Number Theory
Lecture Notes

Vahagn Aslanyan[1]

[1] www.math.cmu.edu/~vahagn/

Contents

Preface

Dominus Illuminatio Mea.
Psalm 27:1

These are lecture notes for the Number Theory course taught at CMU in Fall 2017 and Fall 2018. I used several texts when preparing these notes. In particular, most of the material can be found in [Bak12, Gre17, HW80]. The books [Bak12, HW80] go way beyond the material of these notes and the reader is referred to those books for more advanced topics.

Please email me if you notice any mistakes or typos.

Synopsis

Divisibility in the ring of integers, primes, the fundamental theorem of arithmetic. Multiplicative functions, the Möbius inversion formula. Modular arithmetic, Wilson's theorem, Fermat's little theorem, Euler's theorem, the Chinese Remainder Theorem. Primitive roots. Quadratic residues, Gauss's law of quadratic reciprocity. Fermat's theorem on sums of two squares, Lagrange's theorem on sums of four squares. Some classical Diophantine equations. Continued fractions, Pell's equation. Diophantine approximations, Liouville numbers, algebraic and transcendental numbers. Quadratic number fields, Gaussian integers. Chebyshev's theorem, a weak version of the prime number theorem.

Notations

- The sets of natural numbers[1] (positive integers), integers, rationals, reals and complex numbers will be denoted by $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ respectively.

[1] The standard convention is that $0 \in \mathbb{N}$ but in this course it is more convenient to assume 0 is not a natural number.

- For a field (or ring) K the ring of polynomials in indeterminate X with coefficients from K is denoted by $K[X]$, while $K(X)$ is the field of rational functions.
- The degree of a polynomial $f(X)$ is denoted by $\deg(f)$.
- The greatest common divisor of two integers a and b is denoted by $\gcd(a, b)$.
- For a real number x its integral part, denoted $[x]$, is the greatest integer which does not exceed x. The fractional part of x is defined as $\{x\} = x - [x]$.

More notations will be introduced throughout the text.

Chapter 1

Divisibility

1.1 Greatest common divisors

Definition 1.1. For two integers a and b with $a \neq 0$ we say that *a divides b* or *b is divisible by a* and write $a \mid b$, iff there is an integer c such that $b = ca$.

Definition 1.2. The *greatest common divisor* of two integers a and b, denoted $\gcd(a, b)$, is defined as a positive number d which divides a and b and is divisible by every common divisor of a and b.

Proposition 1.3. *The greatest common divisor of any two numbers a and b, which are not simultaneously zero, exists and is unique. It is the biggest among the common divisors of a and b.*

Proof. Denote $d := \min\{ax + by : x, y \in \mathbb{Z},\ ax + by > 0\}$. We claim that $d = \gcd(a, b)$.

If $d' \mid a, b$ then clearly $d' \mid ax + by$ for all integers x, y and hence $d' \mid d$. Further, if $a = dq + r$ for some r with $0 \leq r < d$, then it is easy to see that $r = a - dq = ax_0 + by_0$ for some $a_0, b_0 \in \mathbb{Z}$. Now if $r > 0$ then we get a contradiction to minimality of d.

Uniqueness of d is evident. □

Lemma 1.4. *For any integers a, b there are integers x, y such that*

$$\gcd(a, b) = ax + by.$$

Proof 1. This follows immediately from the proof of the previous result. □

Proof 2. (Euclid's algorithm)

Let $a = bq_0 + r_0$ for some $0 \leq r_0 < b$. If $r_0 = 0$ then $\gcd(a, b) = b$. Otherwise let $b = q_1 r_0 + r_1$ with $0 \leq r_1 < r_0$. Now if $r_1 = 0$ then $\gcd(a, b) =$

r_0. Otherwise continue the process and divide r_0 by r_1 with remainder. In the $(k+2)$-th step we get $r_{k-1} = q_{k+1}r_k + r_{k+1}$ with $0 \leq r_{k+1} < r_k$. The sequence of non-negative integers r_i is strictly decreasing so the process must terminate at some point, that is, $r_{k+1} = 0$ for some k. Then $r_k = \gcd(a,b)$ and going up from the bottom we see that r_k can be written as a linear combination of a and b with integer coefficients. □

Remark 1.5. Euclid's algorithm allows one to compute the numbers x and y for which $ax + by = \gcd(a,b)$.

Definition 1.6. Two integers are called *coprime* or *relatively prime* if their greatest common divisor is 1.

Corollary 1.7. *If a and b are coprime and $a \mid bc$ then $a \mid c$.*

Proof. There are integers x, y such that $ax+by = 1$. This implies $acx+bcy = c$. Now a divides the left hand side and so it must divide the right hand side as well. □

1.2 Linear Diophantine equations

Theorem 1.8. *Let a, b, c be integers with $a, b \neq 0$. Then the linear equation*

$$ax + by = c$$

has an integer solution, that is, a solution $(x, y) \in \mathbb{Z}^2$, if and only if $d := \gcd(a,b) \mid c$. Given one solution (x_0, y_0), all solutions are of the form $x = x_0 + k \cdot \frac{b}{d}$, $y = y_0 - k \cdot \frac{a}{d}$ for $k \in \mathbb{Z}$.

Proof. If the equation has a solution (x_0, y_0) then obviously $d \mid ax_0+by_0 = c$. Conversely, if $c = dl$ then since $d = am + bn$ for some integers m, n, we know that (ml, nl) is a solution.

If (x, y) and (x_0, y_0) are two solutions then $a(x - x_0) + b(y - y_0) = 0$. Denote $u = x - x_0$, $v = y_0 - y$. Then $au = bv$. If $a = da'$, $b = db'$ then $\gcd(a', b') = 1$ and $a'u = b'v$ and so $a' \mid v$, $b' \mid u$. Therefore $v = a'k$, $u = b'k$ for some integer k. The result follows. □

1.3 Primes and irreducibles

Definition 1.9. An integer $p \neq 0, \pm 1$ is called *irreducible* if $a \mid p$ implies $a = \pm 1$ or $a = \pm p$. A number $p \neq 0, \pm 1$ is *prime* if whenever $p \mid ab$, we have $p \mid a$ or $p \mid b$.

Here 1 and -1 are the only *units* of the ring $\mathbb{Z}$, that is, integers with a multiplicative inverse in $\mathbb{Z}$.

Lemma 1.10. *An integer p is prime iff it is irreducible.*

Proof. Suppose p is prime and $p = ab$. But then $p \mid ab$, hence $p \mid a$ or $p \mid b$. Assume the former holds. On the other hand a and b divide p. This shows that $b = \pm 1$ and $a = \pm p$.

Now suppose p is irreducible and $p \mid ab$. If $p \nmid a$ then $\gcd(p, a) = 1$. Therefore $p \mid b$. □

Note that primes and irreducibles may not coincide in other rings.

Lemma 1.11. *Every non-zero integer has a prime divisor.*

Proof. The number $\min\{d : 1 < d,\ d \mid n\}$ is a prime divisor of n. □

1.4 The fundamental theorem of arithmetic

Theorem 1.12 (Fundamental theorem of arithmetic)**.** *Every natural number $n > 1$ can be factored into a product of (positive) primes in a unique way (up to the order of the factors).*

Proof. Firs we prove the existence of a factorisation. Take a prime divisor p_1 of n and write $n = p_1 n_1$. Now, by induction, n_1 is a product of primes and hence so is n.

Now let us prove uniqueness. Suppose $p_1 \cdots p_k = q_1 \cdots q_s$ where all p_i and q_j are prime. Since p_1 divides the right hand side, it must divide one of the factors q_j, say q_1 (after reordering q_j's). However, q_1 is prime and hence $p_1 = q_1$. Thus, $p_2 \cdots p_k = q_2 \cdots q_s$ and we can proceed by induction. This shows in particular that $k = s$ and the two collections of primes $p_1, \dots, p_k$ and $q_1, \dots, q_k$ coincide up to reordering. □

Proposition 1.13 (Euclid)**.** *There are infinitely many primes.*

Proof. Suppose there are only finitely many primes. Denote those by $p_1, \dots, p_k$. Consider the number $N := p_1 \dots p_k + 1$. If q is a prime divisor of N (which exists!) then it must be different from $p_1, \dots, p_k$ which is a contradiction. □

Proof 2. (Euler) Suppose there are finitely many primes, namely $p_1, \dots, p_k$. Then

$$\prod_{i=1}^{k} \frac{1}{1 - p_i^{-1}} = \prod_{i=1}^{k} \sum_{l=0}^{\infty} p_i^{-l} = \sum_{n=1}^{\infty} \frac{1}{n} = \infty,$$

which is a contradiction. Note that the second equality follows from the fundamental theorem. □

1.5 Exercises

1. Show that

   (i) every two consecutive integers are relatively prime,

   (ii) every two consecutive odd integers are relatively prime.

2. Find all integers x and y for which $84x + 120y = 24$.

3. Let p and q be prime numbers. Suppose that the polynomial $x^2 - px + q$ has an integer root. Find all possible values of p and q.

4. Show that a natural number is divisible by 3 iff the sum of its digits is divisible by 3.

5. Prove that $4 \nmid n^2 + 1$ for any integer n.

6. Show that $\gcd(a^m - 1, a^n - 1) = a^{\gcd(m,n)} - 1$.

7. Show that if for a positive integer n the number $2^n - 1$ is prime then n is prime as well.

8. Find all integers a, b for which $a^2b^2 \mid a^4 + a^3b + ab^3 + b^4$.

9. Prove that for every positive integer n there are n consecutive numbers none of which is prime.

10. Show that if for a positive integer n the number $2^n + 1$ is prime then n must be a power of 2.

11. Let $a, b, c, n \in \mathbb{N}$ with $\gcd(a, b) = 1$ and $ab = c^n$. Prove that both a and b must be n-th powers of some positive integers.

12. Prove that there are infinitely many primes of the form $4k + 3,\ k \in \mathbb{N}$.

13. Let a, b be positive integers. Show that if $\gcd(a, b) = 1$ then $\gcd(a + b, a^2 - ab + b^2)$ is either 1 or 3.

14. Show that if p is an odd prime and $\gcd(a, b) = 1$ then

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

Chapter 2

Multiplicative functions

2.1 Basics

Definition 2.1. A function $f : \mathbb{N} \to \mathbb{C}$ is *multiplicative* if $f(1) = 1$ and whenever m, n are coprime natural numbers, we have $f(mn) = f(m)f(n)$.

Lemma 2.2. *If f is multiplicative and $g(n) = \sum_{d|n} f(d)$ then g is also multiplicative.*

Proof. If $\gcd(m, n) = 1$ then any divisor d of mn can be factored into a product ab with $a \mid m$ and $b \mid n$. The numbers a and b are determined uniquely by d, m, n. Therefore

$$\begin{aligned} g(mn) &= \sum_{d|mn} f(d) = \sum_{a,b:a|m,b|n} f(ab) = \sum_{a,b:a|m,b|n} f(a)f(b) \\ &= \sum_{a|m} f(a) \cdot \sum_{b|n} f(b) = g(m)g(n). \end{aligned}$$

□

Example 2.3. If $\tau(n)$ and $\sigma(n)$ are respectively the number and the sum of all positive divisors of a natural number n, then those are multiplicative. Indeed, $\tau(n) = \sum_{d|n} 1$ and $\sigma(n) = \sum_{d|n} d$ and the functions $f(n) = 1$ and $f(n) = n$ are obviously multipliactive.

An important multiplicative function, namely Euler's totient function, will be introduced in the next chapter.

The following is obvious.

Lemma 2.4. *If f is multiplicative and $n = \prod_i p_i^{k_i}$ is the prime factorisation of n then*

$$f(n) = \prod_i f(p_i^{k_i}).$$

Thus, in order to show that two multiplicative functions are identically equal, it suffices to show they agree on prime powers.

2.2 The Möbius inversion formula

Definition 2.5. The Möbius function $\mu(n)$ is defined as

$$\mu(n) = \begin{cases} 0, & \text{if } n \text{ is divisible by a square,} \\ (-1)^k, & \text{if } n \text{ is a product of } k \text{ distinct primes.} \end{cases}$$

It is obvious that μ is multiplicative. Hence, by Lemma 2.2, the function

$$\nu(n) = \sum_{d|n} \mu(d)$$

is multiplicative as well.

Lemma 2.6.

$$\nu(n) = \begin{cases} 0, & \textit{if } n > 1, \\ 1, & \textit{if } n = 1. \end{cases}$$

Proof. Since both the left hand side and the right hand side are multiplicative functions, it suffices to establish the equality for $n = p^k$ for primes p. This is left as an exercise. □

Proposition 2.7 (Möbius Inversion Formula)**.** *If* $f : \mathbb{N} \to \mathbb{C}$ *and* $g(n) = \sum_{d|n} f(d)$ *then*

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

Note that here f does not have to be multiplicative.

Proof. We have

$$\begin{aligned} \sum_{d|n} \mu(d) g(n/d) &= \sum_{d|n} \sum_{d'|\frac{n}{d}} \mu(d) f(d') = \sum_{d'|n} \sum_{d|\frac{n}{d'}} \mu(d) f(d') \\ &= \sum_{d'|n} \left(f(d') \sum_{d|\frac{n}{d'}} \mu(d) \right) = \sum_{d'|n} f(d') \nu(n/d') = f(n). \end{aligned}$$

□

Proposition 2.8. *If*

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

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

Proof. We have

$$\begin{aligned}\sum_{d|n} f(d) &= \sum_{d|n}\sum_{d'|d} \mu(d')g(d/d') = \sum_{d'|n}\sum_{d:d'|d,d|n} \mu(d')g(d/d') \\ &= \sum_{d'|n}\sum_{a|\frac{n}{d'}} \mu(d')g(a) = \sum_{a|n}\left(g(a)\sum_{d'|\frac{n}{a}}\mu(d')\right) = \sum_{a|n} g(a)\nu(n/a) = g(n).\end{aligned}$$

□

2.3 Exercises

1. For a complex number s denote

   $$\sigma_s(n) := \sum_{d|n} d^s.$$

   Prove that σ_s is multiplicative. Notice that $\sigma_0 = \tau,\ \sigma_1 = \sigma$.

2. If n has k distinct prime factors, show that $\sum_{d|n} |\mu(d)| = 2^k$.

3. Prove that

   $$\sum_{d|n} \frac{\mu(d)^2}{\varphi(d)} = \frac{n}{\varphi(n)}.$$

4. Prove that $\tau(n) \leq 2\sqrt{n}$ for all $n \in \mathbb{N}$.

Chapter 3

Modular arithmetic

3.1 Congruences

Definition 3.1. For $a, b, m \in \mathbb{Z}$ with $m \neq 0$ it is said that *a is congruent to b modulo m*, written $a \equiv b \mod m$, if $m \mid b - a$.

This gives rise to residue classes mod m. More precisely for an integer a we denote $[a]_m := \{b \in \mathbb{Z} : a \equiv b \mod m\}$. Obviously $m\mathbb{Z} = \{mk : k \in \mathbb{Z}\}$ is an ideal of the ring of integers and $[a]_m = a + m\mathbb{Z}$ is just its coset with a representative a. Then the quotient ring $\mathbb{Z}/m\mathbb{Z}$ consists of the residue classes modulo m and the operations are defined by

$$\begin{aligned}(a + \mathbb{Z}/m\mathbb{Z}) + (b + \mathbb{Z}/m\mathbb{Z}) &= (a + b) + \mathbb{Z}/m\mathbb{Z},\\ (a + \mathbb{Z}/m\mathbb{Z}) \cdot (b + \mathbb{Z}/m\mathbb{Z}) &= (ab) + \mathbb{Z}/m\mathbb{Z}.\end{aligned}$$

We will often identify $a + m\mathbb{Z}$ with the integer a and by abuse of notation write $\mathbb{Z}/m\mathbb{Z} = \{0, 1, \ldots, m - 1\}$.

Definition 3.2. For a ring R its *multiplicative group $R^\times$* is the set of all invertible elements of R which is a group under multiplication of the ring.

Lemma 3.3. *For $m \neq 0$ we have*

$$(\mathbb{Z}/m\mathbb{Z})^\times = \{a + m\mathbb{Z} : \gcd(a, m) = 1\}.$$

Note that here $\gcd(a, m)$ does not depend on the choice of the representative a.

Proof. Clearly, $a + m\mathbb{Z}$ is invertible iff $ak \equiv 1 \mod m$ for some integer k. Such a k exists iff $\gcd(a, m) = 1$. □

Corollary 3.4. *If p is prime then every non-zero element in $\mathbb{Z}/p\mathbb{Z}$ is invertible. Hence it is a field and is normally denoted by $\mathbb{F}_p$.*

3.2 Wilson's theorem

Theorem 3.5. *If p is prime then $(p-1)! \equiv -1 \mod p$.*

Proof. Pair each element of $\mathbb{F}_p^\times = \{1, 2, \dots, p-1\}$ with its inverse. If $a = a^{-1}$ then $a^2 \equiv 1 \mod p$ and hence $a \equiv \pm 1 \mod p$. Thus, if $a \neq \pm 1$ then its inverse is different from itself. The product of all those numbers and their inverses is clearly 1 modulo p. Adding 1 and -1 to the product we get the desired result. □

3.3 Fermat's little theorem

Theorem 3.6. *If p is prime and $p \nmid a$ then $a^{p-1} \equiv 1 \mod p$.*

Remark 3.7. This is equivalent to the following: $a^p \equiv a \mod p$ for all integers a.

This is a special case of Euler's theorem that we prove later. The standard proof of Fermat's little theorem presented in textbooks is actually a special case of the proof of Euler's theorem. To avoid repetition, we now give a different proof to Fermat's theorem.

Proof. We can assume $a > 0$ and we use induction on a. For $a = 1$ the theorem is obvious. Now since $\binom{p}{k}$ is divisible by p for $0 < k < p$, we deduce that

$$(a+1)^p = \sum_{k=0}^{p} \binom{p}{k} a^k \equiv a^p + 1^p \equiv a + 1 \mod p,$$

where the last equality follows from the induction hypothesis. □

3.4 The Chinese Remainder Theorem

Theorem 3.8 (Chinese Remainder Theorem)**.** *Let $a_i, m_i \in \mathbb{Z},\ i = 1, \dots, n$, with $m_i \neq 0$ and $\gcd(m_i, m_j) = 1$ for $i \neq j$. Then there is an integer x such that $x \equiv a_i \mod m_i$ for all i.*

This follows immediately from the following result which is also known as the Chinese Remainder Theorem.

Theorem 3.9. *Let $m_1, \dots, m_m$ be pairwise coprime and denote $M := \prod_{i=1}^n m_i$. Then*

$$\mathbb{Z}/M\mathbb{Z} \cong \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_n\mathbb{Z}.$$

Proof. Consider the map

$$\begin{aligned}\psi &: \mathbb{Z}/M\mathbb{Z} \to \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_n\mathbb{Z},\\ \psi &: x + M\mathbb{Z} \mapsto (x + m_1\mathbb{Z}, \ldots, x + m_n\mathbb{Z}) .\end{aligned}$$

Observe that $x + M\mathbb{Z} = y + M\mathbb{Z}$ iff $x \equiv y \mod M$ iff $x \equiv y \mod m_i,\ 1 \leq i \leq n$ iff $x + m_i\mathbb{Z} = y + m_i\mathbb{Z},\ 1 \leq i \leq n$ iff $\psi(x + M\mathbb{Z}) = \psi(y + M\mathbb{Z})$. This shows that ψ is well defined and injective. Hence, it is surjective since its domain and range have the same cardinality M. It is now evident that ψ is also an isomorphism. □

Remark 3.10. For the proof of Theorem 3.8 we need only surjectivity of ψ. However, the isomorphism of the rings is an important result and we will use it later.

The above proof is not constructive so we give a second proof for Theorem 3.8.

Second proof of Theorem 3.8. Denote $M_i := \dfrac{M}{m_i}$ where $M = m_1 \cdots m_n$. It is easy to see that $\gcd(M_i, m_i) = 1$. Hence there are integers k_i with $k_iM_i \equiv 1 \mod m_i$ for all i. Now take

$$x = \sum_{i=1}^{n} a_i k_i M_i.$$

Since $M_i \equiv 0 \mod m_j$ for $i \neq j$, we deduce that $x \equiv a_i \mod m_i$. □

Remark 3.11. Compare this to Lagrange interpolation.

Note that if x solves the above system of congruences then y is also a solution if and only if $x \equiv y \mod M$.

3.5 Euler's totient function

Definition 3.12 (Euler's function)**.** For $m > 0$ define $\varphi(m) = \#\{k : 1 \leq k \leq m,\ \gcd(k, m) = 1\}$.

The following result follows from Lemma 3.3.

Proposition 3.13. *The order of the group* $(\mathbb{Z}/m\mathbb{Z})^\times$ *is* $\varphi(m)$.

Lemma 3.14. *Euler's function* φ *is multiplicative. Hence if* $n = p_1^{k_1} \cdots p_l^{k_l}$ *is the prime factorisation of* n *then*

$$\varphi(n) = \prod_i p_i^{k_i - 1}(p_i - 1).$$

Proof. Theorem 3.9 implies that if $\gcd(m,n)=1$ then

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

and the result follows.

For the second part of the theorem notice that if p is prime and $k>0$ then among the numbers $1,2,\dots,p^k$ all multiples of p and only those are not coprime to p^k. There are p^{k-1} many such numbers. Therefore $\varphi(p^k)=p^k-p^{k-1}$. □

Proposition 3.15. *$\sum_{d|n}\varphi(d)=n$.*

Proof 1. For a divisor d of n denote $A_d := \{a : 1\le a\le n : \gcd(n,a)=d\}$. Then $\{1,\dots,n\}$ is the disjoint union of $(A_d)_{d|n}$ and $|A_d|=\varphi(n/d)$. □

Proof 2. Since φ is multiplicative, $\sum_{d|n}\varphi(d)$ is also multiplicative. So it suffices to establish the equality for $n=p^k$ where p is a prime. In this case

$$\sum_{d|n}\varphi(d)=\sum_{i=0}^{k}\varphi(p^i)=1+(p-1)+(p^2-p)+\cdots+(p^k-p^{k-1})=p^k.$$

□

3.6 Euler's theorem

Theorem 3.16 (Euler)**.** *If* $\gcd(a,m)=1$ *then* $a^{\varphi(m)}\equiv 1 \mod m$.

Proof 1. Denote $k:=\varphi(m)$ and let $m_1,\dots,m_k$ be a *reduced residue system* modulo m, that is, $\gcd(m_i,m)=1$ for all i and $m_i\not\equiv m_j \mod m$ for $i\ne j$. In other words, $(\mathbb{Z}/m\mathbb{Z})^\times=\{m_1,\dots,m_k\}$.

Observe that $am_1,\dots,am_k$ is again a reduced residue system mod m and hence it is a permutation of $m_1,\dots,m_k$ mod m. Therefore

$$\prod_i m_i \equiv \prod_i am_i = a^k\prod_i m_i.$$

Since $\gcd\left(\prod_i m_i, m\right)=1$, we can deduce that $a^k\equiv 1 \mod m$. □

Proof 2. Consider the multiplicative group $(\mathbb{Z}/m\mathbb{Z})^\times$ and its subgroup (a) generated by a. If the latter has order (cardinality) n then $a^n=1$. By Lagrange's theorem n divides the order of the group $(\mathbb{Z}/m\mathbb{Z})^\times$ which is $\varphi(m)$. Thus, $a^{\varphi(m)}=1$ in $(\mathbb{Z}/m\mathbb{Z})^\times$ and we are done. □

Corollary 3.17 (Fermat's little theorem)**.** *If p is prime and $p\nmid a$ then $a^{p-1}\equiv 1 \mod p$.*

Proof. When p is prime, $\varphi(p)=p-1$. □

3.7 Exercises

1. Find a positive integer x such that $x \equiv 2 \mod 4,\ 2x \equiv 3 \mod 9,\ 7x \equiv 1 \mod 11$.

2. Suppose $a^{28}+28^b$ is prime for some positive integers a and $b > 1$. Prove that $2 \mid b$ or $29 \mid a$.

3. Let $n \geq 6$ be composite. Show that $n \mid (n-1)!$.

4. For a composite n show that $\varphi(n) \leq n - \sqrt{n}$.

5. Let $p > 2$ be a prime number with $p \equiv 3 \mod 4$. Show that if $p \mid a^2+b^2$ for some integers a and b then $p \mid a$ and $p \mid b$.

6. Find all integers x, y such that $15x^2 - 7y^2 = 9$.

Chapter 4

Primitive roots

4.1 The order of an element

Definition 4.1. For integers $a, m \neq 0$ with $\gcd(a, m) = 1$ the *order* of a mod m is its order in the multiplicative group $(\mathbb{Z}/m\mathbb{Z})^\times$, that is,

$$\operatorname{ord}_m(a) = \min\{\gamma \in \mathbb{N} : a^\gamma \equiv 1 \mod m\}.$$

Lemma 4.2. *If* $a^n \equiv 1 \mod m$ *then* $\operatorname{ord}_m(a) \mid n$. *In particular* $\operatorname{ord}_m(a) \mid \varphi(m)$.

Proof. Denote $k := \operatorname{ord}_m(a)$ and let $n \equiv l \mod k$ with $0 \leq l < k$. Then $a^l \equiv 1 \mod m$ and by minimality of k we must have $l = 0$.

The second part of the lemma follows from the first part and Euler's theorem (or from Lagrange's theorem directly). □

Definition 4.3. If $\operatorname{ord}_m(a) = \varphi(m)$ then a is called a primitive root modulo m.

In other words, a primitive root is a generator of the multiplicative group $(\mathbb{Z}/m\mathbb{Z})^\times$. So there is a primitive root mod m if and only if $(\mathbb{Z}/m\mathbb{Z})^\times$ is cyclic.

We are going to describe all integers that have a primitive root. More generally, we will prove a structure theorem about the multiplicative group $(\mathbb{Z}/m\mathbb{Z})^\times$. But now we recall two classical results from the theory of finite groups.

Lemma 4.4. *If G is a group and $a, b \in G$ with $ab = ba$ and* $\operatorname{ord}(a) = k$, $\operatorname{ord}(b) = l$ *and* $\gcd(k, l) = 1$ *then* $\operatorname{ord}(ab) = lk$.

Proof. Let $\operatorname{ord}(ab) = m$. Since $(ab)^{kl} = (a^k)^l(b^l)^k = 1$, we have $m \mid kl$. On the other hand $(ab)^m = 1$, hence $1 = (ab)^{ml} = a^{ml}$. This implies $k \mid ml$ and so $k \mid m$. Similarly, $l \mid m$ and $kl \mid m$. □

Lemma 4.5. *Let G be a group and $a \in G$ be an element of order m. Then for every integer k the order of a^k is equal to $\frac{m}{\gcd(m,k)}$.*

Proof. Let $\operatorname{ord}(a^k) =: l$. Then $a^{kl} = 1$ and so $m|kl$. Now if $d := \gcd(m,k)$ then $m = dm',\ k = dk'$ with $\gcd(m',k') = 1$. So $m'|k'l$ and $m'|l$. Thus $\frac{m}{d}|l$. On the other hand $(a^k)^{m/d} = a^{mk'} = 1$. □

4.2 Primitive roots modulo a prime

Let p be a prime number. Since $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ is a field, every polynomial equation $f(x) = 0$ hast at most $\deg(f)$ solutions in $\mathbb{F}_p$ where $f(X) \in \mathbb{F}_p[X] \setminus \{0\}$. Alternatively, we could say the congruence $f(x) \equiv 0 \mod p$ has at most $\deg(f)$ integer solutions incongruent mod p where $f(X) \in \mathbb{Z}[X]$ and p does not divide all coefficients of f.

Lemma 4.6. *For $d \mid p-1$ the polynomial $X^d - 1$ has exactly d roots in $\mathbb{F}_p$.*

Proof. Since $d \mid p-1$ there is a polynomial $g(X) \in \mathbb{F}_p[X]$ such that

$$X^{p-1} - 1 = (X^d - 1)g(X).$$

Obviously, $\deg(g) = p - 1 - d$ and hence $g(X)$ has at most $p - 1 - d$ roots. However, $X^{p-1} - 1$ has exactly $p - 1$ roots and therefore $X^d - 1$ has at least $p - 1 - (p - 1 - d) = d$ roots. On the other hand, it cannot have more than d roots and therefore it has exactly d roots. □

Lemma 4.7. *If p, q are primes and $q^k \mid p - 1$ for some $k \geq 1$, then there is $a \in \mathbb{F}_p$ with $\operatorname{ord}_p(a) = q^k$.*

Proof. By the above lemma the equation $x^{q^k} - 1 = 0$ has exactly q^k solutions in $\mathbb{F}_p$. Suppose none of them has order q^k. This means that for each $a \in \mathbb{F}_p$ with $a^{q^k} = 1$ we have $\operatorname{ord}_p(a) \mid q^{k-1}$. Therefore a is a root of $X^{q^{k-1}} - 1$. But this polynomial has only q^{k-1} roots which is a contradiction. □

Theorem 4.8. *For every prime p there is a primitive root mod p. Equivalently, $\mathbb{F}_p^\times$ is cyclic.*

Proof. Let $p - 1 = q_1^{k_1} \cdots q_s^{k_s}$ be the prime factorisation of $p - 1$ and let $a_i \in \mathbb{F}_p^\times$ be of order $q_i^{k_i}$. Then $g = a_1 \cdots a_s$ has order $p - 1$. □

Remark 4.9. This proof actually shows that every finite subgroup of the multiplicative group of a field is cyclic.

Proof 2. For each divisor $d \mid p-1$ let $A_d := \{a \in \mathbb{F}_p : \operatorname{ord}_p(a) = d\}$ and $\psi(d) := |A_d|$. Clearly, $(A_d)_{d|p-1}$ is a partition of $\mathbb{F}_p^\times$. Therefore

$$\sum_{d|p-1} \psi(d) = p - 1.$$

Now suppose $A_d \neq \emptyset$ and let $a \in A_d$. Then $\operatorname{ord}_p(a^k) = \frac{d}{\gcd(d,k)}$ for each k. In particular, $\operatorname{ord}_p(a^k) = d$ iff $\gcd(d,k) = 1$, hence the set $A_d \cap \{1, a, \ldots, a^{d-1}\}$ has $\varphi(d)$ elements. Thus $\psi(d) \geq \varphi(d)$. We claim that the equality holds. Indeed, let $b \in \mathbb{F}_p^\times$ with $\operatorname{ord}_p(b) = d$. Then b is a root of the polynomial $X^d - 1$. However, the latter has exactly d roots and those are $1, a, \ldots, a^{d-1}$. Hence b is actually a power of k which proves our claim.

Thus, for every d either $\psi(d) = 0$ or $\psi(d) = \varphi(d)$. Since

$$\sum_{d|p-1} \psi(d) = p - 1 = \sum_{d|p-1} \varphi(d),$$

$\psi(d) = \varphi(d)$ for every $d \mid p-1$. In particular $\psi(p-1) = \varphi(p-1) > 0$. □

Corollary 4.10. *There are exactly $\varphi(p-1)$ primitive roots modulo p.*

4.3 Primitive roots modulo prime powers

Theorem 4.11. *If p is an odd prime then there is a primitive root mod p^k for every positive integer k.*

Proof. The proof is split into two steps presented in the following claims.

Claim. If g is a primitive root mod p then either g or $g + p$ must be a primitive root mod p^2.

Let g be a primitive root mod p. Denote $k := \operatorname{ord}_{p^2}(g)$. We have $k \mid \varphi(p^2) = p(p-1)$ and $g^k \equiv 1 \mod p^2$. Since $\operatorname{ord}_p(g) = p - 1$, k must be divisible by $p - 1$. Thus $k = p(p-1)$ or $k = p - 1$. In the former case g is a primitive root modulo p. So assume $k = p - 1$, that is, $g^{p-1} \equiv 1 \mod p^2$. Then

$$(g+p)^{p-1} \equiv g^{p-1} + p(p-1)g^{p-2} \equiv 1 + p(p-1)g^{p-2} \not\equiv 1 \mod p^2.$$

As $g+p$ is a primitive root mod p, by the above argument it must be primitive mod p^2.

Claim. If g is a primitive root mod p^2 then it is primitive mod p^n for all $n \geq 2$.

We prove by induction on n that $\operatorname{ord}_{p^n}(g) = p^{n-1}(p-1)$. It is true for $n = 2$. Now assume it is true for n and prove it for $n+1$.

Let $l := \operatorname{ord}_{p^{n+1}}(g)$ so that $l \mid p^n(p-1)$. On the other hand, since $\operatorname{ord}_{p^n}(g) = p^{n-1}(p-1)$, either $l = p^n(p-1)$ or $l = p^{n-1}(p-1)$. In the former case we are done, so assume the latter is the case.

Now $g^{p^{n-2}(p-1)} \equiv 1 \mod p^{n-1}$ and therefore $g^{p^{n-2}(p-1)} = 1 + p^{n-1}t$ for some integer t. Furthermore, $p \nmid t$ as otherwise we would have $g^{p^{n-2}(p-1)} \equiv 1 \mod p^n$ which contradicts the induction hypothesis. Thus,

$$g^{p^{n-1}(p-1)} = (1 + p^{n-1}t)^p \equiv 1 + p^n t \not\equiv 1 \mod p^{n+1}.$$

Note that here we used the fact that $\binom{p}{i}$ is divisible by p for $0 < i < p$ and that $p^{n+1} \mid p^{2n-1}$ for $n \geq 2$. For the last term, where the binomial coefficient is 1, we actually have $p^{n+1} \mid p^{p(n-1)}$ for $p > 2$ and $n \geq 2$. □

There is no primitive root mod 2^k for $k > 2$ according to the following result the proof of which is left to the reader as an exercise. Actually, to show that $(\mathbb{Z}/2^k\mathbb{Z})^\times$ is not cyclic it suffices to notice that it has a non-cyclic subgroup, namely $(\mathbb{Z}/8\mathbb{Z})^\times$.

Lemma 4.12. *For $k > 2$ and odd a we have $a^{2^{k-2}} \equiv 1 \mod 2^k$.*

Theorem 4.13. *There is a primitive root mod m iff $m = 2, 4, p^k, 2p^k$ where p is an odd prime and k is a positive integer.*

Proof. Clearly, 1 is primitive mod 2 and -1 is primitive mod 4.

Now let $m = p_1^{k_1} \cdots p_s^{k_s}$ be the prime factorisation of m. Then

$$(\mathbb{Z}/m\mathbb{Z})^\times \cong (\mathbb{Z}/p_1^{k_1}\mathbb{Z})^\times \times \cdots \times (\mathbb{Z}/p_s^{k_s}\mathbb{Z})^\times,$$

which is cyclic iff m is not divisible by 8 and the numbers $\varphi(p_1^{k_1}), \ldots, \varphi(p_s^{k_s})$ are pairwise coprime. Since $\varphi(p^k)$ is even for an odd p, the result follows.[1] □

4.4 The structure of $(\mathbb{Z}/2^k\mathbb{Z})^\times$

Now we give a characterisation of the group $(\mathbb{Z}/2^k\mathbb{Z})^\times$ for $k > 2$ which will allow us to understand the structure of any group $(\mathbb{Z}/m\mathbb{Z})^\times$ using the Chinese Remainder Theorem.

[1] We can also show directly that if g is a primitive root mod p^k then either g (if it is odd) or $g + p^k$ (if g is even) is a primitive root mod $2p^k$.

Notation. For a positive integer n denote by C_n the cyclic group of order n (which is unique up to isomorphism).

Proposition 4.14. *For* $k > 2$ *we have* $(\mathbb{Z}/2^k\mathbb{Z})^\times \cong C_2 \times C_{2^{k-2}}$.

Definition 4.15. For a prime number p, the *p-adic valuation* of a non-zero integer n, denoted $v_p(n)$, is the biggest integer γ for which p^γ divides n.

Proof of Proposition 4.14. We claim that $\operatorname{ord}_{2^k}(5) = 2^{k-2}$. By Lemma 4.12 it suffices to prove that $5^{2^{k-3}} \not\equiv 1 \mod 2^k$. We have

$$5^{2^{k-3}} = (1+2^2)^{2^{k-3}} = 1 + 2^{k-1} + \sum_{i=2}^{2^{k-3}} \binom{2^{k-3}}{i} 2^{2i}.$$

So it is enough to show that $2^k \mid \sum_{i=2}^{2^{k-3}} \binom{2^{k-3}}{i} 2^{2i}$. For this we need to prove that $v_2\left(\binom{2^{k-3}}{i}\right) \geq k - 2i$ for $i \geq 2$. Observe that for $m < n$

$$\binom{n}{m} = \frac{n}{m}\binom{n-1}{m-1}$$

hence

$$v_2\left(\binom{2^{k-3}}{i}\right) \geq k - 3 - v_2(i) \geq k - 2i,$$

as clearly $v_2(i) < i$. This proves our claim.

Now let $(5) = \{1, 5, 5^2, \ldots, 5^{2^{k-2}-1}\}$ be the subgroup of $(\mathbb{Z}/2^k\mathbb{Z})^\times$ generated by 5. We claim that none of those elements is congruent to -1 mod 2^k. Indeed, suppose $5^l \equiv -1 \mod 2^k$ for some $l \in \{0, 1, \ldots, 2^{k-2} - 1\}$. Then $5^{2l} \equiv 1 \mod 2^k$ and so $2^{k-2} \mid 2l$. Obviously $l \neq 0$ as otherwise we would have $1 \equiv -1 \mod 2^k$ which is wrong. Thus, $l = 2^{k-3}$. However, we saw above that

$$5^{2^{k-3}} \equiv 1 + 2^{k-1} \not\equiv -1 \mod 2^k.$$

Thus, the subgroup $\{\pm 1\} \cdot (5) \leq (\mathbb{Z}/2^k\mathbb{Z})^\times$ has the same cardinality as the whole group $(\mathbb{Z}/2^k\mathbb{Z})^\times$. Hence $(\mathbb{Z}/2^k\mathbb{Z})^\times = \{\pm 1\} \cdot (5)$. So the intersection of the subgroups (5) and $\{\pm 1\}$ is the trivial subgroup $\{1\}$ and their product is equal to the whole group $(\mathbb{Z}/2^k\mathbb{Z})^\times$. Therefore

$$(\mathbb{Z}/2^k\mathbb{Z})^\times \cong \{\pm 1\} \times (5).$$

□

Example 4.16. Let us decompose the group $(\mathbb{Z}/360\mathbb{Z})^\times$ into a direct product of cyclic groups. By the Chinese Remainder Theorem we have

$$(\mathbb{Z}/360\mathbb{Z})^\times \cong (\mathbb{Z}/2^3\mathbb{Z})^\times \times (\mathbb{Z}/3^2\mathbb{Z})^\times \times (\mathbb{Z}/5\mathbb{Z})^\times \cong C_2 \times C_2 \times C_6 \times C_4.$$

4.5 Exercises

1. Let $n > 0$ be an integer. Prove that $n \mid \varphi(2^n - 1)$.

2. Show that if g_1 and g_2 are primitive roots modulo an odd prime p, then $g_1 g_2$ is not a primitive root modulo p.

3. Find all positive integers n for which the congruence $a^{25} \equiv a \mod n$ holds for all integers a.

4. Show that a primitive root modulo p^2 is also a primitive root modulo p, where p is an odd prime.

5. Prove that for $k > 2$ and $a \in \mathbb{Z}$, if a is odd, then
$$a^{2^{k-2}} \equiv 1 \mod 2^k.$$

6. Show that if m has a primitive root then it has exactly $\varphi(\varphi(m))$ of them.

7. Let p be an odd prime. Prove that there is a number $1 < g < p$ which is a primitive root modulo p^n for every positive integer n.

Chapter 5

Quadratic residues

5.1 The Legendre symbol and Euler's criterion

Throughout this chapter p is going to be an odd prime unless explicitly stated otherwise.

Definition 5.1. An integer $a \in \mathbb{Z}$ (or its residue class mod p) with $p \nmid a$ is a *quadratic residue* mod p iff there is $x \in \mathbb{Z}$ such that $x^2 \equiv 1 \mod p$, and a *quadratic non-residue* otherwise.

In other words, quadratic residues are the non-zero squares in the field $\mathbb{F}_p^\times$.

Lemma 5.2. *There are exactly $\frac{p-1}{2}$ quadratic residues and $\frac{p-1}{2}$ non-residues.*

Proof. In $\mathbb{F}_p$ $x^2 = y^2$ iff $x = \pm y$. Thus $1^2, 2^2, \dots, \left(\frac{p-1}{2}\right)^2$ are all quadratic residues. □

Definition 5.3 (Legendre symbol)**.** For $a \in \mathbb{Z}$ we define

$$\left(\frac{a}{p}\right) = \begin{cases} 1, & \text{if } a \text{ is a quadratic residue mod } p, \\ -1, & \text{if } a \text{ is a quadratic non-residue mod } p, \\ 0, & \text{if } p \mid a. \end{cases}$$

Proposition 5.4 (Euler's criterion)**.** *If $p \nmid a$ then $a^{\frac{p-1}{2}} \equiv \left(\frac{a}{p}\right) \mod p$.*

Proof. By Fermat's little theorem, all quadratic residues are roots of the polynomial $X^{\frac{p-1}{2}} - 1$ in $\mathbb{F}_p$. There are exactly $\frac{p-1}{2}$ quadratic residues and the polynomial has at most (in fact, exactly) $\frac{p-1}{2}$ roots, hence quadratic residues are all the roots of the above polynomial, that is, if a is a quadratic non-residue then $a^{\frac{p-1}{2}} \neq 1$ in $\mathbb{F}_p$. However, we know that $a^{p-1} = 1$ and since $\mathbb{F}_p$ is a field, $a^{\frac{p-1}{2}} = \pm 1$. Thus, if a is a quadratic non-residue then $a^{\frac{p-1}{2}} = -1$. □

Corollary 5.5. *-1 is a quadratic residue mod p iff $p \equiv 1 \mod 4$.*

Lemma 5.6. *The Legendre symbol has the following properties.*

- *If $a \equiv b \mod p$ then $\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$,*
- $\left(\frac{a^2}{p}\right) = 1$,
- $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$.

Proof. The first two are evident and the third one follows from Euler's criterion. □

Remark 5.7. The third property states that the product of two quadratic residues is a quadratic residue, the product of a quadratic residue and a non-residue is a non-residue, and the product of two quadratic non-residues is a quadratic residue. These could be deduced directly from definitions. Indeed, the first two statements are obvious, while the third one follows from the first two and the pigeonhole principle.

5.2 Gauss's lemma

Proposition 5.8 (Gauss's Lemma)**.** *Let $I \subseteq \mathbb{F}_p^\times$ be such that $\mathbb{F}_p^\times$ is the disjoint union of I and $-I$. Then for $a \in \mathbb{F}_p^\times$ we have $\left(\frac{a}{p}\right) = (-1)^t$ where $t = \#\{i \in I : ai \in -I\}$.*

Proof. Let $I_+ := \{i \in I : ai \in I\}$ and $I_- := \{i \in I : ai \in -I\}$. Then I is the disjoint union of I_+ and I_-.

Now we show that I is also the disjoint union of aI_+ and $-aI_-$. First, if $ai = -aj$ for some $i \in I_+, j \in I_-$ then $i = -j \in I \cap -I$ which is a contradiction. Further, $|aI_+| = |I_+|$ and $|-aI_-| = |I_-|$ therefore $|aI_+ \cup -aI_-| = |I_+| + |I_-| = |I|$. By definition $aI_+ \cup -aI_- \subseteq I$ and so $aI_+ \cup -aI_- = I$ by the pigeonhole principle.

Thus,

$$\prod_{i\in I} i = \prod_{i\in I_+} ai \cdot \prod_{i\in I_-} (-ai) = (-1)^t \cdot a^{\frac{p-1}{2}} \cdot \prod_{i\in I} i$$

and the result follows. □

Note that in applications one always takes $I = \{1, 2, \dots, \frac{p-1}{2}\}$. In particular, it follows immediately from Gauss's lemma that $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$ giving a second proof of this result. Now we give a slightly more sophisticated application.

Proposition 5.9. $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$*, that is,* 2 *is a quadratic residue mod* p *iff* $p \equiv \pm 1 \mod 8$.

Proof. We need to find the number t of all elements $i \in \{1, \dots, \frac{p-1}{2}\}$ for which $2i \in \{\frac{p+1}{2}, \dots, p-1\}$. If $p \equiv 1 \mod 4$ then $\frac{p-1}{4} < i \leq \frac{p-1}{2}$ and so $t = \frac{p-1}{4}$. This is even iff $p \equiv 1 \mod 8$. If $p \equiv -1 \mod 4$ then $\frac{p+1}{4} \leq i \leq \frac{p-1}{2}$ and $t = \frac{p+1}{4}$. This is even iff $p \equiv -1 \mod 8$. □

5.3 The quadratic reciprocity law

We will now formulate and prove one of the most prominent theorems in elementary number theory established by Gauss, known as the law of quadratic reciprocity.

Theorem 5.10 (Quadratic Reciprocity Law)**.** *For distinct odd primes* p, q

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

Proof. Consider the rectangle

$$E := \left\{(x, y) \in \mathbb{Z}^2 : 0 < x < \frac{p}{2},\ 0 < y < \frac{q}{2}\right\}.$$

Denote

$$\begin{aligned} E_1 &:= \left\{(x, y) \in E : qx - py < -\frac{p}{2}\right\}, \\ E_2 &:= \left\{(x, y) \in E : -\frac{p}{2} < qx - py < 0\right\}, \\ E_3 &:= \left\{(x, y) \in E : 0 < qx - py < \frac{q}{2}\right\}, \\ E_4 &:= \left\{(x, y) \in E : \frac{q}{2} < qx - py\right\}, \end{aligned}$$

and $t_i := \#E_i,\ i = 1, 2, 3, 4$. It is clear that E is the disjoint union of $E_1, \dots, E_4$.

By Gauss's lemma we know that $\left(\frac{q}{p}\right) = (-1)^t$ where

$$t = \#\left\{x : 0 < x < \frac{p}{2},\ qx \equiv a \mod p \text{ for some } a \in \left(-\frac{p}{2}, 0\right)\right\}.$$

Given x as above, there is a unique $a \in \left(-\frac{p}{2}, \frac{p}{2}\right)$ with $qx \equiv a \mod p$. Furthermore, there is a unique integer y such that $a = qx - py$. Moreover, if $-\frac{p}{2} < a < 0$ then $y = \frac{qx-a}{p} \in \left(0, \frac{q}{2}\right)$ (since it is an integer). Therefore $t = \#E_2 = t_2$ and $\left(\frac{q}{p}\right) = (-1)^{t_2}$.

Similarly, $\left(\frac{p}{q}\right) = (-1)^{t_3}$. Thus, to prove the theorem we need to show that

$$\frac{(p-1)(q-1)}{4} - (t_2 + t_3)$$

is even. But this number is equal to $\#E - \#E_2 - \#E_3 = t_1 + t_4$. Observe that the map $(x, y) \mapsto \left(\frac{p+1}{2} - x, \frac{q+1}{2} - y\right)$ is a bijection from E_1 onto E_4. Hence $t_1 = t_4$ and $t_1 + t_4$ is indeed even. □

The law of quadratic reciprocity, along with the results on the quadratic nature of -1 and 2, gives an algorithm for determining whether a given number is a quadratic residue modulo a prime.

Example 5.11. Is -6 a quadratic residue modulo $p = 113$?

We have $\left(\frac{-6}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{2}{p}\right)\left(\frac{3}{p}\right)$.

Since $p \equiv 1 \mod 4$, $\left(\frac{-1}{p}\right) = 1$. Also, $p \equiv 1 \mod 8$ and so $\left(\frac{2}{p}\right) = 1$.

By the reciprocity law

$$\left(\frac{3}{113}\right) = (-1)^{\frac{(3-1)(113-1)}{4}}\left(\frac{113}{3}\right) = \left(\frac{113}{3}\right) = \left(\frac{2}{3}\right) = -1.$$

Thus, $\left(\frac{-6}{113}\right) = -1$.

5.4 Composite moduli

So far we have only studied quadratic residues modulo primes. Now we consider the case of composite moduli.

We define a quadratic residue modulo a number m to be an integer a which is coprime to m and for which the congruence $x^2 \equiv a \mod m$ has a solution.[1]

[1] Given a description of coprime quadratic residues one can easily get a description of non-coprime quadratic residues as well.

Proposition 5.12. *Let $m = \prod_i p_i^{k_i}$ be the prime factorisation of m. Then a number a is a quadratic residue mod m iff it is a qaudratic residue mod $p_i^{k_i}$ for each i.*

Proof. Let x_i be an integer with $x_i^2 \equiv a \mod p_i^{k_i}$. By the Chinese Remainder Theorem there is an integer x such that $x \equiv x_i \mod p_i^{k_i}$. Now it is easy to see that $x^2 \equiv a \mod m$. □

Proposition 5.13. *Let p be an odd prime. If a is a quadratic residue modulo p then it is a quadratic residue modulo p^k for every $k > 0$.*

Proof. We induct on k. Assume a is a quadratic residue mod p^k. This means that for some integers x and b we have

$$x^2 = a + p^k b.$$

Let $y := x + p^k z$ where z is yet to be determined. We want the congruence

$$y^2 \equiv a \mod p^{k+1}$$

to hold. This is equivalent to

$$p^{k+1} \mid (2xz - b)p^k.$$

Since $\gcd(2x, p) = 1$, we can find z such that $2xz \equiv b \mod p$. □

Now let us explore quadratic residues modulo powers of 2. All odd numbers are quadratic residues modulo 2, and quadratic residues modulo 4 are exactly the numbers of the form $4n + 1$.

Proposition 5.14. *An odd integer a is a quadratic residue modulo 2^k for $k \geq 3$ iff $a \equiv 1 \mod 8$.*

Proof. If a is a quadratic residue modulo 8 then it must be 1 mod 8.

Now assume $a \equiv 1 \mod 8$ and prove by induction that a is a quadratic residue mod 2^k for $k \geq 3$. To this end we replicate the above argument, only in this case y is taken to be of the form $a + 2^{k-1}b$. We leave the details to the reader. □

5.5 Exercises

1. Show that a primitive root modulo an odd prime is a quadratic nonresidue.

2. Let p be an odd prime. Prove that the product of quadratic residues mod p is $\equiv (-1)^{\frac{p+1}{2}} \mod p$.

3. Does the congruence $x^2 - 20x + 10 \equiv 0 \mod 50893$ have a solution? Note that 50893 is prime.

4. Show that there are infinitely many primes of the form $4k+1,\ k \in \mathbb{N}$.

5. Describe all primes p (in terms of congruences) for which 7 is a quadratic residue.

6. Suppose $p \equiv 1 \mod 4$ is prime and $2p+1$ is prime. Show that 2 is a primitive root mod $2p+1$.

7. Show that there are infinitely many primes of the form $3k+1,\ k \in \mathbb{N}$.

Chapter 6

Representation of integers by some quadratic forms

6.1 Sums of two squares

In this section we are going to prove a result of Fermat describing all natural numbers that can be represented as sums of two squares.

Proposition 6.1. *An odd prime number p can be expressed as a sum of two squares if and only if $p \equiv 1 \mod 4$.*

Proof. If $p = x^2 + y^2$ then obviously $p \equiv 1 \mod 4$.

Now let $p \equiv 1 \mod 4$. Then $\left(\frac{-1}{p}\right) = 1$ and hence there is an integer u such that $u^2 \equiv -1 \mod p$. Consider the numbers

$$au + b;\ 0 \leq a, b \leq [\sqrt{p}].$$

This list contains more than p elements, hence there are $a_1, a_2, b_1, b_2 \in [0, \sqrt{p})$ such that $(a_1, b_1) \neq (a_2, b_2)$ and

$$a_1 u + b_1 \equiv a_2 u + b_2 \mod p.$$

Denote $x := a_1 - a_2,\ y := b_2 - b_1$. Then $x^2 + y^2 \neq 0$ and $xu \equiv y \mod p$. Therefore

$$y^2 \equiv x^2 u^2 \equiv -x^2 \mod p.$$

Thus, $p \mid x^2 + y^2$ and $0 < x^2 + y^2 \leq 2[\sqrt{p}]^2 < 2p$. Hence $x^2 + y^2 = p$. □

Theorem 6.2. *Let*

$$n = 2^l \cdot \prod_i p_i^{r_i} \cdot \prod_j q_j^{s_j}$$

be the prime factorization of n *where* $p_i \equiv 1 \mod 4$ *and* $q_j \equiv 3 \mod 4$ *and* $l, r_i, q_j \geq 0$. *Then* $n = x^2 + y^2$ *for some integers* x *and* y *iff* s_j *is even for all* j.

Proof. First observe that for all a, b, c, d

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

Therefore, if all s_j's are even then n is the sum of two squares (note that $2 = 1^2 + 1^2$).

Conversely, suppose $n = x^2 + y^2$ and $q \mid n$ is a prime divisor with $q \equiv 3 \mod 4$ with $v_q(n) = s$, that is, $q^s \mid n$ and $q^{s+1} \nmid n$. We claim that $q \mid x, y$. Indeed, otherwise let $y' \in \mathbb{Z}$ be such that $yy' \equiv 1 \mod q$. Then since $q \mid x^2 + y^2$, we have $q \mid (xy')^2 + 1$, i.e. $\left(\frac{-1}{q}\right) = 1$ which contradicts $q \equiv 3 \mod 4$.

Thus, $q \mid x, y$ and $x = qx_1, y = qy_1$. Now we deduce $q^{s-2} \mid x_1^2 + y_2^2$. If s were odd, this would lead to a contradiction (after applying the same argument repeatedly). □

6.2 Sums of four squares

Theorem 6.3 (Lagrange). *Every natural number can be expressed as a sum of four squares.*

Proof. Notice that

$$\begin{aligned}&(x^2 + y^2 + z^2 + w^2)(x_1^2 + y_1^2 + z_1^2 + w_1^2) = \\ &(xx_1 + yy_1 + zz_1 + ww_1)^2 + (xy_1 - yx_1 + wz_1 - zw_1)^2 + \\ &(xz_1 - zx_1 + wy_1 - yw_1)^2 + (xw_1 - wx_1 + zy_1 - yz_1)^2. \qquad (2.1)\end{aligned}$$

Thus, it is enough to prove the theorem for prime numbers. Let p be an odd prime. Consider the $p + 1$ numbers

$$x^2, \ -y^2 - 1; \ 0 \leq x, y \leq \frac{p-1}{2}.$$

If $0 \leq x_1, x_2 \leq \frac{p-1}{2}$ and $x_1 \neq x_2$ then $x_1^2 \not\equiv x_2^2 \mod p$. Hence for some x and y we must have $x^2 \equiv -y^2 - 1 \mod p$.

Thus, there are integers $x, y \in \left[0, \frac{p-1}{2}\right]$ such that $p \mid x^2 + y^2 + 1$. So $mp = x^2 + y^2 + 1$ for some positive integer m where $m \leq \frac{1}{p}\left(2\left(\frac{p-1}{2}\right)^2 + 1\right) < p$.

Let l be the smallest positive integer for which there are integers x, y, z, w such that

$$lp = x^2 + y^2 + z^2 + w^2.$$

Such an l exists and $1 \leq l \leq m < p$. We will show that $l = 1$. Assume for contradiction that $l > 1$.

If l is even then an even number of x, y, z, w are odd and, assuming $x \equiv y \mod 2$ and $z \equiv w \mod 2$, we get

$$\frac{l}{2}p = \left(\frac{x+y}{2}\right)^2 + \left(\frac{x-y}{2}\right)^2 + \left(\frac{z+w}{2}\right)^2 + \left(\frac{z-w}{2}\right)^2$$

which contradicts minimality of l.

Therefore l must be odd. Let $x_1 \in \left(-\frac{l}{2}, \frac{l}{2}\right)$ be such that $x \equiv x_1 \mod l$. Define y_1, z_1, w_1 similarly. Then

$$n := x_1^2 + y_1^2 + z_1^2 + w_1^2 \equiv 0 \mod l.$$

Moreover, as $l \nmid p$ we must have $0 < n < 4 \cdot (l/2)^2 = l^2$. Hence $n = kl$ with $0 < k < l$.

Now

$$(lp)(kl) = (x^2 + y^2 + z^2 + w^2)(x_1^2 + y_1^2 + z_1^2 + w_1^2) = A^2 + B^2 + C^2 + D^2,$$

where A, B, C, D are determined by the identity (2.1). In particular, it is clear that $B, C, D \equiv 0 \mod l$. Furthermore, $A = xx_1 + yy_1 + zz_1 + ww_1 \equiv x^2 + y^2 + z^2 + w^2 \equiv 0 \mod l$. Thus,

$$kp = \left(\frac{A}{l}\right)^2 + \left(\frac{B}{l}\right)^2 + \left(\frac{C}{l}\right)^2 + \left(\frac{D}{l}\right)^2$$

which contradicts the minimality of l. □

6.3 Exercises

1. Describe (in terms of congruences) all primes p which can be expressed as $p = x^2 + 2y^2$ for some $x, y \in \mathbb{Z}$.

Chapter 7

Diophantine equations

Equations of the form

$$f(x_1, \ldots, x_n) = 0,$$

where $f(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n]$ is a polynomial with integer coefficients, and where we are interested in integer (or rational) solutions, are known as Diophantine equations. We studied linear Diophantine equations in Section 1.2. In this chapter we are going to consider some classical non-linear Diophantine equations.

Note that Hilbert's tenth problem asks to find a general algorithm which, given a Diophantine equation, would decide whether there is an integer solution. It was proven by M. Davis, Y. Matiyasevich, H. Putnam, J. Robinson (the theorem being completed by Matiyasevich in 1970) that such an algorithm does not exist.

7.1 The Pythagorean equation

Theorem 7.1. *All solutions of the equation*

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

are of the form $(d(a^2 - b^2), 2dab, d(a^2 + b^2))$ *or* $(2dab, d(a^2 - b^2), d(a^2 + b^2))$ *where* a, b, c, d *are arbitrary integers (and all those triples are solutions).*

Proof. First, if $\gcd(x, y) = d$ then $d \mid z$ and $x = dx', y = dy', z = dz'$ with $(x')^2 + (y')^2 = (z')^2$. Hence, dividing by d we can assume without loss of generality that $\gcd(x, y) = 1$ which implies that x, y, z are pairwise coprime. A solution to the Pythagorean equation with this property is known as a primitive solution.

Reducing the equation modulo 4 we see that z is odd and exactly one of x, y is even. Since the equation is symmetric with respect to x and y, we can assume that y is even. Furthermore, it suffices to find positive solutions and so we can assume $x, y, z > 0$.

Thus, the theorem follows from the following claim.

Claim. All primitive positive solutions of the Pythagorean equation with y even are given by $x = a^2 - b^2, y = 2ab, z = a^2 + b^2$ where $a, b \in \mathbb{N}$ with $\gcd(a, b) = 1$ and $a > b$.

Let $y = 2w$. We write the equation in the form

$$w^2 = \frac{z-x}{2} \cdot \frac{z+x}{2}.$$

Since $\gcd(\frac{z-x}{2}, \frac{z+x}{2}) = 1$, and the product of two coprime numbers is a square iff each of them is, $\frac{z-x}{2}$ and $\frac{z+x}{2}$ must be squares themselves. Thus, there must be coprime numbers a and b such that

$$\frac{z-x}{2} = b^2, \frac{z+x}{2} = a^2.$$

This yields $z = a^2 + b^2, x = a^2 - b^2, y = 2ab$. Conversely, for any a, b we have

$$(a^2 - b^2)^2 + (2ab)^2 = (a^2 + b^2)^2.$$

□

We will shortly give another proof of this theorem.

7.2 A special case of Fermat's last theorem

Fermat's last theorem states that the equation

$$x^n + y^n = z^n$$

with $n > 2$ does not have any positive integer solutions. This turned out to be an extremely difficult problem solved by A. Wiles in 1995 (after being unsuccessfully attacked by many mathematicians for about 350 years).

It is easy to see that in order to prove Fermat last theorem, it is enough to prove it for $n = 4$ and odd prime values of n. We consider the case $n = 4$ below. It is also interesting to note that the equation for $n = 2$ has infinitely many solutions while for $n > 2$ it does not have any solutions.

Theorem 7.2. *The equation*

$$x^2 + y^4 = z^4 \tag{2.1}$$

has no solutions in positive integers.

Proof. If $\gcd(y, z) = d$ then $d^2 \mid x$ and $(x/d^2)^2 + (y/d)^4 = (z/d)^4$. So we can assume x, y, z are pairwise coprime.

Case 1. First suppose y is even. By Theorem 7.1 there are coprime integers a, b such that

$$x = a^2 - b^2,\ y^2 = 2ab,\ z^2 = a^2 + b^2.$$

Assume a is even and b is odd. From the third equation we deduce that $a = 2uv,\ b = u^2 - v^2$ for some relatively prime integers u, v. Now $(y/2)^2 = uv(u^2 - v^2)$ hence $u = p^2,\ v = q^2,\ u^2 - v^2 = r^2$ for some integers p, q, r (since $u, v, u^2 - v^2$ are mutually coprime). Therefore

$$r^2 + q^4 = p^4.$$

Thus, assuming (x, y, z) is a positive solution to (2.1), we found another positive solution (r, q, p) with $p < z$. Applying the same procedure to this new solution, we will find yet another solution the third coordinate of which is smaller than p. This can be repeated indefinitely which will yield an infinite strictly decreasing sequence of positive integers which is a contradiction.

Case 2. Now suppose x is even. Then by Theorem 7.1 there are coprime integers a, b such that

$$x = 2ab,\ y^2 = a^2 - b^2,\ z^2 = a^2 + b^2.$$

Therefore $(yz)^2 + b^4 = a^4$ and $a < z$. This leads to a contradiction as above. □

Remark 7.3. The method of the proof, invented by Fermat, is known as the method of infinite descent.

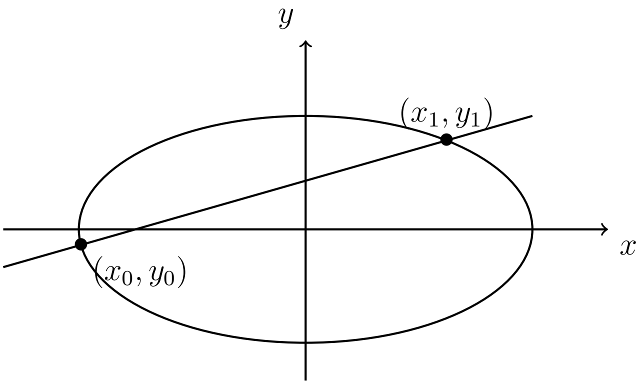
Corollary 7.4. *The equation $x^4 + y^4 = z^4$ has no positive integer solutions.*

7.3 Rational points on quadratic curves

Let $f(X, Y) \in \mathbb{Q}[X, Y]$ be a quadratic polynomial with rational coefficients. We will see in this section that we can find all rational solutions of the equation

$$f(x, y) = 0 \tag{3.2}$$

provided we know one rational solution (x_0, y_0).



Denote $C(\mathbb{R}) := \{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}$[1] and $C(\mathbb{Q}) := C(\mathbb{R}) \cap \mathbb{Q}^2$. Thus $(x_0, y_0) \in C(\mathbb{Q})$. Now pick an arbitrary point $(x_1, y_1) \in C(\mathbb{Q})$ and consider the line l passing through (x_0, y_0) and (x_1, y_1). Its equation is

$$\frac{y - y_0}{x - x_0} = \frac{y_1 - y_0}{x_1 - x_0}$$

and it has rational slope $t = \frac{y_1 - y_0}{x_1 - x_0}$.

Conversely, let l be a line with a rational slope t passing through (x_0, y_0). Its equation is

$$y = t(x - x_0) + y_0.$$

The line l intersects the quadratic curve C in two points one of which is (x_0, y_0). Let $(x_1, y_1) \in \mathbb{R}^2$ be the other point of intersection. It may actually coincide with (x_1, y_1) if l is tangent to C. The abscissa x_1 is a root of the quadratic polynomial $f(X, t(X - x_0) + y_0)$. Since one of the roots of this polynomial, namely x_0, is rational, and its coefficients are rational, the second root must be rational as well. Therefore x_1 is rational and hence $y_1 = t(x_1 - x_0) + y_0$ is rational too.

Thus, there is a one-to-one correspondence between rational points on C and lines with rational slope passing through (x_0, y_0). This gives an algorithm for finding all rational solutions of (3.2). We just need to solve the equation $f(x, t(x - x_0) + y_0)$ for x. Note that since x_0 is a root, the root x_1 will be given by a rational function of t, that is, $x_1 = r(t)$ for some $r(X) \in \mathbb{Q}(X)$. So for each rational value of t the point $(r(t), t(r(t) - x_0) + y_0)$ is a solution of (3.2) and all solutions are of that form. In other words we found a rational parametrisation of the quadratic curve C. One says in this case that C is rational.

Example 7.5. Let us find all rational solutions of the equation

$$x^2 + y^2 = 1.$$

Obviously, $(-1, 0)$ is a solution. So consider the line $y = t(x + 1)$. Solving $x^2 + t^2(x + 1)^2 = 1$ we get $x = -1$ or $x = \frac{1 - t^2}{1 + t^2}$. Thus, all rational points on

[1] One says that $C(\mathbb{R})$ is the set of $\mathbb{R}$-rational points of the curve given by the equation (3.2)

the unit circle are of the form

$$\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right), \ t \in \mathbb{Q}.$$

This can be used to find all integer solutions of the Pythagorean equation

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

As we have seen before, we may assume x, y, z are pairwise coprime. Then since $(x/z)^2 + (y/z)^2 = 1$, we must have

$$\frac{x}{z} = \frac{1-t^2}{1+t^2}, \ \frac{y}{z} = \frac{2t}{1+t^2}.$$

Let $t = \frac{b}{a}$ with $\gcd(a, b) = 1$. Then

$$\frac{x}{z} = \frac{a^2-b^2}{a^2+b^2}, \ \frac{y}{z} = \frac{2ab}{a^2+b^2}.$$

Hence

$$x = a^2 - b^2, \ y = 2ab, \ z = a^2 + b^2.$$

Remark 7.6. Here we implicitly assumed that a and b have different parities in order to deduce the last qualities from the previous ones. We leave it as an exercise for the reader to prove that when both a and b are odd, x, y and z can be written as $x = 2mn, \ y = m^2 - n^2, \ z = m^2 + n^2$ for some integers m, n with $\gcd(m, n) = 1$.

7.4 Exercises

1. Find all integer solutions of the equation $y^2 = x^3 + 16$.

2. Find all pairs (x, y) of positive rational numbers such that $x^2 + 3y^2 = 1$.

3. Find all integers x, y for which $x^4 - 2y^2 = 1$.

Chapter 8

Continued fractions

8.1 Finite continued fractions

Definition 8.1. A *finite continued fraction* is a function $[a_0, a_1, \dots, a_N]$ of variables $a_0, a_1, \dots, a_N$ of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \dots + \cfrac{1}{a_N}}}.$$

We call $a_0, \dots, a_N$ the *partial quotients* of the continued fraction, and $[a_0, \dots, a_n]$, $n \le N$, is called the n-th *convergent* to $[a_0, \dots, a_N]$.

Lemma 8.2. *Define the sequences p_n and q_n recursively by*

- $p_0 = a_0,\ p_1 = a_1 a_0 + 1,\ p_n = a_n p_{n-1} + p_{n-2},\ 2 \le n \le N,$
- $q_0 = 1,\ q_1 = a_1,\ q_n = a_n q_{n-1} + q_{n-2},\ 2 \le n \le N.$

Then $[a_0, \dots, a_n] = \frac{p_n}{q_n}$.

Remark 8.3. Note that p_n and q_n are functions of $a_0, \dots, a_n$ and hence they depend only on these variables.

Proof. We induct on n. For $n = 0, 1$ the equality holds obviously. Now

$$[a_0, \dots, a_n, a_{n+1}] = \left[a_0, \dots, a_{n-1}, a_n + \frac{1}{a_{n+1}}\right] = \frac{\left(a_n + \frac{1}{a_{n+1}}\right) p_{n-1} + p_{n-2}}{\left(a_n + \frac{1}{a_{n+1}}\right) q_{n-1} + q_{n-2}}$$

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

In the second equality we used the induction hypothesis and the above remark. □

Lemma 8.4. *The following identities hold.*

- $p_n q_{n-1} - p_{n-1} q_n = (-1)^{n-1}$ *for* $n > 0$,
- $p_n q_{n-2} - p_{n-2} q_n = (-1)^n a_n$ *for* $n > 1$.

Proof. For the first identity we use induction on n.

$$\begin{aligned} p_n q_{n-1} - p_{n-1} q_n &= (a_n p_{n-1} + p_{n-2}) q_{n-1} - p_{n-1}(a_n q_{n-1} + q_{n-2}) \\ &= -(p_{n-1} q_{n-2} - p_{n-2} q_{n-1}) = -(-1)^{n-2} = (-1)^{n-1}. \end{aligned}$$

The second identity follows easily from the first one. □

Corollary 8.5. *The functions* p_n *and* q_n *satisfy*

- $\dfrac{p_n}{q_n} - \dfrac{p_{n-1}}{q_{n-1}} = \dfrac{(-1)^{n-1}}{q_{n-1} q_n}$,
- $\dfrac{p_n}{q_n} - \dfrac{p_{n-2}}{q_{n-2}} = \dfrac{(-1)^n a_n}{q_{n-2} q_n}$.

Now we assign numerical values to the quotients a_n. We will assume that $a_n \in \mathbb{Z}$ for all n and $a_n > 0$ whenever $n > 0$. Observe that then $\gcd(p_n, q_n) = 1$.

Let $x_n = \dfrac{p_n}{q_n}$ be the n-th convergent.

Lemma 8.6. *The sequence* x_{2n} *is strictly increasing and* x_{2n+1} *is strictly decreasing. Furthermore,* $x_{2n} < x_{2m+1}$ *for every* m, n.

Proof. The first part follows immediately from the second identity of Corollary 8.5 (note that q_n is positive).

For the second part, first botice that $x_{2n} < x_{2n+1}$ according to the first identity of Corollary 8.5. If $n \leq m$ then $x_{2n} \leq x_{2m} < x_{2m+1}$, and if $n \geq m$ then $x_{2n} < x_{2n+1} \leq x_{2m+1}$. □

Thus, if the value of the continued fraction is x then $x_{2n} \leq x$ and $x_{2m+1} \geq x$, and equality holds only once when the index is equal to N.

8.2 Representation of rational numbers by continued fractions

Theorem 8.7. *Every finite continued fraction represents a rational number and every rational number can be represented by a finite continued fraction.*

Proof. It is obvious that finite continued fractions represent rational numbers, since the partial quotients are integers. We show now that every rational number does have such a representation.

Let $x \in \mathbb{Q}$ and denote $x_0 = x$. Define the sequences a_n, x_n by $a_n = [x_n]$ and $x_{n+1} = \frac{1}{x_n - a_n}$. Continue this process as long as $x_n \neq a_n$. Now we show that the process must terminate, that is, $x_k = a_k$ for some k.

We use Euclid's algorithm. If $x = \dfrac{a}{b}$ with $\gcd(a, b) = 1$ then write

$$\begin{aligned} a &= a_0 \cdot b + r_0, \\ b &= a_1 \cdot r_0 + r_1, \\ r_0 &= a_2 \cdot r_1 + r_2, \\ &\cdots \\ r_{N-2} &= a_N \cdot r_{N-1} + 0. \end{aligned}$$

We know Euclid's algorithm terminates. We see that $x_0 = \frac{a}{b}, x_1 = \frac{b}{r_0}, x_n = \frac{r_{n-1}}{r_n}, 2 \leq n < N$. Then obviously $x = [a_0, \ldots, a_N]$. □

If $x = [a_0, \ldots, a_N]$ then also $x = [a_0, \ldots, a_N - 1, 1]$ and thus the continued fraction representation is not unique. But these are the only two representations of x according to the following result.

Proposition 8.8. *If $x = [a_0, \ldots, a_n] = [b_0, \ldots, b_m]$ with $a_n > 1, b_m > 1$ then $m = n$ and $a_i = b_i$ for every i.*

Proof. Note that $[a_1, \ldots, a_n] > 1$ and $x = [a_0, \ldots, a_n] = a_0 + \dfrac{1}{[a_1, \ldots, a_n]}$, therefore $a_0 = [x]$. Similarly, $b_0 = [x] = a_0$ and $[a_1, \ldots, a_n] = [b_1, \ldots, b_m]$. Now the result follows by induction. □

8.3 Infinite continued fractions

Let $(a_n)_{n \geq 0}$ be a sequence of integers with $a_n > 0$ for $n > 0$. Denote

$$x_n := [a_0, \ldots, a_n] = \frac{p_n}{q_n}.$$

Lemma 8.9. *The sequence x_n is convergent.*

Proof. Our results on finite continued fractions show that x_{2n} is strictly increasing, x_{2n+1} is strictly decreasing and $x_{2n} < x_{2m+1}$ for all n, m. Hence the limits

$$x' := \lim_{n\to\infty} x_{2n} \text{ and } x'' := \lim_{n\to\infty} x_{2n+1}$$

exist and $x_{2n} < x' \le x'' < x_{2m+1}$.

Now Corollary 8.5 implies

$$|x' - x''| \le |x_{2n} - x_{2n+1}| = \frac{1}{q_{2n+1}q_{2n}} \to 0.$$

Therefore $x' = x'' =: x$ and $x_n \to x$. □

Definition 8.10. We define the infinite continued fraction $[a_0, a_1, \ldots]$ as the limit x of its convergents $x_n = [a_0, \ldots, a_n]$. One also writes

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ldots}}.$$

Remark 8.11. The proof shows actually that

$$\left|x - \frac{p_n}{q_n}\right| \le |x_{n+1} - x_n| = \frac{1}{q_{n+1}q_n} < \frac{1}{q_n^2}.$$

Proposition 8.12. *Every irrational number can be represented by an infinite continued fraction.*

Proof. Let $x \in \mathbb{R} \setminus \mathbb{Q}$ and denote $u_0 := x$. Define the sequences a_n, u_n by

$$a_n = [u_n],\ u_{n+1} = \frac{1}{u_n - a_n},\ n \ge 0.$$

This process does not terminate, i.e. $u_n \ne a_n$, since otherwise x would be rational.

We claim that $x = [a_0, a_1, \ldots]$. To this end notice that

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

Hence

$$x - \frac{p_n}{q_n} = \frac{p_{n-1}q_n - p_nq_{n-1}}{(u_{n+1}q_n + q_{n-1})q_n} = \frac{(-1)^n}{(u_{n+1}q_n + q_{n-1})q_n} \to 0$$

as $u_n > 1$ whenever $n > 0$. □

The proof of Proposition 8.8 generalises to infinite continued fractions.

Proposition 8.13. *Two infinite continued fractions are equal if and only if all their corresponding partial quotients are equal. Furthermore, an infinite continued fraction cannot be equal to a finite one.*

Corollary 8.14. *The value of an infinite continued fraction is irrational.*

Proof. Indeed, if the value is rational then the continued fraction algorithm terminates. This contradicts uniqueness. □

We sum up the above results in the following theorems.

Theorem 8.15. *Every real number can be represented by a continued fraction (uniquely for irrational numbers). The continued fraction representation of a number is finite iff it is rational.*

Theorem 8.16. *If x is an irrational real number then there are infinitely many rational numbers $\frac{p}{q}$ such that*

$$\left| x - \frac{p}{q} \right| < \frac{1}{q^2}.$$

Proof. Take $\frac{p}{q} = \frac{p_n}{q_n}$. □

This is known as Dirichlet's theorem on diophantine approximations. We will give another proof (independent of continued fractions) and establish stronger results in the next chapter.

Example 8.17. Let us find the continued fraction representation of $\sqrt{2}$. We have

$$\begin{aligned} \sqrt{2} &= 1 + (\sqrt{2} - 1), \\ \frac{1}{\sqrt{2} - 1} &= \sqrt{2} + 1 = 2 + (\sqrt{2} - 1). \end{aligned}$$

Noticing the repeating pattern we conclude that

$$\sqrt{2} = [1, 2, 2, \dots] = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \dots}}.$$

8.4 Periodic continued fractions

Definition 8.18. A continued fraction $[a_0, a_1, a_2, \ldots]$ is called (*ultimately*) *periodic* if there are integers $n_0 \geq 0$ and $l > 0$ such that $a_{n+l} = a_n$ for all $n > n_0$. When $n_0 = 0$, it is called *purely periodic.*

A periodic continued fraction is denoted (using the above notation)

$$[a_0, \ldots, a_{n_0}, \overline{a_{n_0+1}, \ldots, a_{n_0+l}}],$$

and $[\overline{a_{n_0+1}, \ldots, a_{n_0+l}}]$ is called the purely periodic part.

Definition 8.19. A *quadratic irrational* is an irrational root of a quadratic equation with rational coefficients.

In other words a complex number α is a quadratic irrational iff the field extension $\mathbb{Q}(\alpha) \supseteq \mathbb{Q}$ has degree 2.

We will only consider real numbers so by a quadratic irrational we will normally understand a real one. The following is obvious.

Lemma 8.20. *A real number is a quadratic irrational iff it is of the form $\frac{a+b\sqrt{d}}{c}$ for $a, b, c, d \in \mathbb{Z}$ with $d > 0$ non-square and $c \neq 0$.*

Proposition 8.21. *A number $u \in \mathbb{R}$ is a quadratic irrational iff its continued fraction representation is periodic.*

Proof. Let $u = [a_0, \ldots, a_n, \overline{a_{n+1}, \ldots, a_{n+l}}]$. If $\frac{p_m}{q_m}$ is the m-th convergent to u and $v := [\overline{a_{n+1}, \ldots, a_{n+l}}]$ then

$$u = \frac{p_n v + p_{n-1}}{q_n v + q_{n-1}}$$

and hence it suffices to show that y is a quadratic irrational.

Observe that

$$v = [a_{n+1}, \ldots, a_{n+l}, v] = \frac{pv + p'}{qv + q'}$$

for some integers p, p', q, q'. This shows that v is a root of a quadratic equation. It is irrational since its continued fraction representation is infinite.

Conversely, let u be a root of a polynomial $aX^2 + bX + c$ with $a, b, c \in \mathbb{Z}$ with $d = b^2 - 4ac > 0$ non-square. Consider the quadratic form

$$f(X, Y) = aX^2 + bXY + cY^2.$$

If $\frac{p_n}{q_n}$ is the n-th convergent to u then the substitution

$$X = p_n X' + p_{n-1} Y', \; Y = q_n X' + q_{n-1} Y'$$

takes $f(X,Y)$ into a form

$$f_n(X',Y') = a_nX'^2 + b_nX'Y' + c_nY'^2$$

where

$$\begin{aligned}
a_n &= ap_n^2 + bp_nq_n + cq_n^2 = f(p_n,q_n),\\
c_n &= ap_{n-1}^2 + bp_{n-1}q_{n-1} + cq_{n-1}^2 = f(p_{n-1},q_{n-1}) = a_{n-1},\\
b_n &= 2ap_np_{n-1} + 2cq_nq_{n-1} + b(p_nq_{n-1} + p_{n-1}q_n).
\end{aligned}$$

It is easy to see that $b_n^2 - 4a_n^2c_n^2 = d$. This also follows from the fact that the above transformation has determinant $p_nq_{n-1} - p_{n-1}q_n = (-1)^{n-1}$ and hence it does not change the discriminant of the form.

Since $f(u,1) = 0$, we have

$$\frac{a_n}{q_n^2} = a\left(\frac{p_n^2}{q_n^2} - u^2\right) + b\left(\frac{p_n}{q_n} - u\right).$$

We know that $\left|u - \frac{p_n}{q_n}\right| < \frac{1}{q_n^2} < 1$, hence $\left|u + \frac{p_n}{q_n}\right| \le |u| + \left|\frac{p_n}{q_n}\right| < (2|u|+1)$. Therefore

$$\frac{|a_n|}{q_n^2} < |a|\frac{2|u|+1}{q_n^2} + |b|\frac{1}{q_n^2}.$$

Thus, $|a_n| < |a|(2|u|+1) + |b|$ and the right hand side does not depend on n. So we showed that the sequence a_n is bounded, hence $c_n = a_{n-1}$ is bounded. Also, $b_n^2 = 4a_n^2c_n^2 + d$ and b_n is bounded as well.

Now let $u = [a_0, a_1, \ldots]$ and $u_n = [a_n, a_{n+1}, \ldots]$ be the n-th complete quotient of u. Then

$$u = \frac{p_nu_{n+1} + p_{n-1}}{q_nu_{n+1} + q_{n-1}}$$

and

$$f_n(u_{n+1},1) = f(p_nu_{n+1}+p_{n-1}, q_nu_{n+1}+q_{n-1}) = (q_nu_{n+1}+q_{n-1})^2 \cdot f(u,1) = 0.$$

Thus, u_{n+1} is a root of the quadratic polynomial $f_n(X,1)$ which has bounded coefficients. Therefore there are only finitely many possibilities for u_n. So $u_{n_0} = u_{n_0+l}$ for some $n_0 \ge 0$ and $l > 0$. By uniqueness of the continued fraction representation $a_n = a_{n+l}$ for all $n \ge n_0$. □

For a quadratic irrational $x = a + b\sqrt{d}$ where $a, b \in \mathbb{Q}$ and $d \in \mathbb{Z}$ is not a square, its *conjugate* is the number $\bar{x} := a - b\sqrt{d}$. It is the other root of the quadratic polynomial which vanishes at x.

Proposition 8.22. *The continued fraction representation of a quadratic irrational x is purely periodic iff $x > 1$ and $-1 < \bar{x} < 0$.*

Proof. Assume $x > 1$ and $-1 < \bar{x} < 0$. Let $x = [a_0, a_1, \ldots]$ and let $x_n = [a_n, a_{n+1}, \ldots]$ be the n-th complete quotient.

First, $a_0 > 0$ as $x > 1$. Further, we show by induction that $-1 < \bar{x}_n < 0$. Since $x_n = a_n + \frac{1}{x_{n+1}}$, we have $\bar{x}_n = a_n + \frac{1}{\bar{x}_{n+1}}$. Hence $-1 < \bar{x}_{n+1} < 0$ by the induction hypothesis.

Now $-\frac{1}{\bar{x}_{n+1}} = a_n - \bar{x}_n$ and $0 < -\bar{x}_n < 1$, hence $a_n = \left[-\frac{1}{\bar{x}_{n+1}}\right]$.

We know that $x_n = x_{n+l}$ for some n, l. Hence $\frac{1}{\bar{x}_n} = \frac{1}{\bar{x}_{n+l}}$. This implies $a_{n-1} = a_{n+l-1}$ which shows in its turn that $x_{n-1} = x_{n-1+l}$. Repeating this argument we get $x_0 = x_l$ and we are done.

Conversely, if $x = [a_0, a_1, \ldots]$ is purely periodic then $a_0 = a_l \geq 1$ for some $l > 0$. So $x > a_0 \geq 1$.

Moreover,

$$x = \frac{p_n x + p_{n-1}}{q_n x + q_{n-1}}$$

and therefore

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

The roots of the above quadratic equation are x and $\bar{x}$, hence $x \cdot \bar{x} = -\frac{p_{n-1}}{q_n} < 0$. In particular $\bar{x} < 0$ as $x > 0$.

Recall that even convergents of a continued fraction are less than its value while odd convergents are greater. So if n is even, then $\frac{p_{n-1}}{q_n} < \frac{p_n}{q_n} < x$ since p_n is increasing. And if n is odd, then $\frac{p_{n-1}}{q_n} < \frac{p_n}{q_{n-1}} < x$ since q_n is increasing. Therefore $x \cdot \bar{x} > -x$ and $\bar{x} > -1$. □

Example 8.23. Let $d > 0$ be a non-square and $x = \frac{1}{\sqrt{d} - [\sqrt{d}]}$. Then $x > 1$ and $\bar{x} = \frac{1}{-\sqrt{d} - [\sqrt{d}]} \in (-1, 0)$. Thus x has a purely periodic continued fraction representation, $x = [\overline{a_1, \ldots, a_l}]$. Since $\sqrt{d} = [\sqrt{d}] + \frac{1}{x}$, the continued fraction representation of $\sqrt{d}$ is almost purely periodic, i.e.

$$\sqrt{d} = [a_0, \overline{a_1, \ldots, a_l}]$$

where $a_0 = [\sqrt{d}]$.

8.5 Pell's equation

Let $d > 0$ be a non-square positive integer. The equation

$$x^2 - dy^2 = 1 \tag{5.1}$$

is known as Pell's equation. We are going to apply the results of the last section to solve Pell's equation.

First, notice that for all d there is a trivial solution $(\pm 1, 0)$. We will first show that for all d there is a non-trivial solution.

Lemma 8.24. *The equation* (5.1) *has a non-trivial solution.*

Proof. We know that the continued fraction representation of $\sqrt{d}$ is "almost" purely periodic, that is,

$$\sqrt{d} = [a_0, \overline{a_1, \ldots, a_m}].$$

Denote $\theta := [\overline{a_1, \ldots, a_m}]$. Let n be an even multiple of m. Then

$$\sqrt{d} = \frac{\theta p_n + p_{n-1}}{\theta q_n + q_{n-1}}$$

where $\frac{p_k}{q_k}$ is the k-th convergent to $\sqrt{d}$.

On the other hand, $\sqrt{d} = a_0 + \frac{1}{\theta}$. Substituting $\theta = \frac{1}{\sqrt{d} - a_0}$ in the previous equation we get

$$\sqrt{d} = \frac{p_n + p_{n-1}(\sqrt{d} - a_0)}{q_n + q_{n-1}(\sqrt{d} - a_0)}.$$

This yields

$$q_n - a_0 q_{n-1} = p_{n-1}, \; p_n - a_0 p_{n-1} = q_{n-1} d$$

and therefore

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

since n is even.

Thus, for every even multiple n of m the pair (p_{n-1}, q_{n-1}) is a solution to Pell's equation. So there are infinitely many non-trivial solutions. □

Solutions (x, y) to Pell's equation are in one-to-one correspondence with the numbers $z := x + y\sqrt{d}$. Often we will say that z is a solution to Pell's equation. Recalling that the conjugate is defined by $\bar{z} = x - y\sqrt{d}$ we can rewrite (5.1) in the form

$$z \cdot \bar{z} = 1. \tag{5.2}$$

Lemma 8.25. *If $z = x + y\sqrt{d}$ with $z\bar{z} = 1$ then $z > 1$ iff $x > 0, y > 0$.*

Proof. If $x > 0, y > 0$ then $z > x \geq 1$. Conversely, if $z > 1$ then $0 < \bar{z} < 1$, hence $x > 0, -y < 0$. □

Definition 8.26. The number $z_1 = \min\{z \in \mathbb{Z}[\sqrt{d}] : z > 1,\ z\bar{z} = 1\}$ is called the *fundamental solution* of the equation (5.2).

Proposition 8.27. *All solutions of the equation* (5.2) *are given by* $\pm z_1^m$, $m \in \mathbb{Z}$.

Proof. Since $z_1\bar{z}_1 = 1$, $z_1^m \cdot \overline{z_1^m} = (z_1\bar{z}_1)^m = 1$. Now, taking into account that $z_1^{-1} = \bar{z}_1$, it suffices to show that for every solution z with $z > 1$ there is a positive integer m such that $z = z_1^m$. As $z_1 > 1$ there is $m > 0$ such that $z_1^m \leq z < z_1^{m+1}$. If the first inequality is strict then $z' := \frac{z}{z_1^m}$ satisfies $z' \cdot \bar{z'} = 1$ and $1 < z' < z_1$ which contradicts the definition of z_1. This finishes the proof. □

Remark 8.28. The above proposition shows that if $z_1 = x_1 + y_1\sqrt{d}$ then all solutions of Pell's equation are given by

$$x = \pm\frac{(x_1 + y_1\sqrt{d})^n + (x_1 - y_1\sqrt{d})^n}{2},\ y = \pm\frac{(x_1 + y_1\sqrt{d})^n - (x_1 - y_1\sqrt{d})^n}{2\sqrt{d}},$$

where n is a positive integer.

Alternatively, if $x_n, y_n, n > 0$, are determined from the equation $x_n + y_n = (x_1 + y_1\sqrt{d})^n$ then all solutions are given by $(\pm x_n, \pm y_n)$. The trivial solutions correspond to $n = 0$.

Remark 8.29. In the proof of Lemma 8.24 if we choose $n = (m, 2)$, that is, $m = n$ if n is even and $m = 2n$ if n is odd, then (p_{n-1}, q_{n-1}) is the fundamental solution of Pell's equation. Though we will not prove this, it gives an algorithm for solving the equation. Moreover, all solutions of Pell's equation are of the form (p_{n-1}, q_{n-1}) where n is an even multiple of m.

Example 8.30. Consider the equation

$$x^2 - 2y^2 = 1.$$

Since $\sqrt{2} = [1, \overline{2}]$, $(3, 2)$ is the fundamental solution. So all solutions are of the form

$$\left(\pm\frac{(3+2\sqrt{2})^n + (3-2\sqrt{2})^n}{2},\ \pm\frac{(3+2\sqrt{2})^n - (3-2\sqrt{2})^n}{2\sqrt{2}}\right).$$

8.6 Exercises

1. Find the continued fraction representation of the numbers $2, -\frac{23}{6}, \sqrt{5}, \sqrt{7}, \frac{1}{\sqrt{3}}$.

2. Evaluate the continued fraction $[1, 2, 3, \overline{1, 4}]$.

3. Find all integer solutions of the following equations:

   (i) $x^2 - 3y^2 = 1$,

   (ii) $x^2 - 6y^2 = 1$.

4. Prove that the sum of the first n natural numbers is a perfect square for infinitely many n.

5. (i) Show that $x^2 - 7y^2 = -1$ has no integer solutions.

   (ii) Show that if d is divisible by a prime number p with $p \equiv 3 \mod 4$ then the equation $x^2 - dy^2 = -1$ has no integer solutions.

6. Let $n \neq 0$ be an integer. Show that if $x^2 - dy^2 = n$ has an integer solution (d is a non-square) then it has infinitely many integer solutions.

Chapter 9

Diophantine approximations

9.1 Dirichlet's theorem

Theorem 9.1. *Let x be a real number. For every integer $Q > 1$ there are integers p, q with $0 < q < Q$ such that $|qx - p| \leq 1/Q$.*

Proof. Consider the numbers $\{0x\}, \{x\}, \{2x\}, \ldots, \{(Q-1)x\}, 1$.[1] By the pigeonhole principle, we can choose two of these numbers the distance of which is at most $1/Q$. Assume first for some $0 \leq i < j < Q$ we have $|\{ix\} - \{jx\}| \leq 1/Q$. This means

$$|(j-i)x + ([ix] - [jx])| \leq 1/Q$$

and we can choose $q = j - i,\ p = [ix] - [jx]$.

If one of those two numbers whose distance is $\leq 1/Q$ is 1 then we have $|\{ix\} - 1| \leq 1/Q$ for some $0 < i < Q$ and we choose $q = i,\ p = [ix] + 1$. □

Corollary 9.2. *If x is irrational then there are infinitely many fractions p/q such that $|x - p/q| < 1/q^2$.*

Proof. For an arbitrary Q let p, q be such that $|qx - p| \leq 1/Q < 1/q$. Then $|x - p/q| < 1/q^2$ so there is at least one such p/q. Now let $Q > \frac{1}{qx-p}$ be an arbitrary integer and let p', q' be such that

$$|q'x - p'| \leq 1/Q < |qx - p|.$$

This means that $p'/q' \neq p/q$ and $|x - p'/q'| \leq 1/q'Q < 1/q'^2$. □

[1] For a real number r its fractional part, denoted $\{r\}$, is defined by $\{r\} = r - [r]$ where $[r]$ is the integral part of r.

The latter is known as Dirichlet's theorem. We have already proved it in the previous chapter using continued fractions. In the next section we will use the properties of continued fractions to strengthen Dirichlet's theorem.

Remark 9.3. If x is rational then there are finitely many such fractions p/q. Indeed, if $x = a/b$ and $|a/b - p/q| < 1/q^2$ with $p/q \neq a/b$ then $q < b$ so there are only finitely many possibilities for q. Further, for every q there are only finitely many possibilities for p.

9.2 Better approximations

Proposition 9.4. *For an irrational number x there are infinitely many rational numbers p/q such that $|x - \frac{p}{q}| < \frac{1}{2q^2}$.*

Proof. We show that at least one of any two consecutive convergents to (the continued fraction of) x satisfies the desired inequality. Indeed, we have

$$\left|x - \frac{p_n}{q_n}\right| + \left|x - \frac{p_{n+1}}{q_{n+1}}\right| = \left|\frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n}\right| = \frac{1}{q_n q_{n+1}} < \frac{1}{2q_n^2} + \frac{1}{2q_{n+1}^2},$$

where the first equality holds since $x - \frac{p_n}{q_n}$ and $x - \frac{p_{n+1}}{q_{n+1}}$ have opposite signs and the last inequality follows from the obvious fact that for all real numbers $a \neq b$ we have $2ab < a^2 + b^2$.

Now we deduce from the above inequality that either $|x - p_n/q_n| < 1/2q_n^2$ or $|x - p_{n+1}/q_{n+1}| < 1/2q_{n+1}^2$. □

Proposition 9.5. *For an irrational number x there are infinitely many rational numbers p/q such that $|x - \frac{p}{q}| < \frac{1}{\sqrt{5}q^2}$.*

Proof. Here we show that at least one of any three consecutive convergents to x must satisfy the required inequality.

Suppose for some n

$$\left|x - \frac{p_k}{q_k}\right| \geq \frac{1}{\sqrt{5}q_k^2}$$

for $k = n, n+1, n+2$.

Using the proof of the previous proposition we see that

$$\frac{1}{\sqrt{5}q_n^2} + \frac{1}{\sqrt{5}q_{n+1}^2} \leq \frac{1}{q_n q_{n+1}}.$$

Denoting $\lambda = q_{n+1}/q_n$ we get $\lambda + 1/\lambda \leq \sqrt{5}$.

Obviously λ is a rational number, hence the above inequality is strict. Thus, $\lambda^2 - \sqrt{5}\lambda + 1 < 0$ which implies $\lambda < \frac{\sqrt{5}+1}{2}$.

Similarly, for $\mu := q_{n+2}/q_{n+1}$ we have $\mu < \frac{\sqrt{5}+1}{2}$.

However, $q_{n+2} = a_{n+2}q_{n+1} + q_n \geq q_{n+1} + q_n$ hence $\mu \geq 1 + 1/\lambda$. This yields $1 + 1/\lambda < \frac{\sqrt{5}+1}{2}$ and so $\lambda > \frac{2}{\sqrt{5}-1} = \frac{\sqrt{5}+1}{2}$, a contradiction. □

Now we show that $\sqrt{5}$ in the last result is best possible.

Proposition 9.6. *Proposition 9.5 does not remain true if we replace $\sqrt{5}$ by a bigger number.*

Proof. Let $A > \sqrt{5}$. Consider the number $x = \frac{\sqrt{5}-1}{2}$. Assume p/q satisfies $|x - p/q| < 1/Aq^2$.

Write $x = \frac{p}{q} + \frac{\delta}{q^2}$ where $|\delta| < \frac{1}{A} < \frac{1}{\sqrt{5}}$. Then

$$\frac{\delta}{q} - \frac{\sqrt{5}}{2}q = -\frac{q}{2} - p.$$

Taking squares we get

$$\frac{\delta^2}{q^2} - \delta\sqrt{5} = p^2 + pq - q^2.$$

Now $|\delta\sqrt{5}| < \sqrt{5}/A$ which does not depent on q and is less than 1. Since $\delta < 1/\sqrt{5}$, if q is large enough then $\frac{\delta^2}{q^2} - \delta\sqrt{5} \in (-1, 1)$. This means $p^2 + pq - q^2 \in (-1, 1)$ and therefore $p^2 + pq - q^2 = 0$ since it is an integer. This is a contradiction. □

Remark 9.7. It can be shown (though we will not do it) that for an irrational x the convergents p_n/q_n give the best approximations to x among all rational numbers p/q with $1 \leq q \leq q_n$. More precisely, if $n \geq 1,\ 0 < q \leq q_n$ and $p/q \neq p_n/q_n$ then

$$\left| x - \frac{p_n}{q_n} \right| \leq \left| x - \frac{p}{q} \right|$$

and the strict inequality holds for $n > 1$.

9.3 Transcendental numbers and Liouville's theorem

Definition 9.8. A complex number α is called *algebraic* (over $\mathbb{Q}$) if there is a non-zero polynomial $f(X) \in \mathbb{Q}[X]$ such that $f(\alpha) = 0$. If there is no such polynomial, α is called *transcendental.*

Proposition 9.9. *There are countably many algebraic numbers.*

Proof. There are countably many rationals, hence countably many polynomials with rational coefficients. Each of them has finitely many roots, hence there are countably many algebraic numbers. □

Corollary 9.10. *Almost all numbers (complex or real) are transcendental.*

This establishes the existence of transcendental numbers. Nevertheless, proving that a given number, like e or π, is transcendental is much more difficult. Liouville was the first to construct a transcendental number and thus establish their existence (Cantor's set theory (and the above argument) came after his discovery). We will now prove Liouville's result asserting that transcendental numbers have better diophantine approximations than algebraic numbers.

Definition 9.11. A real number ξ is *approximable by rationals to the order* n if there is a constant $c = c(\xi)$, depending only on ξ, and infinitely many rational numbers p/q such that

$$\left|\xi - \frac{p}{q}\right| < \frac{c}{q^n}.$$

We saw for example that irrational numbers are approximable to the order 2 while rationals are not.

Definition 9.12. A real number ξ is a *Liouville number* if for every $n > 0$ there are integers p, q with $q > 1$ such that

$$0 < \left|\xi - \frac{p}{q}\right| < \frac{1}{q^n}.$$

Lemma 9.13. *A real number is Liouville iff it is approximable to the order n for every $n > 0$.*

Proof. Obvious. □

Theorem 9.14. *Let $f(X) \in \mathbb{Z}[X]$ be a polynomial of degree n. If ξ is a real zero of f then it is not approximable to the order $n + 1$.*

Proof. Assume ξ is approximable to the order $n+1$. Then there are a constant c and a rational number $p/q \in (\xi - 1, \xi + 1)$ such that $|\xi - p/q| < c/q^{n+1}$ and $f(p/q) \neq 0$.

Let $f(X) = a_nX^n + \ldots + a_1X + a_0$. The derivative $f'(X)$ is bounded on every bounded set. Let $|f'(x)| < M$ for all $x \in (\xi - 1, \xi + 1)$. Now

$$\left|f\left(\frac{p}{q}\right)\right| = \frac{|a_np^n + a_{n-1}p^{n-1}q + \ldots + a_0q^n|}{q^n} \geq \frac{1}{q^n}$$

since the numerator is a non-zero integer. On the other hand by the mean value theorem we have

$$f\left(\frac{p}{q}\right) = f\left(\frac{p}{q}\right) - f(\xi) = \left(\frac{p}{q} - \xi\right) \cdot f'(x)$$

for some number x between ξ and p/q. Combining what we obtained above we get

$$\frac{c}{q^{n+1}} > \left|\xi - \frac{p}{q}\right| = \frac{|f(p/q)|}{|f'(x)|} > \frac{1}{Mq^n}.$$

This shows that $q < cM$, i.e. there are finitely many possibilities for q, which is a contradiction. □

Corollary 9.15. *Liouville numbers are transcendental.*

Example 9.16. Consider the number

$$\xi := \sum_{n=1}^{\infty} \frac{1}{10^{n!}}.$$

We will show that this is a Liouville number. Fix a positive integer N and let

$$\xi_N := \sum_{n=1}^{N} \frac{1}{10^{n!}} = \frac{p}{q}$$

where $q = 10^{N!}$. Clearly

$$0 < \xi - \frac{p}{q} = \sum_{n=N+1}^{\infty} \frac{1}{10^{n!}} < \frac{2}{10^{(N+1)!}} = \frac{2}{q^{N+1}} < \frac{1}{q^N}.$$

9.4 Transcendence of e

Theorem 9.17. *The number e is transcendental.*

Proof. Consider the integral

$$I(t) := \int_0^t e^{t-x} f(x) dx$$

for $t > 0$ where $f(X) \in \mathbb{R}[X]$ with $\deg(f) = m$. Integrating by parts m times we get

$$I(t) = e^t \sum_{j=0}^{m} f^{(j)}(0) - \sum_{j=0}^{m} f^{(j)}(t).$$

Now suppose e is algebraic, i.e. for some integers $a_0, \ldots, a_n$ with $a_0 \neq 0$

$$a_n e^n + \ldots + a_1 e + a_0 = 0.$$

Let f be the polynomial

$$f(X) = X^{p-1}(X-1)^p \cdots (X-n)^p$$

where p is a prime to be determined later. Also, denote

$$g(X) = X^{p-1}(X+1)^p \cdots (X+n)^p.$$

Obviously $|f(x)| \leq g(x)$ for all real values of x and g is increasing on the positive half-line hence

$$|I(t)| \leq te^t g(t).$$

Further, denote

$$J := \sum_{j=0}^{n} a_j I(j).$$

If $m = \deg(f) = (n+1)p - 1$ then

$$J = -\sum_{j=0}^{m} \sum_{k=0}^{n} a_k f^{(j)}(k).$$

Observe that

- if $1 \leq k \leq n$ and $j < p$, then $f^{(j)}(k) = 0$,
- if $1 \leq k \leq n$ and $j \geq p$, then $f^{(j)}(k)$ is an integer divisible by $p!$,
- $f^{(j)}(0) = 0$ for $j < p - 1$,
- if $j > p - 1$, then $f^{(j)}(0)$ is an integer divisible by $p!$,
- $f^{(p-1)}(0)$ is an integer divisible by $(p-1)!$ but not by p for $p > n$.

Therefore, $J \neq 0$ and $(p-1)!|J$, so $|J| \geq (p-1)!$. On the other hand $g(k) \leq (2n)^m \leq (2n)^{2np}$ for all $1 \leq k \leq n$. Thus

$$|J| \leq \sum_{j=1}^{m} |a_j| \cdot |I(j)| \leq \sum_{j=1}^{m} |a_j| \cdot je^j g(j) \leq (2n)^{2np} \sum_{j=1}^{m} |a_j| je^j.$$

Since $\sum_{j=1}^{m} |a_j| je^j$ does not depend on p, the right hand side can be bounded by c^p where c is a constant independent of p. This implies $(p-1)! < c^p$ which cannot hold if p is big enough. □

9.5 Exercises

1. Find two rational numbers a/b such that $|\sqrt{2} - \frac{a}{b}| < \frac{1}{\sqrt{5}b^2}$.

2. Let $b > 1$ be an integer and $(a_n)_{n \geq 1}$ be a sequence of integers with $1 \leq a_n < b$ for all n. Show that the number

$$\sum_{n=1}^{\infty} \frac{a_n}{b^{n!}}$$

is transcendental.

Chapter 10

Quadratic number fields

Definition 10.1. A field extension K of $\mathbb{Q}$ is *quadratic* if it has degree 2 over $\mathbb{Q}$, that is, if the dimension of K as a $\mathbb{Q}$-vector space is 2.

It is clear that quadratic extensions of $\mathbb{Q}$ are obtained by adjoining a quadratic irrational to $\mathbb{Q}$. Since those are of the form $a\sqrt{d}+b$ for $a, b, d \in \mathbb{Q}$ with d non square, $K = \mathbb{Q}(\sqrt{d})$.

More explicitly, $\mathbb{Q}(\sqrt{d})$ can be seen as a subfield of complex numbers consisting of the elements $a+b\sqrt{d}$ for rational a, b. When $d > 0$ the field can be embedded into the field of real numbers. The subset $\mathbb{Z}[\sqrt{d}] = \{a + b\sqrt{d} : a, b \in \mathbb{Z}\}$ is obviously a subring. We are going to study some properties of those rings. Note however that when $d \equiv 1 \mod 4$ this is not the ring of algebraic integers[1] of $\mathbb{Q}(\sqrt{d})$.

Definition 10.2. The *conjugate* of an element $\alpha = a + b\sqrt{d} \in \mathbb{Q}(\sqrt{d})$ is $\bar{\alpha} = a - b\sqrt{d}$. The *norm* of $\alpha = a+b\sqrt{d}$ is defined as $N(\alpha) = \alpha\bar{\alpha} = a^2 - b^2 d$.

The following is evident.

Lemma 10.3. *The norm is multiplicative, that is, $N(\alpha\beta) = N(\alpha)N(\beta)$ for all $\alpha, \beta \in \mathbb{Q}(\sqrt{d})$.*

10.1 Units

In this section we will describe the multiplicative group of the ring $\mathbb{Z}(\sqrt{d})$. Recall that it consists of the invertible elements (units) of the ring, and is denoted by $\left(\mathbb{Z}[\sqrt{d}]\right)^{\times}$.

[1] An element $\alpha \in \mathbb{Q}(\sqrt{d})$ is an algebraic integer if it is a root of a quadratic polynomial $X^2 + bX + c$ with $b, c \in \mathbb{Z}$.

Lemma 10.4. *An element $\alpha \in \mathbb{Z}[\sqrt{d}]$ is a unit iff $N(\alpha) = \pm 1$.*

Proof. If $N(\alpha) = \pm 1$ then $\alpha \cdot \bar{\alpha} = 1$ or $\alpha \cdot (-\bar{\alpha}) = 1$, hence α is invertible.

Conversely, if α is a unit then there is $\beta \in \mathbb{Z}[\sqrt{d}]$ such that $\alpha\beta = 1$. This yields $N(\alpha)N(\beta) = 1$. Since $N(\alpha)$, $N(\beta)$ are integers, $N(\alpha) = \pm 1$. □

Now let $\alpha = x + y\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$ be a unit. We distinguish three cases.
Case 1. $d < -1$
In this case $x^2 - dy^2 = \pm 1$ has only two integer solutions in x, y, namely, $(1, 0)$ which correspond to $\alpha = \pm 1$. So the only units are ± 1.

Case 2. $d = -1$
In this case the field $\mathbb{Q}(i)$ (where $i^2 = -1$) is known as the *Gaussian field* and the elements of $\mathbb{Z}(i)$, which correspond to lattice points on the complex plane, are called *Gaussian integers.*

The equation $x^2 + y^2 = \pm 1$ has four solutions, $(\pm 1, \pm 1)$. Thus, there are four units in the ring of Gaussian integers, namely, $\pm 1, \pm i$.

Case 2. $d > 0$
This case is non-trivial unlike the case $d < 0$. We need to solve the equation

$$x^2 - dy^2 = \pm 1.$$

We have already studied the equation $x^2 - dy^2 = 1$ in Section 8.5 and seen that it has infinitely many solutions. In particular, the ring $\mathbb{Z}[\sqrt{d}]$ has infinitely many units.

Let $\alpha_0 := \min\left\{\alpha \in \left(\mathbb{Z}[\sqrt{d}]\right)^\times : \alpha > 1\right\}$. This is called the *fundamental unit.* As in Section 8.5 one can show easily that all positive units must be powers of α_0. Moreover, if $N(\alpha_0) = -1$ then all odd powers of α_0 will have norm -1 (those will be all positive solutions of the equation $x^2 - dy^2 = -1$) and all even powers will have norm 1 (those will be all solutions of Pell's equation). If $N(\alpha) = 1$ then all powers will have norm 1 and will be solutions of Pell's equation. In particular, the equation $x^2 - dy^2 = -1$ will have no solutions in this case.

Taking into account negative solutions of the above equations as well, we conclude that

$$\left(\mathbb{Z}[\sqrt{d}]\right)^\times = \{\pm\alpha_0^m : m \in \mathbb{Z}\} = (-1, \alpha_0),$$

i.e. the group of units is generated by -1 and the fundamental unit.

10.2 Gaussian integers

In this section we will study prime elements of the ring $\mathbb{Z}[i]$. First, notice that it is a Euclidean domain with the norm being a Euclidean function.

Lemma 10.5. *If $\alpha, \beta \in \mathbb{Z}[i]$ with $\beta \neq 0$ then there are γ and ρ in $\mathbb{Z}[i]$ such that $\alpha = \beta\gamma + \rho$ and $N(\rho) < N(\beta)$.*

Proof. Let $\frac{\alpha}{\beta} = x + yi \in \mathbb{Q}(i)$. If u and v are the closest integers to x and y respectively then $|x-u|, |y-v| \leq \frac{1}{2}$ and

$$N\left(\frac{\alpha}{\beta} - (u+vi)\right) = (x-u)^2 + (y-v)^2 \leq \frac{1}{2} < 1.$$

So we can take $\gamma = u + vi,\ \rho = \alpha - \beta\gamma$. □

Thus, $\mathbb{Z}[i]$ is a unique factorisation domain[2] and so irreducibles and primes[3] coincide in $\mathbb{Z}[i]$. In order to distinguish between primes of $\mathbb{Z}[i]$ and $\mathbb{Z}$, we will call the latter rational primes.

Two elements $\alpha, \beta \in \mathbb{Z}[i]$ are *associates*, written $\alpha \sim \beta$, if $\alpha|\beta$ and $\beta|\alpha$, i.e. β/α is a unit of $\mathbb{Z}[i]$. Obviously, an associate of an irreducible element is itself irreducible and actually the factorisation into a product of irreducibles is unique up to associates.

Lemma 10.6. *If $N(\alpha)$ is a rational prime then α is prime in $\mathbb{Z}[i]$.*

Proof. If $\alpha = \beta \cdot \gamma$ then $N(\alpha) = N(\beta)N(\gamma)$, hence $N(\beta) = \pm 1$ or $N(\gamma) = \pm 1$. Thus, either β or γ must be a unit. □

Now let $\pi \in \mathbb{Z}[i]$ be prime. Then $\pi | \pi \cdot \bar{\pi} = N(\pi)$. The latter is an integer and so it can be factored into a product of rational primes. Then π must divide one of those rational primes. It cannot divide two distinct rational primes p, q since we know that $px + qy = 1$ for some integers x, y (in other words, if two integers are coprime then they are coprime as Gaussian integers as well). Thus, every Gaussian prime divides a unique rational prime.

Let $\pi = a + bi | p$. Then $N(\pi) | N(p) = p^2$. As $N(\pi)$ is an integer, it must be either ± 1 or $\pm p$ or $\pm p^2$. The first case is impossible since π is a prime and hence not a unit. So there are three possibilities.

[2] It is a general result that unique factorisation holds in Euclidean domains and the proof is basically the same as for $\mathbb{Z}$.

[3] These are defined as in Section 1.3, that is, an element $\pi \neq 0, \pm 1, \pm i$ is *irreducible* if $\pi = \alpha\beta$ implies that either α or β is a unit, and it is a *(Gaussian) prime* if $\pi \mid \alpha\beta$ implies $\pi \mid \alpha$ or $\pi \mid \beta$.

Case 1. $N(\pi) = p$ is an odd rational prime.
Then $p = \pi\bar{\pi} = a^2 + b^2$, hence $p \equiv 1 \mod 4$. Conversely, assume p is a rational prime with $p \equiv 1 \mod 4$. Then $\left(\frac{-1}{p}\right) = 1$ and so $p \mid x^2 + 1$ for some integer x. This means that $p \mid (x+i)(x-i)$. But clearly $\frac{x}{p} \pm \frac{1}{p}i \notin \mathbb{Z}[i]$, hence $p \nmid x \pm i$. So p is not a Gaussian prime, hence it is not irreducible in $\mathbb{Z}[i]$. Thus, $p = \alpha\beta$ for some Gaussian integers α, β which are not units. Then $p^2 = N(p) = N(\alpha)N(\beta)$ and $N(\alpha) = N(\beta) = p$ (since the norms are positive). Thus, $p = \alpha\bar{\alpha}$ and as $N(\alpha) = p$ is a rational prime, α is a Gaussian prime. Moreover, it is clear that $\alpha \nsim \bar{\alpha}$. One says in this case that p *splits* in $\mathbb{Z}[i]$.

Case 2. $N(\pi) = 2$.
We notice that $2 = (1+i)(1-i) = -i(1+i)^2$. Here $1+i$ is a Gaussian prime and $2 \sim (1+i)^2$. One says that 2 *ramifies* in $\mathbb{Z}[i]$.

Case 3. $N(\pi) = p^2$.
Then $N(\pi/p) = 1$, hence $\pi \sim p$ and p is a Gaussian prime. It is said to be *inert* in $\mathbb{Z}[i]$. We already saw that $p = 2$ are $p \equiv 1 \mod 4$ cannot be Gaussian primes, hence in this case $p \equiv 3 \mod 4$.

Thus, we obtain the following characterisation of Gaussian primes.

Theorem 10.7. *A Gaussian integer π is a Gaussian prime if and only if it is of one of the following forms.*

- *$\pi = a + bi$ with $a^2 + b^2 = p$ a rational prime with $p \equiv 1 \mod 4$.*
- *$\pi \sim p$ where $p \equiv 3 \mod 4$ is a rational prime.*
- *$\pi \sim 1 + i$.*

As a consequence we get another proof of Fermat's theorem stating that primes of the form $4k + 1$ can be represented as a sum of two squares. This approach also gives the uniqueness of such a representation.

Proposition 10.8. *If $p \equiv 1 \mod 4$ then $p = a^2 + b^2$ for some integers a, b. Moreover, a and b are unique up to signs, that is, if $p = c^2 + d^2$ then $c = \pm a, d = \pm b$ or $c = \pm b, d = \pm a$.*

Proof. Existence of a, b follows immediately from the above analysis. Uniqueness follows from the fact that $\mathbb{Z}[i]$ is a unique factorisation domain. The details are left to the reader as an exercise. □

10.3 Fermat's little theorem for Gaussian integers

Theorem 10.9. *Let $\pi \in \mathbb{Z}[i]$ be a Gaussian prime which does not divide $\alpha \in \mathbb{Z}[i]$. Then*

$$\alpha^{N(\pi)-1} \equiv 1 \mod \pi.$$

We sketch two proofs below leaving the details to the reader to complete.

Proof 1. We will show that $\alpha^{N(\pi)} \equiv \alpha \mod \pi$. Let $\alpha = x + yi$.

If $N(\pi) = p \equiv 1 \mod 4$ then

$$(x+yi)^p \equiv x^p + y^p i^p \equiv x + yi \mod p.$$

If $\pi = q \equiv 3 \mod 4$ then

$$\alpha^q = (x+yi)^q \equiv x^q + y^q i^q \equiv x - yi \equiv \bar{\alpha} \mod q$$

and so

$$\alpha^{q^2} \equiv \bar{\bar{\alpha}} \equiv \alpha \mod q.$$

□

Proof 2. We will construct a complete residue system modulo π consisting of $N(\pi)$ elements.[4]

If $\pi = q \equiv 3 \mod 4$ is a rational prime then $\{x + yi : 0 \leq x, y < q\}$ is a complete residue system mod π.

If $N(\pi) = p \equiv 1 \mod 4$ then for some integer $u \in \mathbb{Z}$ we have $p \mid u + i$. So the system $\{0, 1, \ldots, p-1\}$ is a complete residue system mod π. □

10.4 Using Gaussian integers to solve Diophantine equations

In this section we demonstrate how Gaussian integers can be applied to solve Diophantine equations.

Let us consider the equation

$$y^2 = x^3 - 1. \tag{4.1}$$

We write it as

$$(y+i)(y-i) = x^3.$$

[4] A set $S \subseteq \mathbb{Z}[i]$ is a complete residue system modulo π if every Gaussian integer is congruent to an element of S mod π and no two elements of S are congruent mod π.

Denote $\delta := \gcd(y+i, y-i)$. Then $\delta \mid 2i$ so either $\delta \sim 2$ or $\delta \sim 1+i$ or $\delta \sim 1$.

Obviously $2 \nmid y+i$ so $\delta \nsim 2$. If $\delta \sim 1+i$ then $1+i$ divides $y+i$, i.e. $(y-1)+(y+1)i$ is divisible by 2, hence y is odd. However, reducing (4.1) mod 4 we see that y cannot be odd.

Thus δ is a unit, that is, $y+i$ and $y-i$ are coprime. But since their product is a prefect cube, they must be associates of perfect cubes. Since all units in $\mathbb{Z}[i]$ are cubes, $y+i$ and $y-i$ must be prefect cubes. Hence $y+i = (a+bi)^3$ and taking conjugates we get $y-i = (a-bi)^3$. This implies $3a^2b - b^3 = 1$ and $a^3 - 3ab^2 = y$. The former equality forces b to be ± 1. Computing the value of a we see that the only possibility is $b = -1, a = 0$. This corresponds to $y = 0, x = 1$.

Thus, the only integer solutions of the equation (4.1) is $(1, 0)$.

10.5 Exercises

1. Prove Proposition 10.8.

2. Factorise the numbers -14 and $5i$ in $\mathbb{Z}[i]$.

3. Show that $2, 3, 1+\sqrt{-5}, 1-\sqrt{-5}$ are irreducible in $\mathbb{Z}[\sqrt{-5}]$. Use this to show that the latter is not a unique factorisation domain.

Chapter 11

Chebyshev's theorem

The following conjecture was posed by Bertrand in 1845 and proved by Chebyshev in 1852.

Theorem 11.1. *For every $n > 1$ there is a prime number p with $n < p < 2n$.*

We are going to estimate the binomial coefficient $\binom{2n}{n} = \frac{(2n)!}{(n!)^2}$ from above and from below. It is quite easy to get a lower bound which we will do shortly. To get an upper bound we will factorise it into a product of primes and get estimates for prime factors. In particular, if Chebyshev's theorem is false then all prime factors of $\binom{2n}{n}$ are at most n. Then we will see that if n is large enough then the two bounds are inconsistent. This will prove the theorem for sufficiently large n and small values of n will be dealt with separately by exhibiting prime numbers in each interval $(n, 2n)$ (we will do this in a smart way and not consider all possible values of n).

The proof that we are going to present is due to Erdős, and is taken from [AZ14].

11.1 Basic estimates

Lemma 11.2. *For $n \geq 1$ we have*

$$\binom{2n}{n} \geq \frac{4^n}{2n}.$$

Proof. Observe that

$$4^n = (1+1)^{2n} = \sum_{k=0}^{2n} \binom{2n}{k} = 2 + \sum_{k=1}^{2n-1} \binom{2n}{k} \leq 2n \cdot \binom{2n}{n}.$$

The last inequality follows from the fact that $\binom{2n}{n} \geq \binom{2n}{k}$ for all k and $\binom{2n}{n} \geq 2$. □

Recall that for a positive integer n and a prime number p the p-adic valuation of n is defined by

$$v_p(n) := \max\{\gamma \in \mathbb{Z} : p^\gamma \mid n\}.$$

Then the fundamental theorem of arithmetic implies

$$n = \prod_{p|n} p^{v_p(n)} = \prod_p p^{v_p(n)},$$

where the last product is over all primes p. It is actually a finite product since if $p \nmid n$ then $v_p(n) = 0$.

Lemma 11.3. *For $n \geq 1$ we have*

$$v_p(n!) = \sum_{k=1}^{\infty} \left[\frac{n}{p^k}\right].$$

Proof. There are $\left[\frac{n}{p}\right]$ numbers between (and including) 1 and n divisible by p. Each of those numbers contributes at least 1 to $v_p(n!)$. Now, $\left[\frac{n}{p^2}\right]$ many of those numbers is divisible by p^2 and they contribute two. Repeating this argument for each p^k we get the desired result. Note that when $k > \log_p n$, the terms $\left[\frac{n}{p^k}\right]$ vanish and hence the sum is actually a finite sum. □

Denote $N := \binom{2n}{n}$.

Lemma 11.4. *For each prime $p \leq 2n$ we have*

(i) $p^{v_p(N)} \leq 2n$,

(ii) *if $p > \sqrt{2n}$ then $v_p(N) \leq 1$,*

(iii) *if $\frac{2n}{3} < p \leq n$ then $v_p(N) = 0$ (for $n \geq 3$).*

Proof. (i) First notice that for any real number x the difference $[2x] - 2[x]$ is either 0 or 1. Further, by the previous lemma,

$$v_p(N) = v_p((2n)!) - 2v_p(n!) = \sum_{k=1}^{[\log_p 2n]} \left(\left[\frac{2n}{p^k}\right] - 2\left[\frac{n}{p^k}\right]\right) \leq \log_p 2n.$$

(ii) If $\sqrt{2n} < p \leq 2n$ then $\log_p 2n = 1$.

(iii) If $\frac{2n}{3} < p \leq n$ then $3p > 2n$ and $2p > n$ and $p^2 \geq 3p > 2n$. So $v_p((2n)!) = 2$ and $v_p(n!) = 1$.

□

Lemma 11.5. *For any real number $x > 1$*

$$\prod_{p \leq x} p \leq 4^{x-1},$$

where the product is over all prime numbers that do not exceed x.

Proof. Replacing x by $[x]$ we may assume that x is a positive integer and we will proceed to the proof by induction on x. Further, we can assume that $x = 2m+1$ is an odd integer (and a prime number).

First, notice that $\binom{2m+1}{m} \leq 4^m$ since

$$2^{2m+1} = (1+1)^{2m+1} = \sum_{k=0}^{2m+1} \binom{2m+1}{k} \geq \binom{2m+1}{m} + \binom{2m+1}{m+1} = 2\binom{2m+1}{m}.$$

Second,

$$\prod_{m+1<p\leq 2m+1} p \,\Big|\, \binom{2m+1}{m} = \frac{(2m+1)!}{m!(m+1)!}$$

since each prime p with $m+1 < p \leq 2m+1$ divides $(2m+1)!$ and does not divide $m!(m+1)!$.

Now, using the induction hypothesis for m, we get

$$\prod_{p \leq 2m+1} p = \prod_{p \leq m+1} p \cdot \prod_{m+1<p\leq 2m+1} p \leq 4^m \cdot \binom{2m+1}{m} \leq 4^m \cdot 4^m = 4^{2m}.$$

□

11.2 Proof of Chebyshev's theorem

Now we are ready to prove Chebyshev's theorem.

Proof of Theorem 11.1. First we will show that Chebyshev's theorem is true for $n > 5000$. Assume for contradiction it is false for some such n, i.e. there are no primes between n and $2n$. Then from Lemmas 11.2, 11.4 and 11.5 we have

$$\frac{4^n}{2n} \leq \binom{2n}{n} = \prod_{p \leq 2n} p^{v_p(N)} = \prod_{p \leq \sqrt{2n}} p^{v_p(N)} \cdot \prod_{\sqrt{2n}<p\leq \frac{2n}{3}} p \cdot \prod_{n<p\leq 2n} p \leq (2n)^{\sqrt{2n}} \cdot 4^{\frac{2n}{3}}.$$

Thus,

$$4^{n/3} \leq (2n)^{\sqrt{2n}+1},$$

and so

$$2n \cdot \frac{\log 2}{3} \leq (\sqrt{2n}+1)\log 2n.$$

This is obviously wrong for sufficiently large n since the left hand side grows much faster. We will now show that it is in fact false for $n > 5000$. Indeed, it suffices to show that for $x > 10000$ we have

$$x \cdot \frac{\log 2}{3} > (\sqrt{x}+1)\log x.$$

First, it is easy to check that $\frac{\log 2}{3} > 0.2$ and $\log x < x^{\frac{1}{4}}$ if $x > 10000$. We claim that the stronger inequality $0.2x > 2\sqrt{x} \cdot x^{\frac{1}{4}}$ holds. But this is obviously equivalent to $x > 10000$.

Thus, we proved Chebyshev's theorem for $n > 5000$. For smaller n we argue as follows. Consider the following sequence of primes:

$$2, 3, 5, 7, 13, 23, 43, 83, 163, 317, 631, 1259, 2503, 5003.$$

Every prime in this sequence is smaller than twice the previous one. Hence for any $1 \leq n \leq 5000$ one of those primes is between n and $2n$. □

11.3 On the prime number theorem

For a positive real number x let $\pi(x)$ be the number of primes that do not exceed x (π is called the prime counting function).

Theorem 11.6 (Prime number theorem). *$\pi(x) \sim \frac{x}{\log x}$ as $x \to \infty$.*[1]

This is a classical result in analytic number theory first proved by Hadamard and de la Vallée-Poussin in 1896. The proof is beyond the scope of this course. However, we will use the estimates established in the previous sections to prove a weak version of the prime number theorem.

Theorem 11.7. *There are constants $0 < A < 1 < B$ such that*

$$A\frac{x}{\log x} \leq \pi(x) \leq B\frac{x}{\log x}$$

for all sufficiently large x.

[1] The notation $f(x) \sim g(x)$ means that $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$.

The following proof is from [Gre16] and the reader can also find a proof of the prime number theorem there.

Proof. First, we will establish the upper bound. Lemma 11.5 implies

$$\prod_{x/2<p\leq x} p \leq 4^x,$$

and hence

$$\left(\frac{x}{2}\right)^{\pi(x)-\pi(x/2)} \leq 4^x.$$

Taking logarithms we get

$$\pi(x) \leq \pi(x/2) + \frac{x \log 4}{\log(x/2)}.$$

Applying this inequality repeatedly for $\frac{x}{2}, \frac{x}{4}, \ldots$ for each positive integer m we get

$$\pi(x) \leq \pi\left(\frac{x}{2^m}\right) + 2\log 4 \sum_{k=1}^{m} \frac{x/2^k}{\log(x/2^k)}.$$

Now let m be the biggest integer for which $2^m \leq \sqrt{x}$. Then $2^{m+1} > \sqrt{x}$ and $x/2^m < 2\sqrt{x}$. Also, for each $1 \leq k \leq m$ we have $x/2^k \geq \sqrt{x}$. Hence

$$\pi(x) < 2\sqrt{x} + 2\log 4 \sum_{k=1}^{m} \frac{x/2^k}{\log\sqrt{x}} < 2\sqrt{x} + \frac{x}{\log x} \cdot 4\log 4 \sum_{k=1}^{\infty} \frac{1}{2^k}.$$

Obviously, $2\sqrt{x} < \frac{x}{\log x}$ for sufficiently large x and so we can take $B = 1 + 4\log 4$.[2]

Now we turn to the lower bound.

Claim. There is a constant $C > 1$ for which $\prod_{p\leq 2n} p \geq C^n$ for all sufficiently large n.

Indeed, we saw in the proof of Chebyshev's theorem that

$$\frac{4^n}{2n} < (2n)^{\sqrt{2n}} \cdot \prod_{p\leq 2n} p.$$

So any constant $1 < C < 4$ works.

[2] Actually, any number $B > 4\log 4$ would work. Moreover, it can be shown that $\prod_{x/2<p\leq x} p \leq C^x$ for any real number $C > 2$. Hence taking $B > 4\log 2$ would suffice.

Now the claim implies

$$n \log C \leq (2n)^{\pi(2n)}$$

and

$$\pi(2n) \geq \frac{\log C}{2} \cdot \frac{2n}{\log 2n}.$$

This establishes the lower bound for $x = 2n$. The general case follows easily from this. □

11.4 Exercises

1. How many zeroes are there at the end of the decimal representation of 99! ?

2. Show that $\prod_{x/2<p\leq x} p \leq C^x$ for any real number $C > 2$.

Bibliography

[AZ14] Martin Aigner and Günter M. Ziegler. *Proofs from THE BOOK.* Springer-Verlag Berlin Heidelberg, 5 edition, 2014.

[Bak12] Alan Baker. *A Comprehensive Course in Number Theory.* Cambridge University Press, 2012.

[Gre16] Ben Green. *Analytic Number Theory: Lecture Notes.* University of Oxford, Oxford, 2016. Available at `http://people.maths.ox.ac.uk/greenbj/papers/primenumbers.pdf`.

[Gre17] Ben Green. *Introduction to Number Theory: Lecture Notes.* University of Oxford, Oxford, 2017. Available at `http://people.maths.ox.ac.uk/greenbj/papers/numbertheory-2017.pdf`.

[HW80] Godfrey Hardy and Edward Wright. *An Introduction to the Theory of Numbers.* Oxford University Press, 1980.