Part II - Number Theory

Lectured by Prof. T. A. Fisher

Artur Avameri

Lent 2022

Contents

0 Introduction

06 Oct 2022, Lecture 1

Books:

- A. Baker, *A concise introduction to the theory of numbers*, CUP 1984
- N. Koblitz, *A course in number theory & cryptography*, Springer 1994
- H. Davenport, *The higher arithmetic*, CUP 2008

Number theory studies the hidden and mysterious properties of the integers and the rational numbers.

It has always been an experimental science. Examining numerical data leads to **conjectures**, many of which are very old and still unproven today.

Example 0.1. (i) Let $N \geq 1$ be an integer of the form $8n+5, 8n+6$ or $8n+7$. Does there exist a right-angled triangle of area N, all of whose sides have rational length? We don't know.

(ii) Let $\pi(x)$ be the number of primes less than or equal to x and define $\mathrm{li}(x) = \int_2^x \frac{dt}{\log t}$. Then for all $x \geq 3$, $|\pi(x) - \mathrm{li}(x)| \leq \sqrt{x} \log x$. This is in fact equivalent to the Riemann hypothesis.

(iii) There are infinitely many twin primes. We now know there is an integer $N \leq 246$ such that there are infinitely many pairs of primes the form $p, p+N$.

1 Euclid's algorithm and factoring

Definition 1.1 (Division algorithm)**.** Given $a, b \in \mathbb{Z}$, with $b > 0$, there exist $q, r \in \mathbb{Z}$ such that $a = qb + r$, and $0 \leq r < b$.

Notation. If $r = 0$, then we write $b|a$, else $b \nmid a$.

Proof. Let $S = \{a - nb \mid n \in \mathbb{Z}\}$. This certainly contains integers ≥ 0, so take the smallest one r. We claim $r < b$. Indeed, if not, then $r - b \geq 0$, contradicting minimality. □

Given $a_1, \ldots, a_n \in \mathbb{Z}$ not all zero, let $I = \{\lambda_1 a_1 + \ldots + \lambda_n a_n \mid \lambda_i \in \mathbb{Z}\}$.

Lemma 1.1. $I = d\mathbb{Z}$ for some $d > 0$.

Proof. I certainly contains integers ≥ 0. Let d be the least positive element of I. We claim it works. Take $a \in I$, then $a = qd + r$ with $0 \leq r < d$. But $r = a - qd \in I \implies r = 0$. □

Remark. We get from this that d divides each a_i, and any common divisor of the a_i must divide d. Why?

We write $d = \gcd(a_1, \ldots, a_n)$ for the **greatest common divisor** (or **highest common factor**), or just use the shorthand $d = (a_1, \ldots, a_n)$.

Corollary 1.2. Let $a, b, c \in \mathbb{Z}$. Then there exist $x, y \in \mathbb{Z}$ such that $ax + by = c$ if and only if $(a, b)|c$.

The division algorithm gives a very efficient way to compute (a, b). Assume $a > b > 0$. Apply the division algorithm recursively to get

$$\begin{aligned}
a &= q_1 b + r_1 & 0 &\leq r_1 < b \\
b &= q_2 r_1 + r_2 & 0 &\leq r_2 < r_1 \\
r_1 &= q_3 r_2 + r_3 & 0 &\leq r_3 < r_2 \\
&\vdots \\
r_{k-2} &= q_k r_{k-1} + r_k & 0 &\leq r_k < r_{k-1}, r_k \neq 0 \\
r_{k-1} &= q_{k+1} r_k + 0
\end{aligned}$$

Claim. $r_k = (a, b)$. Indeed, $(a, b) = (b, r_1) = (r_1, r_2) = \ldots = (r_{k-1}, r_k) = r_k$. This is called **Euclid's algorithm**.

Remark. If $d = (a, b)$, then by Lemma 1.2, there exist $r, s \in Z$ such that $ra + st = d$. Euclid's algorithm gives us a way to find r and s.

In the following table, x and y stand for 34 and 25, and we then compute remainders as linear combinations of them.

We can use a trick here to speed this up: find each row as $q\cdot$ the row before it + the second row before it, then figure out signs at the end. (In fact, the minus signs zigzag down).

	x	y
$a = 34$	1	0
$b = 25$	0	1
$34 = 1 \cdot 25 + 9$	1	-1
$25 = 2 \cdot 9 + 7$	-2	3
$9 = 1 \cdot 7 + 2$	3	-4
$7 = 3 \cdot 2 + 1$	-11	15

We hence get $-11 \cdot 34 + 15 \cdot 25 = 1$.

Definition 1.2. An integer $n > 1$ is **prime** if its only positive divisors are 1 and n. Otherwise n is **composite**.

Lemma 1.3. *Let p be a prime, and $a, b \in \mathbb{Z}$. If $p|ab$, then $p \mid a$ or $p \mid b$.*

Proof. Assume $p \nmid a$. Then $(a, p) = 1$. By Lemma 1.2, $\exists r, s \in \mathbb{Z}$ such that $ra + sp = 1 \implies rab + spb = b$. Since $p \mid ab$, $p \mid b$ follows. $\square$

Theorem 1.4 (**Fundamental Theorem of Arithmetic**)**.** *Every integer $n > 1$ can be written as a product of primes. This representation is unique up to reordering.*

Proof. Existence is obvious. For uniqueness, suppose $n = p_1 p_2 \ldots p_r = q_1 q_2 \ldots q_s$ for p_i, q_i primes. We have $p_1 \mid q_1 q_2 \ldots q_r$, so by Lemma 1.5, $p_1 \mid q_j$ for some j, so $p_1 = q_j$. Now cancel these out and induct. $\square$

Remark. If $m = \prod_{i=1}^k p_i^{\alpha_i}$ and $n = \prod_{i=1}^k p_i^{\beta_i}$ for p_i distinct primes and $\alpha_i, \beta_i \geq 0$, then

$$(m, n) = \prod_{i=1}^k p_i^{\min(\alpha_i, \beta_i)}.$$

However, if m and n are large, it is more efficient to compute (m, n) using Euclid's algorithm. 08 Oct 2022, Lecture 2

Suppose we have some large positive integer N. An obvious algorithm for factoring N is to trial divide by 2 and the odd integers up to $\sqrt{N}$.

Definition 1.3. An algorithm with input a positive integer N is **polynomial** or a **polynomial time** algorithm if it takes $\leq c(\log N)^b$ **elementary operations** for some constants b and c.

Remark. An elementary operation is just adding/multiplying two numbers in $\{0, 1, \ldots, 9\}$.

Remark. "Polynomial" makes sense here as it takes $\log N$ digits to write N.

Polynomial algorithms are known for:

- Adding and multiplying integers (the usual way);
- Computing gcd's (via Euclid's algorithm);
- Detecting n^{th} powers (compute $\sqrt[n]{}$ numberically and round)
- More remarkably, primality testing (Agrawal, Kayal, Saxena in 2002)

But trial division up to $\sqrt{N}$ is not polynomial.

Fundamental question: Is there a polynomial time algorithm for factoring? This is unknown.

Later in this course we study the distribution of the prime numbers, in particular the function $\pi(x)$, the number of primes $\leq x$.

Theorem 1.5. There are infinitely many prime numbers, i.e. $\lim_{x\to\infty} \pi(x) \to \infty$.

Proof. Suppose there are only finitely many, say $p_1, \ldots, p_k$. Consider $N = \prod_{i=1}^{k} p_k + 1$. Then N must be divisible by some prime other than the p_i, so we're done. □

All the largest known primes are of the form $2^n - 1$ for n a prime. These are called **Mersenne primes**. 51 of them are known, the largest being $2^{82589933} - 1$.

2 Congruences

Fix a positive integer n (the modulus).

Definition 2.1. We say $a \equiv b \pmod{n}$, or that a is congruent to $b \pmod{n}$ if n divides $a - b$.

This defines an equivalence relation on $\mathbb{Z}$, and we write $\mathbb{Z}/n\mathbb{Z}$ for the set of equivalence classes. We can denote these by $a + n\mathbb{Z}$, or (more commonly) by $a \pmod{n}$. We can check that addition and multiplication are well-defined.

Remark. $n\mathbb{Z}$ is a subgroup/ideal of $\mathbb{Z}$ and $\mathbb{Z}/n\mathbb{Z}$ is the quotient group/ring.

Lemma 2.1. Let $a \in \mathbb{Z}/n\mathbb{Z}$. Then the following are equivalent:

(i) $(a, n) = 1$

(ii) $\exists b \in \mathbb{Z}$ such that $ab \equiv 1 \pmod{n}$

(iii) a is a generator for $\mathbb{Z}/n\mathbb{Z}$.

Proof. (i) $\implies$ (ii): $(a,n)=1 \implies \exists r,s\in\mathbb{Z}$ such that $ra+sn=1$, so $ra\equiv 1 \pmod n$.

(ii) $\implies$ (i): $ab\equiv 1 \pmod n \implies ab+kn=1$ for some $k\in\mathbb{Z} \implies (a,b)=1$.

(ii) $\iff$ (iii): $\exists b\in\mathbb{Z}$ s.t. $ab\equiv 1 \pmod n \iff$ 1 belongs to the subgroup of $\mathbb{Z}/n\mathbb{Z}$ generated by a. □

Notation. $(\mathbb{Z}/n\mathbb{Z})^\times$ is the group of **units** in $\mathbb{Z}/n\mathbb{Z}$, i.e. the elements with an inverse under multiplication.

Definition 2.2. $\phi(n)=|(\mathbb{Z}/n\mathbb{Z})^\times|$ is called the **Euler totient function**. We also have $\phi(n)=|\{1\le a\le n \mid (a,n)=1\}|$.

Remark. $\mathbb{Z}/n\mathbb{Z}$ is a field $\iff \phi(n)=n-1 \iff n$ is prime.

Theorem 2.2 (Euler-Fermat theorem)**.** If $(a,n)=1$, then $a^{\phi(n)}\equiv 1 \pmod n$.

Proof. Apply Lagrange's theorem to the group $G=(\mathbb{Z}/n\mathbb{Z})^\times$. Then for $a\in G$, its order divides $|G|=\phi(n)$. □

As a corollary:

Theorem 2.3 (Fermat's little theorem)**.** If $p\nmid a$, then $a^{p-1}\equiv 1 \pmod p$.

Lemma 2.4. Let G be a cyclic group of order n. We have

$$|\{g\in G \mid \mathrm{order}(g)=d\}| = \begin{cases}\phi(d) & \text{if } d\mid n\\ 0 & \text{otherwise}\end{cases}$$

In particular, $\sum_{d|n}\phi(d)=n$.

Proof. WLOG let $G=(\mathbb{Z}/n\mathbb{Z},+)$. We have $|\{g\in G \mid \mathrm{order}(g)=n\}| \overset{(*)}{=} \phi(n)$ by Lemma 2.2. If $d\mid n$, say $n=dk$, then the elements of order dividing d are the classes $0,k,2k,\dots,(d-1)k \pmod n$. These form a cyclic subgroup of order d. Applying $(*)$ to this cyclic subgroup shows that there are $\phi(d)$ elements of order d. □

Example 2.1. Consider the simultaneous linear congruences $x\equiv 7 \pmod{10}$ and $x\equiv 3 \pmod{13}$. Suppose we can find $u,v\in\mathbb{Z}$ such that

$$\begin{cases}u\equiv 1 & \pmod{10}\\ u\equiv 0 & \pmod{13}\end{cases}, \begin{cases}v\equiv 0 & \pmod{10}\\ v\equiv 1 & \pmod{13}\end{cases}.$$

Then $x=7u+3v$ is a solution. But $(10,13)=1 \implies \exists r,s\in\mathbb{Z}$ such that $10r+13s=1$, and we can just take $u=13s, v=10r$. To find r,s, we can use Euclid's algorithm to get $r=4, s=-3$, so $u=-39, v=40$, and so $x\equiv 7\cdot(-39)+3\cdot 40\equiv 107 \pmod{130}$.

11 Oct 2022, Lecture 3

Theorem 2.5 (Chinese Remainder Theorem)**.** Let $m_1, \ldots, m_k$ be pairwise coprime integers greater than 1. Let $a_1, \ldots, a_k \in \mathbb{Z}$. Let $M = m_1 m_2 \ldots m_k$. Then $\exists x \in \mathbb{Z}$ satisfying

$$\begin{cases} x \equiv a_1 \pmod{m_1} \\ \vdots \\ x \equiv a_k \pmod{m_k} \end{cases}.$$

Moreover, the solution is unique mod M.

Proof. Uniqueness: Suppose $x \equiv x' \pmod{m_i}$ $\forall i$. Then by considering the prime factorization of $x - x'$ and using the fact that the m_i are pairwise coprime, we get $x \equiv x' \pmod{M}$.

Existence: Put $M_i = \frac{M}{m_i}$, so $(M_i, m_i) = 1$ $\forall i$. Hence we can find $u_i \in \mathbb{Z}$ such that $u_i M_i \equiv 1 \pmod{m_i}$ $\forall i$. Let $x = \sum_{j=1}^{k} a_j u_j M_j$. Then $x \equiv a_i u_i M_i \equiv a_i \pmod{m_i}$. □

We can write this theorem in one ling using ring theory.

Definition 2.3. Let $R_i = \mathbb{Z}/m_i\mathbb{Z}$, and define $R_1 \times \ldots \times R_k = \{(r_1, \ldots, r_k) \mid r_i \in R_i\}$ with addition and multiplication defined componentwise. This is a ring.

Theorem 2.6 (CRT, ring-theoretic version)**.** Let $m_1, \ldots, m_k$ be pairwise coprime integers greater than 1 and put $M = m_1 \ldots m_k$. Then the map

$$\begin{aligned} \theta : \mathbb{Z}/M\mathbb{Z} &\to \mathbb{Z}/m_1\mathbb{Z} \times \ldots \times \mathbb{Z}/m_k\mathbb{Z} \\ a + M\mathbb{Z} &\mapsto (a + m_1\mathbb{Z}, \ldots, a + m_k\mathbb{Z}) \end{aligned}$$

is an isomorphism of rings.

Proof. θ is a well defined ring homomorphism since $m_i | M$ $\forall i$. Injectivity of θ follows from uniqueness in CRT, and surjectivity of θ follows from existence in CRT. □

Corollary 2.7. θ induces an isomorphism of groups under multiplication

$$\begin{aligned} (\mathbb{Z}/m\mathbb{Z})^\times &\cong (\mathbb{Z}/m_1\mathbb{Z})^\times \times \ldots \times (\mathbb{Z}/m_k\mathbb{Z})^\times \\ a + M\mathbb{Z} &\mapsto (a + m_1\mathbb{Z}, \ldots, a + m_k\mathbb{Z}). \end{aligned}$$

Remark. If $a \in \mathbb{Z}$, then $(a, M) = 1 \iff (a, m_i) = 1$ $\forall i$.

In particular, by looking at orders of the LHS and the RHS above, we get $\phi(M) = \phi(m_1) \ldots \phi(m_k)$, i.e. the Euler phi function is multiplicative.

Definition 2.4. A function $f : \mathbb{Z}^+ \to \mathbb{C}$ is **multiplicative** if $f(m) = f(m)f(n)$ whenever $(m, n) = 1$.

Examples:

- $\phi(n) = |(\mathbb{Z}/n\mathbb{Z})^\times|$;
- $\tau(n) = \sum_{d|n} 1$, the number of divisors of n;
- $\sigma(n) = \sum_{d|n} d$, the sum of divisors of n;
- more generally, $\sigma_k(n) = \sum_{d|n} d^k$, so $\sigma_0 = \tau$ and $\sigma_1 = \sigma$.

To prove this:

Lemma 2.8. If $f : \mathbb{Z}^+ \to \mathbb{C}$ is multiplicative, then so is $g : \mathbb{Z}^+ \to \mathbb{C}$, defined by $g(n) = \sum_{d|n} f(d)$.

Proof. Let m, n be coprime. Note that every divisor d of mn can be written as $d = d_1 d_2$, where $d_1 \mid m$, $d_2 \mid n$ and $(d_1, d_2) = 1$. Thus

$$g(mn) = \sum_{d|mn} f(d) = \sum_{d_1|m} \sum_{d_2|n} f(d_1 d_2) = \sum_{d_1|m} \sum_{d_2|n} f(d_1) f(d_2) = g(m)g(n).$$

□

Lemma 2.9. (i) For p a prime, $\phi(p^k) = p^{k-1}(p-1) = p^k(1 - \frac{1}{p})$.

(ii) $\phi(n) = n \prod_{p|n} (1 - \frac{1}{p})$.

Proof. (i): $\phi(p^k)$ counts the number of integers a between 1 and p^k such that $(p^k, a) = (p, a) = 1$. So we have p^a numbers, and we don't count the multiples of p, so $\phi(p^k) = p^k - p^{k-1}$.

(ii): Follows from the fact that ϕ is multiplicative. □

Alternative proof that $\sum_{d|n} \phi(d) = n$ (cf Lemma 2.6).

Proof. Obviously the RHS is multiplicative. Since $\phi(n)$ is multiplicative, the LHS is multiplicative by Lemma 2.13, so it suffices to check for n a prime power, say $n = p^k$. To this end, compute

$$\sum_{d|p^k} \phi(d) = \phi(1) + \phi(p) + \ldots + \phi(p^k) = 1 + (p-1) + (p^2 - p) + \ldots + (p^k - p^{k-1}) = p^k.$$

□

2.1 Polynomial congruences

Let $R = \mathbb{Z}, \mathbb{Q}, \mathbb{Z}/n\mathbb{Z}$ (or more generally any commutative ring). Set $R[X] = \{$**polynomials** with coefficients in $R\}$, i.e. $a_n X^n + a_{n-1} X^{n-1} + \ldots + a_1 X + a_0$ for $a_i \in R$.

By definition, two polynomials are equal if and only if they have the same coefficients. We can check that $R[X]$ is a ring (with usual $+$ and $\times$).

Warning. The map $R[X] \to \{\text{functions } R \to R\}$ by $f \mapsto (\alpha \mapsto f(\alpha))$ is not always injective. For example, if $R = \mathbb{Z}/p\mathbb{Z}$ for p a prime, and $f(X) = X^p - X$, then $f(\alpha) = 0 \ \forall \alpha \in R$, but f is not the zero function.

Question. Can we show thatif $f \in R[X]$ has degree n, then f has at most n roots in R?

Answer. No. For example, take $R = \mathbb{Z}/8\mathbb{Z}$, then $f(X) = X^2 - 1$ has 4 solutions in $\mathbb{Z}/8\mathbb{Z}$.

13 Oct 2022, Lecture 4

Let $R = \mathbb{Z}, \mathbb{Q}, \mathbb{Z}/n\mathbb{Z}$ (or any commutative ring).

We have a **division algorithm** on $R[X]$:

Let $f, g \in R[X]$ and suppose the leading coefficient of g is a unit. Then $\exists q, r \in R[X]$ such that $f(X) = Q(X)g(X) + r(X)$ and $\deg(r) < \deg(g)$.

Proof. By induction on $\deg(f)$. If $\deg(f) < \deg(g)$, take $q = 0, r = f$. Otherwise, let $f(X) = aX^m + \ldots$ and $g(X) = bX^n + \ldots$ with $m \geq n$ and b a unit.

Let $f_1(X) = f(X) - ab^{-1}X^{m-n}g(X)$. Then $\deg(f_1) < \deg(f)$, so by the induction hypothesis, $f_1(x) = q_1(x)g(x) + r_1(x)$ for some $q_1, r_1 \in R[X]$ and $\deg(r_1) < \deg(g)$. Now take $q(X) = ab^{-1}X^{m-n} + q_1(X)$ and $r = r_1$, so we're done. $\square$

Corollary 2.10. If $f \in R[X]$ and $\alpha \in R$ is such that $f(\alpha) = 0$, then $f(X) = (X - \alpha) f_1(X)$ for some $f_1 \in R[X]$.

Proof. By the division algorithm, $f(X) = (X - \alpha) f_1(X) + r$ for some $r \in R$ (as $\deg(r) < \deg(X - \alpha)$). Plug in $X = \alpha$ to get $r = 0$. $\square$

Definition 2.5. R is an **integral domain** if R has no zero divisors, i.e. $\alpha, \beta \in R$, $\alpha\beta = 0 \implies \alpha = 0$ or $\beta = 0$.

Note. Let $n > 1$. Then $\mathbb{Z}/n\mathbb{Z}$ is an integral domain $\iff$ n is prime.

Theorem 2.11. If R is an integral domain, then any polynomial $f \in R[X]$ of degree n has at most n roots.

Proof. By induction on n, the degree of f. If $n = 0$, then our polynomial is a nonzero constant and we're done. Now suppose $\exists \alpha \in R$ such that $f(\alpha) = 0$ (otherwise we're done). By Corollary 2.15, $f(X) = (X - \alpha) f_1(X)$. Since R is an integral domain, every root of f, except possibly α is a root of f_1. By induction, f_1 has at most $n - 1$ roots, hence f has at most n roots and we're done. $\square$

Corollary 2.12 (Lagrange's Theorem)**.** Let p be a prime and $a_0, \dots, a_n \in \mathbb{Z}$ with $p \nmid a_n$. Then the congruence

$$a_n x^n + a_{n-1}x^{n-1} + \dots + a_1 x + a_0 \equiv 0 \pmod p$$

has at most n solutions mod p.

Proof. Take $R = \mathbb{Z}/p\mathbb{Z}$ in Theorem 2.17. □

Remark. In this course, we will refer to the above theorem as Lagrange's Theorem.

Example 2.2. Let p be a prime. We will factor $X^{p-1} - 1 \pmod p$. Let $f(X) = X^{p-1} - 1 - \prod_{a=1}^{p-1}(X - \alpha)$ in $\mathbb{Z}/p\mathbb{Z}[X]$. By Fermat's Little Theorem, f has at least $p-1$ roots mod p. But $\deg(f) < p - 1$, since the X^{p-1} terms cancel out, so by Lagrange's Theorem, $f = 0$, i.e. $X^{p-1} - 1 = \prod_{a=1}^{p-1}(X - a)$ in $\mathbb{Z}/p\mathbb{Z}[X]$. Plugging in $X = 0$ gives $(p-1)! \equiv -1 \pmod p$, i.e. Wilson's Theorem.

Example 2.3. Working mod 7, the powers of 3 (starting from 0) are $1, 3, 2, 6, 4, 5$. So $(\mathbb{Z}/7\mathbb{Z})^\times$ is cyclic, generated by 3.

Theorem 2.13. Let p be a prime. Then $(\mathbb{Z}/p\mathbb{Z})^\times$ is cyclic.

Proof. Let $S_d = \{a \in (\mathbb{Z}/p\mathbb{Z})^\times \mid \operatorname{ord}(a) = d\}$. Suppose $S_d \neq \emptyset$, say $a \in S_d$. Then $1, a, a^2, \dots, a^{d-1}$ are distinct elements in $\mathbb{Z}/p\mathbb{Z}$ and they are solutions of $x^d \equiv 1 \pmod p$. By Lagrange's theorem, this has at most d solutions, and we found d solutions, so those are all of them, i.e. $S_d \subseteq \{1, a, a^2, \dots, a^{d-1}\}$. Note that the LHS is a cyclic group of order d, so this has $\phi(d)$ elements of order d.

We conclude that for every d, $|S_d| = 0$ or $|S_d| = \phi(d)$. In particular, $|S_d| \leq \phi(d)$. Hence

$$p - 1 \overset{(\star)}{=} \sum_{d|(p-1)} |S_d| \leq \sum_{d|(p-1)} \phi(d) = p - 1,$$

where $(\star)$ follows since we just count all the elements in $(\mathbb{Z}/p\mathbb{Z})^\times$. Hence $|S_d| = \phi(d)\ \forall d \mid (p-1)$. In particular, $S_{p-1} \neq \emptyset$, i.e. $(\mathbb{Z}/p\mathbb{Z})^\times$ contains elements of order $p-1$, i.e. $(\mathbb{Z}/p\mathbb{Z})^\times$ is cyclic. □

Remark. The same argument shows that any finite subgroup of the multiplicative group of a field is cyclic.

Definition 2.6. An integer a such that $a \pmod n$ generates $(\mathbb{Z}/n\mathbb{Z})^\times$ is called a **primitive root** mod n.

Theorem 2.21 showed that primitive roots exist mod p.

Example 2.4. Let $p = 19$. Let d be the order of 2 in $(\mathbb{Z}/19\mathbb{Z})^\times$. We know $d \mid 18$, so we work out

$$\begin{aligned} 2^3 &\equiv 8 \pmod{19} \\ 2^6 &\equiv 7 \not\equiv 1 \pmod{19} \implies d \nmid 6 \\ 2^9 &\equiv -1 \not\equiv 1 \pmod{19} \implies d \nmid 9, \end{aligned}$$

so $d = 18$ and hence 2 is a primitive root mod 19.

In general, $g \in \mathbb{Z}$ (coprime to p) is a primitive root mod p if and only if $g^{\frac{p-1}{q}} \not\equiv 1 \pmod{p}$ $\forall$primes $q \mid (p-1)$.

15 Oct 2022, Lecture 5

Remark. The number of primitive roots mod p is $\phi(p-1) = \phi(\phi(p))$.

Here are some (open) problems concerning primitive roots:

(i) Artin's conjecture (1927) – Let $a > 1$ be an integer which is not a square. Then a is a primitive root mod p for infinitely many primes p. This is unknown for $a = 2$. Hooley (1967) proved this assuming GRH. Heath-Brown (1986) proved that Artin's conjecture holds for at least one of $2, 3$ or 5. In fact, he proved something stronger: he proved the conjecture fails for at most 2 prime values of a.

(ii) How large is the smallest primitive root mod p? Burgess (1962) showed it is $\leq cp^{1/4+\epsilon}$ $\forall \epsilon > 0$ and some constant $c = c(\epsilon)$. Shoup (1992) showed it is $\leq c(\log p)^6$ assuming GRH.

We now consider $\mathbb{Z}/p^n\mathbb{Z}$ for $n > 1$. For $n \geq 3$, there is a surjective group homomorphism from $(\mathbb{Z}/2^n\mathbb{Z})^\times \to (\mathbb{Z}/8\mathbb{Z})^\times = \{\pm 1, \pm 3\} \cong C_2 \times C_2$, so $(\mathbb{Z}/2^n\mathbb{Z})^\times$ is not cyclic (since generators map to generators).

Theorem 2.14. Let p be an odd prime. Then $(\mathbb{Z}/p^n\mathbb{Z})^\times$ is cyclic $\forall n \geq 1$.

We divide the proof into 3 lemmas.

Lemma 2.15. Let $n \geq 2$. Then g is a primitive root mod p^n if and only if the following two conditions hold:

$$\begin{cases} g \text{ is a primitive root mod } p \\ g^{p^{n-2}(p-1)} \not\equiv 1 \pmod{p^n} \end{cases}.$$

Proof. ($\implies$) is clear, as $\phi(p^n) = p^{n-1}(p-1)$.

($\impliedby$): Let d be the order of g in $(\mathbb{Z}/p^n\mathbb{Z})^\times$. Then $d \mid \phi(p^n) = p^{n-1}(p-1)$. Since $g^d \equiv 1 \pmod{p^n}$, we have $g^d \equiv 1 \pmod{p}$. Hence by assumption 1, we have $(p-1) \mid d$. Say $d = p^j(p-1)$ for some $0 \leq j \leq n-1$. If $j \leq n-2$, then this contradicts assumption 2. Hence $j = n-1$, so $d = \phi(p^n)$ is a primtive root mod p^n. $\square$

Next we show $\exists g \in \mathbb{Z}$ satisfying conditions 1 and 2 in the case $n = 2$.

Lemma 2.16. $\exists g \in \mathbb{Z}$ a primitive root mod p such that $g^{p-1} \not\equiv 1 \pmod{p^2}$.

Proof. Let g be a primtive root mod p. If $g^{p-1} \equiv 1 \pmod{p^2}$, then consider $g+p$, which is still a primtive root mod p, but

$$(g+p)^{p-1} = g^{p-1} + (p-1)g^{p-2}p + \ldots \equiv 1 + (p-1)g^{p-2}p \pmod{p^2},$$

where the second term is not divisible by p^2, so $(g+p)^{p-1} \not\equiv 1 \pmod{p^2}$. $\square$

Next we show that if g is a primitive root mod p^2, then it is a primitive root mod p^n $\forall n \geq 2$.

Lemma 2.17. If $g^{p-1} \not\equiv 1 \pmod{p^2}$, then $g^{p^{n-2}(p-1)} \not\equiv 1 \pmod{p^n}$ $\forall n \geq 2$.

Proof. By induction on n, the case $n = 2$ being given. Suppose the result is true for n. By Euler-Fermat, $g^{p^{n-2}(p-1)} \equiv 1 \pmod{p^{n-1}}$, so $g^{p^{n-2}(p-1)} = 1 + bp^{n-1}$ for some $b \in \mathbb{Z}$, where $p \nmid b$ by the induction hypothesis. Taking p^{th} powers gives

$$g^{p^{n-1}(p-1)} = (1+bp^{n-1})^p = 1 + bp^n + \binom{p}{2} b^2 p^{2(n-1)} + \ldots \equiv$$
$$1 + bp^n + \binom{p}{2} b^2 p^{2(n-1)} \overset{\star}{\equiv} 1 + bp^n \pmod{p^{n+1}},$$

where $\star$ follows since p is odd, so $p \mid \binom{p}{2}$ (and also we use $3(n-1) \geq n+1$ and $2(n-1)+1 \geq n+1$). Thus $g^{p^{n-1}(n-1)} \equiv 1 + bp^n \not\equiv 1 \pmod{p^{n+1}}$, so the result follows for $n+1$. $\square$

This completes the proof of Theorem 2.24.

Example 2.5. We saw 3 is a primitive root mod 7. We calculate $3^3 = -1 + 4 \cdot 7$, so $3^6 \equiv 1 - 8 \cdot 7 \not\equiv 1 \pmod{7^2}$. Hence 3 is a primitive root mod 7^n $\forall n$.

For the case $p = 2$, let $G = \{a \in (\mathbb{Z}/2^n\mathbb{Z})^\times \mid a \equiv 1 \pmod 4\}$. Then $(\mathbb{Z}/2^n\mathbb{Z})^\times \cong \{\pm 1\} \times G$ by $a + 2^n\mathbb{Z} \mapsto \begin{cases} (1, a + 2^n\mathbb{Z}) & \text{if } a \equiv 1 \pmod 4 \\ (-1, -a + 2^n\mathbb{Z}) & \text{if } a \equiv 3 \pmod 4. \end{cases}$

Exercise. Show that G is cyclic (and generated by 5).

Exercise. For which n is $(\mathbb{Z}/n\mathbb{Z})^\times$ cyclic? 18 Oct 2022, Lecture 6

3 Quadratic residues

Let p be an odd prime and $a \in \mathbb{Z}$. By Lagrange's theorem, the congruence $x^2 \equiv a \pmod p$ has at most 2 solutions. If $a \not\equiv 0 \pmod p$, then there are either 0 or 2 solutions. Indeed, if x is a solution, then so is $-x \not\equiv x \pmod p$.

Definition 3.1. Suppose $a \not\equiv 0 \pmod p$. We say a is a **quadratic residue** (QR) if $x^2 \equiv a \pmod p$ is soluble. We say a is a **quadratic nonresidue** (NQR) if $x^2 \equiv a \pmod p$ is unsoluble.

Example 3.1. $p = 7$. $1, 2, 4$ are QRs and $3, 5, 6$ are QNRs.

Lemma 3.1. Let p be an odd prime. Then there are $\frac{p-1}{2}$ quadratic residues mod p (and hence also $\frac{p-1}{2}$ quadratic nonresidues).

Proof 1. Let $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ (a field with p elements). We show that the map $\mathbb{F}_p^\times \to \mathbb{F}_p^\times$ by $x \mapsto x^2$ is exactly 2–to–1.

Indeed, if $x^2 \equiv y^2 \pmod p$, then $p \mid x^2 - y^2$, so $p \mid (x - y)$ or $p \mid (x + y)$, so $x \equiv \pm y \pmod p$. □

Proof 2. Let g be a primitive root mod p. Then $\mathbb{F}_p^\times = \{1, g, g^2, \dots, g^{p-2}\}$.

We claim that g^i is a QR $\iff$ i is even.

$\Longleftarrow$ is clear. For $\Longrightarrow$, suppose $g^i \equiv x^2 \pmod p$. Then we can write $x = g^j \pmod p$, so $g^i \equiv g^{2j} \pmod p \implies i \equiv 2j \pmod{p-1}$. But $p - 1$ is even, so $i = 2j + k(p - 1)$ is even. □

Definition 3.2 (Legendre symbol)**.** Let p be an odd prime, $a \in \mathbb{Z}$. Then

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{if } p \mid a \\ 1 & \text{if } a \text{ is a QR mod } p \\ -1 & \text{if } a \text{ is a QNR mod } p \end{cases}$$

Theorem 3.2 (Euler's Criterion)**.** Let p be an odd prime and $a \in \mathbb{Z}$. Then

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

Proof. This is obvious if $p \mid a$, so suppose $(a, p) = 1$. By Fermat's little theorem, $a^{p-1} \equiv 1 \pmod p \implies a^{\frac{p-1}{2}} \equiv \pm 1 \pmod p$.

If $\left(\frac{a}{p}\right) = 1$, then $a \equiv b^2 \pmod p$ for some $b \in \mathbb{Z}$, but then $a^{\frac{p-1}{2}} \equiv b^{p-1} \equiv 1 \pmod p$. This gives $\frac{p-1}{2}$ solutions to the congruence $x^{\frac{p-1}{2}} \equiv 1 \pmod p$. By Lagrange's theorem, these are all the solutions. Hence if $\left(\frac{a}{p}\right) = -1$, then $a^{\frac{p-1}{2}} \not\equiv 1 \pmod p$, so $a^{\frac{p-1}{2}} \equiv -1 \pmod p$ and we're done. □

Corollary 3.3. $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$.

Proof.

$$\left(\frac{ab}{p}\right) \equiv (ab)^{\frac{p-1}{2}} \equiv a^{\frac{p-1}{2}} b^{\frac{p-1}{2}} \equiv \left(\frac{a}{p}\right)\left(\frac{b}{p}\right) \pmod p.$$

Since $0, \pm 1$ are distinct mod p, we have equality in the above. □

The corollary is equivalent to the statements:

- $\mathcal{X} : \mathbb{F}_p^\times \to \{\pm 1\}$ by $a \mapsto \left(\frac{a}{p}\right)$ is a group homomorphism.
- (i) QR $\cdot$ QR $=$ QR
  (ii) QR $\cdot$ QNR $=$ QNR
  (iii) QNR $\cdot$ QNR $=$ QR

We can give an alternative proof for this:

(i) $a \equiv x^2 \pmod p, b \equiv y^2 \pmod p \implies ab \equiv (xy)^2 \pmod p$.

(ii) If $a \equiv x^2$ and $ab \equiv z^2 \pmod p$, then $b \equiv (x^{-1}z)^2 \pmod p$, a contradiction.

(iii) Suppose a is a QNR. The map $\mathbb{F}_p^\times \to \mathbb{F}_p^\times$ by $x \mapsto ax$ is a bijection sending QRs to NQRs by (ii). By Lemma 3.1, it sends QNRs to QRs, done.

Remark. We can also prove Euler's criterion using primitive roots.

Corollary 3.4. Let p be a odd prime. Then

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}} = \begin{cases} 1 & \text{if } p \equiv 1 \pmod 4. \\ -1 & \text{if } p \equiv -1 \pmod 4. \end{cases}$$

In the next lecture, we show

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} = \begin{cases} 1 & \text{if } p \equiv \pm 1 \pmod 8. \\ -1 & \text{if } p \equiv \pm 3 \pmod 8. \end{cases}$$

Let p, q be distinct odd primes. The law of quadratic reciprocity gives a relation between $\left(\frac{p}{q}\right)$ and $\left(\frac{q}{p}\right)$. Generalizing this result (in many different ways) has been one of the main goals of number theory ever since.

Theorem 3.5 (Law of quadratic reciprocity)**.** Let p, q be distinct odd primes. Then

$$\left(\frac{q}{p}\right) = \begin{cases} \left(\frac{p}{q}\right) & \text{if } p \equiv 1 \pmod 4 \text{ or } q \equiv 1 \pmod 4. \\ -\left(\frac{p}{q}\right) & \text{if } p \equiv q \equiv 3 \pmod 4. \end{cases}$$

Example 3.2.

$$\left(\frac{19}{73}\right) = \left(\frac{73}{19}\right) = \left(\frac{16}{19}\right) = 1.$$

20 Oct 2022, Lecture 7

Another proof of Fermat's little theorem:

If $(a, p) = 1$, then working mod p, the set $\{a, 2a, 3a, \dots, (p-1)a\}$ is the same as $\{1, 2, \dots, (p-1)\}$. Taking the product gives $a^{p-1}(p-1)! \equiv (p-1)! \pmod p \implies a^{p-1} \equiv 1 \pmod p$ as desired.

We can use the same idea to compute $a^{\frac{p-1}{2}} \bmod p$:

Lemma 3.6 (Gauss' Lemma)**.** Let p be an odd prime, let $a \in \mathbb{Z}$ be coprime to p, and put $m = \frac{p-1}{2}$. For $j = 1, 2, \dots, m$ let a_j be the unique integer such that

(i) $a_j \equiv ja \pmod{p}$

(ii) $-m \leq a_j \leq m$.

Then $\left(\frac{a}{p}\right) = (-1)^\nu$, where $\nu = \{\#1 \leq j \leq m \mid a_j < 0\}$.

Proof. Consider $a_1, \dots, a_m \in \{\pm 1, \pm 2, \dots, \pm m\}$. Can any two of these be the same? No, since $a_i \equiv a_j \implies ai \equiv aj \implies i \equiv j \pmod{p}$.

Can any two differ by a sign? No, since $a_i \equiv -a_j \implies ia \equiv -ja \implies i \equiv -j \pmod{p}$.

Hence $a_1, \dots, a_m$ are $\pm 1, \pm 2, \dots, \pm m$ in some order with some choice of signs. Taking the product gives

$$a_1 \dots a_m \equiv (-1)^\nu 1 \cdot \ldots \cdot m \pmod{p} \implies a^m m! \equiv (-1)^\nu m! \pmod{p}.$$

So by Euler's criterion, $\left(\frac{a}{p}\right) \equiv a^m \equiv (-1)^\nu \pmod{p}$. ☐

Corollary 3.7. Let p be an odd prime. Then

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} = \begin{cases} 1 & \text{if } p \equiv \pm 1 \pmod{8}. \\ -1 & \text{if } p \equiv \pm 3 \pmod{8}. \end{cases}$$

Proof. Let $m = \frac{p-1}{2}$. Then $a_j = \begin{cases} 2j & \text{for } 1 \leq j \leq \frac{m}{2}. \\ 2j - p & \text{for } \frac{m}{2} < j \leq m. \end{cases}$ Hence

$$\nu = m - \left\lfloor \frac{m}{2} \right\rfloor = \begin{cases} \frac{m}{2} \text{ if } m \text{ is even.} \\ \frac{m+1}{2} \text{ if } m \text{ is odd.} \end{cases}$$

It follows that $\left(\frac{2}{p}\right) = 1 \iff \nu \text{ is even} \iff m \equiv 0, 3 \bmod 4 \iff p \equiv \pm 1 \pmod{8}$. ☐

Theorem 3.8 (Law of quadratic reciprocity)**.** Let p, q be distinct odd primes. Then

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

Proof. Step 1: Let a, p, ν be as in Gauss' Lemma (with $a \geq 1$).

Claim:

$$\nu = \sum_{i=1}^{2n} (-1)^i \left\lfloor \frac{ip}{2a} \right\rfloor$$

where $n = \lfloor \frac{a}{2} \rfloor$. Moreover, $\frac{ip}{2a} \notin \mathbb{Z} \; \forall \; 1 \le i \le 2n$.

Proof: Consider all multiples of a less than $\frac{ap}{2}$ ($= np$ or $(n+\frac{1}{2})p$). Hence ν is the number of multiples of a in the intervals

$$\left[\frac{1}{2}p, p\right], \left[\frac{3}{2}p, 2p\right], \ldots, \left[(n-\frac{1}{2})p, np\right].$$

On dividing through by a, we see that ν is the number of integers in

$$\left[\frac{p}{2a}, \frac{2p}{2a}\right], \left[\frac{3p}{2a}, \frac{4p}{2a}\right], \ldots, \left[\frac{(2n-1)p}{2a}, \frac{2np}{2a}\right].$$

The end points are not in $\mathbb{Z}$, since the end points of the original intervals are not multiples of a. Hence $\#([\alpha, \beta] \cap \mathbb{Z}) = \lfloor \beta \rfloor - \lfloor \alpha \rfloor$. This proves the claim.

Step 2: Let p_1, p_2 be primes and $a \in \mathbb{Z}$ coprime to $p_1 p_2$. By Gauss' lemma, $\left(\frac{a}{p_i}\right) = (-1)^{\nu_i}$.

(i) Suppose $p_1 \equiv p_2 \pmod{4a}$. Then $\lfloor \frac{ip_1}{2a} \rfloor \equiv \lfloor \frac{ip_2}{2a} \rfloor \pmod 2$. By Step 1, we have $\nu_1 \equiv \nu_2 \pmod 2$. Hence $\left(\frac{a}{p_1}\right) = \left(\frac{a}{p_2}\right)$.

(ii) Suppose $p_1 \equiv -p_2 \pmod{4a}$. Then $\lfloor \frac{ip_1}{2a} \rfloor \equiv \lfloor \frac{ip_2}{2a} \rfloor + 1 \pmod 2$. (We use the fact that if $\alpha \in \mathbb{R}/\mathbb{Z}$, then $\lfloor -\alpha \rfloor = -\lfloor \alpha \rfloor - 1$). By Step 1, we again deduce $\left(\frac{a}{p_1}\right) = \left(\frac{a}{p_2}\right)$.

Step 3: Conclusion of the proof.

(i) Suppose $p \equiv q \pmod 4$, say $p = 4a + q$. Then $\left(\frac{p}{q}\right) = \left(\frac{4a+q}{q}\right) = \left(\frac{a}{q}\right)$, and $\left(\frac{q}{p}\right) = \left(\frac{p-4a}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{a}{p}\right)$. But $p \equiv q \pmod{4a} \overset{\text{Step 2(i)}}{\Longrightarrow} \left(\frac{a}{p}\right) = \left(\frac{a}{q}\right)$, hence we conclude

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

(ii) Suppose $p \not\equiv q \pmod 4$, say $p + q = 4a$. Then $\left(\frac{p}{q}\right) = \left(\frac{4a-q}{q}\right) = \left(\frac{a}{q}\right)$ and $\left(\frac{q}{p}\right) = \left(\frac{4a-p}{p}\right) = \left(\frac{a}{p}\right)$. But $p \equiv -q \pmod{4a} \overset{\text{Step 2(ii)}}{\Longrightarrow} \left(\frac{a}{p}\right) = \left(\frac{a}{q}\right)$, so $\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right)$, done.

□

22 Oct 2022, Lecture 8

Example 3.3. Compute the Legendre symbol $\left(\frac{7411}{9283}\right)$. In fact, 7411 and 9283 are both prime. Hence

$$\left(\frac{7411}{9283}\right) = -\left(\frac{9283}{7411}\right) = -\left(\frac{1872}{7411}\right).$$

As $1872 = 2^4 \cdot 3^2 \cdot 13$, we get

$$-\left(\frac{1872}{8411}\right) = -\left(\frac{13}{7411}\right) = -\left(\frac{7411}{13}\right) = -\left(\frac{1}{13}\right) = -1.$$

Hence 7411 is not a QR mod 9283.

Recall that the Legendre symbol $\left(\frac{a}{p}\right)$ is only defined for p an odd prime.

Definition 3.3. Let n be an odd positive integer, say $n = p_1 \dots p_k$ for p_i (not necessarily distinct) odd primes. Let $a \in \mathbb{Z}$. We define the **Jacobi symbol** as

$$\left(\frac{a}{n}\right) = \prod_{i=1}^{k} \left(\frac{a}{p_i}\right).$$

Remark. If $(a, n) \neq 1$, then $\left(\frac{a}{n}\right) = 0$.

Proposition 3.9. (i) $\left(\frac{a}{n}\right)$ depends only on a mod n.

(ii) $\left(\frac{ab}{n}\right) = \left(\frac{a}{n}\right)\left(\frac{b}{n}\right)$ and $\left(\frac{a}{mn}\right) = \left(\frac{a}{m}\right)\left(\frac{a}{n}\right)$.

(iii) $\left(\frac{-1}{n}\right) = (-1)^{\frac{n-1}{2}}$.

(iv) $\left(\frac{2}{n}\right) = (-1)^{\frac{n^2-1}{8}}$.

Proof. (i) Clear, since the Legendre symbol only depends on a mod p.

(ii) The first part follows since the Legendre symbol is totally multiplicative, and the second follows from the definition of the Jacobi symbol.

(iii) This holds for $n = p$ a prime by previous results. We will now show that if they hold for odd integers m, n, then they hold for mn. But

$$\left(\frac{-1}{mn}\right) = \left(\frac{-1}{m}\right)\left(\frac{-1}{n}\right) = (-1)^{\frac{m-1}{2}}(-1)^{\frac{n-1}{2}} \overset{\star}{=} (-1)^{\frac{mn-1}{2}},$$

where we can check that $\star$ holds, since $(m-1)(n-1) \equiv 0 \pmod 4$, which gives $mn - 1 \equiv (m-1) + (n-1) \pmod 4$.

(iv) This is analogous to above, except we get

$$(-1)^{\frac{m^2-1}{8}}(-1)^{\frac{n^2-1}{8}} = (-1)^{\frac{(mn)^2-1}{8}},$$

since $(m^2-1)(n^2-1) \equiv 0 \pmod{16}$, so $(mn)^2 - 1 \equiv (m^2-1) + (n^2-1) \pmod{16}$.

$\square$

Theorem 3.10 (Law of Quadratic Reciprocity for Jacobi Symbols)**.** If m, n are odd positive integers, then

$$\left(\frac{m}{n}\right) = (-1)^{\frac{m-1}{2}\frac{n-1}{2}} \left(\frac{n}{m}\right).$$

Remark. If $(m, n) \neq 1$, this says $0 = 0$.

Proof. Again, we deduce this from the corresponding result for the Legendre symbol. Assume $(m, n) = 1$. Write $m = \prod_{i=1}^{k} p_i$ and $n = \prod_{j=1}^{l} q_j$ for p_i, q_j (not necessarily distinct) primes.

Let r count the number of p_i with $p_i \equiv 3 \pmod 4$ and s count the number of q_j with $q_j \equiv 3 \pmod 4$. Then

$$\left(\frac{m}{n}\right) = \prod_{i=1}^{k}\prod_{j=1}^{l}\left(\frac{p_i}{q_j}\right) = \prod_{i=1}^{k}\prod_{j=1}^{l}(-1)^{\frac{p_i-1}{2}\frac{q_j-1}{2}}\left(\frac{q_j}{p_i}\right) =$$
$$(-1)^{rs}\prod_{i=1}^{k}\prod_{j=1}^{l}\left(\frac{q_j}{p_i}\right) = (-1)^{rs}\left(\frac{n}{m}\right).$$

But $m \equiv 1 \pmod 4 \iff r$ is even, and $n \equiv 1 \pmod 4 \iff s$ is even, hence $(-1)^{rs} = (-1)^{\frac{m-1}{2}\frac{n-1}{2}}$. ∎

Remark. The Jacobi symbol $\left(\frac{a}{n}\right)$ tells us surprisingly little about whether the congruence $x^2 \equiv a \pmod n$ is soluble.

If $x^2 \equiv a \pmod n$ is soluble, then so is $x^2 \equiv a \pmod p$ for all primes $p \mid n$. So $\left(\frac{a}{p}\right) = 1\ \forall p \mid n$, hence $\left(\frac{a}{n}\right) = 1$.

But the converse is false. For example, $\left(\frac{2}{15}\right) = \left(\frac{2}{3}\right)\left(\frac{2}{5}\right) = (-1)\cdot(-1) = 1$, yet $x^2 \equiv 2 \pmod{15}$ is not soluble.

The point of the Jacobi symbol is rather that it allows us to compute Legendre symbols without having to factor (except for removing powers of 2).

Example 3.4.

$$\left(\frac{33}{73}\right) = \left(\frac{73}{33}\right) = \left(\frac{7}{33}\right) = \left(\frac{33}{7}\right) = \left(\frac{5}{7}\right) = -1,$$

so 33 is not a QR mod 73.

Three tricks to evaluate Legendre symbols:

Example 3.5. (i) $\sum_{a=1}^{p-1}\left(\frac{a}{p}\right) = 0$

(ii) $\sum_{a=1}^{p-1} a\left(\frac{a}{p}\right) \equiv 0 \pmod p$ if $p > 3$.

(iii) $\sum_{a=1}^{p-1} \left(\frac{a(a+1)}{p}\right) = -1$.

Proof. (i) We have already done this since we have an equal number of QRs and QNRs. However, alternate proof:

Let b be a QNR $\pmod p$. Then

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

so $\sum_{a=1}^{p-1} \left(\frac{a}{p}\right) = 0$.

(ii) Since $p > 3$, we can choose $b \not\equiv 0, \pm 1 \pmod p$, whence

$$\sum_{a=1}^{p-1} a \left(\frac{a}{p}\right) \equiv \sum_{a=1}^{p-1} ab \left(\frac{ab}{p}\right) \equiv \pm b \sum_{a=1}^{p-1} a \left(\frac{a}{p}\right) \pmod p.$$

Since $b \not\equiv \pm 1 \pmod p$, we deduce $\sum_{a=1}^{p-1} a \left(\frac{a}{p}\right) \equiv 0 \pmod p$.

(iii) If $ab \equiv 1 \pmod p$, then

$$\left(\frac{a(a+1)}{p}\right) \equiv \left(\frac{a^2(1+b)}{p}\right) = \left(\frac{b+1}{p}\right).$$

Then

$$\sum_{a=1}^{p-1} \left(\frac{a(a+1)}{p}\right) = \sum_{b=1}^{p-1} \left(\frac{b+1}{p}\right) = -1.$$

□

4 Binary quadratic forms

25 Oct 2022, Lecture 9

Question. Which numbers can be written as the sum of two squares?

Fermat gave an answer around 1630, and Euler published the first proof in 1749.

Theorem 4.1. Let N be a positive integer. Then N is the sum of two squares if and only if every prime $p \equiv 3 \pmod 4$ that divides N divides it to an even power.

Proof of the easy direction. $\Longrightarrow$: Suppose $N = x^2 + y^2$ and $p \mid N$, then $x^2 + y^2 \equiv 0 \pmod p$. If $p \equiv 3 \pmod 4$, then $\left(\frac{-1}{p}\right) = -1$, so we must have $x \equiv y \equiv 0 \pmod p$. Then divide N by p^2 and repeat until $p \nmid N$.

$\Longleftarrow$: Since $(x^2+y^2)(z^2+t^2) = (xz-yt)^2 + (xt+yz)^2$, it suffices to prove the result the case $N = p$ with $p = 2$ or $p \equiv 1 \pmod 4$. $p = 2$ is easy, but $p \equiv 1 \pmod 4$ is a little more involved, and we will prove it a later lecture. $\square$

Euler also studied $x^2 + 2y^2, x^2 + 3y^2$, etc. In this section we study **binary quadratic forms** with integer coefficients, i.e. $f(x, y) = ax^2 + bxy + cy^2$ for $a, b, c \in \mathbb{Z}$.

Definition 4.1. We say f **represents** n if $f(x, y) = n$ for some $x, y \in \mathbb{Z}$.

We may write f as (a, b, c) or in matrix notation as

$$f(x, y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Example 4.1. $f(x, y) = x^2 + y^2$ may be written as $(1, 0, 1)$ or $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

$g(x, y) = 4x^2 + 12xy + 10y^2$ may be written as $(4, 12, 10)$ or $\begin{pmatrix} 4 & 6 \\ 6 & 10 \end{pmatrix}$.

Note that $g(x, y) = (2x + 3y)^2 + y^2 = f(2x + 3y, y)$. Do f and g represent the same numbers? No, as g only represents even numbers.

Let $X = 2x + 3y, Y = y$, then

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \implies \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -3 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}.$$

Note that we can have $X, Y \in \mathbb{Z}$, yet $x, y \notin \mathbb{Z}$.

Definition 4.2. A **unimodular substitution** is one of the form $X = \alpha x + \gamma y, Y = \beta X + \delta Y$ where $\alpha, \beta, \gamma, \delta \in \mathbb{Z}$ and $\alpha\delta - \beta\gamma = 1$.

Definition 4.3. Two BQFs f and g are **equivalent**, written $f \sim g$, if they are related by a unimodular substitution.

Exercise: Check $\sim$ is an equivalence relation (this is on the example sheet).

Note. Equivalent forms represent the same integers.

The group $SL_2(\mathbb{Z}) = \left\{ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \mid \alpha, \beta, \gamma, \delta \in \mathbb{Z}, \alpha\delta - \beta\gamma = 1 \right\}$ acts on the set of BQFs via $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} : f(x, y) \mapsto f(\alpha x + \gamma y, \beta x + \delta y)$. The equivalence classes are the orbits of this action.

To check a group action, we need to check

(i) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} f = f$, which is true.

(ii) $\sigma(\tau f) = (\sigma\tau)f \ \forall \sigma, \tau \in SL_2(\mathbb{Z})$.

Suppose $f = (a, b, c)$ and $g = (a', b', c')$ are equivalent, say $g = \sigma f$ for $\sigma = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$. Then

$$g(x,y) = f(\alpha x + \gamma y, \beta x + \delta y) = \begin{pmatrix} \alpha x + \gamma y & \beta x + \delta y \end{pmatrix} \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix} \begin{pmatrix} \alpha x + \gamma y \\ \beta x + \delta y \end{pmatrix} =$$

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix} \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Hence $\begin{pmatrix} a' & \frac{b'}{2} \\ \frac{b'}{2} & c' \end{pmatrix} = \sigma \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix} \sigma^\top$. Call this $(\star)$.

To check (ii), we note that

$$\sigma \left(\tau \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix} \tau^\top \right) \sigma^\top = (\sigma\tau) \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix} (\sigma\tau)^\top.$$

Definition 4.4. The **discriminant** of $f(x,y) = ax^2 + bxy + cy^2$ is

$$\mathrm{disc}(f) = b^2 - 4ac.$$

Example 4.2. $\mathrm{disc}(1,0,1) = -4, \mathrm{disc}(4,12,10) = -16$.

Lemma 4.2. Equivalent BQFs have the same discriminant.

Proof. Taking determinants in $(\star)$ gives

$$a'c' - \left(\frac{b'}{2}\right)^2 = (\det \sigma)^2 \left(ac - \left(\frac{b}{2}\right)^2 \right).$$

But $\det \sigma = 1$, so multiplying both sides by -4 gives $(b')^2 - 4a'c' = b^2 - 4ac$ as desired. $\square$

Remark. The converse is not true, i.e. there exist BQFs with the same discriminant which are not equivalent.

For example, $(1,0,6)$ and $(2,0,3)$ both have discriminant -24, but $(1,0,6)$ represents 1 (with $x=1, y=0$), but $(2,0,3)$ does not.

Lemma 4.3. There exists a BQF f with $\mathrm{disc}(f) = d \iff d \equiv 0, 1 \pmod 4$.

Proof. $\implies$: $d = b^2 - 4ac \equiv b^2 \equiv 0, 1 \pmod 4$.

$\impliedby$: If $d \equiv 0 \pmod 4$, let $f = (1, 0, -\frac{d}{4})$. If $d \equiv 1 \pmod 4$, take $f = (1, 1, \frac{1-d}{4})$. $\square$

27 Oct 2022, Lecture 10

Definition 4.5. A quadratic form $f(x_1, \ldots, x_n) = \sum_{i \leq j} a_{ij} x_i x_j$ with $a_{ij} \in \mathbb{R}$ is:

- **positive definite** if $f(x) > 0 \ \forall 0 \neq x \in \mathbb{R}^n$.
- **negative definite** if $f(x) < 0 \ \forall 0 \neq x \in \mathbb{R}^n$.
- **indefinite** if $f(x) > 0$ and $f(x') < 0$ for some $x, x' \in \mathbb{R}^n$.

We are interested in the case $n = 2$ and $a_{ij} \in \mathbb{Z}$.

Lemma 4.4. Let $f(x, y) = ax^2 + bxy + cy^2$ be a BQF which has discriminant $d = b^2 - 4ac$.

(i) If $d < 0$ and $a > 0$, then f is positive definite.

(ii) If $d < 0$ and $a < 0$, then f is negative definite.

(iii) If $d > 0$, then f is indefinite.

(iv) If $d = 0$, then $f = \lambda(mx + ny)^2$ for $\lambda, m, n \in \mathbb{Z}$.

Proof.

$$4af(x, y) = 4a^2x^2 + 4abxy + 4acy^2 = \\ (2ax + by)^2 + (4ac - b^2)y^2 = (2ax + by)^2 - dy^2.$$

(i) and (ii): If $d < 0$ and $a \neq 0$, then it follows that $4af(x, y) \geq 0$ with equality if and only if $x = y = 0$. The cases $a > 0$ and $a < 0$ now show f is either positive or negative definite as desired.

(iii): Suppose $d > 0$. If $a \neq 0$, then the above equation shows us that $4af(1, 0) > 0$ and $4af(-b, 2a) < 0$, so f is indefinite.

If $a = 0$, then replace $f(x, y) \mapsto f(y, x)$. This works unless $a = c = 0$, but then $b \neq 0$, so $f(x, y) = bxy$, which is obviously indefinite.

(iv): Omitted (not interesting nor difficult). □

Remark. It is possible for a BQF (a, b, c) with $a, b, c > 0$ to be indefinite, e.g. $(1, 3, 1)$.

It is also possible for (a, b, c) with $b < 0$ to be positive definite, e.g. $(1, -1, 2)$.

From now on, we will concentrate on positive definite BQFs, i.e. forms (a, b, c) with $d = b^2 - 4ac < 0$ and $a > 0$ (and hence $c > 0$).

We have an equivalence relation $\sim$ on positive definite BQFs, and we want to study the equivalence classes. It will help if we can specify a "simplest" form for each equivalence class.

Example 4.3. Consider $(10, 34, 29)$. The middle coefficient is large – can we decrease it? If $f(x) = ax^2 + bxy + cy^2$, then one substitution we may try is

$$f(x + \lambda y, y) = a(x + \lambda y)^2 + b(x + \lambda y)y + cy^2 = \\ ax^2 + (b + 2\lambda a)xy + (\lambda^2 a + \lambda b + c)y^2.$$

Taking $\lambda = \pm 1$ shows

$$(a, b, c) \sim (a, b \pm 2a, a \pm b + c). \quad (\dagger)$$

In our example, we get $(10, 34, 29) \sim (10, 14, 5) \sim (10, -6, 1)$.

Making the substitution $X = y, Y = -x$ gives

$$(a, b, c) \sim (c, -b, a). \quad (\ddagger)$$

In our example we now get

$$(10, -6, 1) \sim (1, 6, 10) \sim (1, 4, 5) \sim (1, 2, 2) \sim (1, 0, 1).$$

Remark. It is a good idea to check that the discriminant doesn't change (to catch mistakes).

Remark. We can ensure $|b| \le a$ via ($\dagger$), and $a \le c$ via ($\ddagger$).

Definition 4.6. A positive definite BQF is **reduced** if either

$$-a < b \le a < c \text{ , or } \ 0 \le b \le a = c.$$

(Think of this as $|b| \le a \le c$ with some extra conditions).

Lemma 4.5. Every positive definite BQF is equivalent to a reduced form.

Proof. We have operations

$$S : (a, b, c) \mapsto (c, -b, a), \ T_\pm : (a, b, c) \mapsto (a, b \pm 2a, a \pm b + c).$$

If $a > c$, then use S to decrease a while leaving $|b|$ unchanged. If $a \le c$ and $|b| > a$, then use $T_\pm$ to decrease $|b|$ while leaving a unchanged.

Repeat these steps. Each step decreases $a + |b|$, so this procedure must eventually reach a form with $|b| \le a \le c$. Finally, to get the form we want in the lemma:

- If $b = -a$, then apply T_+ to replace $(a, -a, c) \mapsto (a, a, c)$.
- If $a = c$ and $b < 0$, then apply S to get $b > 0$.

□

29 Oct 2022, Lecture 11

Lemma 4.6. Let $f = (a, b, c)$ be a reduced positive definite BQF with discriminant d. Then $|b| \le a \le \sqrt{\frac{|d|}{3}}$ and $b \equiv d \pmod 2$.

Proof. Being reduced implies $|b| \le a \le c$, and $d = b^2 - 4ac \le ac - 4ac = -3ac \le -3a^2 \implies a^2 \le \frac{|d|}{3}$. Also $d = b^2 - 4ac \implies b \equiv d \pmod 2$. □

Example 4.4. Consider $d = -4$. We must have $a = 1$ by the lemma above (as $a > 0$), and $b = 0$ (by parity), so solve for c to get $c = 1$, i.e. $x^2 + y^2$ is the only positive definite reduced BQF with discriminant -4.

We can now return to the beginning of this section and answer our original question: which numbers can be written as the sum of two squares?

Proof of Theorem 4.1 (continued). Let p be a prime, $p \equiv 1 \pmod 4$. We have $\left(\frac{-1}{p}\right) = 1$, so $\exists u \in \mathbb{Z}$ such that $u^2 \equiv -1 \pmod p \implies u^2 = -1 + kp$ for some $k \in \mathbb{Z}$. Let $f = (p, 2u, k)$, so $\operatorname{disc}(f) = 4u^2 - 4pk = -4$.

By Lemma 4.5, $f \sim g$ for some reduced form g, but by our above example, $g(x, y) = x^2 + y^2$. Now f represents p (take $x = 1, y = 0$), so g also represents p, i.e. p is the sum of two squares as required. □

Question. Can reduced forms be equivalent?

Definition 4.7. Let f be a BQF and $n \in \mathbb{Z}$. We say f **represents** n if $n = f(x, y)$ for some $x, y \in \mathbb{Z}$. We say f **properly represents** n if $n = f(x, y)$ for some coprime $x, y \in \mathbb{Z}$.

Remark. Equivalent forms properly represent the same integers, since if $X = \alpha x + \gamma y, Y = \beta x + \delta y$ with $\alpha, \beta, \gamma, \delta \in \mathbb{Z}$, then $\alpha\delta - \beta\gamma = 1$ implies $\gcd(X, Y) = 1 \iff \gcd(x, y) = 1$.

Lemma 4.7. The smallest integers properly represented by a reduced positive definite BQF $f = (a, b, c)$ are $a, c, a - |b| + c$ in that order.[1]

Proof. f reduced $\implies |b| \le a \le c \implies a \le c \le a - |b| + c$. We have $f(1, 0) = a$, $f(0, 1) = c$. If $x = 0$, then $\gcd(x, y) = 1 \implies y = \pm 1$. Likewise, if $y = 0$, then $x = \pm 1$.

So it remains to show that the smallest number represented by f using nonzero x, y is $a - |b| + c$. But if $|x| \ge |y| \ge 1$, then

$$f(x, y) = ax^2 + bxy + cy^2 \ge ax^2 - |b||x||y| + cy^2 \ge (a - |b|)x^2 + cy^2 \ge a - |b| + c.$$

We can achieve equality with $f(1, \pm 1)$. We proceed similarly if $|y| \ge |x| \ge 1$. □

[1] Values on this list are repeated if they are represented in more than one way, not counting repeats of the form $f(x, y) = f(-x, -y)$.

Theorem 4.8. Every positive definite BQF is equivalent to a unique reduced form.

Proof. Existence follows from Lemma 4.5.

Uniqueness: Suppose $f = (a, b, c)$ and $g = (a', b', c')$ are equivalent reduced BQFs. We want to show $a = a', b = b', c = c'$. By Lemma 4.7, $a = a', c = c'$ and $a - |b| + c = a' - |b'| + c'$, so $(a, b, c) = (a', \pm b', c')$.

If $b = 0$, we're done. If $b \neq 0$, can (a, b, c) and $(a, -b, c)$ both be reduced? If yes, then $a < c$ (since $a = c$ requires $b \geq 0$ by definition) and $|b| < a$ (since we can't have $b = -a$). Hence $a < c < a - |b| + c$. By Lemma 4.7 again, $f(x, y) = a \iff (x, y) = (\pm 1, 0)$ and $f(x, y) = c \iff (x, y) = (0, \pm 1)$, and likewise for g.

Suppose $g(x, y) = f(\alpha x + \gamma y, \beta x + \delta y) = f(X, Y)$. Then

$$(X, Y) = (\pm 1, 0) \iff (x, y) = (\pm 1, 0)$$
$$(X, Y) = (0, \pm 1) \iff (x, y) = (0, \pm 1),$$

i.e. $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}$. But $\alpha\delta - \beta\gamma = 1$, so $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, so $f = g$ as required. □

01 Nov 2022, Lecture 12

Question. How many reduced forms are there with a given discriminant?

Example 4.5. Consider $d = -24$. We want to find $f = (a, b, c)$ reduced with $b^2 - 4ac = -24$. By Lemma 4.6, $|b| \leq a \leq \sqrt{8}$ and b is even.

- If $a = 1$, then $b = 0$ and hence $c = 6 \implies (1, 0, 6)$. We can check that this is reduced.
- If $a = 2$, then $c = \frac{b^2+24}{8}$.
  - If $b = 0$, then $c = 3$. This is reduced.
  - If $b = \pm 2$, then $c \notin \mathbb{Z}$.

So the only reduced forms with discriminant -24 are $(1, 0, 6)$ and $(2, 0, 3)$.

More generally, Lemma 4.6 shows that for every d, there are only finitely many reduced forms with discriminant d.

Definition 4.8. The **class number** of d, denoted $h(d)$ is the number of equivalence classes of positive definite BQFs with discriminant d.

By Theorem 4.8, this is the number of reduced forms with discriminant d, hence finite by the last remark.

Example 4.6. As we have already seen, $h(-4) = 1, h(-24) = 2$.

Definition 4.9. $d \equiv 0, 1 \pmod 4$ is a **fundamental discriminant** if it is not of the form $d = k^2 d_1$ for some integer $k \geq 1$ and $d_1 \equiv 0, 1 \pmod 4$.

Aside:

Remark. Let $d < 0$ be a fundamental discriminant. Gauss defined a group law on the set of equivalence classes of positive definite BQFs with discriminant d. The abelian group obtained in this way is the same as the class group of the field $\mathbb{Q}(\sqrt{d})$ (see Part II Number Fields). We insisted that $\alpha\delta - \beta\gamma = 1$ in the definition of equivalence (not just $= \pm 1$), since otherwise inverse elements in the class group would be the same element, hence it is no longer a group. End of aside.

Some theorems about class numbers.

(i) (Mertens 1874).

$$\sum_{-X<d<0} h(d) \sim \frac{\pi}{18} X^{\frac{3}{2}} \text{ as } X \to \infty.$$

(ii) (Heilbronn 1934) $h(d) \to \infty$ as $|d| \to \infty$.

(iii) (Siegel 1935) For every $\epsilon > 0, \exists c > 0$ such that $h(d) > c|d|^{\frac{1}{2}-\epsilon}$.

(iv) (Baker-Stark 1967) $h(d) = 1 \iff d \in \{-3, -4, -7, -8, -11, -19, -43, -67, -163\}$.

End of aside.

Lemma 4.9. Let f be a BQF and $n \in \mathbb{Z}$. Then f properly represents n if and only if f is equivalent to a form with first coefficient n.

Proof. $\Longleftarrow$: Suppose $f \sim g(n, b, c)$. Then $g(1, 0) = n \implies g$ properly represents n, so f properly represents n.

$\Longrightarrow$: $f(\alpha, \beta) = n$ for some $\alpha, \beta \in \mathbb{Z}$ coprime. By Euclid's algorithm, $\exists \gamma, \delta \in \mathbb{Z}$ such that $\alpha\delta - \beta\gamma = 1$. Then f is equivalent to $g(x, y) = f(\alpha x + \gamma y, \beta x + \delta y)$ with first coefficient $g(1, 0) = f(\alpha, \beta) = n$. □

Theorem 4.10. Let n be a positive integer and $d < 0$ a discriminant. Then n is properly represented by some positive definite BQF with discriminant d if and only if the congruence

$$x^2 \equiv d \pmod{4n}$$

is soluble.

Proof. $\Longrightarrow$: Lemma 4.9 shows $f \sim g$ with $g = (n, b, c)$. Then

$$d = \operatorname{disc}(f) = \operatorname{disc}(g) = b^2 - 4nc \equiv b^2 \pmod{4n}.$$

$\Longleftarrow$: We are given $b, c \in \mathbb{Z}$ such that $b^2 = d + 4nc$. Then $f = (n, b, c)$ is a form of discriminant d and it properly represents n (with $x = 1, y = 0$). $\square$

Example 4.7. Which integers are properly represented by $f(x, y) = x^2 + xy + 2y^2$?

We have $\mathrm{disc}(f) = -7$, so f is positive definite. By Lemma 4.6, any reduced form with discriminant -7 satisfies $|b| \leq a \leq 1$ and b is odd. Hence $(a, b, c) = (1, 1, 2)$ or $(a, b, c) = (-1, -1, 2)$. But the second one is not reduced, hence $h(-7) = 1$ and all positive definite BQFs with discriminant -7 are equivalent.

Hence n is properly represented by $x^2 + xy + 2y^2$ if and only if $x^2 \equiv -7 \pmod{4n}$ is soluble.

03 Nov 2022, Lecture 13

Assume $n = p$ is prime and $p \neq 2, 7$. By CRT, the above is equal to

$$\begin{cases} x^2 \equiv -7 \pmod 4. \text{ This is soluble.} \\ x^2 \equiv -7 \pmod p. \text{This is soluble } \iff \left(\frac{-7}{p}\right) = 1. \end{cases}$$

But $\left(\frac{-7}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{7}{p}\right) = (-1)^{\frac{p-1}{2}}(-1)^{\frac{p-1}{2}}\left(\frac{p}{7}\right) = \left(\frac{p}{7}\right)$.

We conclude that $p = x^2 + xy + 2y^2$ for some $x, y \in \mathbb{Z}$ means that $p \equiv 1, 2, 4 \pmod 7$ or $p = 2, 7$ (we check $p = 2, 7$ separately).

Lemma 4.11. Let p be an odd prime and $a \in \mathbb{Z}$. If $\left(\frac{a}{p}\right) = 1$, then the congruence $x^2 \equiv a \pmod{p^n}$ is soluble $\forall n \geq 1$.

Proof. Induction on n. The case $n = 1$ is clear.

Now let $n \geq 1$ and suppose $x^2 \equiv a \pmod{p^n}$, i.e. $x^2 = a + kp^n, k \in \mathbb{Z}$. For $t \in \mathbb{Z}$, we have $(x + tp^n)^2 \equiv x^2 + 2xtp^n \equiv a + (2xt + k)p^n \pmod{p^{n+1}}$. Now we have $(2x, p) = 1$, so we can solve $2xt + k \equiv 0 \pmod p$, so we're done. $\square$

Remark. A similar argument shows that $a \in \mathbb{Z}$ with $a \equiv 1 \pmod 8$, then $x^2 \equiv a \pmod{2^n}$ is soluble $\forall n \geq 1$.

Above example continued: Write $n = 2^\alpha 7^\beta p_1^{\gamma_1} \dots p_r^{\gamma_r}$ for p_i distinct powers. Then

$$x^2 \equiv -7 \pmod{4n} \text{ is soluble} \iff \begin{cases} x^2 \equiv -7 \pmod{2^{\alpha+2}} \text{ is soluble.} \\ x^2 \equiv -7 \pmod{7^\beta} \text{ is soluble.} \\ x^2 \equiv -7 \pmod{p_i^{\gamma_i}} \text{ is soluble } \forall 1 \leq i \leq r. \end{cases}$$

The first condition is always true by the remark above. The second one has no solutions mod 49, so hence $\beta \leq 1$. For the last condition, use the above lemma to get that we need $\left(\frac{-7}{p_i}\right) = 1 \ \forall 1 \leq i \leq r$.

Hence we want $7^2 \nmid n$ and all primes $p \mid n$ with $p \neq 7$ satisfy $p \equiv 1, 2, 4 \pmod 7$.

The integers represented by $x^2 + xy + y^2$ (not necessarily properly) are then of the form $k^2 n$ for $k \in \mathbb{Z}$ and n as described above.

Conclusion. $n = x^2 + xy + 2y^2$ for some $x, y \in \mathbb{Z} \iff$ every prime $p \equiv 3, 5, 6$ which divides n divides it to an even power.

Remarks.

(i) If $h(d) = 1$, we have shown how to solve the problem of which integers are represented by a given form of discriminant $d < 0$.

If $h(d) > 1$, we can determine which integers are represented by *some* form of discriminant d. For some values of d we can still distinguish which forms represent which numbers using congruence conditions.

(ii) What about quadratic forms in more variables?

Theorem 4.12 (Lagrange 1770)**.** Every positive integer is a sum of four squares.

Theorem 4.13 (Legendre 1797)**.** A positive integer n is a sum of 3 squares if and only if $n \neq 4^a(8b + 7)$ for some integers $a, b \geq 0$.

(iii) A geometric way to think about reduction: Let $f(x, y) = ax^2 + bxy + cy^2$ be a positive definite BQF, so $d = b^2 - 4ac < 0$. Let $\tau \in \mathbb{C}$ with $f(\tau, 1) = 0$ and $\mathrm{Im}(\tau) > 0$, so $\tau = \frac{-b \pm \sqrt{|d|} i}{2a}$, and $|\tau|^2 = \frac{b^2 - d}{4a^2} = \frac{c}{a}$.
So $|b| \leq a \leq c \iff |\mathrm{Re}(\tau)| \leq \frac{1}{2}$ and $|\tau| \geq 1$. Let $\mathcal{F}$ be this subregion of $\mathbb{C}$. Then $SL_2(\mathbb{Z}) \subset SL_2(\mathbb{R})$ acts on $\mathcal{H} = \{\tau \in \mathbb{C} \mid \mathrm{Im}(\tau) > 0\}$ via $\begin{pmatrix} a & b \\ c & d \end{pmatrix} : \tau \to \frac{a\tau + b}{c\tau + d}$, and the operations S and $T_\pm$ in the proof of Lemma 4.5 correspond to the Möbius maps $S : \tau \mapsto \frac{-1}{\tau}$ and $T_\pm : \tau \mapsto \tau \pm 1$. So we just start somewhere in the complex plane and apply these transformations until we end up in $\mathcal{F}$.

(iv) Extra conditions in the definition of a reduced form correspond to conditions concerning the boundary of $\mathcal{F}$.

5 The distribution of primes

Define $\pi(x)$ to the number of primes $\leq x$. In lecture 2, we saw that $\pi(x) \to \infty$ as $x \to \infty$ (by Euclid). On Example Sheet 1, we saw $\pi(x) \geq \frac{\log x}{\log \log x}$ if $x \geq 8$. 05 Nov 2022, Lecture 14

Lemma 5.1. $\exists c > 0$ such that $\pi(x) > c \log x$.

Proof. For $n \leq x$ we can write $n = k^2 p_1^{\alpha_1} p_r^{\alpha_r}$ with $k \leq \sqrt{x}$, p_i all the primes $\leq x$ and $\alpha_i \in \{0, 1\}$ (so $p_1^{\alpha_1} p_r^{\alpha_r}$ is squarefree).

There are $\leq \sqrt{x}$ choices for k and $\leq 2^r$ choices for $\alpha_1, \ldots, \alpha_r$, so

$$x \leq \sqrt{x} 2^{\pi(x)} \implies \pi(x) \geq \frac{\log x}{2 \log 2}.$$

□

The following result gives another proof of the infinitude of primes.

Theorem 5.2. $\sum_p \frac{1}{p}$ diverges and $\prod_p (1 - \frac{1}{p})^{-1}$ diverges.

Proof. For $x \geq 2$, we define $P(x) = \prod_{p \leq x} (1 - \frac{1}{p})^{-1}$ and $S(x) = \sum_{p \leq x} \frac{1}{p}$. We show that $P(x) \to \infty$ and $S(x) \to \infty$ as $n \to \infty$.

(i) Let $p_1, \ldots, p_r$ be the primes $\leq x$. Then

$$P(x) = \prod_{i=1}^{r} (1 + \frac{1}{p_i} + \frac{1}{p_i^2} + \ldots) = \sum_{\alpha_1 = 0}^{\infty} \cdots \sum_{\alpha_r = 0}^{\infty} \frac{1}{p_1^{\alpha_1} \ldots p_r^{\alpha_r}} \geq \sum_{n=1}^{\lfloor x \rfloor} \frac{1}{n} \overset{n \to \infty}{\to} \infty.$$

(ii)

$$\log P(x) = - \sum_{i=1}^{r} \log \left(1 - \frac{1}{p_i} \right) \overset{(\star)}{=} \sum_{i=1}^{r} \sum_{m=1}^{\infty} \frac{1}{m p_i^m} = S(x) + \sum_{i=1}^{r} \sum_{m=2}^{\infty} \frac{1}{m p_i^m}$$

where $(\star)$ follows from the Taylor series expansion of $\log(1 + x)$. But $\sum_{m=2}^{\infty} \frac{1}{p^m} = \frac{p^{-2}}{1 - p^{-1}} = \frac{1}{p(p-1)}$, so

$$0 < \log P(x) - S(x) < \frac{1}{2} \sum_{i=1}^{r} \frac{1}{p_i (p_i - 1)} \leq \frac{1}{2} \sum_{n=2}^{\infty} \frac{1}{n(n-1)} = \frac{1}{2}.$$

Thus $S(x) \to \infty$ as $n \to \infty$.

□

Remark. $\sum_{n=1}^{\lfloor x \rfloor} \frac{1}{n} > \int_1^{\lfloor x \rfloor + 1} \frac{\mathrm{d}u}{u} = \log(\lfloor x \rfloor + 1) \geq \log x$. So the proof of (i) shows $P(x) > \log(x)$ and the proof of (ii) shows $S(x) > \log \log x - \frac{1}{2}$. This is a rather good approximation:

Theorem 5.3 (Mertens 1874)**.** There exists a constant B such that $S(x) = \log \log x + B + O(\frac{1}{\log x})$.

Proof. Omitted, but a key ingredient is the following theorem which we will later prove. □

Theorem 5.4 (Tchebychev 1852)**.** There exist constants $a, b > 0$ such that $\frac{ax}{\log x} < \pi(x) < \frac{bx}{\log x}$.

Lemma 5.5. If $\frac{\pi(x)\log x}{x}$ tends to a limit as $x \to \infty$, then that limit must be 1.

Proof.

$$S(x) = \sum_{p\leq x}\frac{1}{p} = \sum_{n\leq x}\frac{\pi(n)-\pi(n-1)}{n} = \sum_{n=2}^{\lfloor x\rfloor -1}\pi(x)\left(\frac{1}{n}-\frac{1}{n-1}\right)+\frac{\pi(x)}{\lfloor x\rfloor} =$$
$$\sum_{n=2}^{\lfloor x\rfloor -1}\int_n^{n+1}\frac{\pi(u)}{u^2}\mathrm{d}u + \int_{\lfloor x\rfloor}^x\frac{\pi(u)}{u^2}\mathrm{d}u + \frac{\pi(x)}{x} = \frac{\pi(x)}{x}+\sum_2^x\frac{\pi(u)}{u^2}\mathrm{d}u.$$

If $\frac{\pi(x)\log x}{x} \to \alpha$ as $x \to \infty$, then we get

$$S(x) \sim \alpha\int_2^x\frac{\mathrm{d}u}{u\log u} = \alpha\left[\log\log u\right]_2^x \implies S(x)\sim\alpha\log\log x.$$

By Theorem 5.2, $\alpha \geq 1$, but by Mertens (Theorem 5.3), $\alpha = 1$. $\square$

Theorem 5.6 (Prime Number Theorem)**.**

$$\pi(x) \sim \frac{x}{\log x}.$$

Remarks.

- Equivalently, this says $\frac{\pi(x)\log x}{x} \to 1$ as $x \to \infty$.
- This was proved independently by Hadamard and de la Vallee Poussin.
- The proof uses the Riemann zeta function and complex analysis.

Definition 5.1 (Riemann zeta function)**.** For $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$, we say

$$\zeta(s) = \sum_{n=1}^{\infty}\frac{1}{n^s}.$$

Remark. In this context, the convention is to write $s = \sigma + it$.

Lemma 5.7. For $\mathrm{Re}(s) > 1$, the series $\sum_{n=1}^{\infty}\frac{1}{n^s}$ converges absolutely. Moreover, for any $\delta > 0$, it converges uniformly on $\mathrm{Re}(s) \geq 1+\delta$ (and hence is analytic on $\mathrm{Re}(s) > 1$).

Proof. For $s = \sigma + it$, we have

$$|n^s| = |n^{\sigma+it}| = |e^{(\sigma+it)\log n}| = e^{\sigma\log n} = n^{\sigma}.$$

But $\sum_{n=1}^{\infty} \frac{1}{n^\sigma}$ converges for $\sigma > 1$, and it converges uniformly for $\sigma \geq 1 + \delta$ (by IA Analysis). ☐

The following result links ζ to the primes.

Proposition 5.8 (Euler product for ζ)**.** For $\mathrm{Re}(s) > 1$, we have

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

Proof. The rough idea:

$$\prod_p \left(1 - \frac{1}{p^s}\right)^{-1} = \prod_p \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \dots\right) \stackrel{(\star)}{=} \sum_{n=1}^{\infty} \frac{1}{n^s}$$

where $(\star)$ follows from the Fundamental Theorem of Arithmetic.

In detail: Fix s with $\mathrm{Re}(s) > 1$. If $M > \frac{\log N}{\log 2}$, then $p^M > N$ $\forall$primes p. Now:

$$\prod_{p \leq N} \sum_{j=0}^{M} \frac{1}{p^{js}} = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \dots + \frac{1}{N^s} + \left(\text{extra terms } \frac{1}{n^s} \text{ for } n > N\right).$$

Hence

$$\left| \sum_{n=1}^{\infty} \frac{1}{n^s} - \prod_{p \leq N} \sum_{j=0}^{m} \frac{1}{p^{js}} \right| \leq \sum_{n=N+1}^{\infty} \frac{1}{n^\sigma}.$$

Take the limit as $M \to \infty$ to get

$$\left| \zeta(s) - \prod_{p \leq N} \left(1 - \frac{1}{p^s}\right)^{-1} \right| = \sum_{n=N+1}^{\infty} \frac{1}{n^\sigma} \stackrel{N \to \infty}{\to} 0.$$

☐

08 Nov 2022, Lecture 15

Corollary 5.9. If $\mathrm{Re}(s) > 1$, then $\zeta(s) \neq 0$.

Proof. If $\mathrm{Re}(s) > 1$, then

$$\left[\prod_{p \leq N} \left(1 - \frac{1}{p^s}\right)\right] \zeta(s) = \prod_{p > N} \left(1 - \frac{1}{p^s}\right)^{-1} = \prod_{p > N} \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \dots\right)$$

$$\implies \left| \prod_{p \leq N} \left(1 - \frac{1}{p^s}\right) \zeta(s) \right| \geq 1 - \sum_{n=N+1}^{\infty} \frac{1}{n^\sigma} \stackrel{N \to \infty}{\to} 1.$$

Hence $\zeta(s) \neq 0$. $\square$

Theorem 5.10. $\zeta(s) - \frac{1}{s-1}$ has an analytic continuation to $\mathrm{Re}(s) > 0$.

Proof. If $\mathrm{Re}(s) > 2$ we have

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{n-(n-1)}{n^s} = \sum_{n=1}^{\infty} n \left(\frac{1}{n^s} - \frac{1}{(n+1)^s} \right) = s \sum_{n=1}^{\infty} n \int_n^{n+1} \frac{\mathrm{d}x}{x^{s+1}} =$$
$$= s \int_1^{\infty} \frac{\lfloor x \rfloor}{x^{s+1}} \mathrm{d}x = s \int_1^{\infty} \frac{\mathrm{d}x}{x^s} - s \int_1^{\infty} \frac{\{x\}}{x^{s+1}} \mathrm{d}x = \frac{s}{s-1} - s \int_1^{\infty} \frac{\{x\}}{x^{s+1}} \mathrm{d}x.$$

Since $\{x\}$ is bounded, the second integral converges to an analytic function for $\mathrm{Re}(s+1) > 1$, i.e. $\mathrm{Re}(s) > 0$. $\square$

For $\mathrm{Re}(s) > 0$, the **Gamma function** is defined as

$$\Gamma(s) = \int_0^{\infty} e^{-x} x^{s-1} \mathrm{d}x.$$

This can be extended to a meromorphic[2] function on $\mathbb{C}$ with simple poles at $s = 0, -1, -2, \ldots$ using the rule $s\Gamma(s) = \Gamma(s+1)$. For an integer $n \geq 1$, $\Gamma(n) = (n-1)!$.

Theorem 5.10 tells us that ζ extends to a meromorphic function on the set $\{s \in \mathbb{C} \mid \mathrm{Re}(s) > 0\}$ with just one pole at $s = 1$ with residue 1. In fact, ζ extends to a meromorphic function on $\mathbb{C}$ and there are no further poles.

Moreover, the completed zeta function

$$\Xi(s) = \pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) \zeta(s)$$

satisfies the functional equation $\Xi(1-s) = \Xi(s)$.

ζ has trivial zeroes at $s = -2, -4, -6, \ldots$. By Corollary 5.9 and the functional equation, any further zeroes lie in the critical strip $0 \leq \mathrm{Re}(s) \leq 1$.

The key step in the proof of the Prime Number Theorem is showing that $\zeta(s) \neq 0$ for $\mathrm{Re}(s) = 1$.

Theorem 5.11 (The Riemann Hypothesis)**.** All zeroes of ζ in the critical strip lie on the line $\mathrm{Re}(s) = \frac{1}{2}$.

Proof. lol $\square$

RH is equivalent to

$$|\pi(x) - \mathrm{li}(x)| \leq \sqrt{x} \log x \ \forall x \geq 3,$$

[2] Analytic except on a set of isolated points.

where $\mathrm{li}(t) = \int_2^x \frac{\mathrm{d}t}{\log t}$. Integrating by parts shows $\mathrm{li}(x) \sim \frac{x}{\log x}$. Numerical evidence suggested to Gauss that $\mathrm{li}(x)$ is a better approximation to $\pi(x)$ than $\frac{x}{\log x}$. We have $\pi(x) < \mathrm{li}(x)\ \forall x \leq 10^{21}$, but Littlewood showed that $\pi(x) - \mathrm{li}(x)$ changes sign infinitely often.

A **Dirichlet series** is a series of the form

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

where $(a_i) \in \mathbb{C}$.

A useful tool for manipulating the all the aforementioned series is the Möbius function. Let $f : \mathbb{N} \to \mathbb{C}$ be any function. Define $g : \mathbb{N} \to \mathbb{C}$ by

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

Question. How do we compute f from g?

Let's compute $f(6)$. We have

$$\begin{aligned} g(1) &= f(1) \\ g(2) &= f(1) + f(2) & \implies f(6) = g(6) - g(3) - g(2) + g(1). \\ g(3) &= f(1) + f(3) \\ g(6) &= f(1) + f(2) + f(3) + f(6) \end{aligned}$$

Definition 5.2. The **Möbius function** $\mu : \mathbb{N} \to \{0, \pm 1\}$ is defined by

$$\mu(n) = \begin{cases} (-1)^k & \text{if } n = p_1 p_2 \dots p_k \text{ is a product of distinct primes.} \\ 0 & \text{if } n \text{ is not square-free.} \end{cases}$$

Remark. We have $\mu(1) = 1$.

Exercise. μ is a multiplicative function. (This is on ES3).

Let $\nu(n) = \sum_{d \mid n} \mu(d)$. By Lemma 2.8, ν is multiplicative. But $\nu(p^r) = \mu(1) + \mu(p) = 1 - 1 = 0$, so $\sum_{d \mid n} \mu(d) = \begin{cases} 1 & \text{if } n = 1. \\ 0 & \text{otherwise.} \end{cases}$

Proposition 5.12. If $g(n) = \sum_{d \mid n} f(d)$, then $f(n) = \sum_{m \mid n} \mu(m) g\left(\frac{n}{m}\right)$.

Proof.

$$\sum_{m|n}\mu(m)g\left(\frac{n}{m}\right)=\sum_{m|n}\mu(m)\sum_{d|\frac{n}{m}}f(d)=$$

$$=\sum_{d|n}\left(\sum_{m|\frac{n}{d}}\mu(m)\right)f(d)=\sum_{d|n}\nu\left(\frac{n}{d}\right)f(d)=f(n).$$

□

10 Nov 2022, Lecture 16

Notation. Let $n\in\mathbb{N}$ and p a prime. Then $\nu_p(n)$ denotes the exponent of p in the prime factorization of n.

Remarks.

- We can write $n=p^{\nu_p(n)}b$ for $p\nmid b$.
- $\nu_p(mn)=\nu_p(m)+\nu_p(n)\ \forall m,n\in\mathbb{N}$.
- $\nu_p(n!)=\sum_{j=1}^{\infty}\left\lfloor\frac{n}{p^j}\right\rfloor$ (this is also on ES3).

Proposition 5.13. Let $n\in\mathbb{N}$. Let $N=\binom{2n}{n}$.

(i) We have
$$\frac{2^{2n}}{2n}\le N\le 2^{2n}.$$

(ii) If $p^k\mid N$, then $p^k\le 2n$.

(iii) We have
$$n^{\pi(2n)-\pi(n)}\le N\le(2n)^{\pi(2n)}.$$

Proof. (i)
$$(1+1)^{2n}=\sum_{j=0}^{2n}\binom{2n}{j}=2+\sum_{j=1}^{2n-1}\binom{2n}{j}.$$

Hence $N\le 2^{2n}\le 2+(2n-1)N\le 2nN$.

(ii) We have $N=\frac{(2n)!}{(n!)^2}$, so
$$\nu_p(N)=\nu_p((2n)!)-2\nu_p(n!)=\sum_{j=1}^{\infty}\left(\left\lfloor\frac{2n}{p^j}\right\rfloor-2\left\lfloor\frac{n}{p^j}\right\rfloor\right).$$

But for $x\in\mathbb{R}$ we have $\lfloor 2x\rfloor-2\lfloor x\rfloor=\begin{cases}0 & \text{if } \{x\}<\frac{1}{2}.\\ 1 & \text{if } \{x\}\ge\frac{1}{2}.\end{cases}$ If $p^k>2n$, then

$\left\lfloor \frac{2n}{p^k} \right\rfloor = 0$, so

$$\nu_p(N) = \sum_{j=1}^{k-1} \left(\left\lfloor \frac{2n}{p^j} \right\rfloor - 2 \left\lfloor \frac{n}{p^j} \right\rfloor \right) \leq k-1.$$

Thus if $\nu_p(N) \geq k$, then $p^k \leq 2n$.

(iii)

$$N = \frac{(2n)(2n-1)\dots(n+1)}{n(n-1)\dots 1} \geq \prod_{n<p<2n} p \geq n^{\pi(2n)-\pi(n)}.$$

But also

$$N = \prod_{p \leq 2n} p^{\nu_p(n)} \leq (2n)^{\pi(2n)}$$

by part (ii).

$\square$

Theorem 5.14 (Tchebychev)**.** $\exists c_2 > c_1 > 0$ such that $\forall x \geq 4$,

$$c_1 \frac{x}{\log x} \leq \pi(x) \leq c_2 \frac{x}{\log x}.$$

Our proof will give $c_1 = \frac{\log 2}{2} \approx 0.346$ and $c_2 = 6 \log 2 \approx 4.158$.

Proof of the upper bound. By Proposition 5.13, we have

$$n^{\pi(2n)-\pi(n)} \leq N \leq 2^{2n}$$
$$\implies \pi(2n) - \pi(n) \leq 2 \log 2 \frac{n}{\log n} \ (\star).$$

We prove by induction on k that $\pi(2^k) \leq 3\frac{2^k}{k}$ $\forall k \geq 1$ (†). This is obvious for $k \leq 6$ as $\pi(x) \leq \frac{x}{2}$ $\forall x \geq 2$ even. Induction step:

$$\pi(2^{k+1}) \overset{(\star)}{\leq} \pi(2^k) + 2\log 2 \frac{2^k}{\log(2^k)} \overset{(\dagger)}{\leq} 3\frac{2^k}{k} + 2\frac{2^k}{k} \leq 6\frac{2^k}{k+1} = 3\frac{2^{k+1}}{k+1}$$

as $\frac{5}{k} \leq \frac{6}{k+1}$ for $k \geq 5$.

$\frac{x}{\log x}$ is increasing for $\forall x \geq e$ (as its derivative is $\frac{\log x - 1}{(\log x)^2}$), so if $2^k \leq x \leq 2^{k+1}$, then

$$\pi(x) \leq \pi(2^{k+1}) \leq 3\frac{2^{k+1}}{k+1} < 6\frac{2^k}{k} = 6\log 2 \frac{2^k}{\log(2^k)} \leq 6 \log 2 \frac{x}{\log x}.$$

$\square$

Proof of the lower bound. By Proposition 5.13, we have

$$\begin{aligned} & \frac{2^{2n}}{2n} \le N \le (2n)^{\pi(2n)} \\ \implies & 2n \log 2 - \log(2n) \le \pi(2n)\log(2n) \\ \implies & \pi(2n) \ge \log 2 \frac{2n}{\log(2n)} - 1. \end{aligned}$$

So if $2n \le x \le 2n+2$, then

$$\pi(x) \ge \pi(2n) \ge \log 2 \frac{(x-2)}{\log x} - 1.$$

To complete the proof, it is enough to show that

$$\log 2 \frac{(x-2)}{\log x} - 1 \ge \frac{\log 2}{2} \frac{x}{\log x}.$$

This is equivalent to $\frac{\log 2}{2}\frac{x}{\log x} \ge 1 + \frac{2\log 2}{\log x}$, which is true for $x = 16$ and hence for all $x \ge 16$ since the LHS is increasing and the LHS is decreasing.

Finally, if $4 \le x \le 16$, then $\frac{\log 2}{2}\frac{x}{\log x} \le 2 \le \pi(x)$. □

Theorem 5.15 (Bertrand's postulate)**.** If $n > 1$ is an integer, then there exists a prime with $n < p < 2n$.

Proof. Let $N = \binom{2n}{n}$. If $\frac{2n}{3} < p \le n$, then

$$\begin{aligned} & \nu_p((2n)!) = 2 \text{ as } 2p \le 2n < 3p. \\ & \nu_p((n)!) = 1 \text{ as } p \le n < 2p. \end{aligned}$$

Hence $\nu_p(N) = 0$. Suppose Bertrand's postulate is false. Then, using Proposition 5.13 (ii),

$$N = \prod_{p \le \frac{2n}{3}} p^{\nu_p(n)} \le \prod_{p \le \sqrt{2n}} p^{\nu_p(n)} \prod_{p \le \frac{2n}{3}} p \le (2n)^{\sqrt{2n}} \prod_{p \le \frac{2n}{3}} p.$$

On Example Sheet 3 we show that $\prod_{p \le m} p = 4^m$, hence (again by Proposition 5.13)

$$\begin{aligned} & \frac{2^{2n}}{2n} \le N \le (2n)^{\sqrt{2n}} 2^{\frac{4n}{3}} \\ \implies & 2^{\frac{2n}{3}} \le (2n)^{1+\sqrt{2n}} \\ \implies & 2n \log 2 \le 3(1+\sqrt{2n}) \log 2n. \end{aligned}$$

12 Nov 2022, Lecture 17

Choose $2n = 2^{2x}$ (so $x = \frac{\log(2n)}{2\log 2}$), so

$$\begin{aligned} &\implies 2^{2x}\log 2 \le 3(1+2^x)2x\log 2 \\ &\implies 2^x \le 6x(1+2^{-x}) \end{aligned}$$

If $x > 5$, say $x = 5(y+1)$ for some $y > 0$, we get

$$\begin{aligned} &2^{5y} \le \frac{6}{32}5(y+1)(1+\frac{1}{32}) \le y+1 < e^y \\ &\implies 5y\log 2 < y, \end{aligned}$$

contradiction, so $x \le 5$ and so $n \le 2^9 = 512$. For $n < 512$ it suffices to take $2, 3, 5, 7, 13, 23, 43, 83, 163, 317, 631$. $\square$

Legendre's formula. Let p_n be the n^{th} prime.

Definition 5.3. Let $N_r(x) = |\{1 \le n \le x \mid n \text{ is coprime to } a_1, a_2, \ldots, a_r\}|$ and let $A_i = \{1 \le n \le x \mid p_i \mid n\}$, $A_i^c = \{1 \le n \le x \mid p_i \nmid n\}$.

By the inclusion–exclusion principle,

$$N_r(x) = |\bigcap_{i=1}^r A_i^c| = \lfloor x \rfloor - \sum_i |A_i| + \sum_{i<j} |A_i \cap A_j| - \ldots + (-1)^r |A_1 \cap \ldots \cap A_r| =$$

$$= \lfloor x \rfloor - \sum_i \left\lfloor \frac{x}{p_i} \right\rfloor + \sum_{i<j} \left\lfloor \frac{x}{p_i p_j} \right\rfloor - \ldots + (-1)^r \left\lfloor \frac{x}{p_1 \ldots p_r} \right\rfloor.$$

For ease of calculation, remember that $\left\lfloor \frac{x}{p_1 p_2} \right\rfloor = \left\lfloor \frac{\lfloor \frac{x}{p_1} \rfloor}{p_2} \right\rfloor$.

Theorem 5.16 (Legendre's formula)**.** Let $r = \pi(\sqrt{x})$. Then

$$\pi(x) - \pi(\sqrt{x}) + 1 = N_r(x).$$

Proof. Every composite integer $n \le x$ is divisible by some prime $\le \sqrt{x}$. So if $1 \le n \le x$, then

$$n \text{ coprime to } p_1, \ldots, p_r \iff n = 1 \text{ or } n \text{ is a prime with } \sqrt{x} < n \le x.$$

$\square$

Remark. If we set $P = p_1 \dots p_r$, then

$$N_r(x) = |\{1 \le n \le x \mid (n, P) = 1\}| =$$
$$= \sum_{n=1}^{\lfloor x \rfloor} \sum_{d|(n,P)} \mu(d) = \sum_{d|P} \mu(d) \sum_{n=1}^{\lfloor x \rfloor} \mathbb{1}_{\{d|n\}} = \sum_{d|P} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor,$$

which is the same formula as above.

Definition 5.4. A **Dirichlet series** is a series of the form $\sum_{n=1}^{\infty} \frac{a^n}{n^s}$ for some sequence $a_1, a_2, \dots \in \mathbb{C}$.

Remark. If $|a_n| \le \text{const} \cdot n^k$ for all n large enough for some k, then the series converges for $\text{Re}(s) > k + 1$.

Assuming absolute convergence, we can multiply two Dirichlet series:

$$\sum_{m=1}^{\infty} \frac{a_m}{m^s} \sum_{n=1}^{\infty} \frac{b_n}{n^s} = \sum_{N=1}^{\infty} \frac{c_N}{N^s}$$

where $N = mn$ and $c_N = \sum_{d|N} a_d b_{N/d}$.

For example, for $\text{Re}(s) > 2$, we get

$$\zeta(s)\zeta(s-1) = \sum_{m=1}^{\infty} \frac{1}{m^s} \sum_{n=1}^{\infty} \frac{n}{n^s} = \sum_{N=1}^{\infty} \frac{\sigma(N)}{N^s}$$

where $\sigma(N) = \sum_{d|N} d$.

The following until the end of the section is now non-examinable.

Definition 5.5. Define the **von Mangoldt function** as

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^r \text{ is a prime power.} \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 5.17. For $\text{Re}(s) > 1$, we have

$$\frac{\zeta'(s)}{\zeta(s)} = -\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}.$$

Proof.

$$\begin{aligned}
\zeta(s) &= \prod_p \left(1 - \frac{1}{p^s}\right)^{-1} \\
\implies \log\zeta(s) &= -\sum_p \log(1-p^{-s}) \\
\overset{\text{differentiate}}{\implies} \frac{\zeta'(s)}{\zeta(s)} &= -\sum_p \frac{(\log p)p^{-s}}{1-p^{-s}} \\
\implies \frac{\zeta'(s)}{\zeta(s)} &= -\sum_p \log p \sum_{j=1}^{\infty} p^{-js} = -\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}.
\end{aligned}$$

$\square$

Let $\mathbb{1}_{\text{prime}}(n) = \begin{cases} 1 & \text{if } p \text{ is prime.} \\ 0 & \text{otherwise.} \end{cases}$ Then $\pi(x) = \sum_{n \le x} \mathbb{1}_{\text{prime}}(x)$. We should think about the von Mangoldt function as a modified version of this indicator function. Indeed, let

$$\psi(x) = \sum_{n \le x} \Lambda(n).$$

On Example Sheet 3, we will show that $\psi(x) \sim \pi(x)\log x$ as $x \to \infty$. The Prime Number Theorem is then equivalent to $\psi(x) \sim x$ as $x \to \infty$. This is proved by integrating

$$\frac{\zeta'(s)}{\zeta(s)} \frac{x^{s+1}}{s(s+1)}$$

around a suitable contour.

Theorem 5.18 (Dirichlet's theorem on primes in arithmatic progressions, 1839)**.** Let $N > 1$ be an integer and $a \in \mathbb{Z}$ with $(a, N) = 1$. Then there are infinitely many primes p with $p \equiv a \pmod{N}$.

In other words, the arithmetic progression $a, a+N, a+2N, \dots$ contains infinitely many primes.

Let $\chi : (\mathbb{Z}/n\mathbb{Z})^\times \to \mathbb{C}^*$ be a group homomorphism. Define $\overline{\chi} : \mathbb{Z} \to \mathbb{C}$ by

$$a \mapsto \begin{cases} \chi(a) & \text{if } (a, N) = 1. \\ 0 & \text{otherwise.} \end{cases}$$

Let

$$L(s,\chi)=\sum_{n=1}^{\infty}\frac{\overline{\chi}(n)}{n^s},$$

called the Dirichlet L–function.

- It can be shown that if $\chi \neq 1$, then this converges for $\mathrm{Re}(s) > 0$.
- Like ζ, this has an Euler product

  $$L(s,\chi)=\prod_{p\nmid n}\left(1-\frac{\chi(p)}{p^s}\right)^{-1}.$$

  In a neighborhood of $s=1$, we have

  $$\log L(s,\chi)=\sum_{p\nmid N}\frac{\chi(p)}{p^s}+(\text{a function bounded near } s=1).$$

  Taking linear combinations of this formula (fixing N and varying χ), Dirichlet was able to show that

  $$\sum_{p\equiv a \pmod N}\frac{1}{p^s}\to\infty \text{ as } s\to 1,$$

  which then implies the theorem.

The key step in the proof (that we completely glossed over) is to show that $L(1,\chi)\neq 0$ for $\chi \neq 1$.

6 Continued fractions

15 Nov 2022, Lecture 18

The continued fraction algorithm systematically produces the best (for a given size of denominator) rational approximations to a given real number.

Description of the algorithm.

We take $\theta \in \mathbb{R}$ (usually $\theta > 0$), and define integers $a_0, a_1, a_2, \ldots$ as follows:

Let $a_0 = \lfloor \theta \rfloor$. Stop if $\theta = a_0$, otherwise write $\theta_0 = a_0 + \frac{1}{\theta_1}$.

Now let $a_1 = \lfloor \theta_1 \rfloor$. Stop if $\theta_1 = a_1$, otherwise write $\theta_1 = a_1 + \frac{1}{\theta_2}$. Continue analogously.

If the algorithm stops, we get a **finite** continued fraction:

$$\theta = a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\dots + \frac{1}{a_n}}}} \overset{\text{def}}{=} [a_0, a_1, a_2, \dots, a_n].$$

Otherwise, the continued fraction is **infinite** and we write $\theta = [a_0, a_1, a_2, \dots]$.

Definition 6.1. $a_0, a_1, a_2, \dots$ are called **partial quotients**.

Lemma 6.1. The continued fraction of θ is finite $\iff \theta \in \mathbb{Q}$.

Proof. ($\implies$) is clear, multiply out and we get a rational number.

($\impliedby$): Suppose $\theta \in \mathbb{Q}$, say $\theta = \frac{a}{b}$ for $a, b \in \mathbb{Z}, b > 0$. By Euclid's algorithm, write

$$\begin{aligned}
a &= a_0 b + r_1,\ 0 \le r_1 < b & \theta &= \frac{a}{b} = a_0 + \frac{r_1}{b} \\
b &= a_1 r_1 + r_2,\ 0 \le r_2 < r_1 & \theta_1 &= \frac{b}{r_1} = a_1 + \frac{r_2}{r_1} \\
r_1 &= a_2 r_2 + r_3,\ 0 \le r_3 < r_2 & \theta_2 &= \frac{r_1}{r_2} = a_2 + \frac{r_3}{r_2}.
\end{aligned}$$

We eventually get $r_n = 0$, so the algorithm stops. □

Let $[a_0, a_1, a_2, \dots]$ be an infinite continued fraction. This may be approximated by the finite continued fraction $[a_0, a_1, \dots, a_n]$.

Motivation:

$$\begin{aligned}
[a_0] &= a_0 \\
[a_0, a_1] &= a_0 + \frac{1}{a_1} = \frac{a_0 a_1 + 1}{a_1} \\
[a_0, a_1, a_2] &= a_0 + \frac{a_2}{a_1 a_2 + 1} = \frac{a_0 a_1 a_2 + a_0 + a_2}{a_1 a_2 + 1}
\end{aligned}$$

Definition 6.2. Given $a_0, a_1, a_2, \dots$, we define sequences (p_n) and (q_n) as follows:

$$\begin{aligned}
p_0 &= a_0 & q_0 &= 1 \\
p_1 &= a_0 a_1 + 1 & q_1 &= 1 \\
p_n &= a_n p_{n-1} + p_{n-2} & q_n &= a_n q_{n-1} + q_{n-2}\ \forall n \ge 2.
\end{aligned}$$

Remark. The (q_n) are an increasing sequence of positive integers.

Lemma 6.2. (i) For $n \ge 0$, we have $\frac{p_n}{q_n} = [a_0, a_1, \dots, a_n]$.

(ii) Let $\beta > 0$ be a real number. For $n \geq 2$, we have

$$\frac{\beta p_{n-1} + p_{n-2}}{\beta q_{n-1} + q_{n-2}} = [a_0, a_1, \ldots, a_{n-1}, \beta].$$

Proof. (i) Check above for $n = 0, 1$. The general case follows by part (ii) with $\beta = a_n$.

(ii) Induction on n. If $n = 2$, $[a_0, a_1, \beta] = \frac{\beta(a_0 a_1 + 1) + a_0}{\beta a_1 + 1} = \frac{\beta p_1 + p_0}{\beta q_1 + q_0}$.

Suppose the claim is now true for n. Then

$$[a_0, \ldots, a_n, \beta] = [a_0, a_1, \ldots, a_{n-1}, a_n + \frac{1}{\beta}] =$$
$$= \frac{(a_n + \frac{1}{\beta})p_{n-1} + p_{n-2}}{(a_n + \frac{1}{\beta})q_{n-1} + q_{n-2}} = \frac{\beta p_n + p_{n-1}}{\beta q_n + q_{n-1}},$$

so the result is true for $n + 1$ by induction.

□

Definition 6.3. The fraction $\frac{p_n}{q_n}$ is called a **convergent** to θ.

Lemma 6.3. (i) For $n \geq 1$, we have $p_n q_{n-1} - p_{n-1} q_n = (-1)^{n-1}$.

(ii) For $n \geq 2$, we have $p_n q_{n-2} - p_{n-2} q_n = (-1)^n a_n$.

Proof. (i) By induction on n. For $n = 1, p_1 q_0 - p_0 q_1 = (a_0 a_1 + 1) - a_0 a_1 = 1$. Assuming the claim is now true for $n - 1$, we get

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

(ii)

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

by part (i).

□

Remarks.

- Lemma 6.3 (i) shows that p_n and q_n are coprime and that
$$\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} = \frac{(-1)^{n-1}}{q_n q_{n-1}} \overset{n\to\infty}{\to} 0.$$
- Lemma 6.3 (ii) shows that
$$\frac{p_n}{q_n} - \frac{p_{n-2}}{q_{n-2}} = \frac{(-1)^n a_n}{q_n q_{n-2}}.$$
- Therefore
$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \frac{p_4}{q_4} < \ldots < \frac{p_5}{q_5} < \frac{p_3}{q_3} < \frac{p_1}{q_1}.$$
So the limit as $\lim_{n\to\infty} \frac{p_n}{q_n}$ exists. We will now show that this limit is exactly the number θ, which justifies the notation $\theta = [a_0, a_1, \ldots]$ and calling $\frac{p_n}{q_n}$ the convergent fractions.

17 Nov 2022, Lecture 19

Theorem 6.4. Let θ be an irrational number. Define a_n, p_n, q_n as above. Then for all $n \geq 0$, we have
$$\left|\theta - \frac{p_n}{q_n}\right| < \frac{1}{q_n q_{n-1}}.$$