma$. In $\mathbf{Z}[i]$, we write $u|v$ to say that there exists a Gaussian integer $w \in \mathbf{Z}[i]$ such that $v = wu$. Further,

THEOREM 44. *Let $u, v \in \mathbf{Z}[i]$. There exists $q, r \in \mathbf{Z}[i]$ such that $v = qu + r$ with $0 \leq N(r) < N(u)$.*

PROOF. For each point p of the square grid $\mathbf{Z}[i]$, place a unit square (that is, a square whose sides have length 1) on the complex plane such that its center is at p. In this way, the complex plane is covered by unit squares whose interiors are disjoint and who are connected at their sides. (Draw the picture.) Consider the $\frac{v}{u} \in \mathbf{C}$. Then $\frac{v}{u}$ lies in at least one unit square (two if it is on a side, four if it is on a corner). Let $q \in \mathbf{Z}[i]$ be the Gaussian integer closest to $\frac{v}{u}$. (Note that if $\frac{v}{u}$ is on a side, there might be two choices for q, and if $\frac{v}{u}$ is on a corner, there will be four choices for q.) The distance of $\frac{v}{u}$ to q is maximized if $\frac{v}{u}$ lies on a corner; in that case, by Pythagoras theorem

$$N(\tfrac{v}{u} - q) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

Set $r := v - qu$. Then

$$N(r) = N(v - qu) = N(u(\tfrac{v}{u} - q)) = N(u)N(\tfrac{v}{u} - q) < N(u).$$

□

With the Division Theorem in hand, one can prove all results of Chapter 1 for $\mathbf{Z}[i]$. In particular, any Gaussian integer factors in an 'essentially' unique way as a product of **Gaussian primes**.

Definition 132. *A Gaussian integer $u \in \mathbf{Z}[i]$ is called a Gaussian prime if it can not be written as a product $u = v \cdot w$ of two Gaussian integers $v, w \in \mathbf{Z}[i]$ with $N(v)$, $N(w) < N(u)$.*

Proposition 133. *Let $u \in \mathbf{Z}[i]$. If $N(u)$ is prime, then u is a Gaussian prime.*

PROOF. Suppose instead that u factors as a non-trivial product, $u = vw$, then $N(u) = N(v)N(w)$ also factors as a product of smaller numbers. □

Each prime $p \in \mathbf{Z}$ can also be seen as an element of $\mathbf{Z}[i]$, since $\mathbf{Z} \subset \mathbf{Z}[i]$. Not every prime is a Gaussian prime however. For instance,

$$2 = (1 + i)(1 - i)$$

and if $p \equiv 1 \pmod 4$, then p can be written as a sum of two squares, and

$$p = a^2 + b^2 = (a + ib)(a - ib).$$

However, primes that are $\equiv 3 \pmod 4$ are never sums of squares, and therefore also primes in $\mathbf{Z}[i]$.

Example 134. *Consider $u = 10 + 7i$. Then $N(u) = 100 + 49 = 149$ is prime, hence u is a Gaussian prime.*

Example 135. *Consider* $u = 9 + 7i$. *Then* $N(u) = 130 = 2 \cdot 5 \cdot 13$. *We have seen that* $2 = (1 + i)(1 - i)$ *and in particular,* $N(1 \pm i) = 2$. *Since* 5 *and* $13 \equiv 1$ *(mod 4), these primes also "split" over* $\mathbf{Z}[i]$*:*

$$5 = 2^2 + 1^2 = (2 + i)(2 - i)$$
$$13 = 3^2 + 2^2 = (3 + 2i)(3 - 2i)$$

Hence $N(u) = N(1 \pm i)N(2 \pm i)N(3 \pm 2i)$. *To find the correct factorization of* u*, one needs to figure the correct signs out. We end up seeing that*

$$(1 + i)(2 + i)(3 - 2i) = 9 + 7i = u.$$

Gauss studied the arithmetic of $\mathbf{Z}[i]$ in connection to the study of quartic residues, and stated a quartic reciprocity law for $\mathbf{Z}[i]$.

THEOREM 45 (Quartic reciprocity). *Let* u, v *be distinct Gaussian primes, and* $(\frac{u}{v})_4 = 1$ *if* $x^4 \equiv u$ *(mod* v*) has a solution in the Gaussian integers, and* $= -1, \pm i$ *otherwise. Then*

$$\left(\frac{u}{v}\right)_4 \left(\frac{v}{u}\right)_4 = (-1)^{\frac{N(u)-1}{4}\frac{N(v)-1}{4}}.$$