

Number Theory

November 29, 2019

1 Euclid's Algorithm

Theorem 1.1 (Division Algorithm). *Given $a, b \in \mathbb{Z}, b > 0$, we can determine $\exists q, r \in \mathbb{Z}$ s.t. $a = qb + r$ with $0 \leq r < b$.*

Proof. Let $S = \{a - nb : n \in \mathbb{Z}\}$. S contains some non-negative integer. Let r be the least such integer, say $a - qb$. Then $a = qb + r$, so STP $r < b$.

Suppose $b \leq r$. Then $0 < r - b = a - (q + 1)b \in S$, and $r - b < r$. \nmid (choice of r) \square

If $r = 0$, i.e. if $a = qb$ for some $q \in \mathbb{Z}$, then we write $b|a$ and say “ b **divides** a ” or “ b is a **divisor** of a ”. If $r \neq 0$, then we instead write $b \nmid a$ and say “ b does **not divide** a ”.

Given $a_1, \dots, a_n \in \mathbb{Z}$ not all 0, let $I = \{\lambda_1 a_1 + \dots + \lambda_n a_n : \lambda_i \in \mathbb{Z}\}$. Observe if $a, b \in I, \ell, m \in \mathbb{Z}$, then $\ell a + mb \in I$.

Theorem 1.2. $I = d\mathbb{Z} = \{dm : m \in \mathbb{Z}\}$ for some $d > 0$

Proof. I contains some positive integer. Let $d > 0$ be the least such. Then clearly $I \supseteq d\mathbb{Z}$.

Conversely, let $a \in I$ and apply **1.1** to obtain $a = qd + r$ for some $q, r \in \mathbb{Z}, 0 \leq r < d$. Then $r = a - qd \in I \implies r = 0$, so $d\mathbb{Z} \supseteq I$

$\therefore I = d\mathbb{Z}$ \square

Note that $a_i \in I \forall i$, so $d|a_i \forall i$. Conversely, if $c|a_i \forall i$ then c divides every element of I , so in particular $c|d$.

We write $d = \gcd(a_1, \dots, a_n) = (a_1, \dots, a_n)$, and say d is the **greatest common divisor** of the a_i .

Corollary 1.3 (Bézout). *Let $a, b \in \mathbb{Z}$, a, b not both 0. Then $\exists x, y \in \mathbb{Z}$ s.t. $ax + by = c \iff (a, b)|c$.*

The division algorithm gives an efficient method for computing (a, b) .

Theorem 1.4 (Euclid's Algorithm). *Suppose $a > b > 0$. Then:*

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

and $r_k = (a, b)$

Proof. We have $r_k | r_{k-1} \implies \dots \implies r_k | a, r_k | b \implies r_k | (a, b)$, so $r_k \leq (a, b)$. Note also that any m s.t. $m | a$ and $m | b$ also divides r_k . In particular, $(a, b) | r_k$, and thus $(a, b) \leq r_k$, hence $r_k = (a, b)$. \square

Additionally, by working back up the algorithm, we can obtain a representation $(a, b) = \lambda a + \mu b$ where $\lambda, \mu \in \mathbb{Z}$

An integer $n > 1$ is **prime** if its only positive divisors are 1 and n . Otherwise, we say n is **composite**.

Corollary 1.5. *Let p be a prime, $a, b \in \mathbb{Z}$. Then $p | ab \iff p | a$ or $p | b$*

Proof. It is clear that if $p | a$ or $p | b$, then $p | ab$. Conversely, suppose $p | ab$ but $p \nmid a$. Then $(a, p) \neq p$. By definition, $(a, p) | p \implies (a, p) \in \{1, p\}$, so $(a, p) = 1$. Now by **1.3** we can find $x, y \in \mathbb{Z}$ s.t. $1 = ax + by \implies b = b(ax + py) = x(ab) + (by)p$, so $p | b$. \square

Theorem 1.6 (The Fundamental Theorem of Arithmetic). *Every integer $n > 1$ can be written as a product of primes uniquely up to reordering*

Proof. We have existence by strong induction.

For uniqueness, n is the least integer with two distinct such representations, say $n = p_1 \dots p_s = q_1 \dots q_r$ for p_i, q_j primes.

Then $p_1 | q_1 \dots q_r \implies p_1 | q_j$ for some j . WLOG $j = 1$. Since $p_1 > 1$ as 1 is non-primes, $n/p_1 < n$, and $n/p_1 = p_2 \dots p_s = q_2 \dots q_r$ can be written in two distinct ways as a product of primes. \nmid (choice of n) \square

If $m = \prod_{i=1}^k p_i^{\alpha_i}, n = \prod_{i=1}^k p_i^{\beta_i}$ where p_i are distinct primes, $\alpha_i, \beta_i \geq 0$, then $(m, n) = \prod_{i=1}^k p_i^{\gamma_i}$ with $\gamma_i = \min\{\alpha_i, \beta_i\}$. However, if m, n are large, it is much more "efficient" to compute the gcd via Euclid's algorithm.

An algorithm with input $N > 0$ is said to run in **polynomial time** if it takes at most $c(\log N)^k$ elementary operations to complete, where $c, k > 0$ are constants independent of N . If the algorithm takes inputs N_1, N_2, \dots, N_s , the polynomial time means $c(\max \log N_i)^k$.

Examples of polynomial time algorithms:

- Adding and multiplying integers
- The gcd of two numbers via Euclid's algorithm

- Testing of primality

On the other hand, factoring a number into prime factors does not have a polynomial time algorithm, and it is conjectured that one does not exist. For instance, if $N = p \cdot q$ with p, q primes of ~ 50 digits each, to do trial division up to \sqrt{N} at a rate of 2^9 divisions per second, it would take approximately $\sqrt{10^{100}}/2^9$ seconds, or about 6×10^{39} years. However, we can compute the gcd in milliseconds using Euclid's algorithm.

Theorem 1.7. *There are infinitely many primes. i.e. $\pi(x) \rightarrow \infty$ as $x \rightarrow \infty$*

Proof. Fix $N > 1$, let p be the largest prime $\leq N$. Let q be a prime factor of $M = (2 \times 3 \times 5 \times \dots \times p) + 1$. Then $q > N$ since $q \notin \{2, 3, \dots, p\}$, but N was arbitrary. \square

2 Congruences

Let $n \geq 1$ be an integer. We write $a \equiv b \pmod{n}$ if $n \mid a - b$. This defines an equivalence relation on \mathbb{Z} , and we will write $\mathbb{Z}/n\mathbb{Z}$ for the equivalence classes induced by this relation, which are $a + n\mathbb{Z}$ for $0 \leq a < n$. It is easy to check that $(a + n\mathbb{Z}) + (b + n\mathbb{Z}) = (a + b + n\mathbb{Z})$ and that $(a + n\mathbb{Z}) \times (b + n\mathbb{Z}) = (ab + n\mathbb{Z})$ are well defined operations, i.e. $n\mathbb{Z}$ is an ideal, and $\mathbb{Z}/n\mathbb{Z}$ is the quotient ring.

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

1. $(a, n) = 1$
2. $\exists b \in \mathbb{Z}$ s.t. $ab \equiv 1 \pmod{n}$
3. The equivalence class of a generates the group $(\mathbb{Z}/n\mathbb{Z}, +)$

Proof.

- (1) \implies (2): $(a, n) = 1 \implies \exists b, c \in \mathbb{Z}$ s.t. $ab + cn = 1$ by **1.3**, and hence $ab \equiv 1 \pmod{n}$.
- (2) \implies (1): $ab \equiv 1 \pmod{n} \iff ab - 1 = kn$ for some $k \in \mathbb{Z}$, and so by **1.3** $(a, n) = 1$.
- (2) \iff (3): $ab \equiv 1 \pmod{n} \iff 1 \in \langle a \rangle \leq \mathbb{Z}/n\mathbb{Z} \iff \langle a \rangle = \mathbb{Z}/n\mathbb{Z}$

\square

We write $(\mathbb{Z}/n\mathbb{Z})^\times$ for the set of **units** (the elements with a multiplicative inverse) of $\mathbb{Z}/n\mathbb{Z}$. By **2.1**, $(\mathbb{Z}/n\mathbb{Z})^\times$ contains precisely those classes $a + n\mathbb{Z}$ such that $(a, n) = 1$. Note that if $n > 1$ then $\mathbb{Z}/n\mathbb{Z}$ is a field precisely when n is prime.

Let **Euler's φ function** be $\varphi(n) := |(\mathbb{Z}/n\mathbb{Z})^\times|$ for $n > 1$, and let $\varphi(1) = 1$. Observe that $\varphi(p) = p - 1$ for p prime. Moreover, φ is a multiplicative function: $(m, n) = 1 \implies \varphi(mn) = \varphi(m)\varphi(n)$.

Corollary 2.2. *Let G be a cyclic group of order $n \geq 1$. Then $\varphi(n) = |\{g \in G : \text{ord}(g) = n\}|$, the number of generators of G .*

Theorem 2.3 (Euler-Fermat). *IF $(a, n) = 1$, $a, n \in \mathbb{Z}$, then $a^{\varphi(n)} \equiv 1 \pmod{n}$*

Proof. By Lagrange's Theorem, the order of a in the group $(\mathbb{Z}/n\mathbb{Z})^\times$ divides the order of $(\mathbb{Z}/n\mathbb{Z})^\times$, which is $\varphi(n)$ \square

Theorem 2.4 (Fermat's Little Theorem). *If $a, p \in \mathbb{Z}$ and p is prime, then $a^p \equiv a \pmod{p}$.*

Proof. If $p|a$, then this holds trivially. If $p \nmid a$, $(a, p) = 1$ and so by **2.3** we have $a^{\varphi(p)} \equiv a^{p-1} \equiv 1 \pmod{p}$ \square

Multiple Congruences

Can we find all $x \in \mathbb{Z}$ s.t. $x \equiv 4 \pmod{7}$ and $x \equiv 5 \pmod{12}$?

Suppose we can find $u, v \in \mathbb{Z}$ s.t. $\begin{cases} u \equiv 1 \pmod{7}; & u \equiv 0 \pmod{12} \\ v \equiv 0 \pmod{7}; & v \equiv 1 \pmod{12} \end{cases}$. Then we can write down

that $x = 4u + 5v$. Since $(7, 12) = 1$, by **1.3** there are some $m, n \in \mathbb{Z}$ with $7m + 12n = 1$, and from Euclid's algorithm we can determine these to be $m = -5, n = 3$. Then we can find $u = 12n = 1 - 7m; v = 7m = 1 - 12n$, and substitution gives $u = 36, v = -35$, and so a solution to the original problem is $4 \times 36 - 5 \times 35 = -31$. Now the lowest common multiple of 7 and 12 is 84, and so our solution set is: $\{x \in \mathbb{Z} : x \equiv -31 \pmod{84}\}$.

We can in fact generalise this process:

Theorem 2.5 (Chinese Remainder Theorem). *Let m_1, \dots, m_k be pairwise coprime positive integers, and let $M = \prod_{i=1}^k m_i$. Then given any integers a_1, \dots, a_k there is a solution x to the system of congruences:*

$$\begin{aligned} x &\equiv a_1 \pmod{m_1} \\ x &\equiv a_2 \pmod{m_2} \\ &\vdots \\ x &\equiv a_k \pmod{m_k} \end{aligned}$$

Moreover, this solution is unique modulo M .

Note that if x satisfies this system of equations, then so does $x + tM$ for any $t \in \mathbb{Z}$, and so the complete set of solutions is $x + M\mathbb{Z}$.

Proof.

Uniqueness: If x, y satisfy the system, then $m_i|x - y$ for all $i = 1, \dots, k$. Since no prime divides any two of the m_i , $M|x - y$ and hence $x \equiv y \pmod{M}$.

Existence: Write $M_i = \frac{M}{m_i} = \prod_{j \neq i} m_j$ for each $i = 1, 2, \dots, k$. Since $(m_i, m_j) = 1 \forall i \neq j$, $(m_i, M_i) = 1$ for all $i = 1, 2, \dots, k$. Therefore, for each $i = 1, 2, \dots, k$ we can find $b_i \in \mathbb{Z}$ such that $M_i b_i \equiv 1 \pmod{m_i}$ and $M_i b_i \equiv 0 \pmod{m_j}$ for $j \neq i$. Then $x = \sum_{i=1}^k a_i b_i M_i$ solves the system of congruences. \square

If m_1, \dots, m_k are pairwise coprime, and $M = \prod m_i$, then map $\theta : \mathbb{Z}/M\mathbb{Z} \rightarrow \mathbb{Z}/m_1\mathbb{Z} \times \dots \times \mathbb{Z}/m_k\mathbb{Z}$, taking $x \pmod{M} \mapsto (x \pmod{m_1}, \dots, x \pmod{m_k})$ is an isomorphism of rings. To see this, note that if $m_i|M$ then $x \pmod{m_i}$ is determined by $x \pmod{M}$ which implies that θ is well-defined. It is a homomorphism by the properties of $+, \times$ in $\mathbb{Z}/n\mathbb{Z}$, and **2.5** implies that θ is a bijection. In particular, if $n = p_1^{\alpha_1} \dots p_k^{\alpha_k}$ for distinct primes p_i , then $\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}/p_1^{\alpha_1}\mathbb{Z} \times \dots \times \mathbb{Z}/p_k^{\alpha_k}\mathbb{Z}$.

Corollary 2.6. *If m_1, \dots, m_k are pairwise coprime and $M = \prod_{i=1}^k m_i$ and $a_1, \dots, a_k \in \mathbb{Z}$ are such that $(a_i, m_i) = 1$ for each $i = 1, 2, \dots, k$, then there is a solution to the system of congruences in **2.5**, and any such solution is in fact coprime to M .*

Proof. **2.5** gives us a solution, say $x \in \mathbb{Z}$. Suppose $(x, M) > 1$. Then there is a prime p such that $p|x$ and $p|M$ simultaneously. p prime, so WLOG suppose that p divides m_1 . Since $x \equiv a_1 \pmod{m_1}$, we must have $p|a_1$, and so $p|(a_1, m_1) \nmid$. \square

Corollary 2.7. *If m_1, \dots, m_k are pairwise coprime with $M = \prod_{i=1}^k m_i$, then $\varphi(M) = \varphi(m_1) \cdot \dots \cdot \varphi(m_k)$*

A **multiplicative function** is a function $f : \mathbb{N} \rightarrow \mathbb{C}$ such that, for all $m, n \in \mathbb{N}$ coprime, $f(mn) = f(m)f(n)$. A function $f : \mathbb{N} \rightarrow \mathbb{C}$ is said to be **totally multiplicative** if for all $m, n \in \mathbb{N}$, $f(m, n) = f(m)f(n)$.

Some multiplicative functions are:

- $\varphi(m)$
- $\tau(n)$ = the number of positive divisors of n
- $\sigma(n)$ = the sum of the positive divisors of n
- $\sigma_k(n) = \sum_{d|n} d^k$, so that $\sigma_0(n) = \tau(n), \sigma_1(n) = \sigma(n)$.

None of these are totally multiplicative however.

Lemma 2.8. *Let f be a multiplicative function. Then so is g , where $g(n) = \sum_{d|n} f(d)$.*

Proof. Let $m, n \in \mathbb{N}, (m, n) = 1$. Then the divisors of mn are precisely the integers of the form $d_1 d_2$ where $d_1|m, d_2|n$ and $(d_1, d_2) = 1$. This means that we can write down

$$\begin{aligned} g(mn) &= \sum_{d|mn} f(d) \\ &= \sum_{d_1|m} \sum_{d_2|n} f(d_1 d_2) \\ &= \sum_{d_1|m} \sum_{d_2|n} f(d_1) f(d_2) \\ &= \sum_{d_1|m} f(d_1) \sum_{d_2|n} f(d_2) \\ &= g(m)g(n) \end{aligned}$$

\square

Then if we let $f(n) = n^k$ for some $k \in \mathbb{N}$. Then $g(n) = \sum_{d|n} d^k = \sigma_k(n)$. Later on, we shall see that we can recover f from g via Möbius inversion.

Theorem 2.9.

1. *If p is a prime and $m \in \mathbb{N}$ then $\varphi(p^m) = p^{m-1}(p-1) = p^m \left(1 - \frac{1}{p}\right)$*
2. *$\forall n \in \mathbb{N}, \varphi(n) = n \cdot \prod_{p|n} \left(1 - \frac{1}{p}\right)$*
3. *$\sum_{d|n} \varphi(d) = n$*

Proof.

1.

$$\begin{aligned}\varphi(p^m) &= |\{1 \leq a \leq p^m : (a, p^m) = 1\}| \\ &= p^m - p^{m-1} \\ &= p^m \left(1 - \frac{1}{p}\right)\end{aligned}$$

2. Let $n = \prod_{i=1}^k p_i^{\alpha_i}$ for p_i distinct primes, $\alpha_i \geq 1$. Then:

$$\begin{aligned}\varphi(n) &= \prod_{i=1}^k \varphi(p_i^{\alpha_i}) \\ &= \prod_{i=1}^k p_i^{\alpha_i} \left(1 - \frac{1}{p_i}\right) \\ &= n \cdot \prod_{p|n} \left(1 - \frac{1}{p}\right)\end{aligned}$$

3. φ is multiplicative and so is $n \mapsto n$, so it suffices to check that both sides agree when n is a prime power. Let p be a prime $m \in \mathbb{N}$. Then:

$$\begin{aligned}\sum_{d|p^m} \varphi(d) &= \varphi(1) + \varphi(p) + \dots + \varphi(p^m) \\ &= 1 + (p-1) + (p^2-p) + \dots + (p^m - p^{m-1}) \\ &= p^m\end{aligned}$$

□

Polynomials over $\mathbb{Z}/n\mathbb{Z}$ can have varying numbers of solutions, e.g.:

1. $x^2 + 2 \equiv 0 \pmod{5}$ has no solutions
2. $x^3 + 1 \equiv 0 \pmod{7}$ has three solutions
3. $x^2 - 1 \equiv 0 \pmod{8}$ has four solutions

Let $R \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{Z}/n\mathbb{Z}\}$ be a ring. Then we define $R[x]$ to be the ring of polynomials with coefficients in R , with addition and multiplication given in the usual way.

WARNING: Two polynomials are *equal* if their coefficients are all equal, however the map from $R[x]$ to the set of all functions $R \rightarrow R$ is not necessarily injective. For instance, if $R = \mathbb{Z}/p\mathbb{Z}$ for some prime \mathbb{Z} , then under this map $x^p - x$ is the zero function by Fermat's little theorem.

Theorem 2.10 (Division Algorithm for Polynomials). *Let $f, g \in R[x]$, and suppose that the leading coefficient of g is a unit (i.e. has a multiplicative inverse) in R . Then $\exists q, r \in R[x]$ such that $f = q \cdot g + r$ where $\deg r < \deg g$.*

Proof. We prove this by induction on $n = \deg f$. If $\deg f < \deg g$, then just take $q = 0, r = f$. Otherwise, $f(x) = ax^n + \dots; g(x) = bx^m + \dots$ for $a, b \neq 0, n \geq m, b = c^{-1}$ for some $c \in R$.

Then define $f'(x) = f(x) - acx^{n-m}g(x)$ has degree $< n$. By the induction hypothesis, there is some $q, r \in R[x]$ such that $f'(x) = q(x)g(x) + r(x)$, with $\deg r < \deg g$.

But now $f(x) = (q(x) + acx^{n-m})g(x) + r(x)$, and we are done. □

Theorem 2.11 (Remainder Theorem). *let $f \in R[x], \alpha \in R$. Then there is some $q \in R[x]$ such that:*

$$f(x) = (x - \alpha)q(x) + f(\alpha)$$

Proof. By 2.10 with $g(x) = x - \alpha$, there is some $q \in R[x]$ and $r \in R$ such that $f(x) = (x - \alpha)q(x) + r$. But now $f(\alpha) = r$, and the required equality holds. \square

A (non-zero) ring R is said to be an **integral domain** if it doesn't have any zero divisors, i.e. $ab = 0 \iff a = 0$ or $b = 0$. Note that \mathbb{Z} and \mathbb{Q} are integral domains, whilst $\mathbb{Z}/n\mathbb{Z}$ is an integral domain if and only if n is prime (if $n = pq$ is a proper factorization, then $pq = 0$ in $\mathbb{Z}/n\mathbb{Z}$).

Theorem 2.12. *Let R be an integral domain, and let $f \in R[x]$ be a non-zero polynomial of degree $n \geq 0$. Then f has at most n roots in R .*

Theorem 2.13 (Lagrange). *Let p be a prime, and let $f(x) = a_n x^n + \dots + a_1 x + a_0 \in \mathbb{Z}[x]$ be a polynomial of degree n such that $p \nmid p$. Then the congruence $f(x) \equiv 0 \pmod{p}$ has at most n distinct solutions.*

Of 2.12. By induction on n . Check $n = 0$ - trivial.

Suppose $n > 0$. If f has no roots then we're done. Otherwise there exists $\alpha \in R$ such that $f(\alpha) = 0$, and so by the remainder theorem, $f(x) = (x - \alpha)q(x)$ with $\deg q < \deg f = n$. So by the induction hypothesis, we know that q has at most $n - 1$ roots. But if $\beta \in R$ is such that $f(\beta) = 0$, then $0 = (\beta - \alpha)q(\beta)$, and since R is an integral domain, we must have $\beta = \alpha$ or $q(\beta) = 0$, and so f has at most n roots. \square

Example: Let p be a prime, $G = \mathbb{Z}/p\mathbb{Z}$, and let $f(x) = x^{p-1} - 1 - \prod_{\alpha \in G} (x - \alpha)$. Observe that $\alpha = 1, 2, \dots, p-1$, then $f(\alpha) = \alpha^{p-1} - 1 \equiv 0 \pmod{p}$, so f has at least $p-1$ roots.

But $\deg f < p-1$ because the coefficient of $x^{p-1} = 0$. This means that f must be the zero polynomial, and hence $0 = f(0) = -1 - (p-1)! \pmod{p}$, and we have Wilson's theorem, that $(p-1)! \equiv -1 \pmod{p}$.

Example: Consider $(\mathbb{Z}/7\mathbb{Z})^\times$.

$3 \in (\mathbb{Z}/7\mathbb{Z})^\times$ since $3 \cdot 5 \equiv 1 \pmod{7}$, $3^6 \equiv 1 \pmod{7}$, so 3 generates $(\mathbb{Z}/7\mathbb{Z})^\times$, and $(\mathbb{Z}/7\mathbb{Z})^\times$ is cyclic.

Theorem 2.14. *If p is a prime, then $G = (\mathbb{Z}/p\mathbb{Z})^\times$ is cyclic and of order $p-1$.*

Proof. $|G| = \varphi(p) = p-1 = \sum_{d|p-1} \varphi(d)$

By Lagrange's Theorem, $|G| = \sum_{a|G} N_a$ where $N_a = |\{g \in G : \text{ord}(g) = d\}|$. Suppose G is not cyclic, so G does not contain an element of order $p-1$, so $N_{p-1} = 0 < \varphi(p-1)$, and so there must be some d for which $N_d > \varphi(d)$. Let α be an element of order for such a d . Then $\langle \alpha \rangle \leq G$ is cyclic of order d , so it has precisely $\varphi(d)$ elements of order d . Since $N_d > \varphi(d)$, $\exists \beta \notin \langle \alpha \rangle$ s.t. $\text{ord}(\beta) = d$. This implies that the polynomial $x^d - 1$ has $d+1$ roots, namely $1, \alpha, \dots, \alpha^{d-1}, \beta$ \nmid 2.12. \square

A positive integer is said to be a **primitive root modulo n** if $\langle g \rangle = (\mathbb{Z}/n\mathbb{Z})^\times$. Hence 2.14 says that primitive roots exist modulo p for all primes p .

For instance, take $p = 19$, and let $d = \text{ord}(2)$ in $(\mathbb{Z}/19\mathbb{Z})^\times$. Then $d \mid \varphi(19) = 18$, so $d = 18$ or $d \mid 6$ or $d \mid 9$. $2^6 = 64 \not\equiv 1 \pmod{19}$, and $2^9 = 512 \not\equiv 1 \pmod{19}$, so $d = 18$, and 2 is a primitive root modulo 19.

There are many open problems concerning primitive roots:

1. Artin's Primitive Root Conjecture:

Given $g \geq 1$ does there exist infinitely many primes p such that g is a primitive root modulo p . We do know that there are infinitely many primes for which one of $\{2, 3, 5\}$ is a primitive root.

2. How large is the smallest primitive root modulo p ?

We can prove that it is $\leq cp^{1/4+\epsilon}$ for some constant $c > 0$ and for any $\epsilon > 0$. However, conditional on the Generalised Riemann Hypothesis (GRH), it is $\leq c \log^6 p$ for constant $c > 0$

Now consider $(\mathbb{Z}/8\mathbb{Z})^\times = \{\pm 1, \pm 3\}$. All of these have order 1 or 2, and hence $(\mathbb{Z}/8\mathbb{Z})^\times$ is not cyclic. In fact, let $\theta : (\mathbb{Z}/2^k\mathbb{Z})^\times \rightarrow (\mathbb{Z}/8\mathbb{Z})^\times$ map $x \pmod{2^k}$ onto $x \pmod{8}$. Since $(a, 2^k) = 1 \iff (a, 8) = 1$, θ is surjective. Hence, for $k \geq 3$ we have that $(\mathbb{Z}/2^k\mathbb{Z})^\times$ is not cyclic, since a generator would map to a generator.

Theorem 2.15. *If $p > 2$, $(\mathbb{Z}/p^k\mathbb{Z})^\times$ is cyclic for $k \geq 1$.*

Lemma 2.16. *Let $p > 2, k \geq 1, y \in \mathbb{Z}$. Then*

1. *If $x \equiv 1 + p^k y \pmod{p^{k+1}}$, then $x^p \equiv 1 + p^{k+1} y \pmod{p^{k+2}}$*
2. *$(1 + yp)^{p^k} \equiv 1 + p^{k+1} y \pmod{p^{k+2}}$*

Proof.

$$1. \ x^p = (1 + p^k y)^p = \sum_{j=0}^p \binom{p}{j} (p^k y)^j = 1 + p^{k+1} y + \dots + p^{pk} y^p.$$

For $2 \leq j \leq p-1$, $p \mid \binom{p}{j}$, so $\binom{p}{j} (p^k y)^j \equiv 0 \pmod{p^{2k+2}}$, and so $\equiv 0 \pmod{p^{k+2}}$.

Since $p \geq 2, pk \geq k+2$, so $p^{pk} y^p \equiv 0 \pmod{p^{k+2}}$, and therefore $x^p \equiv 1 + p^{k+1} y \pmod{p^{k+2}}$

2. Let $x = 1 + py$ and apply part 1 k times.

□

Lemma 2.17. *Let $p > 2, k \geq 1$. If g is a primitive root \pmod{p} , and $g^{p-1} \not\equiv 1 \pmod{p^2}$, then g generates $(\mathbb{Z}/p^k\mathbb{Z})^\times$ for all $k \geq 1$.*

Proof. Let $d = \text{ord } g$ as a member of $(\mathbb{Z}/p^k\mathbb{Z})^\times$. Note that $\varphi(p^k) = p^{k-1}(p-1)$, and so $d \mid p^{k-1}(p-1)$.

If g is not a generator of $(\mathbb{Z}/p^k\mathbb{Z})^\times$, then one of the following holds:

1. $d \mid p^{k-2}(p-1)$
2. $d = p^{k-1}e$ where $e \mid p-1, e \neq p-1$

We tackle each of these cases individually, and will see that they cannot be the case:

1. We thus have $g^{p^{k-2}(p-1)} \equiv 1 \pmod{p^k}$. We've already seen that $g^{p-1} \equiv 1 \pmod{p}$ and $g^{p-1} \not\equiv 1 \pmod{p^2}$, and so there exists some $y \not\equiv 0$ such that $x := g^{p-1} = 1 + py$.

Then we have $x^{p^{k-2}} \equiv 1 + p^{k-1}y \pmod{p^k} \implies g^{p^{k-2}(p-1)} \equiv 1 + p^{k-1}y \pmod{p^k} \not\equiv 1 \pmod{p^k} \nmid$.

2. Here, we have $g^{p^{k-1}e} \equiv 1 \pmod{p^k}$. Fermat tells us that $g^p \equiv g \pmod{p}$, and so $g^{p^{k-1}} \equiv g \pmod{p} \implies g^{p^{k-1}e} \equiv g^e \pmod{p}$. However, $e < p$, and so this is not $1 \pmod{p}$, and hence $g^{p^{k-1}e} \not\equiv 1 \pmod{p^k} \nmid$.

Hence the only case left is that g is a generator of $(\mathbb{Z}/p^k\mathbb{Z})^\times$. \square

Proof of 2.15. Let g be a primitive root modulo p . If $g^{p-1} \not\equiv 1 \pmod{p^2}$, then $(\mathbb{Z}/p^k\mathbb{Z})^\times = \langle g \rangle \forall k \geq 1$.

Otherwise, $g^p \equiv g \pmod{p^2}$. Let $h = (1+p)g$, so that $h^p \equiv (1+p)^p g^p \equiv g \pmod{p^2}$. Observe that $g \not\equiv h \pmod{p^2}$, as g is a primitive root modulo p , so that $(g, p) = 1$.

So $h^p \not\equiv h \pmod{p^2}$, and so $\langle h \rangle = (\mathbb{Z}/p^k\mathbb{Z})^\times \forall k \geq 1$. \square

2.16 fails for $p = 2$ because of the $k = 1$ case in 1. However, it does hold if $p = 2, k \geq 2$. In particular, $(1+4)^{2^{k-1}} \equiv 1 + 2^{k+1} \pmod{2^{k+2}}$. So we have $(\mathbb{Z}/2^k\mathbb{Z})^\times = \langle -1, 5 \rangle \cong \mathbb{Z}/2^{k-2}\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ for $k \geq 3$.

3 Quadratic Residues

Let p be an odd prime, and let $a \in \mathbb{Z}$ such that $a \not\equiv 0 \pmod{p}$. We say that a is a **quadratic residue modulo p** if the congruence $x^2 \equiv a \pmod{p}$ has a solution. Otherwise, we say that a is a **quadratic non-residue modulo p** . So a is a quadratic residue mod p if and only if its residue class in $(\mathbb{Z}/p\mathbb{Z})^\times$ is a square.

Conjecture 3.1 (Open). *Let $n(p)$ be the least quadratic non-residue modulo p . We can show that $n(p) \leq cp^\theta$ for any $\theta > \frac{1}{4}\sqrt{e}$ for some constant $c > 0$, and, conditional on GRH, $n(p) \leq c \log^2 p$ for some $c > 0$*

For instance, let $p = 7$. We have $\frac{x}{x^2} \begin{array}{c|c|c|c|c|c|c} 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 1 & 4 & 2 & 2 & 4 & 1 \end{array}$.

and so the quadratic residues module 7 are $\{1, 2, 4\}$, whilst $\{3, 5, 6\}$ are non-residues.

Lemma 3.2. *Let p be an odd prime. Then there are precisely $\frac{p-1}{2}$ quadratic residues modulo p .*

Proof. Let $\sigma : (\mathbb{Z}/p\mathbb{Z})^\times \rightarrow (\mathbb{Z}/p\mathbb{Z})^\times; x \mapsto x^2$.

It suffices to show that σ is 2-to-1:

$$x^2 \equiv y^2 \pmod{p} \iff (x+y)(x-y) \equiv 0 \pmod{p} \iff x \equiv \pm y \pmod{p}$$

as p is prime and $\mathbb{Z}/p\mathbb{Z}$ is an integral domain. Hence there are precisely $\frac{p-1}{2}$ elements in the image of σ \square

Alternative. Let g be a primitive root mod p , i.e. $(\mathbb{Z}/p\mathbb{Z})^\times = \{1, g, g^2, \dots, g^{p-2}\}$, and so $\{x^2 : x \in (\mathbb{Z}/p\mathbb{Z})^\times\} = \{1, g^2, g^4, \dots, g^{p-3}, g^{p-1}, g^{p+1}, \dots, g^{2p-4}\}$. But $g^{p-1} \equiv 1 \pmod{p}$, and so the second half of this set is the same as the first half, and hence only half the elements are squares. \square

Let p be an odd prime and $a \in \mathbb{Z}$. We define **Legendre's symbol** “ a on p ” to be:

$$\left(\frac{a}{p}\right) := \begin{cases} 0 & p|a \\ +1 & a \text{ is a quadratic residue mod } p \\ -1 & a \text{ is not a quadratic residue mod } p \end{cases}$$

Theorem 3.3 (Euler's Criterion). *Let $p > 2, a \in \mathbb{Z}$. Then $\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \pmod{p}$*

Since $p > 2$, the elements $0, 1, -1$ are distinct mod p , so this congruence determines $\left(\frac{a}{p}\right)$ uniquely.

Proof. If $a \equiv 0 \pmod{p}$, then the result is trivial. Suppose therefore that $(a, p) = 1$. Then by Fermat, $a^{p-1} \equiv 1 \pmod{p}$, which means that $a^{\frac{p-1}{2}} \equiv \pm 1 \pmod{p}$. Observe further that, if $a = x^2 \pmod{p}$, then $a^{\frac{p-1}{2}} \equiv x^{p-1} \equiv 1 \pmod{p}$. By 3.2, there are precisely $\frac{p-1}{2}$ quadratic residues, so the congruence $a^{\frac{p-1}{2}} \equiv 1 \pmod{p}$ has at least $\frac{p-1}{2}$ solutions. However, this is a polynomial in a of degree $\frac{p-1}{2}$, and so it can only have at most this many solutions, and hence every solution is a quadratic residue. So whenever $\left(\frac{a}{p}\right) = -1$ we must have $a^{\frac{p-1}{2}} \equiv -1 \pmod{p}$. \square

Corollary 3.4. *Let $p > 2, a, b \in \mathbb{Z}$. Then $\left(\frac{a}{p}\right) \cdot \left(\frac{b}{p}\right) = \left(\frac{ab}{p}\right)$*

Proof.

$$\begin{aligned} \left(\frac{ab}{p}\right) &\equiv (ab)^{\frac{p-1}{2}} \pmod{p} \\ &\equiv a^{\frac{p-1}{2}} b^{\frac{p-1}{2}} \pmod{p} \\ &\equiv \left(\frac{a}{p}\right) \cdot \left(\frac{b}{p}\right) \pmod{p} \end{aligned}$$

\square

This implies that

1. The map $(\mathbb{Z}/p\mathbb{Z})^\times \rightarrow \{\pm 1\}; a \mapsto \left(\frac{a}{p}\right)$ is a homomorphism.
2. Let R be any residue, N any non-residue. Then $R \times R = R; N \times N = N; R \times N = N$
3. There is a polynomial time algorithm for computing $\left(\frac{a}{p}\right)$ for odd p , because we can efficiently compute $a^n \pmod{p}$ via binary modular exponentiation.

Corollary 3.5. *Let $p > 2$. Then $\left(\frac{-1}{p}\right) = \begin{cases} +1 & p \equiv 1 \pmod{4} \\ -1 & p \equiv -1 \pmod{4} \end{cases}$*

Proof. $p \equiv \begin{cases} +1 & \pmod{4} \\ -1 & \pmod{4} \end{cases} \iff \frac{p-1}{2} \equiv \begin{cases} 0 & \pmod{2} \\ 1 & \pmod{2} \end{cases}$,

and hence $(-1)^{\frac{p-1}{2}} \equiv \begin{cases} +1 & p \equiv 1 \pmod{4} \\ -1 & p \equiv 3 \pmod{4} \end{cases}$, and so by Euler's criterion, $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}} \quad \square$

We can think about this in another way, by considering an alternative proof of Fermat's little theorem:

Observe that multiplying by a simply permutes the elements of the multiplicative group mod p as a is a generator, hence:

$(p-1)! = \prod_{j=1}^{p-1} j \equiv \prod_{j=1}^{p-1} (aj) = a^{p-1} \prod_{j=1}^{p-1} j = a^{p-1} (p-1)!$, and hence $a^{p-1} \equiv 1 \pmod{p}$.

Similarly, $\prod_{j=1}^{\frac{p-1}{2}} (aj) = a^{\frac{p-1}{2}} \prod_{j=1}^{\frac{p-1}{2}} j = a^{\frac{p-1}{2}} (\frac{p-1}{2})!$. Write $aj \equiv \epsilon_j c_j \pmod{p}$ with $c_j \in \{1, 2, \dots, \frac{p-1}{2}\}$, and $\epsilon_j \in \{\pm 1\}$.

We claim that, if $1 \leq j \leq k \leq \frac{p-1}{2}$, then $c_j \neq c_k$.

Indeed, if $c_j = c_k$, then $\frac{aj}{\epsilon_j} \equiv \frac{ak}{\epsilon_k} \pmod{p}$, i.e. $j\epsilon_k \equiv k\epsilon_j \pmod{p}$ iff $j \equiv \pm k \pmod{p}$.

Hence, $a^{\frac{p-1}{2}} (\frac{p-1}{2})! = \prod_{j=1}^{\frac{p-1}{2}} (aj) \equiv \prod_{j=1}^{\frac{p-1}{2}} (\epsilon_j c_j) \pmod{p} \equiv \left(\prod_{j=1}^{\frac{p-1}{2}} \epsilon_j \right) (\frac{p-1}{2})! \pmod{p}$, and hence $a^{\frac{p-1}{2}} \equiv \prod_{j=1}^{\frac{p-1}{2}} \epsilon_j \pmod{p}$. This brings us onto:

Lemma 3.6 (Gauss's Lemma). *Let $p > 2, a \in \mathbb{Z}$. Then:*

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

Where $\mu = |\{1 \leq j \leq \frac{p-1}{2} : aj \equiv k \pmod{p} \text{ for some } \frac{p+1}{2} \leq k \leq p-1\}|$

Proof. By the above and observe that $\mu = |\{n \leq j \leq \frac{p-1}{2} : \epsilon_j = -n\}| = |\{1 \leq j \leq \frac{p-1}{2} : \epsilon_j = -1\}|$, and hence $(-1)^\mu = \prod_{j=n}^{\frac{p-1}{2}} \epsilon_j$. \square

Examples

1. Let $a = -1$. Then $aj = -j$ for all $j \in \{1, \dots, \frac{p-1}{2}\}$, and so $\mu = \frac{p-1}{2}$, and $\left(\frac{-1}{p} \right) = (-1)^{\frac{p-1}{2}}$ as we saw in **3.5**.
2. Let $a = 2$. Then $\{2j : 1 \leq j \leq \frac{p-1}{2}\} = \{2, 4, \dots, p-1\}$.
If $0 < j < \frac{p}{4}$, then $2j \in \{1, \dots, \frac{p-1}{2}\}$, so $\epsilon_j = 1$.
If $\frac{p}{4} < j < \frac{p}{2}$ then $\frac{p}{2} < 2j < p$, so $\epsilon_j = -1$.

Hence $\mu = |\{1 \leq j \leq \frac{p-1}{2} : \frac{p}{4} < j < \frac{p}{2}\}| = \lfloor \frac{p}{2} \rfloor - \lfloor \frac{p}{4} \rfloor$.

The precise value depends on $p \pmod{8}$:

Corollary 3.7. *Let p be an odd prime. Then:*

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

3. Let $a = 3$. Consider $\{3j : 1 \leq j \leq \frac{p-1}{2}\}$. We can check that $\mu = |\{1 \leq j \leq \frac{p-1}{2} : \frac{p}{6} < j < \frac{p}{3}\}|$, and that it's parity depends on $p \pmod{12}$.

In general, we see that $\mu = \sum_{m \in \mathbb{Z}} |\{1 \leq j \leq \frac{p-1}{2} : (m - \frac{1}{2} \frac{p}{a} < j < m \frac{p}{a})\}|$. Observe that $0 < m < \frac{1}{2} + \frac{aj}{p} < \frac{a+1}{2}$. In particular, if a is odd, then $m \leq \frac{a-1}{2}$.

Theorem 3.8 (Gauss's Law of Quadratic Reciprocity). *Let p, q be odd primes. Then:*

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

$$\text{Equivalently, } \left(\frac{q}{p}\right) = \begin{cases} -\left(\frac{p}{q}\right) & p \equiv q \equiv 3 \pmod{4} \\ \left(\frac{p}{q}\right) & \text{otherwise} \end{cases}$$

Examples

1. Let $p \geq 5$. Then $\left(\frac{3}{p}\right) = (-1)^{\frac{p-1}{2} \cdot \frac{3-1}{2}} \left(\frac{p}{3}\right)$.

$$\text{Now } (-1)^{\frac{p-1}{2}} = 1 \iff p \equiv 1 \pmod{4}, \text{ and } \left(\frac{p}{3}\right) = 1 \iff p \equiv 1 \pmod{3}, \text{ so } \left(\frac{3}{p}\right) = \begin{cases} +1 & p \equiv \pm 1 \pmod{12} \\ -1 & p \equiv \pm 5 \pmod{12} \end{cases}.$$

2. $\left(\frac{19}{73}\right) = +1 \times \left(\frac{73}{19}\right) = \left(\frac{16}{19}\right) = 1$ as $16 = 4^2$.

3. $\left(\frac{34}{97}\right) = \frac{2}{17} \times \frac{17}{97} = +1 \times \left(\frac{17}{97}\right) = +1 \times \left(\frac{97}{17}\right) = \left(\frac{12}{17}\right) = \left(\frac{3}{17}\right) = -1$ as $17 \equiv 5 \pmod{12}$.

4.

$$\begin{aligned} \left(\frac{7411}{9283}\right) &= -\left(\frac{9283}{7411}\right) \\ &= -\left(\frac{1872}{7411}\right) \\ &= -\left(\frac{13}{7411}\right) \text{ as } 1872 = 2^4 \cdot 3^2 \cdot 13, \text{ and } 2, 4 \text{ are even powers} \\ &= -\left(\frac{7411}{13}\right) \\ &= -\left(\frac{1}{13}\right) \\ &= -1 \end{aligned}$$

Proof of Gauss's Lemma. $\left(\frac{q}{p}\right) = (-1)^\mu$ where $\mu = |\{(j, m) \in S : (m - \frac{1}{2})\frac{p}{q} < j < m\frac{p}{q}\}|$ where $S = \{(j, m) : 1 \leq j \leq \frac{p-1}{2}, 1 \leq m \leq \frac{q-1}{2}\}$.

Rewriting, $\mu = |\{(j, m) \in S : 0 < mp - jq < \frac{p}{2}\}|$.

Similarly, $\left(\frac{p}{q}\right) = (-1)^\eta$, where $\eta = |\{(j, m) \in S : 0 < jq - mp < \frac{q}{2}\}|$.

Observe that $\left(\frac{q}{p}\right) \left(\frac{p}{q}\right) = (-1)^{\mu+\eta}$. Now $|S| = \frac{p-1}{2} \frac{q-1}{2} = \mu + \eta + |A| + |B|$, where:

$$\begin{aligned} A &= \{(j, m) \in S : mp - jq > \frac{p}{2}\} \\ B &= \{(j, m) \in S : jq - mp > \frac{q}{2}\} \end{aligned}$$

Given $(j, m) \in S$, let $j' = \frac{p+1}{2} - j, m' = \frac{q+1}{2} - m$. Note that $(j', m') \in S$. We now claim that $(j, m) \in A \iff (j', m') \in B$.

Indeed, note that $j'q - m'p = (\frac{p+1}{2} - j)q - (\frac{q+1}{2} - m)p = mp - jq - \frac{p}{2} + \frac{q}{2}$. Hence $j'q - m'p > \frac{q}{2} \iff mp - jq > \frac{p}{2}$.

Hence $|A| = |B|$, and so these terms do not affect the sign of $(-1)^{\mu+\eta+|A|+|B|}$, and hence:

$$\left(\frac{p}{q}\right)\left(\frac{q}{b}\right) = (-1)^{\mu+\eta} = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$$

□

This is a very powerful theorem, however in example 4., we still had to factorise 1872 which is, in general, very hard to do. To get around this problem, we will extend the Legendre symbol.

Given $n \geq 1$ an odd number, and $n = p_1 \cdot \dots \cdot p_k$ where the p_k are not necessarily distinct, we can define the **Jacobi symbol** $\left(\frac{a}{m}\right) = \prod_{i=1}^k \left(\frac{a}{p_i}\right)$, and $\left(\frac{a}{1}\right) = 1$. Note that if $(a, n) = 1$ then $\left(\frac{a}{n}\right) = 0$.

Proposition 3.9. *Let $m, n \geq 1$ be odd, and let $a, b \in \mathbb{Z}$. Then:*

1. *If $a \equiv b \pmod{n}$, then $\left(\frac{a}{n}\right) = \left(\frac{b}{n}\right)$*
2. *$\left(\frac{ab}{n}\right) = \left(\frac{a}{n}\right)\left(\frac{b}{n}\right)$ and $\left(\frac{a}{mn}\right) = \left(\frac{a}{m}\right)\left(\frac{a}{n}\right)$*
3. *$\left(\frac{-1}{n}\right) = (-1)^{\frac{n-1}{2}}$*
4. *$\left(\frac{2}{n}\right) = (-1)^{\frac{n^2-1}{8}}$*

Proof.

1. Note that if $a \equiv b \pmod{n}$, then $a \equiv b \pmod{p_i}$ for all i by the Chinese remainder theorem, and hence $\left(\frac{a}{p_i}\right) = \left(\frac{b}{p_i}\right)$.
2. $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$ for each odd prime p , and so the first part follows from the definition. The second part is immediate from the definition.
3. This is true for n prime, and so by 2. it is sufficient to check that if m, n are odd, then $(-1)^{\frac{m-1}{2}}(-1)^{\frac{n-1}{2}} = (-1)^{\frac{mn-1}{2}}$ and $(-1)^{\frac{m^2-1}{8}}(-1)^{\frac{n^2-1}{8}} = (-1)^{\frac{m^2n^2-1}{8}}$.

□

Theorem 3.10 (Quadratic Reciprocity for Jacobi Symbols). *Let $m, n \geq 1$ be odd. Then:*

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

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

Proposition 3.11. *If $(m, n) > 1$, then both sides are 0 and we're done.*

Suppose $(m, n) = 1$, with $m = \prod_{i=1}^k p_i$, $n = \prod_{j=1}^\ell q_j$, with the $\{p_i\}, \{q_j\}$ not necessarily distinct but $\{p_i\} \cap \{q_j\} = \emptyset$.

Then:

$$\begin{aligned}
jacobimn &= \prod_{i=1}^k \prod_{j=1}^\ell \left(\frac{p_i}{q_j} \right) \\
&= \prod_{i=1}^k \prod_{j=1}^\ell (-1)^{\frac{p_i-1}{2} \frac{q_j-1}{2}} \left(\frac{q_j}{p_i} \right) \\
&= (-1)^{rs} \prod_{i=1}^k \prod_{j=1}^\ell \left(\frac{q_j}{p_i} \right) \\
&= (-1)^{rs} \left(\frac{n}{m} \right)
\end{aligned}$$

where $r = |\{i : p_i \equiv 3 \pmod{4}\}|$, $s = |\{j : q_j \equiv 3 \pmod{4}\}|$.

But $\frac{m-1}{2} \frac{n-1}{2} \equiv 1 \pmod{2} \iff m \equiv n \equiv 3 \pmod{4} \iff r \equiv s \equiv 1 \pmod{2}$, and hence the result follows.

IMPORTANT REMARK: Let p be an odd prime and $a \in \mathbb{Z}$ with $(a, p) = 1$. Then $\left(\frac{a}{p}\right) = 1$ if and only if a is a square mod p , however this is not the case in general for composite n :

Let n be an odd integer, and let $a \in \mathbb{Z}$ with $(a, n) = 1$. Then if a is a square mod n , say $a = x^2 \pmod{n}$, then $\left(\frac{a}{n}\right) = \left(\frac{x^2}{n}\right) = \left(\frac{x}{n}\right)^2 = 1$. The converse is NOT the case: $\left(\frac{2}{15}\right) = \left(\frac{2}{3}\right) \cdot \left(\frac{2}{5}\right) = 1$, but 2 is not a square mod 15.

More generally, let $p \neq q$ be odd primes, and $a \in \mathbb{Z}$ with $(a, pq) = 1$. Then by the Chinese Remainder Theorem:

a is a square mod pq if and only if a is a square mod p and mod q .

Then if $\left(\frac{a}{pq}\right) = -1$, then a is not a square mod pq , but if $\left(\frac{a}{pq}\right) = +1$ then all we know is $\left(\frac{a}{p}\right) = \left(\frac{a}{q}\right)$.

Examples:

1. $\left(\frac{33}{73}\right) = \left(\frac{73}{33}\right) = \left(\frac{7}{33}\right) = \left(\frac{33}{7}\right) = \left(\frac{5}{7}\right) = -1$.
2. $\left(\frac{66}{73}\right) = \left(\frac{2}{73}\right) \left(\frac{33}{73}\right) = -1$

4 Binary Quadratic Forms

Which integers can be represented as the sum of two integer squared?

Theorem 4.1. *Let $N \in \mathbb{N}$. Then N is the sum of two squares if and only if every prime $p \equiv 3 \pmod{4}$ divides N to some even power.*

Proof. Suppose $N = x^2 + y^2$, and $p|N, p \equiv 3 \pmod{4}$.

Then $x^2 \equiv -y^2 \pmod{p}$ and $\left(\frac{-1}{p}\right) = -1$, and so we must have $x \equiv y \equiv 0 \pmod{p}$. Hence $p^2|N$. Now repeat the argument with $(N/p^2, x/p, y/p)$ until $p \nmid N$.

Conversely, write $N = M^2 K$ where K is the product of distinct primes, each either 2 or congruent to 1 mod 4. Then it suffices to show that we can represent $K = x^2 + y^2$ with $x, y \in \mathbb{Z}$ since then $N = (Mx)^2 + (My)^2$. Now observe $(x^2 + y^2)(z^2 + w^2) = (xz - yw)^2 + (xw + yz)^2$, and

$2 = 1^2 + 1^2$, it is enough to show that a prime $p \equiv 1 \pmod{4}$ can be written as a sum of two squares. We will show this next week. \square

A **binary quadratic form** with integer coefficients is of the form $f(x, y) = ax^2 + bxy + cy^2$ for $a, b, c \in \mathbb{Z}$. We say that **f represents n** for $n \in \mathbb{N}$ if there are integers x, y such that $n = f(x, y)$. We sometimes write (a, b, c) in place of f .

It can also be helpful to regard f as a matrix: $f(x, y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$

A **unimodular substitution** is a change of variable of the form $X = \alpha x + \gamma y$; $Y = \beta x + \delta y$ with $\alpha, \beta, \gamma, \delta \in \mathbb{Z}$ and $\alpha\delta - \beta\gamma = 1$. We say that two binary quadratic forms f and g are **equivalent** if $f(X, Y) = g(x, y)$ for some unimodular substitution $(x, y) \mapsto (X, Y)$. Note that equivalence of binary quadratic forms is an equivalence relation (example sheet).

Note that the set of unimodular substitutions is $SL_2(\mathbb{Z})$, a group under matrix multiplications acting on the set of binary quadratic forms via unimodular substitutions. The orbits of this action are precisely the equivalence classes of binary quadratic forms. To see that this is a group action, we need to check that $I \cdot f = f$ and $A \cdot (B \cdot f) = (AB) \cdot f$.

Note that if $f = (a, b, c)$ and $Af = g = (a', b', c')$, one can check via manual substitution that $\begin{pmatrix} a' & b'/2 \\ b'/2 & c' \end{pmatrix} = A \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} A^T$.

The **discriminant** of a binary quadratic form (a, b, c) is defined to be $\text{disc } f = b^2 - 4ac = -4 \det \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$.

Lemma 4.2. *Equivalent BQFs have the same discriminant.*

Proof. If $g = Af$, then:

$$\begin{aligned} \text{disc } g &= -4 \det(AMA^T) \\ &= -4 \det M \det A \det A^T \\ &= -4 \det M \\ &= \text{disc } f \end{aligned}$$

\square

Note that BQFs can have the same discriminant and not be equivalent, e.g. $x^2 + 6y^2$ and $2x^2 + 3y^2$ both have discriminant -24 .

Lemma 4.3. *There exists a BQF f with $\text{disc } f = d$ if and only if $d \equiv 0$ or $1 \pmod{4}$.*

Proof. If $d \equiv 0 \pmod{4}$ take $f = (1, 0, -d/4)$. If $d \equiv 1 \pmod{4}$ take $f = (1, 1, \frac{1-d}{4})$. For the other direction note $0, 1$ are the only squares mod 4. \square

A real quadratic form given by $f(x_1, \dots, x_n) = \sum_{i \leq j} a_{ij} x_i x_j$ with $a_{ij} \in \mathbb{R}$ is said to be:

- **Positive definite** if $f(x) > 0$ for $x \in \mathbb{R}^n \setminus \{0\}$
- **Negative definite** if $f(x) < 0$ for $x \in \mathbb{R}^n \setminus \{0\}$
- **Indefinite** if there are $x, y \in \mathbb{R}^n$ such that $f(x) > 0 > f(y)$.

We will stick to $n = 2, a_{ij} \in \mathbb{Z}$, but the definitions are just the same.

Lemma 4.4. *Let $f = (a, b, c)$ be a BQF, and let $d = \text{disc } f = b^2 - 4ac$. Then the following hold:*

1. *If $d < 0$ and $a > 0$ then f is positive definite.*
2. *If $d < 0$ and $a < 0$ then f is negative definite.*
3. *If $d > 0$ then f is indefinite.*
4. *If $d = 0$ then $f(x, y) = \ell(mx + ky)^2$ for some $\ell, m, k \in \mathbb{Z}$.*

Proof.

- 1., 2. $4af(x, y) = 4a^2x^2 + 4abxy + 4acy^2 = (2ax + by)^2 - b^2y^2 + 4acy^2 = (2ax + by)^2 - dy^2$, where $d = \text{disc } f$. So if $d < 0$ then $4af(x, y) \geq 0$ for all $x, y \in \mathbb{R}$. If it is 0 then $x = y = 0$, and so if $a > 0$ then f is positive definite and if $a < 0$ then f is negative definite.
3. Suppose $d > 0$. As above, $4af(x, y) = (2ax + by)^2 - dy^2 = (2ax + by - \sqrt{d}y)(2ax + by + \sqrt{d}y) = 4a^2(x + \frac{b-\sqrt{d}}{2a}y)(x + \frac{b+\sqrt{d}}{2a}y)$. Hence we can get positive or negative values for $4af$, and hence for f .
4. Write $a = a_1a_2^2$ with a_1 a product of distinct primes. Then $ax^2 + bxy + cy^2 = a_1(a_2x + \frac{b}{2a_1a_2}y)^2$ is of the desired form.

□

Note that there are indefinite forms all of whose coefficients are positive, e.g. $\text{disc}(1, 3, 1) = 520 > 0$. There are also positive definite forms not all of whose coefficients are positive. So $\text{disc}(1, -1, 2) = -7 < 0$. Note that $d < 0$ and $a > 0$ implies that $c > 0$, so the only coefficient that can be negative is the middle one.

Example: Consider $(10, 34, 29)$. The coefficients here are large and awkward - it would be nice to move to an equivalent BQF with smaller coefficients. If $f(x, y) = ax^2 + bxy + cy^2$, consider what happens with the substitution given by $x' = x + \lambda y, y' = y$. Then we have $f(x', y') = g(x, y) = ax^2 + (b + 2\lambda a)xy + (\lambda^2 a + \lambda b + c)y^2$.

Hence $(a, b, c) \sim (a, b \pm 2a, a \pm b + c)$ with $\lambda = \pm 1$, and in our case $(10, 34, 29) \sim (10, 14, 5) \sim (10, -6, 1)$.

Another nice substitution is $x' = y, y' = -x$. This gives $(a, b, c) \sim (c, -b, a)$

So in our example, we get $(10, 34, 29) \sim (1, 6, 10) \sim (1, 4, 5) \sim (1, 2, 2) \sim (1, 0, 1)$, which is very nice.

A positive BQF is said to be **reduced** or **in reduced form** when either $-a < b \leq a < c$ or $0 \leq b \leq a = c$. Note that in either case $|b| \leq a \leq c$.

The form that we ended up with at the end of the previous discussion, $(1, 0, 1)$ is reduced by the second criterion.

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

Proof. Consider the operations:

$$\begin{aligned} T_{\pm} : (a, b, c) &\mapsto (a, b \pm 2a, a \pm b + c) \\ S : (a, b, c) &\mapsto (c, -b, a) \end{aligned}$$

If $a > c$, we use S . Otherwise, if $a \leq c$ and $c \leq |b|$, then use T_{\pm} to reduce $|b|$. At each stage, observe that $|a| + |b|$ decreases by at least 1, so this algorithm will terminate. As such, eventually we must have that $a \leq c$ and $|b| \leq a$. Then, if $a < c$ then $b = -a$ if a is not reduced. Hence $(a, b, c) = (a, -a, c) \sim (a, a, c)$, reduced. If $a = c$, then $-a \leq b < 0$, and so $(a, b, a) \sim (a, -b, a)$, reduced. \square

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

Proof. $b \equiv d \pmod{2}$ is trivial, as $4 \equiv 0 \pmod{2}$. So f reduced implies that $|b| \leq a \leq c \implies d = b^2 - 4ac \leq -3ac \leq -3a^2$, and hence $a \leq \sqrt{\frac{|d|}{3}}$. \square

For instance, if f is a PDBQF with discriminant -4 , then $f \sim (1, 0, 1)$. This seemingly simple lemma will prove to be immensely useful.

Proof of 4.1, continued. We need to show that, if $p \equiv 1 \pmod{4}$, then we can write $p = x^2 + y^2$ for some $x, y \in \mathbb{Z}$.

Observe that $\left(\frac{-1}{p}\right) = 1$, and hence there are $m, k \in \mathbb{Z}$ with $m^2 = -1 + pk$. Now if we let $f = (p, 2m, k)$, then $f(1, 0) = p$ so f represents p , and $\text{disc } f = -4$, so $f \sim (1, 0, 1)$, and hence $x^2 + y^2$ also represents p . \square

We say that a BQF **properly represents** $n \in \mathbb{Z}$ if $\exists x, y \in \mathbb{Z}$ s.t. $\gcd(x, y) = 1$ and $n = f(x, y)$. Observe that if there is a unimodular substitution $(x, y) \leftrightarrow (X, Y)$, then $d|(x, y) \iff d|(X, Y)$, and hence equivalent BQFs properly represent the same integers.

Lemma 4.7. *The least properly represented positive integers by a PDBQF are $a, c, a - |b| + c$.*

Proof. If $f(x, y) = n$ properly, then $(x, y) = 1 \implies x = \pm 1$ and $f(\pm 1, 0) = a$, and similarly $f(0, \pm 1) = c$.

Suppose then that $|x| \geq |y| > 0$. Then $f(x, y) = ax^2 + bxy + cy^2 \geq a - |b| + c \geq c$. We have a similar result if $|y| \geq |x| > 0$.

But then for a suitable choice of $\epsilon \in \{\pm 1\}$, we see $f(1, \epsilon) = a - |b| + c$. \square

Note that values in the above are repeated if represented by more than one way not counting the trivial $f(x, y) = f(-x, -y)$.

Theorem 4.8. *Every PDBQF is equivalent to a unique reduced form. (i.e. each reduced form represents a distinct equivalence class).*

Proof. We already have existence, so it is sufficient to prove that no two reduced forms are equivalent, or that if two reduced forms are equivalent they must be the same, so suppose f, g are both reduced, and $(a, b, c) = f \sim g = (a', b', c')$.

Note that the least value properly represented by f, g is a, a' say. Hence $a = a'$ and similarly $c = c'$. Now by considering discriminants we have $b = \pm b'$. If $b = 0$, we are immediately done, so suppose $0 < b = -b'$. Note that we must have the $-a < b < a < c$, by considering the reduced conditions for both BQFs.

So we have $a < c < a - |b| + c$, and there are no repeats. Since $g(x, y) = f(X, Y)$ for some unimodular substitution represented by $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SL_2(\mathbb{Z})$, we manually have $\alpha, \delta = \pm 1; \beta, \gamma = 0$, and so $A = \pm I$. Observe that we cannot have the $-$ case, and so $g = f$ as required. \square

Let $d < 0, d \equiv 0 \text{ or } 1 \pmod{4}$, we define the **class number** $h(d)$ as the number of reduced BQFs of discriminant d . For instance, we have seen that $h(-4) = 1$. Consider $d = -24$: the forms $(1, 0, 6)$ and $(2, 0, 3)$ have discriminant -24 and are reduced. One can show that there are no others, and so $h(-24) = 2$.

One might ask how $h(d)$ behaves as a function. We know that:

1. $h(d) \rightarrow \infty$ as $d \rightarrow \infty$
2. $h(d) \approx \sqrt{d}$, i.e. $\frac{1}{N} \sum_{3 \leq -d \leq N} h(d) \sim \frac{\pi}{18} N^{1/2}$ as $N \rightarrow \infty$.
3. $h(d) - 1 \implies d \in \{-3, -4, -7, -8, -11, -19, -43, -67, -163, -12, -16, -27, -28\}$
4. Given $h \geq 1$, there is an explicit upper bound the set of all $\{d : h(d) = h\}$. So in principle we can compute all d with a given class number h .

Lemma 4.9. *Let $n \in \mathbb{N}$, and f a BQF. then n is properly represented by f if and only if $f \sim g = (a, b, c)$ with $a = n$*

Proof. If $f \sim g = (n, b, c)$, then $g(1, 0) = n$

Conversely, if $f(\alpha, \beta) = n$ for some $\alpha, \beta \in \mathbb{Z}, (\alpha, \beta) = 1$, then there exists some $\delta, \gamma \in \mathbb{Z}$ s.t. $\alpha\delta - \beta\gamma = 1$. Hence f is equivalent to $g(x, y) = f(\alpha x + \gamma y, \beta x + \delta y)$. But the first coefficient of g is $g(1, 0) = f(\alpha, \beta) = n$. \square

Theorem 4.10. *Let $n \in \mathbb{N}$ and $d < 0$ with $d \equiv 0 \text{ or } 1 \pmod{4}$. Then n is properly represented by some BQF of discriminant d if and only if $x^2 \equiv d \pmod{4n}$ has a solution.*

Proof. If $n = f(x, y)$ with $x, y \in \mathbb{Z}, (x, y) = 1, \text{disc } f = d$ then the previous lemma gives $f \sim g = (n, b, c)$. Then $b^2 - 4nc = d$, and so $b^2 \equiv d \pmod{4n}$.

Conversely, if $\exists b \in \mathbb{Z}$ s.t. $b^2 \equiv d \pmod{4n}$, then $\exists c \in \mathbb{Z}$ s.t. $b^2 = d + 4nc \implies d = b^2 - 4nc$, so (n, b, c) is as required. \square

Example: Which integers are properly represented by $x^2 + xy + 2y^2$? Note that $\text{disc } f = -7 < 0$. If (a, b, c) is a reduced form for discriminant -7 , then $|b| \leq a \leq 1$ and b odd, hence $a = |b| = 1$, and $c = \frac{b^2 - d}{4} = 2$. Hence $(a, b, c) = (1, 1, 2)$ as $(1, -1, 2)$ is not reduced. So f properly represents n if and only if $x^2 \equiv -7 \pmod{4n}$ has solutions.

Suppose n is prime, i.e. $n = p \notin \{2, 7\}$. Then by $x^2 \equiv -7 \pmod{4p}$ if and only if $x^2 \equiv -7 \pmod{4}$ and $x^2 \equiv -7 \pmod{p}$. Then this system holds if and only if $\left(\frac{-7}{p}\right) = 1$. Now $\left(\frac{-7}{p}\right) = \left(\frac{p}{7}\right)$, i.e. $p \equiv 1, 2, 4 \pmod{7}$. Note that $f(0, 1) = 2$ and $f(1, -2) = 7$.

Suppose n is composite, i.e. $n = \prod p^{\alpha_p}$ for p distinct primes, $\alpha_p \geq 0$. Then $x^2 \equiv -7 \pmod{4n}$ is soluble if and only if $x^2 \equiv -7 \pmod{2^{2+\alpha_2}}$ and $x^2 \equiv -7 \pmod{p^{\alpha_p}}$ for odd p is soluble.

Lemma 4.11 (Hensel, only a special case). *Let $a \in \mathbb{Z}$.*

1. If $a \equiv 1 \pmod{8}$ then $x^2 \equiv a \pmod{2^k}$ for all $k \geq 1$ is soluble.

2. If $p > 2$ is prime and $\left(\frac{a}{p}\right) = 1$ then $x^2 \equiv a \pmod{p^k}$ is soluble for all $k \geq 1$.

Note then that $-7 \equiv 1 \pmod{8}$, so the first equation is soluble. If p is odd, then $p \equiv 1, 2, 4 \pmod{7}$ then $\left(\frac{-7}{p}\right) = 1$, and so $x^2 \equiv -7 \pmod{p^k}$ is soluble for all $k \geq 1$. If $p = 7$, then $x^2 \equiv -7 \pmod{7^k}$ is soluble when $k \leq 1$ but not when $k \geq 2$.

In conclusion, $f = (1, 1, 2)$ properly represents n if and only if $n = 7^{\alpha_7} \prod_{p \equiv 1, 2, 4 \pmod{7}} p^{\alpha_p}$ where $\alpha_p \geq 0$, and $\alpha_7 \in \{0, 1\}$.

The numbers represented by f not necessarily properly are then of the form $k^2 n$ with $k \geq 1$ and n as above. In other words, $n = x^2 + xy + 2y^2$ for some $x, y \in \mathbb{Z}$ if and only if every prime $p \equiv 3, 5, 6 \pmod{7}$ dividing n divides it to an even power.

Proof of Hensel. 1. By induction on k . The case $k \leq 3$ is clear - taking $x = 1$ will do. So suppose there exists some x such that $x^2 \equiv a \pmod{2^k}$, i.e. $x^2 = a + 2^k m$ for some $m \in \mathbb{Z}, k \geq 3$. Note that:

$$(x + 2^{k-1})^2 = x^2 + 2^k x + 2^{2k-2} \equiv a + 2^k(x + m) \pmod{2^{k+1}}$$

If m is even then $x^2 \equiv a \pmod{2^{k+1}}$

If m is odd then since x is odd, $(x + 2^{k-1})^2 \equiv a \pmod{2^{k+1}}$.

2. Since $\left(\frac{a}{p}\right) = 1$, $x^2 \equiv a \pmod{p}$ is soluble, so we have a base case. Assume that there is some x such that $x^2 \equiv a \pmod{p^k}$ for some $k \geq 1$. Then for $t \in \mathbb{Z}$, we have:

$$\begin{aligned} (x + tp^k)^2 &= x^2 + 2txp^k + t^2p^{2k} \\ &= a + (2tx + m)p^k + t^2p^{2k} \\ &\equiv a + (2tx + m)p^k \pmod{p^{k+1}} \end{aligned}$$

Since $(x, p) = 1$, there is some t such that $2xt + m \equiv 0 \pmod{p}$, and for this t we have $(x + p^k t)^2 \equiv a \pmod{p^{k+1}}$.

□

Remarks:

1. If $h(d) = 1$ we have completely solved the problem of which integers are represented by a given form of discriminant $d \leq 0$. If $h(d) > 1$, we can determine which integers are represented by some form of discriminant d , but in general we cannot do better. For some d we can distinguish using additional congruence conditions, but there is no congruence condition to tell if e.g. $p = x^2 + 23y^2$.
2. An integer $d \equiv 0$ or $1 \pmod{4}$ is said to be a ***fundamental discriminant*** if it is not of the form $k^2 d'$ for some $k > 1, d' \equiv 0$ or $1 \pmod{4}$. For a fundamental discriminant, Gauss defined a law of composition which makes the set of BQFs of divisor d into an abelian group. The group is known as the ***ideal class group*** of the field $\mathbb{Q}(\sqrt{d})$. More will come later on this in Number Fields.
3. If f is definite, then there are only a finite number of representations of n by f . e.g. $n = x^2 + y^2 \implies |x|, |y| \leq \sqrt{n}$. If f is indefinite, there can be infinitely many representations.

What about quadratic forms in more than two variables?

Theorem 4.12 (Lagrange, 1770). *Every possible integer can be written as the sum of four squares*

Theorem 4.13 (Legendre, 1797). *A positive integer n can be written as the sum of three squares if and only if $n \not\equiv 4^a \pmod{8b+7}$ for some $a, b \geq 0$.*

In 1770 Waring posed the following problem: For every k , let $g(k)$ be the least integer integer such that every positive integer can be written as a sum k -th powers. Hilbert showed in 1909 that $g(k)$ exists for all k . Lagrange showed $g(2) = 4$, and it is known that $g(3) = 9, g(4) = 19$. The asymptotic behaviour remains an open question.

5 The Distribution of Primes

We now turn to questions about how primes occur. How many are there? How quickly to they appear? How densely do they come? How regular is the sequence of primes?

Theorem 5.1 (Prime Number Theorem). *Let $\pi(x)$ be the number of primes less than or equal to x . Then $\pi(x) \sim \frac{x}{\log x}$ as $x \rightarrow \infty$. More precisely, $\pi(x) = \frac{x}{\log x} + O(x \exp(-\sqrt{\log x}))$*

This is sometimes stated in terms of the logarithmic integral $Li(x) = \int_2^x \frac{dt}{\log t}$: $\pi(x) \sim Li(x)$.

Theorem 5.2 (Dirichlet's Theorem on Primes in Arithmetic Progression). *Let a, q be coprime integers. Then there are infinitely many primes $p \equiv a \pmod{q}$.*

Theorem 5.3 (Siegel-Walfisz). *Let $a, q \in \mathbb{Z}, (a, q) = 1$, and let $\pi(x; q, a)$ be the number of primes $p \leq x$ such that $p \equiv a \pmod{q}$.*

Then for all real $B > 0$ there is a constant $C_B > 0$ such that:

$$\pi(x; q, a) = \frac{1}{\varphi(q)} \frac{x}{\log x} + O_B(x \exp(-C_B \sqrt{\log x}))$$

whenever $q \leq (\log x)^B$.

We will not prove any of these theorems, but will study instead the Riemann zeta function and prove Chebyshev's theorem:

Theorem 5.4 (Chebyshev). *There is some $c_2 > c_1 > 0$ such that for all $x > 1$, we have:*

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

This will come later though. For now:

Lemma 5.5. *Let $x \in \mathbb{N}$. Then $\pi(x) \geq \frac{\log x}{2 \log 2}$*

Proof. Let p_1, \dots, p_r be the primes up to x , so that $r = \pi(x)$. Every $n \leq x$ can be written in the form $n = k^2 \prod_{i=1}^r p_i^{\alpha_i}$ with $\alpha_i \in \{0, 1\}$ and $1 \leq k \leq \sqrt{x}$.

There are $\leq 2^r$ choices for the α_i , and $\leq \sqrt{x}$ choices for k , hence $x \leq 2^{\pi(x)} \sqrt{x}$.

Rearranging, we have $\pi(x) \geq \frac{\log(x)}{2 \log 2}$ □

Here let p denote primes numbers. Then:

Theorem 5.6.

1. For $x \geq 5$ we have $\sum_{p \leq x} \frac{1}{p} \geq \log \log x - \frac{1}{2}$. In particular, this sum diverges.
2. For $x \geq 2$, the product $\prod_{p \leq x} (1 - \frac{1}{p})^{-1} \geq \log x$.

Proof. For the second part, observe:

$$\begin{aligned} \prod_{p \leq x} (1 - \frac{1}{p})^{-1} &= \prod_{p \leq x} (1 + \frac{1}{p} + \frac{1}{p^2} + \dots) \\ &\geq \sum_{n \leq x} \frac{1}{n} \\ &\geq \int_1^{x+1} \frac{1}{t} dt \\ &\geq \log x \end{aligned}$$

Then for the first part, we want to estimate the value of $\log \left(\prod_{p \leq x} (1 - \frac{1}{p})^{-1} \right) - \sum_{p \leq x} \frac{1}{p}$, or $\sum_{p \leq x} \left(\log(1 - \frac{1}{p})^{-1} - \frac{1}{p} \right)$.

But note that:

$$\begin{aligned} \log(1 - \frac{1}{p})^{-1} - \frac{1}{p} &= -\log(1 - \frac{1}{p}) - \frac{1}{p} \\ &= (\frac{1}{p} + \frac{1}{2p^2} + \dots) - \frac{1}{p} \\ &= \frac{1}{2p^2} + \frac{1}{3p^3} + \dots \\ &\leq \frac{1}{2p^2} + \frac{1}{2p^3} + \dots \\ &= \frac{1}{2p^2} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \dots \right) \\ &= \frac{1}{2p^2} \frac{1}{1 - \frac{1}{p}} \\ &= \frac{1}{2p(p-1)} \end{aligned}$$

Hence:

$$\begin{aligned}
\log \left(\prod_{p \leq x} \left(1 - \frac{1}{p} \right)^{-1} \right) - \sum_{p \leq x} \frac{1}{p} &\leq \sum_{p \leq x} \frac{1}{2p(p-1)} \\
&\leq \sum_{2 \leq n \leq x} \frac{1}{2n(n-1)} \\
&= \frac{1}{2} \sum_{2 \leq n \leq x} \left(\frac{1}{n-1} - \frac{1}{n} \right) \\
&= \frac{1}{2} \left(1 - \frac{1}{x} \right) \\
&\leq \frac{1}{2}
\end{aligned}$$

And so $\sum_{p \leq x} \frac{1}{p} \geq \log \prod_{p \leq x} \left(1 - \frac{1}{p} \right)^{-1} - \frac{1}{2} \geq \log \log x - \frac{1}{2}$ \square

For a complex number s with $\Re s > 1$, we define the **Riemann zeta function** ζ as:

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

Conventionally, we write $s = \sigma + it$ for $\sigma, t \in \mathbb{R}$.

Lemma 5.7. *For $s \in \mathbb{C}$ with $\Re s > 1$, the series $\sum_{n=1}^{\infty} \frac{1}{n^s}$ converges absolutely. In fact, it converges uniformly on $\Re s \geq 1 + \delta$ for $\delta > 0$, and is thus analytic on $\Re s > 1$.*

Proof. Note $|n^s| = n^\sigma$, and so $\zeta(s)$ converges absolutely if and only if $\sum_{n=1}^{\infty} \frac{1}{n^\sigma}$ converges, which happens precisely when $\sigma > 1$. \square

Theorem 5.8. *For $s \in \mathbb{C}$ with $\Re s > 1$, we have that $\zeta(s) = \prod_{p \text{ prime}} \left(1 - \frac{1}{p^s} \right)^{-1}$.*

The right hand side is known as the **Euler product** for ζ .

Proof (“Hand-waving”).

$$\begin{aligned}
\prod_p \left(1 - \frac{1}{p^s} \right)^{-1} &= \prod_p (1 + p^{-s} + p^{-2s} + \dots) \\
&= \sum_{n=1}^{\infty} n^{-s}
\end{aligned}$$

since every n occurs exactly once by unique factorisation. \square

Proof. Let $s = \sigma + it, \sigma > 1$ and consider:

$$\begin{aligned}
\zeta(s) - \prod_{p \leq x} \left(1 - \frac{1}{p^s} \right)^{-1} &= \sum_{n=1}^{\infty} n^{-s} - \prod_{p \leq x} (1 + p^{-s} + p^{-2s} + \dots) \\
&= \sum_{n \in \mathbb{N}_x} n^{-s}
\end{aligned}$$

where $\mathbb{N}_x = \{n \in \mathbb{N} : \text{at least one prime factor of } n \text{ is } > x\}$. So:

$$\begin{aligned} |\zeta(s) - \prod_{p \leq x} (1 - p^{-s})^{-1}| &\leq \sum_{n \in \mathbb{N}_x} n^{-\sigma} \\ &\leq \sum_{n > x} n^{-\sigma} \\ &\rightarrow 0 \text{ as } x \rightarrow \infty \end{aligned}$$

since $\zeta(s)$ converges. □

Corollary 5.9. *For $s \in \mathbb{C}$, $\Re s > 1$ we have $\zeta(s) \neq 0$*

Proof. Clearly, $\prod_{p \leq X} (1 - p^{-s})^{-1} \neq 0$ for $X \geq 2$.

Now note that $\zeta(s) \prod_{p \leq X} (1 - p^{-s}) = \prod_{p > X} (1 - p^{-s})^{-1} = 1 + \sum_{n \in \mathbb{N}_X} n^{-s}$.

Hence $|\zeta(s) \prod_{p \leq X} (1 - p^{-s})| \geq 1 - \sum_{n > X} n^{-\sigma} \rightarrow 1$. □

It turns out that ζ can be extended to the whole of \mathbb{C} using an analytic continuation. First, one extends ζ to $\Re(s) > 0$ using an integral: we define the **Γ -function** to be $\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$ for $\Re(z) > 0$. This can be continued to a meromorphic function on \mathbb{C} with simple poles at $z = 0, -1, -2, \dots$ and no zeroes. The Γ function satisfies that $z\Gamma(z) = \Gamma(z+1)$.

We can then define the **completed ζ -function** to be

$$\Xi(s) = \pi^{-s/2} \Gamma(s/2) \zeta(s)$$

for which we can prove a **functional equation**

$$\Xi(s) = \Xi(1-s)$$

This allows us to extend ζ to the whole of \mathbb{C} . ζ has just one pole, a simple pole at $s = 1$, and it has trivial zeroes at $s = -2, -4, \dots$. Since $\zeta(s) \neq 0$ for $\Re(s) > 1$, and the zeroes of ζ must lie in the **critical strip** where $0 \leq \Re(s) \leq 1$. To prove the prime number theorem one has to show that $\zeta(s) \neq 0$ for $\Re(s) = 1$.

The **Riemann Hypothesis** asserts that in fact all of the zeroes of $\zeta(s)$ lie on the line $\Re(s) = \frac{1}{2}$, which is backed up by extensive numerical evidence.

Define the **Möbius function** $\mu : \mathbb{N} \rightarrow \{-1, 0, +1\}$ by defining:

$$\begin{aligned} \mu(n) &= \begin{cases} (-1)^k & n = p_1 p_2 \dots p_k; p_i \text{ all distinct} \\ 0 & \text{otherwise} \end{cases} \\ \mu(1) &= 1 \end{aligned}$$

On the example sheet we will show that μ is a multiplicative function and will prove the Möbius inversion formula:

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

for functions f, g defined on \mathbb{N}

Lemma 5.10. Let $u(n) = \sum_{d|n} \mu(d)$. Then $u(n) = \begin{cases} 1 & n = 1 \\ 0 & n > 1 \end{cases}$.

Proof. Assuming μ is multiplicative, we write $v(n) = \prod_{i=1}^k u(p_i^{\alpha_i})$ if $n = \prod_{i=1}^k p_i^{\alpha_i}$.

Observe that for $m \geq 0$, we have:

$$\begin{aligned} v(p^m) &= \mu(1) + \mu(p) + \dots + \mu(p^m) \\ &= \begin{cases} \mu(1) = 1 & m = 0 \\ \mu(1) + \mu(p) = 0 & m > 0 \end{cases} \end{aligned}$$

□

The Riemann Hypothesis is known to be equivalent to the assertion that for all $\epsilon > 0$ there is a constant c_ϵ such that $|\sum_{n \leq x} \mu(n)| \leq c_\epsilon x^{1/2+\epsilon}$.

A **Dirichlet series** is of the form $\sum_{n=1}^{\infty} \frac{a_n}{n^s}$ for some sequence $a_n \in \mathbb{C}$.

1. $\zeta(s)$
2. $\sum_{n=1}^{\infty} \left(\frac{n}{p}\right) n^{-s}$, a Dirichlet L-function
3. Products of Dirichlet series, since assuming absolute convergence:

$$\begin{aligned} \left(\sum_{m=1}^{\infty} \frac{a_m}{m^s}\right) \left(\sum_{n=1}^{\infty} \frac{b_n}{n^s}\right) &= \sum_{m=1, n=1}^{\infty} (a_m b_n) \frac{1}{(mn)^s} \\ &= \sum_{N=1}^{\infty} c_N \frac{1}{N^s} \end{aligned}$$

where $c_N = \sum_{d|N} a_d b_{N/d}$. Concerning convergence, note that if $|a_n| \leq kn^\ell$ for some $k, \ell \in \mathbb{N}$, then $\sum_{n=1}^{\infty} \frac{a_n}{n^s}$ converges absolutely for $\Re(s) > k + 1$.

For example. $\zeta(s)\zeta(s-1) = \left(\sum_{m=1}^{\infty} \frac{1}{m^s}\right) \left(\sum_{m=1}^{\infty} \frac{n}{n^s}\right) = \sum_{N=1}^{\infty} \frac{c_N}{N^s}$ where $c_N = \sum_{d|N} \frac{N}{d} = \sum_{d|N} d = \sigma(N)$ for $\Re(s) > 2$.

Define the **von Mangoldt function** $\Lambda : \mathbb{N} \rightarrow \mathbb{R}$ by:

$$\Lambda(n) = \begin{cases} \log p & n = p^k, k \geq 1, p \text{ prime} \\ 0 & \text{otherwise} \end{cases}$$

there are relatively few prime powers, so Λ is to be thought of as a weighted indicator function of the primes.

One can show that $\pi(x) = \sum_{1 \leq n \leq x} \mathbf{1}_{\text{primes}}(n) \sim \frac{x}{\log x}$, which is equivalent to seeing that $\Psi(x) := \sum_{1 \leq n \leq x} \Lambda(n) \sim x$.

Proposition 5.11. Let $s \in \mathbb{C}, \Re(s) > 1$. Then:

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

Proof. For $\Re(s) > 1$, we have that $\zeta(s) = \prod_p (1 - p^{-s})^{-1}$, and so $\log \zeta(s) = -\sum_p \log(1 - p^{-s})$. Now note that $\frac{d}{ds} p^{-s} = \frac{d}{ds} e^{-s \log p} = -\log(p) p^{-s}$.

Therefore, differentiating both sides, we have:

$$\begin{aligned} \frac{\zeta'(s)}{\zeta(s)} &= -\sum_p \frac{(\log p) p^{-s}}{1 - p^{-s}} \\ &= -\sum_p (\log p) \sum_{j=1}^{\infty} p^{-js} \\ &= -\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} \end{aligned}$$

□

Then integrating $\frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s}$ around a contour, one can show that:

$$\Psi(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \left(-\frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s} \right) ds$$

Then, shifting the contour to the left, this is equal to:

$$= x - \sum_{p:\zeta(p)=0} \frac{x^p}{p} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2} \log \left(1 - \frac{1}{x^2} \right)$$

Theorem 5.12. Let $X > 1$ and let $P = \prod_{p \leq \sqrt{X}} p$. Then $\pi(X) - \pi(\sqrt{X}) + 1 = \sum_{d|P} \mu(d) \lfloor \frac{X}{d} \rfloor$

Proof. By the sieve of Eratosthenes, if we delete from $\{2, 3, \dots, X\}$ all multiples of primes $\leq \sqrt{X}$, we are left with $\{p : \sqrt{X} < p \leq X\}$.

Hence we have that $\{1 \leq n \leq X : (n, P) = 1\} = \{1\} \cup \{p : \sqrt{X} < p \leq X\}$.

So $\pi(X) - \pi(\sqrt{X}) + 1 = \#\{1 \leq n \leq X : (n, P) = 1\}$. Also, by lemma 5.10, $\sum_{d|n} \mu(d) =$

$$\begin{cases} 1 & n = 1 \\ 0 & n > 1 \end{cases}. \text{ Hence:}$$

$$\begin{aligned} \#\{1 \leq n \leq X : (n, P) = 1\} &= \sum_{1 \leq n \leq X} \sum_{d|(n, P)} \mu(d) \\ &= \sum_{d|P} \sum_{1 \leq n \leq X, n \equiv 0 \pmod{d}} \mu(d) \\ &= \sum_{d|P} \mu(d) \sum_{1 \leq n \leq X, n \equiv 0 \pmod{d}} 1 \\ &= \sum_{d|P} \mu(d) \lfloor \frac{X}{d} \rfloor \end{aligned}$$

□

The next two results use divisibility properties of binomial coefficients

Theorem 5.13 (Tchebychev). $\exists c_2 > c_1 > 0$ s.t. $\forall x \geq 2$:

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

Where we can take $c_1 = \frac{\log 2}{2} \approx 0.346 \dots$, and $c_2 = 6 \log 2 \approx 4.158 \dots$ for $x \geq 4$.

Proof of upper bound. We shall sandwich $n \in \mathbb{N}$ between two powers of 2.

We first prove by induction on k that $\pi(2^k) \leq 3 \cdot \frac{2^k}{k}$ (\dagger) for $k \geq 1$.

Observe that, for even $x \geq 4$, $\pi(x) \leq \frac{x}{2}$, so $\pi(2^k) \leq 2^{k-1} \leq 3 \cdot \frac{2^k}{k}$ whenever $k \leq 6$.

Let $n \in \mathbb{N}, N = \binom{2n}{n} < 2^{2n}$. Then $N \geq \prod_{n < p \leq 2n} p$ since every prime in this product divides the numerator once and doesn't divide the denominator, and so $\geq n^{\pi(2n) - \pi(n)}$

So $\pi(2n) - \pi(n) \leq \frac{\log(2^{2n})}{\log n} = 2 \log 2 \frac{n}{\log n}$. Assuming the inductive hypothesis (\dagger) for k

$$\begin{aligned} \pi(2^{k+1}) &\leq \pi(2^k) + 2 \log 2 \frac{2^k}{\log 2^k} \\ &= \pi(2^k) + \frac{2^{k+1}}{k} \\ &\leq 3 \cdot \frac{2^k}{k} + \frac{2^{k+1}}{k} \\ &= 5 \cdot \frac{2^k}{k} = 5 \cdot \frac{2^{k+1}}{2k} \\ &\leq 3 \cdot \frac{2^{k+1}}{k+1} \end{aligned}$$

for $k \geq 5$. Hence (\dagger) holds for all $k \geq 1$. Moreover, $\frac{d}{dx} \left(\frac{x}{\log x} \right) = \frac{\log x - 1}{(\log x)^2} > 0$, so $\frac{x}{\log x}$ is monotonic increasing for $x \geq e$. So if $4 \leq 2^k < x \leq 2^{k+1}$, then $\pi(x) \leq \pi(2^{k+1}) \leq 6 \cdot \frac{2^k}{k+1} < 6 \log 2 \frac{2^k}{\log(2^k)} < 6 \log 2 \frac{x}{\log x}$. \square

Let $n \geq 1$, p a prime. Define the ***p-adic valuation of n*** as $v_p(n)$, the exponent of p in the factorisation of n .

Clearly $n = p^{v_p(n)} n_0$ with $(n_0, p) = 1$, and $v_p(nm) = v_p(n) + v_p(m)$. Question 6 on sheet 3 will show that $v_p(n!) = \sum_{i=1}^{\infty} \lfloor \frac{n}{p^i} \rfloor$

Lemma 5.14. Let $n \in \mathbb{N}, N = \binom{2n}{n}$. Then:

1. If $\frac{2n}{3} < p \leq 2n$, then $(N, p) = 1$
2. If $p^k | N$ for some $k \geq 1$ then $p^k \leq 2n$.

Proof.

1. If $p \in (\frac{2n}{3}, n]$, then $p \leq n < 2p \leq 2n < 3p$. Hence both the denominator and numerator of $N = \binom{2n}{n}$ are divisible by exactly one copy of p

\square

Proof of lower bound. Let $n \in \mathbb{N}$, $N = \binom{2n}{n} = \prod_{p \leq 2n} p^{v_p(N)}$. By the previous lemma, we have $p^{v_p(N)} \leq 2n$, and so $v_p(N) \leq \frac{\log 2n}{\log p}$.

It then follows from this that:

$$\begin{aligned} \log N &= \sum_{p \leq 2n} v_p(N) \log p \\ &\leq \sum_{p \leq 2n} \frac{\log 2n}{\log p} \log p \\ &= \pi(2n) \log 2n \end{aligned}$$

So we need a lower bound on N . Note that $2^{2n} = (1+1)^{2n} = \sum_{j=0}^{2n} \binom{2n}{j} = 2 + \sum_{j=1}^{2n-1} \binom{2n}{j} \leq 2 + (2n-1)N \leq 2nN$

So we have $\pi(2n) \geq \frac{\log N}{\log 2n} \geq \frac{\log(2^{2n}) - \log 2n}{\log 2n} = \log 2 \frac{2n}{\log 2n} - 1$

So if $2n < x \leq 2n+2$, we have $\pi(x) \geq \pi(2n) \geq \log 2 \frac{x-2}{\log x} - 1 \geq \frac{\log 2}{2} \frac{x}{\log x}$ \square

Theorem 5.15 (Bertrand's Postulate). *For all $n > 1$ there is a prime $p \in [n, 2n]$.*

For this, we need a lemma: for $x \geq 1$, let $P(x) = \prod_{p \leq x} p$. Then $P(x) \leq 4^x$ for all $x \geq 1$.

Proof. By induction on n , noting that $P(x) = P(\lfloor x \rfloor)$. Firstly, note that $P(2) \leq 4^2, P(3) \leq 4^3$.

For $n \geq 1$, $P(2n+1) = P(2n+1) = \prod_{p \leq n+1} p \cdot \prod_{n+1 \leq p \leq 2n+1} p = P(n+1) \cdot \prod_{n+1 \leq p \leq 2n+1} p$.

Now clearly $\prod_{n+2 \leq p \leq 2n+1} p$ divides $\binom{2n+1}{n}$ since each prime $p \in [n+2, 2n+1]$ divides $\binom{2n+1}{n}$.

Moreover, $2\binom{2n+1}{n} = \binom{2n+1}{n} + \binom{2n+1}{n+1} < \sum_{j=0}^{2n+1} \binom{2n+1}{j} = 2^{2n+1}$, so $\binom{2n+1}{n} \leq 4^n$.

It then follows that $P(2n+2) = P(2n+1) \leq P(n+1) \cdot 4^n \leq 4^{n+1} \cdot 4^n = 4^{2n+1}$. \square

Proof of Bertrand's Postulate. Suppose $\nexists p \in [n, 2n]$. Let $N = \binom{2n}{n}$. Then if $p|N$, we have $p \leq \frac{2n}{3}$ by 5.14, and also $p^{v_p(N)} \leq 2n$ for all p .

Write $N = N_1 \cdot N_2$ with $N_1 = \prod_{v_p(N)=1} p$, $N_2 = \prod_{v_p(N) \geq 2} p$.

Then $N_1 \leq \prod_{p \leq 2n/3} p = P(2n/3) \leq 4^{2n/3}$

And $N_2 = \prod_{v_p(N) \geq 2} p^{v_p(N)} \leq \prod_{p \leq \sqrt{2n}} 2n \leq (2n)^{\sqrt{2n}}$

Hence $\frac{2^{2n}}{2n} \leq N \leq 4^{2n/3} \cdot (2n)^{\sqrt{2n}}$, and so

$$2^{2n/3} \leq (2n)^{1+\sqrt{2n}} \implies \frac{\log 2}{3}(2n) \leq (1+\sqrt{2n}) \log(2n)$$

which is obviously false for large values of n , e.g. $n = 2^9$.

For smaller values of n , consider the list of primes $p = 2, 5, 11, 23, 47, 89, 179, 359, 719$, where each p_i is the largest prime $\leq 2(p_{i-1} + 1)$. \square

6 Continued Fractions

We want to find good rational approximations for irrational numbers, where $|\pi - \frac{p}{q}|$ small relative to q , e.g. $335/113$. We define a possibly finite sequence of integers a_0, a_1, a_2, \dots with $a_n \geq 1$ for $n \geq 1$ as follows:

1. Let $a_0 = \lfloor \theta \rfloor$. If $\theta = a_0$, HALT. Otherwise, let $\theta_1 = 1/\theta - a_0 > 1$.
2. Let $a_1 = \lfloor \theta_1 \rfloor \geq 1$. Stop if $\theta_1 = 1$. Otherwise let $\theta_2 = \frac{1}{\theta_1 - a_1} > 1$, and repeat.

If this algorithm stops at step n , then we write $\theta = [a_0; a_1, a_2, \dots, a_n]$, otherwise, we write $\theta = [a_0; a_1, a_2, \dots, a_{n-1}, \theta_n]$.

We call this the **continued fraction expansion** of θ . The a_i are known as the **partial quotients** and the θ_i are the **complete quotients** for θ .

For instance, $\theta = \frac{59}{13} = 4 + \frac{7}{13} = 4 + \frac{1}{1+\frac{6}{7}} = 4 + \frac{1}{1+\frac{1}{\frac{7}{6}}}$

Proposition 6.1. *The CFE of $\theta \in \mathbb{R}$ terminates iff $\theta \in \mathbb{Q}$.*

Proof. Only if is clear.

Suppose $\theta = \frac{b_0}{c_0} \in \mathbb{Q}$. Then each θ_i is also in \mathbb{Q} , say $\theta_i = \frac{b_i}{c_i}$ in lowest terms. Then $\theta_i = a_i + \frac{1}{\theta_{i+1}}$, so $b_{i+1} = c_i < b_i$, and so the b_i strictly decrease but are positive, so it terminates. \square

Note that if $\frac{b_i}{c_i} = a_i + \frac{c_{i+1}}{b_{i+1}}$ then $b_i = a_i b_{i+1} + b_{i+2}$, so all we're really doing is Euclid's algorithm on (b_0, c_0) .

Let a_0, a_1, \dots be integers with $a_n \geq 1$. We define the **convergents** of the continued fraction expansion $[a_0; a_1, a_2, \dots]$ as pairs of integers (p_n, q_n) for $n \geq 0$ as follows:

$$p_0 = a_0, p_1 = a_0 a_1 + 1, p_n = a_n p_{n-1} + p_{n-2}$$

$$q_0 = 1, q_1 = a_1, q_n = a_n q_{n-1} + q_{n-2}$$

Note that $1 \leq q_1 < q_2 < \dots$ is an increasing sequence of positive integers tending to ∞ . We can also write this in matrix form, and sometimes let $p_{-1} = 1, q_{-1} = 0$, giving:

$$\begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \dots \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}$$

Proposition 6.2.

1. $(p_n, q_n) = 1$, and in fact $p_n q_{n-1} - p_{n-1} q_n = (-1)^{n-1}$
2. $[a_0; a_1, a_2, \dots, a_n] = \frac{p_n}{q_n}$
3. If $\alpha = [a_0; a_1, \dots, a_{n-1} \beta]$ for some real $\beta > 0$, then α lies strictly between $\frac{p_n}{q_n}$ and $\frac{p_{n-1}}{q_{n-1}}$

Proof.

1. Take the determinant of the matrix and this appears
2. By induction on n - base case is immediate. For induction step, just do manual manipulation.

3. By induction on n - base case is immediate. Let $n > 0$ and set $\gamma = a_n + \frac{1}{\beta}$. By the induction hypothesis:

$$\begin{aligned}\alpha &= [a_0; a_1, \dots, a_{n-1}, \gamma] = \frac{p_{n-1}\gamma + p_{n-2}}{q_{n-1}\gamma + q_{n-2}} \\ &= \frac{p_{n-1}(a_n\beta + 1) + \beta p_{n-2}}{q_{n-1}(a_n\beta + 1) + \beta q_{n-2}} \\ &= \frac{(a_n p_{n-1} + p_{n-2})\beta + p_{n-1}}{(a_n q_{n-1} + q_{n-2})\beta + q_{n-1}} \\ &= \frac{p_n\beta + p_{n-1}}{q_n\beta + q_{n-1}}\end{aligned}$$

Finally, if $\frac{x}{y} < \frac{x'}{y'}$ with $y, y' > 0$ then $\frac{x}{y} < \frac{x+x'}{y+y'} < \frac{x'}{y'}$, and so we are done. \square

Theorem 6.3. *Let $\theta \in \mathbb{R}$ be irrational with $[a_0; a_1, a_2, \dots]$. Then $\forall n \geq 0, |\theta - \frac{p_n}{q_n}| < \frac{1}{q_n q_{n-1}}$. In particular $\frac{p_n}{q_n} \rightarrow \theta$ as $n \rightarrow \infty$.*

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

Then write $\theta = [a_0; a_1, a_2, \dots, a_{n+1}, \theta_{n+2}]$, and we see that θ lies between $\frac{p_n}{q_n}$ and $\frac{p_{n+1}}{q_{n+1}}$.

Hence $|\theta - \frac{p_n}{q_n}| < \frac{1}{q_n q_{n+1}}$. Then observing that $q_n \rightarrow \infty$ gives the last part of the statement. \square

As a result of this theorem, we can think of $\theta = [a_0; a_1, \dots] = \lim_{n \rightarrow \infty} [a_0; a_1, \dots, a_n]$.

If we fix a denominator $q \geq 1$ and try to minimise $|\theta - \frac{p}{q}|$, the best we can expect in general is $\frac{1}{2q}$, where p is the closest integer to θq .

Theorem 6.4. *Let $\theta \in \mathbb{R}$, and suppose that (p_n, q_n) are convergents for the CFE of θ .*

Suppose $q \in \mathbb{N}$ with $1 \leq q < q_{n+1}$ for some $n \in \mathbb{N}$. Then for all $p \in \mathbb{Z}, |q\theta - p| \geq |q_n\theta - p_n|$.

Moreover, if $q \in \mathbb{N}, p \in \mathbb{Z}$ are such that $|\theta - \frac{p}{q}| < |\theta - \frac{p_n}{q_n}|$, then $q > q_n$.

Proof. For the “moreover” part, suppose that $q \leq q_n$. Then the first part of the theorem implies that $|\theta - \frac{p}{q}| \geq \frac{1}{q} |q_n\theta - p_n| = \frac{q_n}{q} |\theta - \frac{p_n}{q_n}| \geq |\theta - \frac{p_n}{q_n}|$.

Now let $q \in \mathbb{N}, p \in \mathbb{Z}$ with $1 \leq q < q_{n+1}$ for some $n \in \mathbb{N}$. Consider the system of equations $p_n u + p_{n+1} v = p; q_n u + q_{n+1} v = q$.

Then we have that $\det \begin{pmatrix} p_n & p_{n+1} \\ q_n & q_{n+1} \end{pmatrix} = (-1)^n$, so this system has a unique solution (u, v) with $u, v \in \mathbb{Z}$ not both 0.

Thus $q\theta - p = q_n u + q_{n+1} v \theta - (p_n u + p_{n+1} v) = (q_n\theta - p_n)u + (q_{n+1}\theta - p_{n+1})v$.

If $v = 0$, then $|q\theta - p| = |q_n\theta - p_n||u| \geq |q_n\theta - p_n|$.

We may assume $v \neq 0$. By the second equation, $u \neq 0$ and has the opposite sign to v .

Also, $\theta - \frac{p_n}{q_n}$ and $\theta - \frac{p_{n+1}}{q_{n+1}}$ have opposite signs, and it follows that $(q_n\theta - p_n)u$ and $(q_{n+1}\theta - p_{n+1})v$ have the same sign.

So $|q\theta - p| = |(q_n\theta - p_n)u| + |(q_{n+1}\theta - p_{n+1})v| \geq |(q_n\theta - p_n)u| \geq |q_n\theta - p_n|$. \square

Theorem 6.5. *Let $\theta \in \mathbb{R} \setminus \mathbb{Q}$, and suppose that (p_n, q_n) are the convergents for the CFE of θ . Then:*

1. *At least one of any pair of consecutive convergents satisfies $|\theta - \frac{p}{q}| < \frac{1}{2q^2}$.*
2. *If $\frac{p}{q} \in \mathbb{Q}$ is such that $|\theta - \frac{p}{q}| < \frac{1}{2q^2}$, then $p/q = p_n/q_n$ for some n .*

Proof.

1. Suppose that for some $n \in \mathbb{N}$, $|\theta - \frac{p_n}{q_n}| \geq \frac{1}{2q_n^2}$ and $|\theta - \frac{p_{n+1}}{q_{n+1}}| \geq \frac{1}{2q_{n+1}^2}$.

Then $\frac{1}{2} \left(\frac{1}{q_n^2} + \frac{1}{q_{n+1}^2} \right) \leq |\theta - \frac{p_n}{q_n}| + |\theta - \frac{p_{n+1}}{q_{n+1}}| = |\frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}}| = \frac{1}{q_n q_{n+1}}$, contradicting AMGM \nmid .

2. Suppose $\frac{p}{q} \in \mathbb{Q}$ with $|\theta - \frac{p}{q}| < \frac{1}{2q^2}$. Then there is some $n \geq 0$ such that $q_n \leq q < q_{n+1}$. We claim that $\frac{p}{q} = \frac{p_n}{q_n}$, because if not, then $\frac{1}{qq_n} \leq |\frac{p}{q} - \frac{p_n}{q_n}| \leq |\theta - \frac{p}{q}| + |\theta - \frac{p_n}{q_n}| = \frac{1}{q} |\theta q - p| + \frac{1}{q_n} |\theta q_n - p_n|$, and so $q < q_n$ \nmid .

\square

A CFE of the form $[a_0; a_1, \dots, a_n, \dots, a_{n+k}, a_n, \dots, a_{n+k}, a_n, \dots]$ is called **eventually periodic**, and is abbreviated as $[a_0; a_1, \dots, \overline{a_n, \dots, a_{n+k}}]$. It is called **purely periodic** if $n = 0$, and the **period** is the least positive integer $k + 1$ such that this happens.

Theorem 6.6 (Lagrange). *The CFE of $\theta \in \mathbb{R} \setminus \mathbb{Q}$ is eventually periodic if and only if θ is a quadratic irrational.*

Proof. Suppose $\phi = [\overline{a_0; a_1, \dots, a_n}]$ is purely periodic. Then $\phi = [a_0, a_1, \dots, a_n, \phi] = \frac{p_n \phi + p_{n-1}}{q_n \phi + q_{n-1}}$. So $\phi^2 + (q_{n-1} - p_n)\phi - p_{n-1} = 0$, and so ϕ is a quadratic irrational.

We then observe quadratic irrationals are closed under division, multiplication, addition, and so if ϕ is eventually periodic, the periodic part is quadratic irrational and so ϕ is also.

Conversely, if $a\theta^2 + b\theta + c = 0$. Then let $f = (a, b, c)$ be a BQF which we can write as $f(x, y) = a(x - \theta y)(x - \theta' y)$, where θ' is the conjugate root of this quadratic.

Then let $f_n(x, y) = f(p_n x + p_{n-1} y, q_n x + q_{n-1} y) = (a_n, b_n, c_n)$. We shall see that the complete quotients θ_n of θ arise as solutions to $f_{n-1}(x, 1) = 0$, and that there are only finitely many forms f_n .

First observe that $\text{disc}(f_n) = \text{disc}(f) = D$. Also $\theta = \frac{p_n \theta_{n+1} + p_{n-1}}{q_n \theta_{n+1} + q_{n-1}}$, so $f_n(\theta_{n+1}, 1) = (q_n \theta_{n+1} + q_{n-1})^2 f(\theta, 1) = 0$. Now observe that $|f(p_n, q_n)| \leq 2a|\theta - \theta'|$, so it follows that $f_n(1, 0) = f(p_n, q_n) = a_n$, and $f_{n+1}(0, 1) = f(p_n, q_n) = c_{n+1}$ are both bounded above. So there are finitely many distinct f_n , and so $f_{n+k} = f_n$. Hence the set of complete quotients is finite, and so the CFE is eventually periodic. \square

Theorem 6.7 (Galois). *θ has a purely periodic CFE if and only if $\theta > 1$ and $\theta' \in (-1, 0)$.*

Corollary 6.8. *If $d \in \mathbb{N}$ is not a square, then $\sqrt{d} = [a_0; \overline{a_1, \dots, a_m}]$.*

Proof. $\theta_1 = \frac{1}{\sqrt{d}-a_0} > 1$, with its conjugate root $\theta'_1 = \frac{1}{-\sqrt{d}-a_0} \in (-1, 0)$. \square

In fact it is the case that $\sqrt{d} = [a_0; \overline{a_1, a_2, \dots, a_2, a_1, 2a_0}]$.

Theorem 6.9 (Pell's Equation). *If $d \in \mathbb{N}$ is not a square then $x^2 - dy^2 = 1$ has a solution in integers $x, y \neq 0$.*

Proof. Let $\theta = \sqrt{d} = [a_0; a_1, \dots, a_n, \theta_{n+1}]$ with $\theta_1 = \theta_{n+1}$.

So $\sqrt{d} = \frac{p_n \theta_{n+1} + p_{n-1}}{q_n \theta_{n+1} + q_{n-1}} = \frac{p_n + p_{n-1}(\sqrt{d} - a_0)}{q_n + q_{n-1}(\sqrt{d} - a_0)}$ and so $q_{n-1}d + (q_n - q_{n-1}a_0)\sqrt{d} = p_n - a_0p_{n-1} + p_{n-1}\sqrt{d}$.

Comparing coefficients, we have $p_{n-1}^2 - dq_{n-1}^2 = q_n p_{n-1} - p_n q_{n-1} = (-1)^n$.

And so if n is even then $(x, y) = (p_{n-1}, q_{n-1})$ is a solution to Pell's equation. If n is odd, then (p_{2n-1}, q_{2n-1}) is a solution. \square