COURSE NOTES ALGEBRAIC NUMBER THEORY

1. Day 1: Quadratic reciprocity

- (a) Quadratic reciprocity
- (b) Elementary proof of quadratic reciprocity (Flath)
- (c) Quadratic fields and integers
- (d) Rings of algebraic integers

Let k be a finite field of characteristic p. So $|k| = p^r$ for some $r > 0$. We recall the following basic fact, to be proved during Week 3:

Proposition 1.1. *The multiplicative group $k^\times$ of k is cyclic of order $p^r - 1$.*

Suppose now that p is odd and $q = p$, so $k = \mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. Then $\mathbb{F}_p^\times$ is a cyclic group of even order $p-1$ and contains a unique subgroup of index 2. This is precisely the image of the map

$$[2] : \mathbb{F}_p^\times \to \mathbb{F}_p^\times; x \mapsto x^2$$

because the kernel has order 2. So half the elements of $\mathbb{F}_p^\times$ are squares and half are not. If $(a, p) = 1$, define the *Legendre symbol* $\left(\frac{a}{p}\right)$ to be 1 if a is a square mod p, -1 if not. The following example is easy:

Proposition 1.2. $\left(\frac{-1}{p}\right) = 1$ *if* $p \equiv 1 \pmod 4$, *and* $\left(\frac{-1}{p}\right) = -1$ *if* $p \equiv 3 \pmod 4$.

Proposition 1.3. *(i) For any $a \in \mathbb{Z}$ prime to p,*

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

(ii) The Legendre symbol is multiplicative.

Proof. It is clear that (i) implies (ii), so we prove (i). The cyclic group $\mathbb{F}_p^\times$ contains a unique subgroup C of order 2. Consider the map

$$e : \mathbb{F}_p^\times \to \mathbb{F}_p^\times; x \mapsto x^{\frac{p-1}{2}}.$$

Now $e \circ [2](x) = x^{p-1} = 1$ (by Fermat's little theorem). So $e(a) = 1$ if $\left(\frac{a}{p}\right) = 1$. On the other hand, if a is a cyclic generator of $\mathbb{F}_p^\times$ then its

order is exactly $p-1$, so $e(a) \neq 1$. Thus e is surjective, so its kernel is exactly the subgroup of index 2, i.e, $e(a) = -1$ if $\left(\frac{a}{p}\right) = -1$. $\square$

Let $q \in \mathbb{Z}$ be a prime different from p. The *quadratic reciprocity theorem* determines $\left(\frac{q}{p}\right)$ algorithmically:

Theorem 1.4. *Suppose q and p are odd primes. Then*

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

Gauss gave many proofs of this theorem, four of which are in Flath's book. The first one, found when he was 19, is "elementary" but in some sense the most complicated. I begin the course with a part of the most elementary proof.

1.1. *Solutions to composite congruences.* First we state a few lemmas about congruences.

Lemma 1.5. *Let $f \in \mathbb{Z}[X]$. Let $m_1, \ldots, m_r$ be positive pairwise relatively prime integers. Then the congruence $f(X) \equiv 0 \pmod{m_1 \ldots m_r}$ has a solution if and only if each of the congruences $f(X) \equiv 0 \pmod{m_i}$ has a solution.*

This is an easy consequence of the Chinese remainder theorem (presented in T.D.?).

Lemma 1.6. *Let $f \in \mathbb{Z}[X]$. Let p be a prime, and suppose $\bar{f}$ has no multiple roots in $\bar{\mathbb{F}}_p$. Then the congruence $f(X) \equiv 0 \pmod{p}$ has a solution if and only if for all $r \geq 1$ the congruence $f(X) \equiv 0 \pmod{p^r}$ has a solution.*

This is a version of Hensel's lemma and it will be proved later. If p is odd, it applies to the polynomial $x^2 - n$ provided $(n, p) = 1$. Finally, the following is an easy exercise.

Lemma 1.7. *Let a be an odd integer. Then a is a square mod 2^r for all $r \geq 3$ if and only if $a \equiv 1 \pmod{8}$.*

This will be explained when we talk about p-adic numbers.

1.2. *An amazing lemma.* At one stage he uses a lemma that Flath says has "the most amazing proof in" his book.

Lemma 1.8. *Let q be a prime congruent to 1 mod 4. Then there exists an odd prime $p' < q$ such that $\left(\frac{q}{p'}\right) = -1$.*

Proof. First suppose $q \equiv 5 \pmod 8$. Then $\frac{q+1}{2} \equiv 3 \pmod 4$. So at least one divisor p' of (the odd number) $\frac{q+1}{2}$ is also $\equiv 3 \pmod 4$. Obviously $p' < q$; but p' divides $q+1$ so $q \equiv -1 \pmod{p'}$. Thus $\left(\frac{q}{p'}\right) = \left(\frac{-1}{p'}\right) = -1$ because $p' \equiv 3 \pmod 4$.

The hard case is $q \equiv 1 \pmod 8$. Let $m \in \mathbb{N}$ satisfy $1 < 2m+1 < q$ and $\left(\frac{q}{p}\right) = 1$ for all odd $p \leq 2m+1$. Thus the equation $X^2 \equiv q \pmod p$ has a solution for all odd divisors p of $(2m+1)!$ and since q is a square mod 8 it is a square modulo every power of 2. Thus by the Chinese remainder theorem, it is a square mod $(2m+1)!$. In other words, there is an x satisfying $x^2 \equiv q \pmod{(2m+1)!}$ and we may assume $x > m$.

Now write

$$\prod_{i=1}^{m}(q-i^2) \equiv \prod_{i=1}^{m}(x^2-i^2) = \frac{(2m+1)!\binom{x+m}{2m+1}}{x} \pmod{(2m+1)!}$$

But the binomial coefficient is an integer, say C, and $q > (2m+1)$ and so q, and therefore x, is prime to $(2m+1)!$, so by Gauss's Lemma x must divide C; say $C = Bx$. Thus

$$\prod_{i=1}^{m}(q-i^2) \equiv (2m+1)!B \equiv 0 \pmod{(2m+1)!}$$

But

$$\frac{\prod_{i=1}^{m}(q-i^2)}{(2m+1)!} = \frac{1}{m+1}\prod_{i=1}^{m}\frac{q-i^2}{(m+1)^2-i^2} \in \mathbb{Z}.$$

But suppose $m^2 < q < (m+1)^2$. Then every factor in the product is a fraction between 0 and 1. In particular, this equation is impossible if $m = [\sqrt{q}]$. But $q \equiv 1 \pmod 8$ so $q \geq 17$, and thus $2[\sqrt{q}]+1 < q$. So the hypothesis was false, and there is some $p' \leq 2[\sqrt{q}]+1$ such that $\left(\frac{q}{p'}\right) = -1$. □

1.3. *Proof of quadratic reciprocity.* Let $p^* = (-1)^{\frac{(p-1)}{2}}p$. Now as an exercise, we state that

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

is equivalent to quadratic reciprocity. Moreover, this is symmetric in p, q.

We can thus assume $p < q$ and prove quadratic reciprocity in this form by induction on q. The first case is $\left(\frac{3^*}{5}\right) = \left(\frac{5}{3}\right) = -1$ which we check byhand. So assume $q > 5$ and the theorem is known for pairs less than 1. There are three cases.

Case (1). $\left(\frac{p*}{q}\right) = 1$. Thus there are integers u, a, $0 < a < q$, with $a^2 = p^* + uq$. Replacing a by $q - a$ if necessary, we may assume a even; then u is odd. Moreover, we have the elementary inequalities (because $q > p$ and $a > 0$, $a < q$)

$$-q < a^2 - p^* \leq (q-1)^2 + p < q^2 - q$$

so $-q < uq < q^2 - q$, i.e. $-1 < u < q$, or $1 \leq u < q$.

First suppose $p \nmid u$; then $p \nmid a$. But $a^2 \equiv uq \pmod{p}$ so $\left(\frac{q}{p}\right) = \left(\frac{u}{p}\right)$, and so to conclude it suffices to show $\left(\frac{u}{p}\right) = 1$. Write $u = \prod_i p_i$ (not necessarily distinct, but all distinct from p). Since $a^2 \equiv p^* \pmod{p_i}$, we have $\left(\frac{p*}{p_i}\right) = 1$ for each i. But since each $p_i < q$, by induction we have $\left(\frac{p_i}{p}\right) = 1$ for all p, and so $\left(\frac{u}{p}\right) = 1$.

Now suppose $p \mid u$, so $p \mid a$. Let $A = \frac{a}{p}$, $U = \frac{u}{p}$. Then $pA^2 = (-1)^{\frac{(p-1)}{2}} + Uq$, and now $p \nmid U$. This implies

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

or

$$\left(\frac{Uq}{p}\right) = \left(\frac{(-1)^{\frac{(p+1)}{2}}}{p}\right) = [(-1)^{\frac{(p+1)}{2}}]^{\frac{(p-1)}{2}} = 1.$$

So $\left(\frac{q}{p}\right) = \left(\frac{U}{p}\right)$. Write $U = \prod_i p_i$. We have $p^*A^2 = 1 + \pm Uq$, so $\left(\frac{p*}{p_i}\right) = 1$ for all i, and by induction as above we find $\left(\frac{U}{p}\right) = 1$.

Case (2). $\left(\frac{p*}{q}\right) = -1$, $q \equiv 3 \pmod{4}$.

Case (3). $\left(\frac{p*}{q}\right) = -1$, $q \equiv 1 \pmod{4}$.

The proofs are similarly elementary. In Case (2) the hypothesis implies that $\left(\frac{-p*}{q}\right) = 1$, and one finds $0 < a < q$ with $a^2 = -p^* + uq$

and continues as before. This is not possible in Case (3). Instead, one begins by using the amazing lemma to find $p' < q$ with $\left(\frac{q}{p'}\right) = -1$. If $p = p'$ we are done; so assume $p' \neq p$. If $\left(\frac{p'}{q}\right) = 1$ then $\left(\frac{p'^{,*}}{q}\right) = 1$ because $\left(\frac{-1}{q}\right) = 1$, and by Case (1) we have $\left(\frac{q}{p'}\right) = 1$, contradiction. So $\left(\frac{p'}{q}\right) = -1$, which means that $\left(\frac{pp'}{q}\right) = 1$, and we find $0 < a < q$ with $a^2 = pp' + uq$ and continue as before.

Instead of presenting the details of Cases (2) and (3) I will give a different (complete) proof in week 3 using Gauss sums. Before then, I will describe the interpretation of this theorem in terms of algebraic number theory.

Algebraic integers. We want to solve the polynomial equation $Q_q(X) = X^2 - q = 0$ in $\mathbb{F}_p$, or equivalently, if α is a root of Q_q we want to decide whether $\mathbb{F}_p(\alpha) = \mathbb{F}_p$ or a quadratic extension. We start by looking at $Q_q \in \mathbb{Z}[X]$. It has a root $\sqrt{q} \in \mathbb{C}$ that is not in $\mathbb{Q}$; so $[\mathbb{Q}(\sqrt{q}) : \mathbb{Q}] = 2$. The intuition is that $\left(\frac{q}{p}\right) = 1$ exactly when $\sqrt{q} \equiv a \pmod{p}$ for some $a \in \mathbb{Z}$. But what does that congruence mean?

It means nothing in the field $\mathbb{Q}(\sqrt{q})$, which has no non-trivial ideals. Instead we work with its subring $\mathcal{O}$ of *algebraic integers*:

Definition 1.9. An algebraic integer is a complex number α such that $P(\alpha) = 0$ for some monic polynomial $P \in \mathbb{Z}[X]$.

Let $\alpha = a + b\sqrt{q}$, with $a, b \in \mathbb{Q}$. When is α an integer? Letting $s(\alpha) = a - b\sqrt{q}$; $s \in Gal(\mathbb{Q}(\sqrt{q}) : \mathbb{Q})$, the minimal monic polynomial of α is

$$P(X) = (X - \alpha)(X - s(\alpha)) = X^2 - 2aX + (a^2 - qb^2).$$

Thus $2a, N = a^2 - qb^2 \in \mathbb{Z}$. So $a = \frac{c}{2}, b = \frac{d}{2}$ with c and $c^2 - qd^2 \in 4\mathbb{Z}$, which implies $d \in \mathbb{Z}$ (if $d = u/v$ in lowest terms, then v^2 divides q, so $v = \pm 1$). If c is even then so is d; if c is odd, then so is d, and then $q \equiv 1 \pmod 4$. Thus

Proposition 1.10. *The ring $\mathcal{O}$ of algebraic integers in $\mathbb{Q}(\sqrt{q})$ is $\mathbb{Z}[\sqrt{q}]$ if $q \equiv 3 \pmod 4$, and is $\mathbb{Z}[\frac{1+\sqrt{q}}{2}]$ if $q \equiv 1 \pmod 4$.*

More generally, if d is a square-free integer, the ring $\mathcal{O}$ of algebraic integers in $\mathbb{Q}(\sqrt{d})$ is $\mathbb{Z}[\sqrt{d}]$ if $d \equiv 2, 3 \pmod 4$, and is $\mathbb{Z}[\frac{1+\sqrt{d}}{2}]$ if $d \equiv 1 \pmod 4$.

Proposition 1.11. *The element $\alpha \in \mathbb{C}$ is integral over $\mathbb{Z}$ if and only if the subring $\mathbb{Z}[\alpha] \subset \mathbb{C}$ is a $\mathbb{Z}$-module of finite type.*

Proof. If α satisfies a monic polynomial over $\mathbb{Z}$ of degree n then $\alpha^n \in \mathbb{Z}+\mathbb{Z}\alpha+\ldots\mathbb{Z}\alpha^{n-1}$ and higher powers are in the same ring by induction. Conversely, if $\mathbb{Z}[\alpha]$ is of finite type then it is contained in a finitely generated $\mathbb{Z}$-module $\mathbb{Z}u_1+\mathbb{Z}u_2+\ldots\mathbb{Z}u_n$ for some n, so that the matrix A of multiplication by α in the basis $(u_1,\ldots,u_n)$ has coefficients in $\mathbb{Z}$. But then α satisfies the monic polynomial $\det(XI_n - A) \in \mathbb{Z}[X]$ by the Cayley-Hamilton theorem. □

This proof works with $\mathbb{Z}$ replaced by any integral domain.

Corollary 1.12. *Let $K/\mathbb{Q}$ be a finite extension. The subset $\mathcal{O}_K$ of algebraic integers in K is a subring.*

Proof. If $\alpha, \beta \in \mathcal{O}_K$, then $\mathbb{Z}[\alpha]$ is finite over $\mathbb{Z}$ and so is $\mathbb{Z}[\beta]$, so $\mathbb{Z}[\alpha,\beta]$ is contained in $\mathbb{Z}(\alpha^i\beta^j)$ for $0 \leq i,j \leq M$ for some M, which implies that every element in $\mathbb{Z}[\alpha,\beta]$ is integral over $\mathbb{Z}$. □

2. Day 2

- (a) Norms and traces, norm of an ideal
- (b) Dedekind properties
- (c) $\sum e_i f_i = g$
- (d) Galois properties, Residue fields, decomposition groups, Frobenius (Samuel)

Say $[K : \mathbb{Q}] = n$. For any $\alpha \in K$, we consider the linear transformation $A_\alpha : K \to K$, $A_\alpha(x) = \alpha \cdot x$. Let $N_{K/\mathbb{Q}}(\alpha) = \det(A_\alpha)$, $Tr_{K/\mathbb{Q}}(\alpha) = Tr(A_\alpha)$. The bilinear form

$$B(\alpha,\beta) = Tr_{K/\mathbb{Q}}(\alpha \cdot \beta)$$

is non-degenerate because if $\alpha \neq 0$ then $B(\alpha,\alpha^{-1}) = n \neq 0$.

Proposition 2.1. *The ring $\mathcal{O}_K$ is a free $\mathbb{Z}$-module of rank n. In particular, $\mathcal{O}_K$ is a noetherian ring. Moreover, any non-zero ideal $I \subset \mathcal{O}_K$ is a free $\mathbb{Z}$-module of rank n.*

Proof. It suffices to show $M \subset \mathcal{O}_K \subset N$ where M and N are free $\mathbb{Z}$-modules of rank n, since $\mathbb{Z}$ is a PID. For any $\alpha \in K$ there exists $d \in \mathbb{Z}$ such that $d\alpha \in \mathcal{O}_K$ (take the minimal monic polynomial of α over $\mathbb{Q}$ and choose d divisible by all its coefficients). So for any basis $e_1,\ldots,e_n$ of $K/\mathbb{Q}$ there exists D such that

$$M = De_1 \oplus \cdots \oplus De_n \subset \mathcal{O}_K.$$

Write $f_i = De_i$ and consider the dual basis f_i^* for the bilinear form B. Every element β such that $B(m, \beta) \in \mathbb{Z}$ for all $m \in M$ is in the $\mathbb{Z}$-span N of the f_i^*. But $B(\mathcal{O}_K, \mathcal{O}_K) \subset \mathbb{Z}$, hence $\mathcal{O}_K \subset N$.

Now let $I \subset \mathcal{O}_K$, $0 \subset \alpha \in I$. Since the principal ideal $(\alpha) = \alpha\mathcal{O}_K \subset I\mathcal{O}_K$, it suffices to replace I by (α). But multiplication by α is invertible, hence injective, and so $(\alpha) \simeq \mathcal{O}_K$ as $\mathbb{Z}$-module. □

Corollary 2.2. *The ring $\mathcal{O}_K$ is integrally closed: any element of K integral over $\mathcal{O}_K$ is already in K.*

Proof. Let $\alpha \in K$ be integral over $\mathcal{O}_K$. Then $\mathcal{O}_K[\alpha]$ is a finite module over $\mathcal{O}_K$. But since $\mathcal{O}_K$ is itself a finite $\mathbb{Z}$-module, any finite $\mathcal{O}_K$ module is also a finite $\mathbb{Z}$-module. Thus $\mathbb{Z}[\alpha] \subset \mathcal{O}_K[\alpha]$ is contained in a $\mathbb{Z}$-module of finite type, hence is itself a $\mathbb{Z}$-module of finite type. Thus α is integral over $\mathbb{Z}$, i.e. belongs to $\mathcal{O}_K$. □

The next corollary is obvious.

Corollary 2.3. *Let $K/\mathbb{Q}$ be an extension of degree n, and let p be a prime number. Then the ring $\mathcal{O}_K/p\mathcal{O}_K$ is an $\mathbb{F}_p$-algebra of dimension n.*

Corollary 2.4. *Let $K/\mathbb{Q}$ be a finite extension, and let $I \subset \mathcal{O}_K$ be a non-zero prime ideal. Then I is maximal.*

Proof. Let $J = I \cap \mathbb{Z}$. Since $\mathbb{Z}/J$ injects into $\mathcal{O}_K/I$, J is a prime ideal. On the other hand, $I \hookrightarrow \mathcal{O}_K$ is an inclusion of $\mathbb{Z}$-modules of the same rank, so $\mathcal{O}_K/I$ is a finite ring. Now any finite integral domain A is a field. Indeed, if $0 \neq \alpha \in A$, multiplication by α is injective since there are no zero-divisors; but since A is finite, multiplication by α is a bijection, so α has a multiplicative inverse. Thus I is a maximal ideal. □

Thus $\mathcal{O}_K$ is a Dedekind ring:

Definition 2.5. A Dedekind ring is an integral domain that is noetherian, integrally closed, and all of whose non-zero prime ideals are maximal.

Return to the case $n = 2$, $K = \mathbb{Q}(\sqrt{q})$, $\mathcal{O} = \mathcal{O}_K$. Let p be an odd prime; then

$$\mathcal{O}/p\mathcal{O} = \mathcal{O}_{(p)}/p\mathcal{O}_{(p)}$$

(localization at p as $\mathbb{Z}$-module), but $\mathcal{O}_{(p)} = \mathbb{Z}_{(p)}[\sqrt{q}]$, so

$$\mathcal{O}/p\mathcal{O} = \mathcal{O}_{(p)}/p\mathcal{O}_{(p)} = \mathbb{Z}_{(p)}[\sqrt{q}]/p(\mathbb{Z}_{(p)}[\sqrt{q}]) = \mathbb{Z}_{(p)}[X]/(p, X^2 - q)$$

which can also be written

$$\mathbb{F}_p[X]/(X^2 - q).$$

Now there are three possibilities:

(1) $X^2 - q$ has a double root over $\mathbb{F}_p$; in other words, $2X$ and $X^2 - q$ have a common factor. This is possible only if $p = 2$ or if $p = q$. The case $p = 2$ requires a separate argument. If $p = q$ then $\mathcal{O}/p\mathcal{O}$ has nilpotents; it's isomorphic to $\mathbb{F}_p[X]/(X^2)$.

(2) $X^2 - q$ is irreducible over $\mathbb{F}_p$; i.e. $\left(\frac{q}{p}\right) = -1$. Then the ideal $p\mathcal{O}$ is prime in $\mathcal{O}$.

(3) $X^2 - q$ is reducible over $\mathbb{F}_p$; i.e. $\left(\frac{q}{p}\right) = 1$. Then $\mathcal{O}/p\mathcal{O} \equiv \mathbb{F}_p \times \mathbb{F}_p$ and the ideal $p\mathcal{O}$ is not prime in $\mathcal{O}$. In fact, $p\mathcal{O}$ is contained in two distinct primes, the kernels of the projection to the two factors $\mathbb{F}_p \times \mathbb{F}_p$.

It is the last interpretation that underlies generalizations of the quadratic reciprocity theorem. Let $I, J \subset R$ be two ideals. Say I and J are *relatively prime* if $I + J = R$. The product $I \cdot J$ is the smallest ideal containing products $a \cdot b$ with $a \in I$, $b \in J$. The product can be extended to fractional ideals:

Definition 2.6. A fractional ideal in K is a non-zero $\mathcal{O}_K$ submodule of finite type.

The definition makes sense in any integral domain. Not all integer rings $\mathcal{O}_K$ have unique factorization; they are not all principal. But they are Dedekind rings, and thus enjoy the following fundamental property:

Theorem 2.7. *Let R be a Dedekind ring. Then any ideal of R admits a unique factorization as a product of prime ideals, and the set of fractional ideals of R forms a group under multiplication.*

Write $\mathcal{O} = \mathcal{O}_K$.

Proof. Suppose $\alpha \in K^\times$ and $I \subset \mathcal{O}$, such that $\alpha I \subset I$. Then $\alpha \in \mathcal{O}$. Indeed, multiplication by α satisfies the characteristic polynomial of its matrix, which is monic and by hypothesis has $\mathcal{O}$-coefficients. (Here we are using that $\mathcal{O}$ is integrally closed in its quotient field.)

Next, if $I, J \subset \mathcal{O}$ and $I = IJ$, then $J = \mathcal{O}$. Indeed, let α_i, $i = 1, \ldots, n$ be a $\mathbb{Z}$-basis for I. The equality $I = IJ$ means that each α_i can be written as

$$\alpha_i = \sum_j b_{ij}\alpha_j; \sum_j (b_{ij} - \delta_{ij})\alpha_j = 0$$

with $b_{ij} \in J$. Thus the matrix $(b_{ij} - \delta_{ij})$ has a non-zero kernel, so $\det(b_{ij} - \delta_{ij}) = 0$, which expanded gives $1 \in J$.

Finally, we have the following relation: if $\alpha \in K^\times$, J, I two ideals, then

$$(\alpha)I = JI \;\Rightarrow\; J = (\alpha) \tag{2.8}$$

Indeed, we already see that $(\alpha)I = JI \subset I$ so the first observation implies that $\alpha \in \mathcal{O}$. Now for every $\beta \in J$ we have $\beta I \subset JI = (\alpha)I$, which means $\beta\alpha^{-1}I \subset I$, so by the above $\beta\alpha^{-1} \in \mathcal{O}$ and $\beta \in (\alpha)$. Thus $J \subset (\alpha)$; on the other hand $J' = \alpha^{-1}J \subset \mathcal{O}$ is an ideal such that $J'I = I$, which means $J' = \mathcal{O}$ by the above. Thus $J = (\alpha)$.

So far, everything is valid for any integrally closed domain. Now we use a shorter proof, valid only for $R = \mathcal{O}$, based on the following important theorem, that will be stated without (?) proof (a proof is given later in the notes):

Theorem 2.9. *Define an equivalence relation on ideals (or fractional ideals) of $\mathcal{O}$ by saying $I \sim J$ if there exists $\alpha \in K^\times$ such that $(\alpha)I = J$. (This is obviously an equivalence relation.) The set of equivalence classes for this relation is finite.*

This implies that, for any non-zero I, there exist $m < n$ such that $I^m \sim I^n$; hence there are $\alpha, \beta \in \mathcal{O}$ such that $\alpha I^m = \beta I^n$. Thus $\alpha/\beta I^m = I^n \subset I^m$. By the first observation, this implies $\gamma = \alpha/\beta \in \mathcal{O}$. So $(\gamma)I^m = I^{n-m}I^m$ and by 2.8 this implies $I^{n-m} = (\gamma)$.

Thus for every $I \neq (0)$ there is $h = h(I) \in \mathbb{N}$ such that I^h is principal. This implies if $IJ = IJ'$ then $J = J'$ – indeed, just multiply both sides by I^{h-1} to get $(\gamma)J = (\gamma)J'$ which implies $J = J'$. Similarly

$$I \subset J \Rightarrow \exists J', I = JJ' \tag{2.10}$$

Indeed, $I \subset J \Rightarrow J^{h-1}I \subset J^h = (\gamma)$, thus $J' = \gamma^{-1}J^{h-1}I \subset \mathcal{O}$ and $J'J = \gamma^{-1}J^hI = I$.

We first prove existence of the prime factorization. Let $I \subset \mathcal{O}$, $i \neq \mathcal{O}$. Then there exists a maximal ideal $\mathfrak{p}_1 \supset I$. By 2.10, $I = \mathfrak{p}_1 \cdot I_1$. If $I_1 \neq \mathcal{O}$, we then have $I = \mathfrak{p}_1\mathfrak{p}_2I_2$. Now $I \subset I_1 \subset I_2 \cdots \subset I_n \ldots$ and since $\mathcal{O}$ is Noetherian, this chain has to stabilize, say $I_n = I_{n+1} = \ldots$. But if $I_n \neq \mathcal{O}$ we can continue the process, so $I_n = \mathcal{O}$, $I = \prod_{j=1}^n \mathfrak{p}_j$.

To prove uniqueness, suppose $\mathfrak{p}$ is prime and $\mathfrak{p}^m = \mathfrak{p}^{m+1}$ for some m; then $\mathfrak{p} \cdot \mathfrak{p}^m = \mathfrak{p}^m$ which implies by the above that $\mathfrak{p} = \mathcal{O}$. Thus the $\mathfrak{p}^i$ are all distinct, and for any I we can define

$$ord_{\mathfrak{p}}(I) = sup_m I \subset \mathfrak{p}^m.$$

Then $I \subset I' = \cap_{\mathfrak{p}}\mathfrak{p}^{ord_{\mathfrak{p}}(I)}$ – which implies that the set of $\mathfrak{p}$ dividing I is finite (because the intersection of infinitely many distinct primes is (0)) so $I = I'J$ for some J, but since ord is a maximum it follows that

$ord_{\mathfrak{p}}(J) = 0$ for all $\mathfrak{p}$, so $I = I'$. Finally it follows from the Chinese Remainder Theorem that

$$\cap_{\mathfrak{p}} \mathfrak{p}^{ord_{\mathfrak{p}}(I)} = \prod_{\mathfrak{p}} \mathfrak{p}^{ord_{\mathfrak{p}}(I)}$$

and we are done. □

Some commutative algebra. We used the following basic results in commutative algebra:

Proposition 2.11. *Let $\mathcal{O}$ be the ring of integers of a number field, $\{\mathfrak{p}_i, i \in \mathbb{N}\}$ a sequence of two-by-two distinct prime ideals. Then $\cap_i \mathfrak{p}_i = \{0\}$.*

(This holds more generally for any Dedekind ring.)

Proof. Let $I = \cap_i \mathfrak{p}_i$. Suppose $[\mathcal{O} : I] < \infty$. The set of ideals between $\mathcal{O}$ and I is then finite. □

Proposition 2.12. *(Chinese Remainder Theorem) Let I_i, ..., I_r be two-by-two relatively prime ideals in a commutative ring R. Then the natural map*

$$R/\prod_j I_j \to \prod_j R/I_j$$

is an isomorphism.

Proof. The kernel of the map is $\cap_j I_j$, so we also need to show that

$$\cap_j I_j = \prod_j I_j.$$

By induction we can assume $r = 2$. Indeed, suppose I, I', I'' are two-by-two relatively prime; we need to show that $I + I' \cdot I" = R$. We know that $I + I' = R$, so there exist $a' + b' = 1$ with $a' \in I$, $b' \in I'$; likewise $a'' + b'' = 1$ with $a'' \in I, b'' \in I''$; then

$$1 = (a' + b')(a'' + b'') = (a'a'' + b'a'' + b''a') + b'b'' \in I + I'I''.$$

So suppose we have two relatively prime ideals I, J. It is obvious that $I \cap J \supset IJ$. For the reverse inclusion we let $1 = a + b$ with $a \in I$, $b \in J$; if now $x \in I \cap J$ then $x = x(a + b) = xa + xb$ is the sum of two elements of IJ.

Finally, the surjectivity is proved the same way: if $(x, y) \in R \times R$ and $1 = a + b$ as above, then $b \equiv 1 \pmod{I}$, $a \equiv q \pmod{J}$, so $z = xb + ya \equiv x \pmod{I}$, $z \equiv y \pmod{J}$. □

Proposition 2.13. *Let R be a Dedekind ring, $\mathfrak{p}$ and $\mathfrak{q}$ distinct prime ideals, $r, s > 1$, then $\mathfrak{p}^r$ and $\mathfrak{q}^s$ are relatively prime.*

Proof. If not, there is a maximal ideal, say $\mathfrak{m} \supset \mathfrak{p}^r + \mathfrak{q}^s$. Either $\mathfrak{m}$ is one of $\mathfrak{p}$ or $\mathfrak{q}$, or it's prime to both of them; in either case $\mathfrak{m}$ is prime to one of the two, say $\mathfrak{p}$, and it suffices to prove that $\mathfrak{m}+\mathfrak{p}^r = \mathcal{O}$ to obtain a contradiction. We know that $\mathfrak{p}$ and $\mathfrak{m}$ are maximal so as above, we can find $a \in \mathfrak{p}$, $b \in \mathfrak{m}$ with $a+b=1$. Then $1 = (a+b)^r = a^r + B$ with $a^r \in \mathfrak{p}^r$ and $B \in \mathfrak{m}$. □

Return to quadratic reciprocity. So to conclude, if p and q are odd, $K = \mathbb{Q}(\sqrt{q})$, $\left(\frac{q}{p}\right) = -1 \Rightarrow p\mathcal{O}$ is prime in $\mathcal{O}$; $\left(\frac{q}{p}\right) = 1 \Rightarrow p\mathcal{O} = \mathfrak{p}_1\mathfrak{p}_2$ in $\mathcal{O}$. Write $k(\mathfrak{p})$ for $\mathcal{O}/\mathfrak{p}$ for any prime (maximal) ideal of $\mathcal{O}$. If $\left(\frac{q}{p}\right) = -1$ then $Gal(k(\mathfrak{p})/\mathbb{F}_p) = \pm 1$; if $\left(\frac{q}{p}\right) = 1$ then $Gal(k(\mathfrak{p}_i)/\mathbb{F}_p) = 1$, $i = 1, 2$. The Legendre symbol is the generator of the Galois group; the next time I make this clear.

We return to norms and traces. Suppose $K = \mathbb{Q}(\alpha)$, and let $f \in \mathbb{Q}[X]$ be the minimal polynomial of α, $n = \deg f = [K : \mathbb{Q}]$. Then it is easy to write the matrix of multiplication by α in the basis $(1, \alpha, \ldots, \alpha^{n-1})$ and we find that the characteristic polynomial of this matrix is just f. Thus if we write

$$f = \prod_{i=1}^{n} (X - \alpha_i)$$

we find that

$$N_{K/\mathbb{Q}}(\alpha) = \prod_i \alpha_i;\ Tr_{K/\mathbb{Q}}(\alpha) = \sum_i \alpha_i.$$

More generally, if $[K : \mathbb{Q}(\alpha)] = d$, then

$$N_{K/\mathbb{Q}}(\alpha) = N_{K/\mathbb{Q}(\alpha)}(N_{\mathbb{Q}(\alpha)/\mathbb{Q}}(\alpha)) = (\prod_i \alpha_i)^d$$

and similarly $Tr_{K/\mathbb{Q}}(\alpha) = d(\sum_i \alpha_i)$. Note that if $\alpha \in \mathcal{O}_K$ then each of the α_i is an algebraic integer, and so its norm and trace are in $\mathbb{Z}$.

If $I \subset \mathcal{O}_K$ is an ideal, then $\mathcal{O}/I$ is a positive integer, denoted $N(I)$.

Proposition 2.14. *Let $\alpha \in \mathcal{O}_K$. Then $N((\alpha)) = |N_{K/\mathbb{Q}}(\alpha)|$. Moreover for any I, J, $N(IJ) = N(I)N(J)$.*

Proof. We know that $N((\alpha)) = [\mathcal{O} : \alpha\mathcal{O}] = [\mathcal{O} : A(\mathcal{O})]$ where A is the linear transformation given by multiplication by α. But for any invertible linear transformation A of a vector space V and any lattice $L \subset V$ fixed by A, we know $[L : A(L)] = |\det(A)|$. The first statement follows.

Now we can write $I = \prod_i \mathfrak{p}_i^{a_i}$, $J = \prod_j \mathfrak{p}_j^{b_j}$ and so it suffices to show that $N(I) = \prod_i N(\mathfrak{p}_i)^{a_i}$. In other words, it suffices to show that the map

$$\mathcal{O}/I \to \prod_i \mathcal{O}/\mathfrak{p}_i^{a_i}$$

is an isomorphism, and that $[\mathcal{O} : \mathfrak{p}^a] = [\mathcal{O} : \mathfrak{p}]^a$ for any prime ideal $\mathfrak{p}$ But the first statement is the Chinese Remainder Theorem, and the second follows easily from localization. Alternatively, it can easily be proved by induction, once one observes that $\mathfrak{p}/\mathfrak{p}^a$ is a principal ideal (generated by any element in $\mathfrak{p}$ not in $\mathfrak{p}^2$).

□

Proposition 2.15. *Let p be a prime number, $[K : \mathbb{Q}] = n$, $p\mathcal{O}_K = \prod_{i=1}^g \mathfrak{p}_i^{e_i}$. Let $f_i = [k(\mathfrak{p}_i) : \mathbb{F}_p]$. Then*

$$\sum_{i=1}^{g} e_i f_i = n.$$

Proof. We know that $N(p\mathcal{O}_K) = N_{K/\mathbb{Q}}(p) = p^n$. On the other hand $N(p\mathcal{O}_K) = \prod_{i=1}^g N(\mathfrak{p}_i)^{e_i}$, and $N(\mathfrak{p}_i) = |k(\mathfrak{p}_i)| = p^{f_i}$. The formula follows easily. □

Say p is *ramified* in K if $e_i > 1$ for some i.

Theorem 2.16. *The set of ramified primes in any finite extension $K/\mathbb{Q}$ is non-empty and finite.*

Suppose henceforward that $K/\mathbb{Q}$ is Galois. This simplifies some of the proofs as well as statements.

Lemma 2.17. *The Galois group $Gal(K/\mathbb{Q})$ preserves $\mathcal{O}_K$ and $\mathcal{O}_K^\times$.*

Proof. If $\alpha \in K$ is the root of a polynomial $P \in \mathbb{Q}[X]$, then $s(\alpha)$ is the root of the same polynomial for any $s \in Gal(K/\mathbb{Q})$. So if α is an integer, we take P monic in $\mathbb{Z}[X]$, which implies $s(\alpha)$ is an integer as well. Then the equation $\alpha\beta = 1$ with $\alpha\beta$ both integers is preserved by $Gal(K/\mathbb{Q})$ as well. □

If $\alpha \in K$, then all the roots of its minimal polynomial are in K, and then

$$N_{K/\mathbb{Q}}(\alpha) = \prod_{\sigma \in Gal(K/\mathbb{Q})} \sigma(\alpha), Tr_{K/\mathbb{Q}} = \sum_{\sigma \in Gal(K/\mathbb{Q})} \sigma(\alpha). \tag{2.18}$$

Lemma 2.19. *Let p be a prime number, $p\mathcal{O} = \prod_{i=1}^g \mathfrak{p}_i^{e_i}$. The group $G = Gal(K/\mathbb{Q})$ acts transitively on the set of $\mathfrak{p}_i$. In particular, there are e, f such that $e_i = e$, $f_i = f$ for all i. The stabilizer of $\mathfrak{p}$ in G is the decomposition group $D_\mathfrak{p}$, and the order of $D_\mathfrak{p}$ is ef.*

Proof. Let $\mathfrak{p}, \mathfrak{p}'$ divide p and suppose $\mathfrak{p}'$ is not a Galois conjugate of $\mathfrak{p}$. By the Chinese remainder theorem, there is $a \in \mathcal{O}$ such that $a \in \mathfrak{p}$ but $a \equiv 1 \pmod{\sigma(\mathfrak{p}')}$ for all $\sigma \in G$. Then $N_{K/\mathbb{Q}}(a) \in \mathfrak{p} \cap \mathbb{Z} = p$ (this intersection is correct because p is prime); but for all σ, $\sigma(a) \notin \mathfrak{p}'$; thus $N_{K/\mathbb{Q}}(a) = \prod_{\sigma \in G} \sigma(a) \notin \mathfrak{p}'$, because $\mathfrak{p}'$ is prime. But this is a contradiction because $p \subset \mathfrak{p}'$.

Now fix $\mathfrak{p}$. Since $G/D_{\mathfrak{p}}$ is in bijection with the set of $\mathfrak{p}_i$, which has $n/(ef)$ elements, it follows that $|D_{\mathfrak{p}}| = ef$. □

Proposition 2.20. *Suppose there is $\alpha \in \mathcal{O}$ such that $\mathcal{O} = \mathbb{Z}[\alpha]$. Let $f \in \mathbb{Z}[X]$ be the minimal polynomial of α. Then p is ramified if and only if the image $\bar{f}$ of f in $\mathbb{F}_p[X]$ has multiple roots.*

Proof. We know that p is ramified if and only if $\mathcal{O}/p\mathcal{O}$ has nilpotents. But $\mathcal{O}/p\mathcal{O} = \mathbb{Z}[X]/(f, p) = \mathbb{F}_p[X]/(\bar{f})$ which has nilpotents if and only if $\bar{f}$ has multiple roots. □

In particular, if $\mathcal{O} = \mathbb{Z}[\alpha]$ then the set of ramified primes is finite. This is in fact always the case. In fact, using localization, it is easy to see that if α is any integral element that generates K, so that $N = [\mathcal{O} : \mathbb{Z}[\alpha]]$ is finite, then the above criterion remains valid for any p prime to N. Let p be an unramified prime in any case and $\mathfrak{p}|p$ in $\mathcal{O}$. Then $|D_{\mathfrak{p}}| = |Gal(k(\mathfrak{p})/\mathbb{F}_p)|$. On the other hand, since $D_{\mathfrak{p}}$ stabilizes both $\mathcal{O}$ and $\mathfrak{p}$, there is a natural map $r_p : D_{\mathfrak{p}} \to Gal(k(\mathfrak{p})/\mathbb{F}_p)$.

Proposition 2.21. *This map is surjective (hence an isomorphism when p is unramified. In particular, there is a unique element $\phi_p \in D_{\mathfrak{p}} \subset Gal(K/\mathbb{Q})$ with the property that $r_p(\phi_p)(x) = x^p$ for all $p \in k(\mathfrak{p})$.*

(If $K/\mathbb{Q}$ is not abelian, then we have to write $\phi_{\mathfrak{p}}$; in general $\phi_{\sigma(\mathfrak{p})} = \sigma\phi_{\mathfrak{p}}\sigma^{-1}$.)

Corollary 2.22. *Suppose $K = \mathbb{Q}(\sqrt{q})$. Then $\phi_p = \left(\frac{q}{p}\right) \in \{\pm 1\} \simeq Gal(K/\mathbb{Q})$.*

Proof. (of proposition) Let L be the fixed field of $D_{\mathfrak{p}}$, $\mathfrak{p}'$ the prime of L contained in $\mathfrak{p}$. Then $\mathfrak{p}$ is the unique prime above $\mathfrak{p}'$; if there were two, say $\mathfrak{p}_1$ and $\mathfrak{p}_2$, the argument above would show that $Gal(K/L) = D_{\mathfrak{p}}$ takes one to the other, but $D_{\mathfrak{p}}$ stabilizes $\mathfrak{p}$. So $f = [K : L] = [k(\mathfrak{p}) : k(\mathfrak{p}')]$ which implies that $k(\mathfrak{p}') = \mathbb{F}_p$.

Now let $x \in \mathcal{O}_K$, $\bar{x} \in k(\mathfrak{p})$ its reduction, and assume $\bar{x}$ generates $k(\mathfrak{p})$ over $\mathbb{F}_p$. Let f be the minimal polynomial of x, $\bar{f}$ its reduction, $\bar{g}$ the minimal polynomial of $\bar{x}$; thus $\bar{g}|\bar{f}$. Since $D_{\mathfrak{p}}$ acts transitively on the roots of f, it also acts transitively on the roots of $\bar{f}$, hence contains

a subgroup acting transitively on those of $\bar{g}$. This implies that r_p is surjective. □

3. Day 3: Pell's equation and units

- (a) Pell's equation (Flath, Hindry)
- (b) Ideals and units

We fix K, $\mathcal{O} = \mathcal{O}_K$. The group of ideals $I(K)$ contains the subgroup of principal ideals $P(K)$, which is isomorphic to $K^\times/\mathcal{O}^\times$, since $(u\alpha) = (\alpha)$ whenever u is a unit. The ideal class group $Cl(K) = I(K)/P(K)$ is a major invariant of the number field K. We have already used the fact that $Cl(K)$ is finite. This week we will consider in detail the case where $[K : \mathbb{Q}] = 2$, so $K = \mathbb{Q}(\sqrt{d})$, where d is a square-free integer. There are two cases: $d > 0$ (K is a real quadratic field) or $d < 0$ (K is an imaginary quadratic field. The TD will prove finiteness of $Cl(K)$ when K is imaginary quadratic.

I will start by talking about units in the imaginary case. The discussion is short.

Lemma 3.1. *Let K be a number field, $u \in \mathcal{O}_K^\times$. Then $N_{K/\mathbb{Q}} = \pm 1$.*

Proof. Write $N = N_{K/\mathbb{Q}}$. Suppose $uv = 1, v \in \mathcal{O}$. Then $1 = N(uv) = N(u)N(v)$. Both $N(u)$ and $N(v)$ are in $\mathbb{Z}$, and the equation says $N(u), N(v) \in \mathbb{Z}^\times$. So it's clear. □

We take $d > 0$ and consider $K = \mathbb{Q}(\sqrt{-d})$; then $\mathcal{O} = \{a + b\sqrt{-d}\}$ or $\mathcal{O} = \{a + b\eta\}$ with $\eta = \frac{1+\sqrt{-d}}{2}$, depending on $-d \pmod 4$. Now $N(a+b\sqrt{-d}) = a^2 + bd^2$ and $N(a+b\eta) = a^2 + ab + b^2(1+d)/4$, in the latter case $d \geq 3$. If either $|a| > 1$ or $|b| > 1$ then the norm is > 1, so the group of units in $\mathcal{O}_K$ is finite; thus $\mathcal{O}^\times$ consists of roots of unity. If there is a primitive nth root of 1 then there is a pth root of 1 for any $p|n$, but we know that $[\mathbb{Q}(\zeta_p) : \mathbb{Q}] = p - 1$, so $p \leq 3$. One can also check that $[\mathbb{Q}(\zeta_8) : \mathbb{Q}] = 4$. On the other hand if $d = 1$ we get 4th roots of 1, and if $d = 3$ we get 6th roots of 1, since $\frac{\sqrt{-3}+1}{2} = e^{\frac{2\pi i}{6}}$. And so we find $|\mathcal{O}^\times| = 4$ if $d = 1$, $|\mathcal{O}^\times| = 6$ if $d = 3$, $|\mathcal{O}^\times| = 2$ otherwise.

The situation is very different for $K = \mathbb{Q}(\sqrt{d})$ with $d > 0$. There is no difficulty in finding solutions to the equation $x^2 - dy^2 = 1$. For example, if $d = 6$, $(5, 2)$ is a solution; if $d = 3$, $(7, 4)$; if $d = 5$, $(9, 4)$. If $N(u) = \pm 1$ with $u = a + b\sqrt{d}$, $a, b \in \mathbb{Z}$ then $\pm 1 = u \cdot su$ with $su = a - b\sqrt{d}$ which implies that u is a unit. Here $u \mapsto su$ is the non-trivial Galois automorphism of K. Each of these examples is of infinite order in the unit group of $\mathcal{O} = \mathcal{O}_K$; indeed, $(a + b\sqrt{d})^2 = (a^2 + b^2d + 2ab\sqrt{d})$ and $a^2 + b^2d > a$.

Theorem 3.2. *Let d be a positive integer that is not a square. Then there exists a non-trivial solution (a_1, b_1), with $a_1, b_1 \in \mathbb{Z}, a_1 > 0, b_1 > 0$, of Pell's equation $x^2 - dy^2 = 1$, such that all positive solutions are given by (a_n, b_n) with $(a_n + b_n\sqrt{d}) = (a_1 + b_1\sqrt{d})^n$, $n = 1, 2, \ldots$; and every solution is of the form $(\pm a_n, \pm b_n)$ for some n.*

We begin by constructing one non-trivial solution.

Lemma 3.3. *Let $\alpha \in \mathbb{R}$ and $N \geq 1$ an integer. Then there exists a rational number $\frac{p}{q} \in \mathbb{Q}$ such that*

$$|\alpha - \frac{p}{q}| \leq \frac{1}{qN} \text{ and } 1 \leq q \leq N.$$

Proof. Cut the interval $[0, 1]$ into N intervals $[\frac{i}{N}, \frac{i+1}{N}]$, $i = 0, \ldots, N-1$. The set $j\alpha - [j\alpha]$, $j = 0, \ldots, N$ contains $N + 1$ elements, so two of them have to be in the same interval, i.e. there exist $0 \leq k < j \leq N$ such that

$$|j\alpha - [j\alpha] - (k\alpha - [k\alpha])| = |(j-k)\alpha - ([j\alpha] - [k\alpha])| \leq \frac{1}{N}.$$

Let $p = [j\alpha] - [k\alpha]$, $q = j - k$; then we have $|\alpha - \frac{p}{q}| \leq \frac{1}{qN}$ and clearly $1 \leq q \leq N$. □

The Corollary is due to Dirichlet:

Corollary 3.4. *Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Then there exist infinitely many $\frac{p}{q} \in \mathbb{Q}$ such that*

$$|\alpha - \frac{p}{q}| \leq \frac{1}{q^2}.$$

Proof. Let $N_1 \geq 1$ and apply the lemma to find a fraction such that

$$|\alpha - \frac{p_1}{q_1}| \leq \frac{1}{q_1 N_1} \leq \frac{1}{q_1^2}.$$

Since α is irrational, the left-hand side is non-zero, so there is an N_2 such that $1/N^2 < |\alpha - \frac{p_1}{q_1}|$. So there is a second fraction such that

$$|\alpha - \frac{p_2}{q_2}| \leq \frac{1}{q_2 N_2} \leq \frac{1}{q_2^2}.$$

By the choice of N_2, we have

$$|\alpha - \frac{p_2}{q_2}| \leq \frac{1}{q_2 N_2} \leq \frac{1}{N_2} < |\alpha - \frac{p_1}{q_1}|$$

so $\frac{p_2}{q_2} \neq \frac{p_1}{q_1}$. We continue by induction. Note that we can choose a subset with $q_1 < q_2 < \ldots$. □

Proposition 3.5. *Pell's equation has a non-trivial solution (r, s).*

Proof. We apply Dirichlet's corollary to $\alpha = \sqrt{d} \neq \mathbb{Q}$. Then there are infinitely many integers (a, b) such that $|\sqrt{d} - \frac{a}{b}| \leq \frac{1}{b^2}$; so

$$|\sqrt{d} + \frac{a}{b}| \leq 2\sqrt{d} + |\frac{a}{b} - \sqrt{d}| \leq 2\sqrt{d} + 1;$$

thus $|a^2 - db^2| \leq b^2 \cdot \frac{1}{b^2} \cdot 2\sqrt{d} + 1 = 2\sqrt{d} + 1$. The set of integers between $-(2\sqrt{d}+1)$ and $2\sqrt{d}+1$ is finite, but there are infinitely many pairs (a, b) such that $a^2 - db^2$ is an integer in this interval; so there is an integer c such that the equation $a^2 - db^2 = c$ has infinitely many distinct solutions. Indeed, there are even infinitely many congruent solutions modulo c (i.e., $a_i \equiv a_j \pmod{c}$, $b_i \equiv b_j \pmod{c}$). Take two such solutions (a_1, b_1) and (a_2, b_2). and set

$$u = r + s\sqrt{d} = \frac{a_1 + b_1\sqrt{d}}{a_2 + b_2\sqrt{d}}.$$

$$r^2 - ds^2 = N(u) = \frac{a_1^2 - db_1^2}{a_2^2 - db_2^2} = \frac{c}{c} = 1.$$

Note that $s \neq 0$; otherwise $a_1 + b_1\sqrt{d} = \pm(a_2 + b_2\sqrt{d})$ and they would not be distinct. So we will be done if we can show that $(r, s) \in \mathbb{Z} \times \mathbb{Z}$. We compute

$$r + s\sqrt{d} = \frac{a_1a_2 - db_1b_2}{c} + \frac{b_1a_2 - a_1b_2}{c}\sqrt{d}.$$

But $a_1a_2 - db_1b_2 \equiv a_1^2 - db_1^2 \equiv 0 \pmod{c}$ and $b_1a_2 - a_1b_2 \equiv b_1a_1 - a_1b_1 \pmod{c}$.

□

Now we prove the theorem. Define

$$L : \mathcal{O}^\times \to \mathbb{R}^2; L(\alpha) = (\log(|\alpha|), \log(|s(\alpha)|)).$$

Proposition 3.6. *The map L has the following properties.*

(a) *L is a homomorphism.*
(b) $\ker L = \pm 1$.
(c) *The image of L is a discrete subgroup*
(d) *The image of L is contained in the line $x + y = 0$.*

Proof. Property (a) follows from the property of log. Next, (d) follows from

$$\log(|\alpha|) + \log(|s(\alpha)| = \log(|\alpha s(\alpha)| = \log(|N(\alpha)|) = \log(1) = 0.$$

The hard properties are (b) and (c). Let B be a ball around 0 in $\mathbb{R}^2$. We show that $L^{-1}(B)$ is finite. This implies that the kernel is finite, hence consists of roots of 1, necessarily ± 1; it also shows that

the image is discrete, since we can shrink the ball to guarantee that $L^{-1}(B) = \ker L$.

Let B be the ball of radius C. Now an element $\alpha \in \mathcal{O}^\times$ satisfies the polynomial $X^2 - Tr(\alpha)X + N(\alpha) \in \mathbb{Z}[X]$. Suppose $L(\alpha) \in B$. Then $|\alpha| < \exp(C)$, $s(|\alpha)| < \exp(C)$, so $|Tr(\alpha)| < 2\exp(C)$, whereas $N(\alpha) = \pm 1$. Thus there are only finitely many possible polynomials, hence finitely many possible α.

□

Finally

Lemma 3.7. *Any discrete subgroup $G \subset \mathbb{R}$ is of the form $\mathbb{Z}\omega$ for some ω.*

This is easy: if $G \neq 0$, take ω to be the smallest positive element. (This exists because otherwise there would be a limit point, hence G would not be discrete.) If $x \in G$ there is $m \in \mathbb{Z}$ such that $x \in [m\omega, (m+1)\omega[$; but then $0 \leq x - m\omega < \omega \Rightarrow x = m\omega$.

Now to prove the theorem, let ω be a smallest element of $Im(L)$ in the line $x + y = 0$, say $\omega = L(u)$. In fact, we can assume $u = a_1 + b_1\sqrt{d}$ which is the smallest positive element of $\mathbb{Z}[\sqrt{d}] \subset \mathbb{R}$; this is not generally going to have $a_1 > 0$, $b_1 > 0$, but the theorem follows easily by considering the four possible pairs of signs.

In general, let K be any number field, $K_\mathbb{R} = K \otimes_\mathbb{Q} \mathbb{R}$, a real vector space of dimension $n = [K : \mathbb{Q}]$. The trace defines a non-degenerate bilinear form $B(x, y) = Tr(xy) : K_\mathbb{R} \otimes K_\mathbb{R} \to \mathbb{R}$, which implies that the nilpotent radical of $K_\mathbb{R}$ is zero; indeed, if $n \in K_\mathbb{R}$ is nilpotent then so is nx for any x, but then $Tr(nx) = 0$. So $K_\mathbb{R}$ is a finite-dimensional semisimple $\mathbb{R}$-algebra, which means it is of the form $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ for some integers with $r_1 + 2r_2 = n$.

Theorem 3.8. *(Dirichlet's unit theorem) The group $\mathcal{O}_K^\times$ is isomorphic to $\mu_K \times \mathbb{Z}^{r_1+r_2-1}$, where μ_K is the finite group of roots of unity in K.*

The proof uses the logarithm map as before:

$$L : \mathcal{O}^\times \to \mathbb{R}^{r_1+r_2}; L(\alpha) = (\log(|\sigma(\alpha)|); \log(|\tau(\alpha)^2|)_\sigma),$$

where σ (resp. τ) runs through the set of real (resp. complex) embeddings of K, and the argument is that the image is discrete and contained in the subspace $\sum x_i = 0$, and the quotient is compact.

4. Day 4: Binary quadratic forms and ideal classes

- (a) Classification of quadratic forms
- (b) principal ideals and class groups

4.1. *Definite binary quadratic forms.* A binary quadratic form is an expression $Q(X,Y) = aX^2+bXY+cY^2$. We will always assume $a, b, c \in \mathbb{Z}$ and are relatively prime, and $Q(X,Y)$ is then called *primitive*. Gauss's *Disquisitiones Arithmeticae* is largely devoted to the theory of these forms, especially in the definite case where $\Delta = b^2 - 4ac < 0$. Let $K = \mathbb{Q}(\sqrt{\Delta})$, $\mathcal{O} = \mathcal{O}_K$. Assume $K = \mathbb{Q}(\sqrt{-d})$ for some square-free positive d. We can write $\mathcal{O} = \mathbb{Z} \oplus \mathbb{Z}\alpha$ where $\alpha = \sqrt{-d}$ if $-d \equiv 2, 3 \pmod 4$ or $\alpha = \frac{1+\sqrt{-d}}{2}$ if $-d \equiv 1 \pmod 4$. The norm $N_{K/\mathbb{Q}}$ defines a primitive binary quadratic form by

$$N(X + \alpha Y) = X^2 + Tr_{K/\mathbb{Q}}(\alpha)XY + N_{K/\mathbb{Q}}(\alpha)Y^2 = X^2 + dY^2$$

$$\text{or } X^2 + XY + \frac{1+d}{4}Y^2.$$

The discriminants are respectively $-4d$ and $-d$. More generally, if $I \subset \mathcal{O}$ is an ideal, it can be written as $\mathbb{Z}\beta \oplus \mathbb{Z}\gamma$ and defines the binary quadratic form $Q_I = N(\beta X + \gamma Y)$. Obviously this depends on the choice of basis; but we can define an equivalence relation:

Definition 4.1. We write $Q \sim Q'$, and say Q is *equivalent* to Q', if there exists $g = \begin{pmatrix} r & s \\ t & u \end{pmatrix} \in GL(2,\mathbb{Z})$ such that $Q'(X,Y) = Q(rX + sY, tX + uY)$. Say Q is *strongly equivalent* to Q' if we can take $g \in SL(2,\mathbb{Z})$, i.e. $\det g = 1$.

Since $GL(2,\mathbb{Z})$ is a group, this is obviously an equivalence relation. If $Q \sim Q'$ then the discriminants of Q and Q' are equal. Indeed, let M be any symmetric 2×2 matrix, and define $Q_M = \begin{pmatrix} X & Y \end{pmatrix} M \begin{pmatrix} X \\ Y \end{pmatrix}$. If $M = \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}$ then $Q_M = aX^2+bXY+cY^2$ has discriminant $b^2-4ac = -4\det(M)$. But if $Q'(X,Y) = Q(rX + sY, tX + uY)$ then

$$Q' = \begin{pmatrix} X & Y \end{pmatrix} \cdot {}^t g \cdot M \cdot g \cdot \begin{pmatrix} X \\ Y \end{pmatrix} = Q_{[g]M},$$

where $[g]M :=^t gMg$. So the discriminant of Q' is $-4\det([g]M) = -4\det M$.

Note that Q_I is not generally primitive. The correct normalization will be discussed in the TD. We start with some elementary remarks about quadratic forms with discriminant $-d < 0$. Note that

$$4aQ(X,Y) = (2aX + bY)^2 + dY^2$$

so since $d > 0$, $4aQ > 0$ unless $Y = 0$ and $2aX + bY = 0$ (so $X = 0$). So we assume $a > 0$ (and so $c > 0$).

Question 4.2. What integers m can be represented by the (primitive, positive-definite) quadratic form Q? That is, for what m does the equation $Q(X,Y) = m$ have a solution?

Obviously m has to be non-negative.

Lemma 4.3. *Suppose Q and Q' are equivalent. Then they represent the same integers.*

Proof. This is obvious. If $Q'(X,Y) = m$, then $Q(rX+sY, tX+uY) = m$, and the reverse is true if we replace the matrix g above by its inverse. □

Obviously, if $\alpha, \gamma \in \mathbb{Z}$ are both divisible by p, then $Q(\alpha,\gamma)$ is divisible by p^2. Say that m is *properly represented* by Q if there exist integers α, γ with no common factors. Obviously, a and c are properly represented (by $(1,0)$ and $(0,1)$ respectively).

Suppose m is properly represented, $m = a\alpha^2 + b\alpha\gamma + c\gamma^2$. Since $GCD(\alpha,\gamma) = 1$, there exist β, δ such that $\alpha\delta - \beta\gamma = 1$; i.e. the matrix $g = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SL(2,\mathbb{Z})$. Let $X' = \alpha X + \beta Y, Y' = \gamma X + \delta Y$. Then writing $Q = Q_M$ as above, Q is strongly equivalent to $Q' = Q_{[g]M}$ and they have the same discriminant Δ. But explicitly,

$$[g]M = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} a' & \frac{b'}{2} \\ \frac{b'}{2} & c' \end{pmatrix}$$

where $a' = \alpha^2 a + \alpha\gamma b + \gamma^2 c = Q(\alpha,\gamma) = m$. This incidentally gives another proof that Q' represents m, which we started by assuming was represented by Q.

But

$$\Delta = -4\det([g]M) = (b')^2 - 4a'c' = (b')^2 - 4mc' \equiv (b')^2 \pmod{4m}.$$

Thus if m is represented by a quadratic form of discriminant Δ, then Δ is a square modulo $4m$. Conversely, if Δ is a square modulo $4m$, say $\Delta = b^2 - 4mc$, for some b and c, then the form $mX^2 + bXY + cY^2$ properly represents m and has discriminant Δ. Thus

Proposition 4.4. *An integer m is properly represented by some form of discriminant Δ if and only if Δ is a square modulo $4m$.*

Example 4.5. *Take $\Delta = -4$. So a prime p is represented by a form of discriminant -4 if and only if -4 is a square modulo $4p$ if and only if -1 is a square modulo p. Indeed if $-4 = b^2 + 4pc$ then b is even, say $b = 2b'$; thus -1 is a square modulo p.*

We work this out. If $p = 2$ then $-4 = 4 \pmod 8$ is a square. Of course -4 is a square modulo 4, so the condition if p is odd is that -1 is a square modulo p. Thus an odd p is represented by a form of discriminant -4 if and only if $p \equiv 1 \pmod 4$.

The form $Q(X, Y) = X^2 + Y^2$ has discriminant -4.

Lemma 4.6. *Every form of discriminant -4 is strongly equivalent to Q.*

Admitting this Lemma, we obtain Gauss's theorem:

Theorem 4.7. *An odd prime p can be written as the sum of two squares if and only if $p \equiv 1 \pmod 4$.*

The lemma is proved as a consequence of Gauss's reduction theorem:

Theorem 4.8. *Each strong equivalence class of definite binary quadratic forms has a unique representative that is reduced in the sense that*

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

The proof will be given later.

Corollary 4.9. *The set of strong equivalence classes of (positive) definite binary quadratic forms with (negative) discriminant Δ is finite.*

Proof. Since $-\Delta = 4ac - b^2 \geq 4a^2 - b^2 \geq 4a^2 - a^2 = 3a^2$ so $a \leq \sqrt{|\Delta|/3}$. The set of such a is finite, and since $|b| \leq a$ the set of possible b's is finite, and for each pair (a, b) $c = (b^2 - \Delta)/4a$ is determined. So the corollary is obvious. □

In particular, if $\Delta = -4$, then $a = 1$ is the only possibility, and therefore $b = 0$ (the case $b = 1$ is inconsistent with $4|b$), which proves the Lemma. It remains to prove the reduction theorem. This is a simple algorithm in linear algebra. Suppose Q is not reduced. First suppose $c < a$ or $c = a$ but $b < 0$. Write $X' = -Y$, $Y' = X$ (the matrix $g = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in SL(2, \mathbb{Z})$). Then $Q \sim Q' = a(-Y)^2 + bX(-Y) + cX^2 = a'X^2 + b'XY + c'Y^2$ with $a' = c, b' = -b, c' = a$. So we have eliminated the condition.

Now suppose $a \leq c$ but b doesn't satisfy $-a < b \leq a$. There is a unique $b' \equiv b \pmod{2a}$ in the range $]-a, a]$, say $b' = b + 2ak$, $c' = Q(k, 1)$, and set $X' = X + kY$, $Y' = Y$ (the matrix $g = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \in SL(2, \mathbb{Z})$). Then $Q'(X, Y) = Q(X', Y') = a(X + kY)^2 + b(X + kY)Y + cY^2 = a'X^2 + b'XY + c'Y^2$ with b', c' as above and $a' = a$. Each time

the algorithm is applied, we get a new triple with $|b'| + a' < |b| + a$, as one checks case by case. If now $c' > a' = a$ then we leave it alone; otherwise we switch them as before and make the new a at most the same size as the old one. So the algorithm can only be applied finitely many times, which means after a finite time the form is reduced.

Suppose $\Delta = -3$. Again we have $a = 1$, but now $b = 1 \pmod 4$, so $b = 1$, and then $c = 1$. The only form of discriminant Δ is the norm form for $\mathcal{O}_K$ where $K = \mathbb{Q}(\sqrt{-3})$. The proposition says that a prime p is a norm of an element of $\mathcal{O}_K$ if and only if -3 is a square mod $4p$. If $p = 2$ the proposition says that 2 is not represented by $X^2 + XY + Y^2$. On the other hand, -3 is a square mod 4, so (by the Chinese Remainder Theorem) if p is odd, the assertion is that p is a norm from $\mathcal{O}_K$ if and only if -3 is a square mod p; i.e. if and only if $\mathbb{F}_p[X]/(X^2+3) = (\mathbb{F}_p)^2$. But if p is odd, $\mathcal{O}_K/p = \mathbb{Z}[\sqrt{-3}]/p\mathbb{Z}[\sqrt{-3}] = \mathbb{F}_p[X]/(X^2+3)$. So the condition is that p is a norm from $\mathcal{O}_K$ if and only if $(p\mathcal{O}_K)$ is a product of two prime ideals $\mathfrak{p}_1$ and $\mathfrak{p}_2$. Say $p = N(x)$; then $xs(x) \in (p) = \mathfrak{p}_1\mathfrak{p}_2 \subset \mathfrak{p}_1$. Thus we can assume $x \in \mathfrak{p}_1$; but $N(x) = N(\mathfrak{p}_1)$ which implies that $(x) = \mathfrak{p}_1$. And indeed, $|Cl(K)| = 1$, in other words $\mathcal{O}_K$ is a principal ideal domain.

This is explained as follows. Let d be a positive square-free integer. If $-d \equiv 1 \pmod 4$, let $\Delta_d = -d$; otherwise, let $\Delta_d = -4d$. The norm form on $\mathcal{O}_K$, where $K = \mathbb{Q}(\sqrt{-d})$, has discriminant Δ_d.

Proposition 4.10. *There is an injection from the ideal class group $Cl(K)$ to the set of strong equivalence classes of binary quadratic forms of discriminant Δ_d. In particular, $Cl(K)$ is finite.*

A part of the proof of this proposition is contained in the homework.

5. Day 5: Cyclotomic fields

- (a) Fermat's last theorem
- (b) Integers in cyclotomic fields
- (c) Cyclotomic reciprocity
- (d) Cyclotomic units

Fermat's Last Theorem was written in the margin of his copy of Diophantus in 1637 and was proved by Andrew Wiles in 1995.

Theorem 5.1. *(Wiles) Let n be an integer greater than 2. Let $x, y, z \in \mathbb{N}$ satisfy*

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

Then $xyz = 0$.

The proof involves every technique presented in this course and many more besides. The methods introduced by Wiles and Taylor have hastened the solution of a number of other problems that appeared intractable. I will use the theorem as an excuse to talk about the algebraic number theory of cyclotomic fields, which brings me to the middle of the 19th century. First, the case proved by Fermat.

Theorem 5.2. *Fermat's last theorem is true for $n = 4$.*

Proof. This is an illustration of the method of infinite descent. It's enough to show that the equation $x^4 + y^4 = z^2$ has no positive integer solutions. So assume (x, y, z) is a solution, with $z > 0$ as small as possible. We are going to construct a new triple with smaller z.

If any pair (x, y), (y, z) or (x, z) has a common divisor d, then d divides the third number; so (by minimality of z) any of the two are relatively prime. Let $a = x^2$, $b = y^2$, so $a^2 + b^2 = z^2$, a relatively prime Pythagorean triple. Either a or b is even, but not both – say b is even. Then it is easy to show (exercise!) that there exist $u > v$ relatively prime integers such that $u^2 - v^2 = a$, $2uv = b$, $u^2 + v^2 = z$. Since a is an odd square, $a \equiv 1 \pmod 4$, so u is odd and v is even, say $v = 2w$, $y^2 = b = 4uw$. But $(u, w) = 1$, so both u and w are squares, say $u = \alpha^2$, $w = \beta^2$. On the other hand, $u^2 - v^2 = a = x^2$ implies that

$$x^2 + v^2 = u^2$$

and we have another relatively prime Pythagorean triple. Hence there are relatively prime integers $e > f$ with

$$e^2 - f^2 = x,\ 2ef = v = 2w = 2\beta^2, e^2 + f^2 = u = \alpha^2.$$

Again, $ef = \beta^2$ implies $e = g^2$, $f = h^2$. And thus we have

$$e^2 + f^2 = u = \alpha^2 \Rightarrow g^4 + h^4 = \alpha^2.$$

But then

$$\alpha \leq \alpha^2 = u \leq u^2 = z - v^2 < z.$$

Thus $\alpha < z$ and we have a new solution with smaller z.

□

Now if Fermat's last theorem is true for m, it is true for any number divisible by m. It follows that in order to prove Fermat's last theorem, it's enough to prove it for all odd primes p. So suppose henceforward $n = p$ is prime. Let $K = K_p$ be the splitting field of $P_p(X) = X^p - 1$, and let $\zeta \neq 1$ be a root of P_p. We know that $P_p = (X - 1)\Phi_p$, with $\Phi_p(X) = (X^{p-1} + X^{p-2} + \cdots + X + 1)$, and that Φ_p is irreducible. Then $K = \mathbb{Q}[\zeta]$, and $Gal(K/\mathbb{Q}) \xrightarrow{\sim} (\mathbb{Z}/p\mathbb{Z})^\times$. Let $\mathcal{O} = \mathcal{O}_K$. Obviously

$\mathbb{Z}[\zeta] \subset \mathcal{O}$, and we will show that the two are equal. Over $\mathcal{O}$, the Fermat equation becomes

$$\prod_{i=0}^{p-1}(X + \zeta^i Y) = Z^p$$

since $X^p + 1 = \prod_{i=0}^{p-1}(x + \zeta^i)$. We return to this equation tomorrow, after developing the basic theory of cyclotomic fields.

First, a fact promised the first day.

Lemma 5.3. *Let $n \in \mathbb{N}$, p a prime not dividing n, k a field of characteristic p. The cyclotomic polynomial $P_n = X^n - 1 \in k[X]$ has no multiple roots. In particular, if $q = p^r$, $n = q - 1$, the $q - 1$ elements of $\mathbb{F}_q$ are the distinct roots of P_{q-1}, and therefore $\mathbb{F}_q^\times$ is a cyclic group.*

The proof is obvious: the derivative $P_n' = nX^{n-1}$ and has no common roots with P_n unless $n = 0$ in k. For the final assertion, $\mathbb{F}_q^\times$ is a group of order $q-1$, hence every element is of order dividing $q-1$ by Lagrange's theorem; but since it is the multiplicative group of a field, this means every element is a root of P_{q-1}. It follows that any primitive root of P_{q-1} is a cyclic generator of $\mathbb{F}_q^\times$.

Define $\lambda = 1 - \zeta$.

Proposition 5.4. *The ring $\mathbb{Z}[\zeta]/(\lambda)\mathbb{Z}[\zeta]$ is isomorphic to $\mathbb{F}_p$. For any $1 \leq i \leq p-1$ the element $\eta_i = \frac{1-\zeta^i}{\lambda}$ is a unit in $\mathbb{Z}[\zeta]$. We have*

$$p = \prod_{i=1}^{p-1}(1 - \zeta^i) = N_{K/\mathbb{Q}}(\lambda)$$

and $(p) = (\lambda)^{p-1}$.

Proof. Of course $\Phi_p = \prod_{i=1}^{p-1}(X - \zeta^i)$, so

$$p = \Phi_p(1) = \prod_{i=1}^{p-1}(1 - \zeta^i) = N_{K/\mathbb{Q}}(\lambda).$$

Now $\zeta \equiv 1 \pmod{\lambda}$, so every element of $\mathbb{Z}[\zeta]$ is congruent to an element of $\mathbb{F}_p$ modulo λ; i.e., the inclusion $\mathbb{F}_p \hookrightarrow \mathbb{Z}[\zeta]/(\lambda)\mathbb{Z}[\zeta]$ is surjective.

As for the η_i, we can write $\eta_i = \sum_{j=0}^{i-1} \zeta^j \in \mathbb{Z}[\zeta]$. But since the group of p-th roots of unity is cyclic of order p, any element other than 1 is a generator. So $\zeta = (\zeta^i)^b$ for some b. Thus $\eta_i^{-1} = \frac{1-(\zeta^i)^b}{1-\zeta^i} = \sum_{j=0}^{b-1} \zeta^{ij} \in \mathbb{Z}[\zeta]$. It follows that $\eta_i \in \mathbb{Z}[\zeta]^\times$ for any i, and the rest follows.

□

It follows that $(\lambda) \cap \mathbb{Z} \supset (p)$, and since (p) is a maximal ideal and $(\lambda) \neq (1)$, we even have $(\lambda) \cap \mathbb{Z} = (p)$.

Proposition 5.5. *$\mathbb{Z}[\zeta]$ is the ring of integers in K. In particular, p is the only ramified prime in K_p.*

Proof. Let $\alpha = \sum_{i=0}^{p-2} a_i \zeta^i \in \mathcal{O}_K$, with $a_i \in \mathbb{Q}$; we show each $a_i \in \mathbb{Z}$. Note that

$$Tr_{K/\mathbb{Q}}\zeta^i = -1, i = 1, \ldots, p-1;\ Tr_{K/\mathbb{Q}}\zeta^0 = p-1.$$

Thus

$$Tr_{K/\mathbb{Q}}(\lambda\alpha) = \sum_{i=0}^{p-2} a_i Tr_{K/\mathbb{Q}}(\zeta^i - \zeta^{i+1})$$

and the a_i terms cancel for $i > 0$, but the a_0 term gives $(p-1-(-1))a_0 = pa_0$. On the other hand, $Tr_{K/\mathbb{Q}}(\lambda\alpha) \in (\lambda) \cap \mathbb{Z} = (p)$, which implies $a_0 \in \mathbb{Z}$. Thus $\alpha_1 = \zeta^{-1}(\alpha - a_0) \in \mathcal{O}_K$, and by induction we show that all the $a_i \in \mathbb{Z}$.

For the last claim, we know that q is ramified if and only if Φ_p has multiple roots modulo q; but this is only possible when $q = p$. □

The Galois group of $K/\mathbb{Q}$ is isomorphic to ${\mathbb{F}_p}^\times$. The element -1 corresponds to the only element $c \in Gal(K/\mathbb{Q})$ of order 2, and since it takes ζ to $\zeta^{-1} = \bar{\zeta}$, it induces complex conjugation on K. Let K^+ be the fixed field of $\{1, c\}$; every element of K^+ is contained in $\mathbb{R}$.

5.1. **Cyclotomic reciprocity.** We have seen that the problem of quadratic reciprocity, namely the determination of whether or not a number a is a square modulo a prime q, can be translated into a question about the decomposition of q into prime factors in the integer ring of $\mathbb{Q}(\sqrt{a})$. Cyclotomic reciprocity determines the decomposition of a prime q in the cyclotomic field $\mathbb{Q}(\zeta_n)$, where ζ_n is a primitive nth root of 1. We will solve this problem when $n = p$ is a prime, and we write $\zeta = \zeta_p = e^{2\pi i p}$.

The integer ring $\mathcal{O}$ is just $\mathbb{Z}[\zeta]$, so we need to study

$$\mathbb{Z}[\zeta]/(q\mathbb{Z}[\zeta]) = \mathbb{Z}[X]/(q, \Phi_p) = \mathbb{F}_q[X]/(\Phi_p).$$

We assume $q \neq p$. Now the polynomial $\Phi_p(X-1) = X^p - 1 = P_p$ has derivative pX^{p-1} in $\mathbb{F}_q[X]$. The only root of P_p' is 0, which is not a root of P_p. It follows that Φ_p has no multiple roots, and therefore the ring $\mathbb{Z}[\zeta]/(q\mathbb{Z}[\zeta])$ has no nilpotents. So if $q = \prod_{i=1}^{g} \mathfrak{q}_i^e$ we see that $e = 1$; no prime other than p is ramified in $\mathbb{Q}(\zeta)$. Thus we have $fg = p - 1 = [\mathbb{Q}(\zeta) : \mathbb{Q}]$, where $f = [k(\mathfrak{q}_i) : \mathbb{F}_q]$ for any i. The problem of cyclotomic reciprocity is thus to determine f, in other words the order of $Gal(k(\mathfrak{q}_i)/\mathbb{F}_q)$.

Let $\phi_q \in Aut(k(\mathfrak{q}_i))$ be the *Frobenius substitution* defined by

$$\phi_q(x) = x^q.$$

This is an automorphism and its set of fixed elements is the set of roots of the polynomial $A(x) = x^q - x$. Again, $A' = 1$ and so A has no multiple roots, so the fixed field is just $\mathbb{F}_q$. One learns in Galois theory that ϕ_q generates $Gal(k(\mathfrak{q}_i)/\mathbb{F}_q)$ in particular that ϕ_q is of order f in the Galois group. On Day 2 we learned that ϕ_q lifts to a unique generator, also called ϕ_q, of the decomposition group $D_q \subset Gal(\mathbb{Q}(\zeta)/\mathbb{Q})$.

Now for all $a \in \mathbb{Z}[\zeta]$ we have

$$\phi_q(a) \equiv a^q \pmod{\mathfrak{q}_i}, i = 1, \dots, g. \tag{5.6}$$

In particular,

$$\phi_q(\zeta) \equiv \zeta^q \pmod{\mathfrak{q}_i}, i = 1, \dots, g. \tag{5.7}$$

On the other hand, in terms of the canonical isomorphism

$$\omega : Gal(\mathbb{Q}(\zeta)/\mathbb{Q}) \xrightarrow{\sim} {\mathbb{F}_p}^\times$$

we have $\phi_q(\zeta) = \zeta^{\omega(\phi_q)}$, so

$$\zeta^{\omega(\phi_q)} \equiv \zeta^q \pmod{\mathfrak{q}_i}, i = 1, \dots, g. \tag{5.8}$$

But since the roots of unity remain distinct modulo $\mathfrak{q}_i$, this means that $q = \omega(\phi_q)$.

We have thus proved the following theorem when n is prime:

Theorem 5.9. *(Cyclotomic reciprocity) Let $K = \mathbb{Q}(\zeta_n)$, q a prime not dividing n. Then q is unramified in K and factors as the product $q = \prod_{i=1}^{g} \mathfrak{q}_i$, with $f = [k(\mathfrak{q}_i) : \mathbb{F}_q]$ for any i and $fg = [K : \mathbb{Q}] = \varphi(n) = |(\mathbb{Z}/n\mathbb{Z})^\times|$. The Frobenius element $\phi_q \in Gal(K/\mathbb{Q})$ is then identified with the image of q modulo n, via the canonical isomorphism*

$$Gal(K/\mathbb{Q}) \xrightarrow{\sim} (\mathbb{Z}/n\mathbb{Z})^\times.$$

The integer f is then the smallest positive integer such that $q^f \equiv 1 \pmod{n}$.

The proof for general n is not much harder than the proof for n prime. The chief difficulty is showing that $\mathbb{Z}[\zeta_n]$ is the full ring of integers in $\mathbb{Q}(\zeta_n)$. It is certainly not difficult but it takes too long for a six-week course.

5.2. **Cyclotomic units.**

Lemma 5.10. *Fix integers m, M. The set $C(m, M)$ of algebraic integers α of degree $\leq m$, all of whose Galois conjugates in $\mathbb{C}$ have absolute value $< M$, is finite. In particular, if α is an algebraic integer all of whose Galois conjugates have absolute value 1, then α is a root of unity.*

Proof. The set $C(m, M)$ is the set of roots of monic polynomials over $\mathbb{Z}$ whose coefficents are bounded; this is a finite set. That gives the first statement. As for the second, if α has the indicated property, then so do all powers α^m. Thus the powers of α belong to a finite set, which means α is a root of unity. □

Lemma 5.11. *The group $\mathcal{O}^\times$ is generated by roots of unity and $[\mathcal{O}^\times]^+$, the units in K^+.*

Proof. Consider the map $a : \mathcal{O}^\times \to \mathcal{O}^\times$ that takes u to $a(u) = u/c(u)$. Since c commutes with all the Galois automorphisms s of K, we see that $|s(a(u))| = |s(u)|/\overline{s(c(u))}| = 1$. Thus the image of a is contained in the group μ of roots of unity. Let $\phi : \mathcal{O}^\times \to \mu/\mu^2$ be the induced map. The kernel contains $[\mathcal{O}^\times]^+$. Suppose $u \in \ker(\phi)$; thus $a(u) = z^2 = a(z)$ for some $z \in \mu$. Thus $\bar{z}u \in \ker(a) = [\mathcal{O}^\times]^+$, and $u \in \mu \cdot [\mathcal{O}^\times]^+$. Similarly, $\mu \cdot [\mathcal{O}^\times]^+ \subset \ker(\phi)$, which implies that $[\mathcal{O}^\times : \mu \cdot [\mathcal{O}^\times]^+] = 1$ (if ϕ is trivial) or 2 (if ϕ is surjective). So it remains to show that ϕ is not surjective. Suppose the contrary; then there exists a unit u with $c(u) = \zeta' u$ and ζ' is not a square. Now $\mu = \{\pm 1\} \cdot < \zeta >$, $\mu^2 =< \zeta >$. In other words, $\zeta' = -\zeta^k$ for some k. If as before we write $u = \sum_{i=0}^{p-2} a_i \zeta^i$, with $a_i \in \mathbb{Z}$, then $c(u) = \sum_{i=0}^{p-2} a_i \bar{\zeta}^i$. We have

$$u \equiv c(u) \equiv \sum_{i=0}^{p-2} a_i \pmod{(\lambda)}$$

which implies that $-1 \equiv 1 \pmod{(\lambda)}$, and that's impossible because $p \neq 2$. □

6. Day 6: Fermat's last theorem

- (a) Class groups of cyclotomic fields
- (b) Regular primes
- (c) Fermat's last theorem for regular primes (first case)
- (d) Gauss's fourth proof of quadratic reciprocity (Flath)

Definition 6.1. The prime p is regular if p does not divide the class number of the cyclotomic field K_p.

Kummer proved Fermat's Last Theorem for regular primes p. He divided the proof into two cases, the *first case* where p is prime to all of x, y, and z where $x^p + y^p = z^p$, and the *second case* where p divides one of the three. The first case is easier and we prove it here. We always assume x, y, z relatively prime.

First, if $p = 3$, the only cubes modulo 9 are $-1, 0, 1$ and the only sums of two cubes are $-2, 0, 2$, so the equation $X^3 + Y^3 = Z^3$ has no

non-trivial solutions mod 3 with all three factors prime to 3. If $p = 5$, the only 5-th powers mod 25 are $-1, 0, 1, 7, -7$, and the only sums of two 5th powers (not including 0) are $-14, -8, -6, -2, 2, 6, 8, 14$, so again there are no solutions. So assume $p > 5$.

Suppose $x \equiv y \equiv -z \pmod{p}$. Then $(-z)^p + (-z)^p \equiv z^p$, hence $p \mid 3z$, contradiction. So one of the congruences can't hold. If $x \equiv y$ then $x \not\equiv -z$ and by rewriting the equation $x^p + (-z)^p = (-y)^p$ we can assume $x \not\equiv y$; i.e $p \nmid x - y$.

We have seen that the equation becomes

$$\prod_{i=0}^{p-1}(x + \zeta^i y) = z^p. \tag{6.2}$$

Let $\mathfrak{p}$ be the unique prime of $\mathcal{O}$ dividing (p); thus $\mathfrak{p} = (1 - \zeta^i)$ for any $1 \leq i \leq p - 1$.

Lemma 6.3. *The factors on the left hand side of 6.2 are relatively prime in pairs.*

Proof. We have to show there is no prime $\mathfrak{q}$ dividing both $x + \zeta^i y$ and $x + \zeta^j y$ for $i \neq j$. Otherwise, $\mathfrak{q} \mid (\zeta^i - \zeta^j)y = \mathfrak{p}y$ and similarly $\mathfrak{q} \mid \mathfrak{p}x$. Since x and y are relatively prime, $\mathfrak{q} = \mathfrak{p}$. Thus

$$x + y \equiv x + \zeta^i y \equiv 0 \pmod{p}.$$

Thus $x + y \in \mathfrak{p} \cap \mathbb{Z} = (p)$. On the other hand

$$z^p = x^p + y^p \equiv x + y \equiv 0 \pmod{p}$$

which implies $p \mid z$, contradiction. □

Lemma 6.4. *For every $\alpha \in \mathcal{O}$, $\alpha^p \in \mathbb{Z} + p\mathcal{O}$.*

Indeed, if $\alpha = \sum_0^{p-2} a_i \zeta^i$ with $a_i \in \mathbb{Z}$ then

$$\alpha^p \equiv \sum_0^{p-2} a_i^p \pmod{p}$$

and the sum on the right is in $\mathbb{Z}$.

Now we return to the proof of the first case. 6.2 gives us an equality of ideals in $\mathcal{O}$:

$$\prod_{i=0}^{p-1}(x + \zeta^i y) = (z)^p.$$

By 6.3, each factor on the left is a p-th power, say

$$(x + \zeta^i y) = I_i^p.$$

Since p is regular and I_i is of order p in the class group, this implies that $I_i = (\alpha_i)$ for some i. Write $\alpha = \alpha_1$. We have $x + \zeta y = u\alpha^p$ for

some unit u. By 5.11 we can write $u = \zeta^r v$ where $v = \bar{v}$. By the previous lemma, we have $\alpha^p \equiv a \pmod{p}$ for some $a \in \mathbb{Z}$. Thus

$$x + \zeta y = u\alpha^p \equiv \zeta^r v a \pmod{p}.$$

Similarly

$$x + \zeta^{-1} y = x + \bar{\zeta} y = \bar{u}\bar{\alpha}^p \equiv \zeta^{-r} v a \pmod{p}.$$

Combining these two, we have

$$\zeta^{-r}(x + \zeta y) \equiv \zeta^r (x + \zeta^{-1} y) \pmod{p},$$

in other words

$$x + \zeta y - \zeta^{2r} x - \zeta^{2r-1} y \in p\mathcal{O}. \tag{6.5}$$

Suppose the four powers of ζ are distinct. Then since $p > 5$, they are part of a $\mathbb{Z}$-basis of $\mathcal{O}$, hence their images mod p are linearly independent over $\mathbb{F}_p$. It follows that $p \mid x$ and $p \mid y$, which is a contradiction. Thus two of the elements of $\{1, \zeta, \zeta^{2r}, \zeta^{2r-1}\}$ are equal, and we know $1 \neq \zeta$. We consider the remaining possibilities:

- If $\zeta^{2r} = 1$, then 6.5 gives $\zeta y - \zeta^{2r-1} y = 0$ which implies $p \mid y$.
- if $\zeta^{2r-1} = 1$ (equivalently $\zeta = \zeta^{2r}$), then 6.5 gives

  $$(x - y) - (x - y)\zeta \equiv 0 \pmod{p},$$

  and as above this implies $p \mid (x - y)$, but we already eliminated this option.
- If $\zeta = \zeta^{2r-1}$ then 6.5 gives $x - \zeta^2 x \equiv 0 \pmod{p}$ which implies $p \mid x$, again a contradiction.

This completes the proof.

Regular primes have been tabulated since the time of Kummer. The first few are

3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 41, 43, 47, 53, 61, 71, 73, 79, 83, 89, 97, 107, 109, 113, 127, 137, 139, 151, 163, 167, 173, 179, 181, 191, 193, 197, 199.

The first irregular primes are 37, 59, 67, 101, 103, 131, 149. The study of class groups of K_p when p is irregular is the beginning of *Iwasawa theory*.

6.1. **Quadratic reciprocity.** The Legendre symbol is a Dirichlet character $a \mapsto \left(\frac{a}{p}\right)$ mod p. All that we are saying is that the product of two squares in ${\mathbb{F}_p}^\times$ is a square, but so is the product of two non-squares; in other words, the squares form a subgroup of index 2. So

$g_p = \sum_{a=1}^{p-1} \left(\frac{a}{p}\right) \zeta^a$, where ζ is the complex number $e^{\frac{2\pi i}{p}}$, is a Gauss sum (see Homework, week 8).

Lemma 6.6. $g_p^2 = p^* = (-1)^{\frac{p-1}{2}} p$.

Proof. This is recognizably a change of variables proof.

$$g_p^2 = \sum_{a=1}^{p-1} \left(\frac{a}{p}\right) \zeta^a \sum_{b=1}^{p-1} \left(\frac{b}{p}\right) \zeta^b = \sum_{a=1}^{p-1} \left(\frac{a}{p}\right) \zeta^a \sum_{b=1}^{p-1} \left(\frac{ab}{p}\right) \zeta^{ab}$$

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

The inner sum is $\Phi_p(\zeta^{b+1}) - 1 = 0 - 1$ if $b+1$ is not divisible by p, or $= p-1$ if $b+1$ is divisible by p, i.e., if $b = p-1$. So we get

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

But $\sum_{b=1}^{p-1} \left(\frac{b}{p}\right) = 0$ because there are equally many squares as non-squares, so this becomes

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

□

Lemma 6.7. *Assume p and q are distinct odd primes. Then* $(g_p)^{q-1} \equiv \left(\frac{q}{p}\right) \pmod{q}$.

Proof. Note that the previous lemma implies that g_p to any even power is in $\mathbb{Z}$, so the statement is meaningful. Write $\mathcal{O} = \mathbb{Z}[\zeta]$, the integers in $\mathbb{Q}(\zeta)$. First, taking qth powers is additive in $\mathcal{O}/q\mathcal{O}$, because of the binomial theorem. So

$$g_p^q = (\sum_{a=1}^{p-1} \left(\frac{a}{p}\right) \zeta^a)^q \equiv \sum_{a=1}^{p-1} (\left(\frac{a}{p}\right))^q \zeta^{aq} = \sum_{a=1}^{p-1} \left(\frac{a}{p}\right) \zeta^{aq} \pmod{q\mathcal{O}}.$$

Now let $uq \equiv 1 \pmod{p}$. Then we can replace a by au, so $auq = a$ in the sum; moreover, $\left(\frac{q}{p}\right) = \left(\frac{u}{p}\right)$ and thus

$$g_p^q \equiv \sum_{a=1}^{p-1} \left(\frac{au}{p}\right) \zeta^a = \left(\frac{u}{p}\right) g_q = \left(\frac{q}{p}\right) g_p. \pmod{q\mathcal{O}}$$

Now $(g_p)^{q+1} = p^*(g_p)^{q-1}$, and by the above this is congruent to $g_p \cdot \binom{q}{p} g_p \pmod{q\mathcal{O}}$. But by the previous lemma again, this is $\binom{q}{p} \cdot p^*$, so dividing by p^* we get the result. □

Now we prove quadratic reciprocity in one line. We have

$$\binom{p^*}{q} \equiv (p^*)^{\frac{q-1}{2}} \equiv (g_p)^{q-1} \equiv \binom{q}{p} \pmod{q}.$$

7. Day 7: Congruences

- (a) Solutions to congruences
- (b) Chevalley-Warning theorem
- (c) p-adic numbers

Let $m \in \mathbb{N}$, $m \geq 1$. Let $P(x_1, x_2, \ldots, x_m) \in \mathbb{Z}[x_1, \ldots, x_m]$ be a polynomial with integer coefficients. Suppose the equation

$$P(x_1, x_2, \ldots, x_m) = 0$$

has a solution in $\mathbb{Z}^m$. Then the congruence

$$P(x_1, x_2, \ldots, x_m) \equiv 0 \pmod{n}$$

has a solution for every integer $n > 1$. Conversely, if this congruence fails to have a solution for some $n > 1$ then the original equation has no integer solution.

Example 7.1. *Let $P(x, y, z) = x^2 + 7xy - 14yz - 3$. The congruence*

$$P(x, y, z) \equiv 0 \pmod{7}$$

is equivalent to the congruence

$$x^2 \equiv 3 \pmod{7}$$

and this has no solution, so the original equation has no solution.

There are, however, examples of equations that have solutions modulo n for all integers n and nevertheless have no integer (or rational) solutions.

Example 7.2. *A famous example was studied by the Norwegian Ernst Selmer in the 1950s. Consider the Diophantine equation*

$$3x^3 + 4y^3 = 5z^3.$$

This equation has solutions modulo every prime power, and therefore modulo every integer. But it has no rational solutions.

I will begin the argument for the former assertion, and complete it on Day 8, and also sketch the argument for the absence of rational solutions. I follow notes of Kevin Buzzard.

First, if $p = 3$, we take $z = 1$, $y = 2$; if $p = 5$ we take $y = 1$, $x = 3$. If $p = 2$ we take $x = y = z = 1$.

Lemma 7.3. *Let p be a prime other than 3 or 5. Then either 3, 5, 15, or 45 is a cube modulo p.*

Proof. Let $k = \mathbb{F}_p$, and consider $C = k^\times/(k^\times)^3$. If $p \equiv 1 \pmod 3$ then $|C| = 3$; if not, then $|C| = 1$. In the latter case, every number is a cube modulo p; in the former case, let a and b denote the images of 3 and 5 in the cyclic group C. If either a or b is trivial, then we are done. So we suppose a and b are both non-trivial. If $a = b$, then $45 = a^2b = 1$; if $a \neq b$, then $15 = ab = 1$. □

In what follows, we seek solutions with $z = 1$. If $x = -y$ then the problem is to solve $y^3 = 5 \bmod p$; so if 5 is a cube, there is a solution. If $y = 1$, the problem is to solve $3x^3 = 1$ modulo p; so if $3 \equiv r^3$, then we take $x \equiv r^{-1}$. If $y = 0$, we have to solve $3x^3 \equiv 5$, or equivalently $(3x)^3 \equiv 45 \pmod p$, so we are done if 45 is a cube.

If $p = 7$ we can take $x = y = 1$, $z = 0$. Finally if $p > 7$ and if 15 is a cube mod p, we take $z = 1$, $y = 5/7$. Then we find

$$3x^3 + 4 \cdot (5/7)^3 \equiv 5 \pmod p \Leftrightarrow 3(7x)^3 \equiv 5 \cdot 243 \Leftrightarrow (7x)^3 \equiv 3^3 \cdot 15$$

which has a solution.

Before returning to this example, I want to talk about congruences modulo a prime p that have many solutions.

Theorem 7.4 (Chevalley-Warning Theorem)**.** *Let $k = \mathbb{F}_q$ be a finite field of characteristic p. Let $P \in k[x_1, \ldots, x_n]$, and suppose the degree is smaller than the number of variables, i.e. $deg(P) < n$. Then the number of solutions in k^n of the congruence $P(x_1, \ldots, x_n) = 0$ is divisible by p.*

Corollary 7.5. *Under the hypotheses of the above theorem, suppose P is a homogeneous polynomial of positive degree. Then $P(x_1, \ldots, x_n) = 0$ has a solution in k^n other than the trivial solution $(0, \ldots, 0)$.*

Proof. (of Corollary). Indeed, when P is homogenous, it has at least the trivial solution. The Chevalley-Warning theorem then says that P has at least p solutions, and $p > 1$. □

Lemma 7.6. *Let $x^m := x_1^{m_1} \cdot \cdots \cdot x_n^{m_n}$ be a monomial. Suppose*

$$A(m_1, \ldots, m_n) = \sum_{x \in k^n} x^m \neq 0.$$

Then every $m_i > 0$ and $q - 1$ divides m_i for all i.

In particular, suppose $m_1 + \cdots + m_n < (q-1)n$. Then $A(m_1, \ldots, m_n) = 0$.

Proof. We can write

$$A(m_1, \ldots, m_n) = \prod_{i=1}^{n} A(m_i) = \prod_i (\sum_{x \in k} x^{m_i}).$$

So we need to show that $A(m) \neq 0$ implies $m > 0$ and is divisible by $q-1$. If $m = 0$, then $A(0) = \sum_{x \in k} x^0 = |k| = q = 0$. Suppose m is not divisible by $q-1$. Then there is $y \in k^\times$ such that $y^m \neq 1$. Then since multiplication by y is a permutation of k,

$$A(m) = \sum_{x \in k} x^m = \sum_{x \in k} (xy)^m = y^m A(m)$$

which implies $A(m) = 0$.

The hypothesis $m_1 + \cdots + m_n < (q-1)n$ implies that at least one of the $m_i < q-1$, which completes the proof. □

Now we prove the Chevalley-Warning theorem. The lemma implies that if $Q \in k[x_1, \ldots, x_n]$ is of degree $< (q-1)n$, then

$$\sum_{x \in k^n} Q(x) = 0.$$

Let P be as in the statement of the Chevalley-Warning theorem, and let $Q = 1 - P^{q-1}$. Then $\deg(Q) = (q-1)\deg(P) < (q-1)n$, so we see that $\sum_{x \in k^n} Q(x) = 0$. But either $P(x) = 0$, in which case $Q(x) = 1$, or $P(x) \in k^\times$, in which case $P(x)^{q-1} = 1$ and $Q(x) = 0$. The lemma thus implies

$$0 = \sum_{x \in k^n} Q(x) = \sum_{x \in k^n; P(x) = 0} 1 = |\{x \in k^n \mid P(x) = 0\}|.$$

This implies that the cardinality on the right is divisible by p, and completes the proof.

There are two ways to define p-adic numbers: by analysis and by algebra. I follow Serre's treatment of p-adic numbers by algebra.

Definition 7.7. Let G_n be a collection of groups indexed by $\mathbb{N}$, and assume that there is a homomorphism $\phi_n : G_n \to G_{n-1}$ for all $n \geq 1$. The *projective limit* (or *inverse limit*) $\varprojlim_n G_n$ is the set of sequences $(a_n \in G_n, n \in \mathbb{N})$ satisfying the relations

$$\phi_n(a_n) = a_{n-1}$$

for all $n \geq 1$. Letting $e_n \in G_n$ denote the identity, the group structure is defined by

$$(a_n)(b_n) = (a_n \cdot b_n);\ (a_n)^{-1} = (a_n^{-1}); e = (e_n \in G_n).$$

If the G_n are rings and the ϕ_n are ring homomorphisms, then $\varprojlim_n G_n$ is a ring.

That the indicated law defines a group structure follows from the elementary properties of homomorphisms. For example, $\phi_n(a_n^{-1}) = \phi_n(a_n)^{-1} = a_{n-1}^{-1}$, for all n.

The example of interest is $G_n = \mathbb{Z}/p^n\mathbb{Z}$, and we write $\mathbb{Z}_p = \varprojlim \mathbb{Z}/p^n\mathbb{Z}$, where $\phi_n(a)$ is the reduction of a modulo p^{n-1}. Suppose $z \in \mathbb{Z}$; we can write z in base p as

$$z = a_0 + pa_1 + p^2a_2 + \cdots + p^ja_j$$

for some j, with $0 \leq a_i \leq p-1$. If $j > n-1$, then the image of z in G_n is the partial sum $\sum_{i=0}^{n-1} a_ip^i$. So the elements of $\mathbb{Z}_p$ are formally infinite sums $\sum_{i=0}^{\infty} a_ip^i$, with $0 \leq a_i \leq p-1$, where this makes sense if we give the set of such sequences the metric topology in which $|p^i| = \epsilon^i$ for some $0 < \epsilon < 1$. The topology doesn't depend on the choice, but one usually takes $|p^i|_p = \frac{1}{p^i}$ for all $i \in \mathbb{Z}$; this defines a metric topology on $\mathbb{Q}$, whose completion is the field $\mathbb{Q}_p$ of p-adic numbers, and the closure of $\mathbb{Z}$ is $\mathbb{Z}_p$.

In Serre's construction, the set $\prod_n G_n$ is given the product topology. Since each G_n is finite, hence compact, the product is also compact. The condition $\phi_n(a_n) = a_{n-1}$ is a closed condition in the product topology for each n, so the intersection defined by the collection of these conditions is a closed subspace, hence is also compact. The maps ϕ_n are all surjective. We begin by observing that, for any pair of integers $m \geq n$, there is a short exact sequence

$$0 \to G_{m-n} \xrightarrow{p^n} G_m \xrightarrow{\varepsilon_n} G_n \to 0$$

where the map labelled p^n takes an element $\bar{a} \in G_{m-n}$, lifts it to $a \in \mathbb{Z}$, and then sends it to the image of $p^n \cdot a$ in G_m. The difference between two lifts a, a' is an element $b = p^{m-n}c \in p^{m-n}\mathbb{Z}$, and the difference $p^n \cdot a - p^n \cdot a' = p^n \cdot p^{m-n}c = p^m \cdot c$ is sent to 0 in G_m.

Moreover, one checks directly that the diagram

$$\begin{array}{ccccccccc}
0 & \longrightarrow & G_{m-n} & \xrightarrow{p^n} & G_m & \xrightarrow{\varepsilon_n} & G_n & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \phi_{m-n}} & & \downarrow{\scriptstyle \phi_m} & & \downarrow{\scriptstyle \phi_n} & & \\
0 & \longrightarrow & G_{m-n-1} & \xrightarrow{p^n} & G_{m-1} & \xrightarrow{\varepsilon_n} & G_{n-1} & \longrightarrow & 0
\end{array}$$

is commutative. It follows easily that

Lemma 7.8. *The sequence*

$$0 \to \mathbb{Z}_p \xrightarrow{p^n} \mathbb{Z}_p \xrightarrow{\varepsilon_n} G_n \to 0$$

is exact.

Proof. First, multiplication by p, and hence by p^n, is injective. Indeed, suppose $p(a_n) = 0$; then $pa_n = 0$ for all n, hence $a_n \in p^{n-1}G_n$ for all n. If $a_n \neq 0$ for some n, then $\phi_{n+1}(a_{n+1}) = a_n \neq 0$, which means that $a_{n+1} \notin p^n G_{n+1}$, contradiction.

Now the image of multiplication by p^n is in $Ker(\varepsilon_n)$. Conversely, if $\varepsilon_n((x_m)) = 0$, then $x_m \equiv 0 \pmod{p^n}$ for all $n \geq m$. Thus the above short exact sequence shows that $x_m \equiv p^n a_m \pmod{p^m}$ for some $a_m \in G_{m-n}$. It then follows from the commutativity of the diagram above that $\phi_m(a_m) = a_{m-1}$ for all $m \geq n$, and so $a = (a_m)$ defines an element of $\mathbb{Z}_p$ (we can take $a_i = 0$ for $i \leq n$) and $p^n a = (x_m)$. Finally, $\mathbb{Z} \subset \mathbb{Z}_p$ by the natural map to G_n for all n, and it follows that ε_n is surjective because its restriction to $\mathbb{Z}$ is already surjective. □

On the other hand, the topology is totally disconnected: for each n the ideal $p^n\mathbb{Z}_p$ is the inverse image of the open subset $\{0\} \subset G_n$ with respect to the (tautologically continuous) map ε_n, and the set of these ideals forms a basis for the neighborhoods of 0 in $\mathbb{Z}_p$. Since each coset of the ideal is homeomorphic to the ideal (translation by an element of the ring is a homeomorphism of the ring) this ideal is closed as well as open.

Proposition 7.9. *The element $u \in \mathbb{Z}_p$ belongs to $\mathbb{Z}_p^\times$ if and only if $u \notin p\mathbb{Z}_p$. If $x \in \mathbb{Z}_p$ is a non-zero element, then there is a unique $r \geq 0$ such that $x = p^r \cdot u$. In other words $\mathbb{Z}_p$ is a discrete valuation ring.*

Proof. Let $I \subset \mathbb{Z}_p$ be an ideal. Let $r = \inf\{i \mid \varepsilon_i(I) \neq \{0\}\}$. Then for all $n \geq r$, $I_n = \varepsilon_n(I) \neq \{0\}$. An element in G_n is invertible if and only if it is not divisible by p. It follows from the lemma that $I \subset p^{r-1}\mathbb{Z}_p$. Let $a \in I$, $\varepsilon_r(a) \neq 0$; then there is $u = (u_n) \in \mathbb{Z}_p$ such that $p^{r-1}u = a$ and $u \notin p\mathbb{Z}_p$. It follows that u_n is invertible mod p^n for all n. Let $v_n = (u_n)^{-1}$; then $\phi_n(v_n) = v_{n-1}$ and so $v = (v_n)$ is the inverse of u in $\mathbb{Z}_p$. It then follows that p^{r-1} generates I; in other words, every ideal is a power of (p). this implies the second statement. □

We write $r = v_p(x)$ in the situation of the proposition. We write $U = \mathbb{Z}_p^\times$. In the next section we study the structure of this group.

8. Day 8: p-adic numbers

- (a) The structure of the p-adic unit group
- (b) Hensel's lemma and diophantine congruences

The map $x \mapsto x^p$ defines a homomorphism $\alpha : U \to U$. $\alpha(x) \equiv x \pmod{p}$. Let $U_n \subset G_n$ be the unit group of G_n; we have $U_n = \varepsilon_n(U)$. For each n there is a homomorphism $\alpha_n : U_n \to U_n$. Now $|U_n| =$

$(p-1)\cdot p^{n-1}$. It follows that for any $u_n \in U_n$, $u_n^{(p-1)\cdot p^{n-1}} = 1$. But this is just $\alpha_n^n(u_n)/\alpha_n^{n-1}(u_n)$. It follows that

Lemma 8.1. *For any $u \in U$ and any n*

$$\alpha^n(u) \equiv \alpha^{n-1}(u) \pmod{p^n}.$$

In particular, $\omega(u) = \lim_n \alpha^n(u)$ is well defined in U. Moreover $\omega(u) \equiv u \pmod{p}$, and $\omega(u)^{p-1} = 1$.

The first part of the lemma follows from the calculation above. The claim that $\omega(u) \equiv u \pmod{p}$ follows by induction from the same property of α. Finally, $\omega(u)^p = \lim_n \alpha^{n+1}(u) = \lim_n \alpha^n(u) = \omega(u)$, which implies that $\omega(u)$ is a $p-1$st root of 1.

The image of ω is thus a subgroup of U isomorphic to ${\mathbb{F}_p}^\times$ under reduction modulo p. The map ω is called the Teichmüller lift. The kernel of ω is $U_1 = \{u \in U \mid u \equiv 1 \pmod{p}\}$.

Now consider the formal power series

$$log(1+x) = \sum_{i\geq 1} (-1)^{i-1} x^i/i.$$

If p divides x then $|x^i/i|_p \geq i - log_p(i)$ which tends to 0 as i tends to infinity. Now

Proposition 8.2. *An infinite series $\sum_n a_n$ of p-adic numbers converges in $\mathbb{Z}_p$ if and only if $\lim_n |a_n|_p = 0$.*

This follows from the (easy) ultrametric property

$$|x+y|_p \leq \sup(|x|_p, |y|_p).$$

It then follows formally that $u \mapsto log(u) : U_1 \to \mathbb{Z}_p$ is a homomorphism and one shows easily that it is in fact an isomorphism if $p > 2$.

I can return to the claims of the first day.

Lemma 8.3. *Let a be an integer prime to p. If p is odd, then a is a square mod p^r for all $r \geq 1$ if and only if a is a square mod p. If $p = 2$, then a is a square mod 2^r for all $r \geq 3$ if and only if $a \equiv 1 \pmod{8}$.*

Proof. First assume p is odd. Write $a = \omega(a)u$ for some $u \in U_1 \subset \mathbb{Z}_p$. We know that a is a square mod p^r if and only if $\omega(a)u$ is a square mod p^r. But if p is odd, then every element of the p-group $U_1/(1+p^r\mathbb{Z}_p)$ is a square, so p is a square mod p^r if and only if $\omega(a)$ is a square mod p^r, but $\omega(a)$ is determined by its image modulo p.

If $p = 2$, then suppose $a \equiv 1 \pmod{8}$. Let U_3 be the subgroup of U of elements congruent to 1 modulo 8. Say $a = 1 + 8x_3$, $b_3 = 1 - 4x_3$. Then $ab_3^2 \equiv 1 \pmod{16}$. Suppose we have found an element

b such that $ab_{r-1}^2 \equiv 1 \pmod{2^r}$, with $r \geq 3$, say $ab_{r-1}^2 = 1 + 2^r x_r$. Let $b_r = b_{r-1} \cdot (1 - 2^{r-1}x_r)$; then $ab_r^2 \equiv 1 \pmod{2^{2r-2}}$. As long as $r > 2$, the approximation is improved, and we find that $lim_r b_r^2 = a$. (Alternatively, we could use that the log converges and is injective on U_3). □

Alternatively, we could use Newton's method, in the following form:

Lemma 8.4. *Let $f \in \mathbb{Z}_p[X]$, f' its derivative. Let $x \in \mathbb{Z}_p$, $n, k \in \mathbb{Z}$ with $0 \leq 2k < n$, $f(x) \equiv 0 \pmod{p^n}$, $v_p(f'(x)) = k$. Then there is $y \in \mathbb{Z}_p$ such that $f(y) \equiv 0 \pmod{p^{n+1}}$, $v_p(f'(y)) = k$, $y \equiv x \pmod{p^{n-k}}$.*

This is a version of *Hensel's Lemma*.

Proof. We consider $y = x + p^{n-k}z$ and try to satisfy these conditions for z. By Taylor's formula,

$$f(y) = f(x) + p^{n-k}zf'(x) + p^{2n-2k}a$$

for some $a \in \mathbb{Z}_p$. Meanwhile, $f(x) = p^n b$ and $f'(x) = p^k c$ with $b \in \mathbb{Z}_p$, $c \in U$. Then we can find $z \in \mathbb{Z}_p$ such that $b + zc \equiv 0 \pmod{p}$. But now

$$f(y) = p^n(b + zc) + p^{2n-2k}a \equiv 0 \pmod{p^{n+1}}$$

because $2n - 2k > n$. When we apply Taylor's formula to f' we get

$$f'(y) = p^k c \pmod{p^{n-k}}$$

but $n - k > k$ so $v_p(f'(y)) = c$. □

So y satisfies the same hypothesis as x but is a better approximation to a root.

Theorem 8.5. *Let $f \in \mathbb{Z}_p[X_1, \ldots, X_m]$, $x = (x_i) \in \mathbb{Z}_p^m$, $n, k \in \mathbb{Z}$, $1 \leq j \leq m$. Suppose $0 \leq 2k < n$, $f(x) \equiv 0 \pmod{p^n}$, $v_p(\partial f/\partial X_j)(x)) = k$. Then f has a zero in $(\mathbb{Z}_p)^m$ congruent to x modulo p^{n-k}.*

Before I prove the theorem, here are a few corollaries. The first one was stated the first day:

Lemma 8.6. *Let $f \in \mathbb{Z}[X]$. Let p be a prime, and suppose $\bar{f}$ has no multiple roots in $\bar{\mathbb{F}}_p$. Then the congruence $f(X) \equiv 0 \pmod{p}$ has a solution if and only if for all $r \geq 1$ the congruence $f(X) \equiv 0 \pmod{p^r}$ has a solution.*

This is the case of the theorem with $m = 1$, $k = 0$. We can then take $n = 1$ and find that f has a zero in $\mathbb{Z}_p$, hence has a solution mod p^r for all r if it has one mod p.

Lemma 8.7. *Selmer's equation $3x^3 + 4y^3 = 5z^3$ has a solution mod p^r for all p and all r.*

Proof. Here $m = 3$. We take $F(x, y, z) = 3x^3 + 4y^3 - 5z^3$. It has a zero modulo p for all p, so we can at least take $n = 1$. Suppose $p > 7$; then there is a solution with $z = 1$, $v_p(\partial F/\partial z)(x, y, z)) = 0$. If $p = 7$ the solution is $x = y = 1$ and the partials with respect to x and y have valuation 0. The primes 2 and 5 pose no problem using $\partial F/\partial x$. Finally for $p = 3$, we need to find a solution with $n = 3$, $k = 1$. Take $x = 0$, $z = 1$, so the equation is $4y^3 \equiv 5 \pmod{27}$, and if we take $y = 2$ we have a solution. □

Now to prove the theorem. Suppose first $m = 1$. Newton's method gives us $x_1 \equiv x_0 = x \pmod{p^{n-k}}$ with $f(x_1) \equiv 0 \pmod{p^{n+1}}$ and no change in $v(f')$. By induction, we have a sequence $x_0, x_1, \ldots x_q, \ldots$ with $x_{q+1} \equiv x_q \pmod{p^{n+q-k}}$, $f(x_q) \equiv 0 \pmod{p^{n+q}}$. Let $y = \lim_q x_q$; this is a Cauchy sequence so the limit exists, and satisfies the requirement.

For $m > 1$, we leave the x_i alone for $i \neq j$ and reduce to the polynomial $\tilde{f}(x_j)$ in one variable obtained by inserting the chosen x_i in the other places.

9. Day 9: Dirichlet series

(Serre, Cours d'arithmetique)

- (a) Dirichlet series
- (b) L-functions of Dirichlet characters
- (c) Factorization of the Dedekind zeta function of cyclotomic fields
- (d) Proof that $L(1, \chi) \neq 0$.

9.1. Basic properties of Dirichlet series.

Lemma 9.1. *(Abel summation) Let (a_n) and (b_n) be two sequences, and define*

$$A(m, p) = \sum_m^p a_i, \;\; S(m, m') = \sum_m^{m'} a_i b_i.$$

Then

$$S(m, m') = \sum_{n=m}^{m'-1} A(m, n)(b_n - b_{n+1}) + A(m, m')b_{m'}.$$

Proof. Rearrangement of the sum □

Proposition 9.2. *Let $f : \mathbb{N}^* \to \mathbb{C}$ be an arithmetic function.*

- *If $f(n) = O(n^\beta)$ for some $\beta \in \mathbb{R}$, then*

$$D_f(s) = \sum_{n \geq 1} \frac{f(n)}{n^s}$$

converges absolutely for $Re(s) > 1 + \beta$, uniformly on compact subsets, and defines a holomorphic function on this region.

- *Conversely, if $D_f(s_0)$ converges (absolutely or not), then there exists $C > 0$ such that $|f(n)| \leq Cn^{Re(s_0)}$ for all n.*
- *If the closed half-plane $Re(s) \geq A$ is contained in the region of absolute convergence, then $D_f(s)$ is uniformly bounded on $Re(s) \geq A$.*

Proof. (1) We always write $s = \sigma + i\tau$. We have

$$|f(n)n^{-s}| = |f(n)|n^{-\sigma} = O(n^{-(\sigma-\beta)}).$$

Then by comparison with the sum $\sum n^{-\sigma}$ we see that $D_f(s)$ is absolutely convergent for $\sigma - \beta > 1$. Since this depends only on $Re(s)$, the uniformity on compact subsets follows from the same property on the line, and it then implies holomorphy.

(2) If $\sum f(n)n^{-s_0}$ converges then $\lim_{n\to\infty} |f(n)|n^{-s_0} = 0$, which implies the claim.

(3) For $\sigma \geq A$,

$$|D_f(s)| \leq \sum |f(n)|n^{-\sigma} \leq \sum |f(n)|n^{-A} = M$$

which gives a uniform bound. □

Proposition 9.3. *Let f be multiplicative and of moderate growth, and say $D_f(s)$ converges absolutely for $Re(s) > \sigma_0$. Then on the half-plane $Re(s) > \sigma_0$, we have the absolutely convergent product decomposition*

$$D_f(s) = \prod_p \sum_{k\geq 0} f(p^k)p^{-ks}.$$

Proof. Since $\sum f(n)n^{-s}$ converges absolutely, the subseries $\sum_{k\geq 0} f(p^k)p^{-ks}$ also converges absolutely.

Now for any number $z \geq 2$ let $\mathbb{N}(z) \subset \mathbb{N}$ be the set of integers all of whose prime factors are $\leq z$. Consider

$$\prod_{p\leq z} \sum_{k\geq 0} f(p^k)p^{-ks} = \sum_{d\in\mathbb{N}(z)} f(d)d^{-s}.$$

Now if $d \notin \mathbb{N}(z)$ then $d > z$. So

$$|D_f(s) - \prod_{p\leq z} \sum_{k\geq 0} f(p^k)p^{-ks}| \leq \sum_{d>z} |f(d))d^{-\sigma}$$

and this tends to zero for $Re(s) > \sigma_0$ because the series $D_f(s)$ converges absolutely. □

Example: Euler product for $\zeta(s)$.

Theorem 9.4 (Landau's Lemma)**.** *Let $D(s) = \sum \frac{a_n}{n^s}$ be a Dirichlet series with $a_n \in \mathbb{R}$, $a_n \geq 0$ for all n. Suppose D converges absolutely for $Re(s) > \sigma_0$ and that D extends to a holomorphic function in a neighborhood of σ_0. Then there exists $\epsilon > 0$ such that D converges absolutely for $Re(s) > \sigma_0 - \epsilon$.*

In other words, if σ_0 is the abcissa of convergence of D, then D has a singularity at the point σ_0.

Proof. First, this statement is invariant under translation, so we may as well assume $\sigma_0 = 0$. By hypothesis, D is holomorphic on a disk around 1 of radius *strictly greater* than 1, say for $|s-1| \leq 1+\epsilon$ for some $\epsilon > 0$. Then the Taylor series converges absolutely in this disk:

$$D(s) = \sum_{i=0}^{\infty} \frac{(s-1)^i}{i!} D^{(i)}(1).$$

On the other hand, the Dirichlet series is absolutely convergent for $Re(s) > 0$, uniformly on compact subsets. So its derivative can be calculated term by term. It follows that

$$D^{(i)}(s) = \sum \frac{a_n(-\log n)^i}{n^s}, \ i \geq 0, Re(s) > 0.$$

In particular

$$D^{(i)}(1) = \sum \frac{a_n(-\log n)^i}{n}.$$

Thus for $|s-1| \leq 1+\epsilon$ we have

$$D(s) = \sum_{i=0}^{\infty} \frac{(s-1)^i}{i!} \sum_n \frac{a_n(-\log n)^i}{n}.$$

We set $s = -\epsilon$. Then

$$\begin{aligned} D(-\epsilon) &= \sum_{i=0}^{\infty} \frac{(-\epsilon-1)^i}{i!} \sum_n a_n \frac{(-\log n)^i}{n} \\ &= \sum_{i=0}^{\infty} \frac{(1+\epsilon)^i}{i!} \sum_n a_n \frac{(-1)^i(-\log n)^i}{n} = \sum_{i=0}^{\infty} \frac{(1+\epsilon)^i}{i!} \sum_n a_n \frac{(\log n)^i}{n}. \end{aligned}$$

But now all the terms in the sum are positive. So we can exchange the order of the summation:

$$\begin{aligned} D(-\epsilon) &= \sum_n \frac{a_n}{n} \sum_{i=0}^{\infty} \frac{(1+\epsilon)^i(\log n)^i}{i!} \\ &= \sum_n \frac{a_n}{n} e^{(1+\varepsilon)\log n} = \sum_n \frac{a_n}{n^{-\epsilon}} \end{aligned}$$

which implies the Dirichlet series converges absolutely for $Re(s) \geq -\epsilon$, and this concludes the proof. □

Proposition 9.5. (a). *The Riemann zeta function has a meromorphic continuation to $Re(s) > 0$ with a simple pole at $s = 1$, with residue $+1$, and no other singularity.*
(b). $\sum_p \frac{1}{p^s} = -log(s-1) + O(1)$.

Proof. Proof of (a): We prove that

$$\zeta(s) = \frac{1}{s-1} + \phi(s)$$

where $\phi(s)$ is holomorphic for $Re(s) > 0$.

Note that

$$s\int_n^{n+1} t^{-s-1}dt = \frac{1}{n^s} - \frac{1}{(n+1)^s}.$$

So

$$\zeta(s) = \sum_{n=1}^{\infty} n{\cdot}s\int_n^{n+1} t^{-s-1}dt = \sum_{n=1}^{\infty} s\int_n^{n+1} \lfloor t\rfloor{\cdot}t^{-s-1}dt = s\int_1^{\infty} \lfloor t\rfloor{\cdot}t^{-s-1}dt.$$

This in turn equals

$$\zeta(s) = s\int_1^{\infty} t^{-s}dt + s\int_1^{\infty} (\lfloor t\rfloor - t)t^{-s-1}dt = \frac{1}{s-1} + 1 + s\int_1^{\infty} (\lfloor t\rfloor - t)t^{-s-1}dt.$$

Now $|(\lfloor t\rfloor - t)| \leq 1$ so the final integral is majorized by $\int_1^{\infty} t^{-s-1}dt$ which converges absolutely and uniformly for $Re(s) > 0$.

Proof of (b): See Proposition 10.1.

□

9.2. **Dirichlet characters and Dirichlet L-functions.**

Definition 9.6. The arithmetic function $f : \mathbb{N}^* \to \mathbb{C}$ is a *Dirichlet character modulo N* if f is multiplicative, $f(a)$ depends only on the image of $a \pmod N$, and $f(a) = 0$ whenever $(a, N) > 1$. If there is no proper divisor $N'|N$ such that $f(a)$ depends only on the value of $a \pmod{N'}$ for $(a, N) = 1$, then f is a primitive Dirichlet character.

We write $L(s, \chi) = D_f(s)$ if f is a Dirichlet character χ. We only consider Dirichlet characters χ modulo a prime p; χ is then primitive if and only if there is a prime to p such that $\chi(a) \neq 1$; the only non-primitive character is χ_0. The restriction of χ to $(\mathbb{Z}/p\mathbb{Z})^{\times}$ is a homomorphism ${\mathbb{F}_p}^{\times} \to \mathbb{C}^{\times}$.

Lemma 9.7. *Euler product for $L(s, \chi)$.*

Lemma 9.8. *Let χ be a Dirichlet character modulo p and let $S(\chi) = \sum_{a=0}^{p-1} \chi(a)$. Then $S(\chi) = 0$ if $\chi \neq \chi_0$; $S(\chi_0) = p - 1$.*

Two proofs: if $\chi \neq \chi_0$ then there is b such that $\chi(b) \neq 1$; but

$$S(\chi) = \sum_{a=0}^{p-1} \chi(ab) = \chi(b)S(\chi)$$

which implies that $S(\chi) = 0$. Alternatively, the values of χ are all dth roots of 1 for some $d > 1$, $d|p-1$. Let a be a cyclic generator of ${\mathbb{F}_p}^\times$, $\zeta = \chi(a) \in \mu_d$ assumed to be a primitive root; and then

$$S(\chi) = \frac{p-1}{d} \sum_{c=1}^{d} \zeta^c = P_d(\zeta) = 0.$$

Proposition 9.9. *Assume $\chi \neq \chi_0$. Then $L(s, \chi)$ extends holomorphically to $Re(s) > 0$.*

For this we use the following lemma:

Lemma 9.10. *Let f be an arithmetic function. Suppose the partial sums $F(m, p) = \sum_{i=m}^{p} f(i)$ are bounded for all m, p. Then D_f converges (conditionally) for $Re(s) > 0$ to a holomorphic function.*

Proof. The proof is by Abel summation. Let (a_n) and (b_n) be two sequences, and define

$$A(m, p) = \sum_{m}^{p} a_i, \;\; S(m, m') = \sum_{m}^{m'} a_i b_i.$$

Then

$$S(m, m') = \sum_{n=m}^{m'-1} A(m, n)(b_n - b_{n+1}) + A(m, m')b_{m'}.$$

as we see by replacing a_n by $A(m, n) - A(m, n-1)$ and regrouping. Apply this to $a_n = f(n)$, $b_n = n^{-s}$. Then there is a constant K such that $|A(m, n)| \leq K$ for all m, n, and we find

$$|S(m, m')(s)| = |\sum_{n=m}^{m'} f(n)n^{-s}| \leq K[\sum_{n=m}^{m'-1} |\frac{1}{n^s} - \frac{1}{(n+1)^s}| + |\frac{1}{n^s}|].$$

If $s \in \mathbb{R}^+$ then the right-hand side is just $\frac{K}{m^s}$ and so $|S(m, m')(s)|$ tends to 0 as m, m' tend to infinity, uniformly on compact subsets of $\mathbb{R}^+$. Another application of Abel summation (see below) then shows that the series converges uniformly on compact subsets of $Re(s) > 0$.

□

We have used the following Lemma:

Suppose $D_f(s)$ converges for $s = s_0$. Then it converges uniformly in every domain of the form $Re(s - s_0) \geq 0$, $|Arg(s - s_0)| \leq \alpha$, provided $\alpha < \pi/2$.

Proof. Of course we may assume $s_0 = 0$. The series $\sum_n f(n)$ then converges, which implies that for any ε there is a constant K such that, for any $m, m' > K$, $|A(m, m')| < \varepsilon$ with $A(m, m') = \sum_m^{m'} f(i)$. We need to prove uniform convergence in any domain of the form $\sigma = Re(s) \geq 0$, $|s|/\sigma \leq k$. Now apply Abel summation, with $b_n = n^{-s}$ as before:

$$S(m, m') = \sum_m^{m'-1} A(m, n)(n^{-s} - (n+1)^{-s}) + A(m, m')(m')^{-s};$$

$$|S(m, m')| \leq \varepsilon \cdot [\sum_m^{m'-1} |n^{-s} - (n+1)^{-s}| + (m')^{-\sigma}]$$

whenever $m, m' > K$. Now there is an elementary inequality (see below):

$$|n^{-s} - (n+1)^{-s}| \leq |s|/\sigma \cdot (n^{-\sigma} - (n+1)^{-\sigma}).$$

Thus since $|(m')^{-\sigma}| \leq 1$ and $|s|/\sigma < k$ we have

$$|S(m, m')| \leq \varepsilon \cdot [1 + k \sum_m^{m'-1} n^{-\sigma} - (n+1)^{-\sigma}] = \varepsilon \cdot [1 + k(m^{-\sigma} - (m')^{-\sigma}] \leq \varepsilon(1+k)$$

for $m, m' >> 0$, and this implies the uniform convergence on the sector. □

The inequality is elementary: for any real numbers $0 < \alpha < \beta$, with $s = \sigma + i\tau$, $\sigma > 0$, we have

$$|e^{-\alpha s} - e^{-\beta s}| \leq \frac{|s|}{\sigma}(e^{-\alpha\sigma} - e^{-\beta\sigma}).$$

Indeed

$$e^{-\alpha s} - e^{-\beta s} = s \int_\alpha^\beta e^{-ts} dt.$$

Thus

$$|e^{-\alpha s} - e^{-\beta s}| \leq |s| \int_\alpha^\beta e^{-t\sigma} dt = \frac{|s|}{\sigma}(e^{-\alpha\sigma} - e^{-\beta\sigma}).$$

9.3. **Dedekind zeta function of $\mathbb{Q}(\zeta_p)$.**

Definition 9.11. Let K be a number field. The *Dedekind zeta function* of K is the Dirichlet series

$$\zeta_K(s) = \sum_{\mathfrak{a} \subset \mathcal{O}_K} N(\mathfrak{a})^{-s} = \sum_{n=1}^{\infty} C_K(n) n^{-s}$$

where $\mathfrak{a}$ runs over ideals of $\mathcal{O}_K$ and $C_K(n)$ is the number of ideals of norm n. ,

The semigroup of integral ideals is generated by prime ideals in $\mathcal{O}$, and so we can write the formal product

$$\zeta_K(s) = \prod_{\mathfrak{p}} (1 - N\mathfrak{p}^{-s})^{-1} = \prod_{p} \prod_{\mathfrak{p}|p} (1 - p^{-f_{\mathfrak{p}} s})^{-1}$$

where $f_{\mathfrak{p}}$ is the residue degree. We simplify by assuming K Galois, so this is just $\prod_p (1-p^{-f_p s})^{-g_p}$. It is a general fact about infinite products that the product converges absolutely (to a holomorphic function) if and only if the sum of the logarithms converges absolutely (uniformly on compact subsets). But the number of primes dividing p is bounded by $[K : \mathbb{Q}]$, and for $Re(s) > 1$ we have $|log(1 - p^{-fs})| \leq |log(1 - p^{-s})|$ for every p. So the sum of logarithms is majorized in absolute value by

$$[K : \mathbb{Q}] \sum_p |log(1 - p^{-s})|$$

and this implies the Euler product converges absolutely whenever the Euler product for $\zeta(s)$ converges absolutely, for $Re(s) > 1$. Then arguing as for $\zeta(s)$, we see $\zeta_K(s)$ converges absolutely for $Re(s) > 1$.

Theorem 9.12. *The Dedekind zeta function $\zeta_K(s)$ has a pole at $s = 1$.*

This is true in general, but the proof requires a lengthy calculation of Dirichlet. Here we prove it for the field $\mathbb{Q}(\zeta_p)$.

Theorem 9.13. *Let $K = \mathbb{Q}(\zeta_p)$. Then*

$$\zeta_K = \zeta(s) \cdot \prod_{\chi \neq \chi_0} L(s, \chi)$$

where the product is over (primitive) characters modulo p.

Proof. The proof is an application of cyclotomic reciprocity. We have to show that for each prime $q \neq p$,

$$(1 - q^{-f_q s})^{-g_q} = \prod_{\chi \neq \chi_0} (1 - \chi(q) q^{-s})^{-1} \cdot (1 - q^{-s})^{-1}$$

or equivalently, setting $T = q^{-s}$

$$(1 - T^{f_q})^{g_q} = \prod_{\chi} (1 - \chi(q)T).$$

By cyclotomic reciprocity, we know that f_q is the order of q in the group ${\mathbb{F}_p}^\times$. Let $X({\mathbb{F}_p}^\times) = Hom({\mathbb{F}_p}^\times, \mathbb{C}^\times)$. Since ${\mathbb{F}_p}^\times$ is cyclic of order $p-1$, choosing a cyclic generator α identifies $X({\mathbb{F}_p}^\times)$ with the group μ_{p-1} of $(p-1)$st roots of 1 in $\mathbb{C}^\times$ by sending χ to $\chi(\alpha)$. Thus $|X({\mathbb{F}_p}^\times)| = |{\mathbb{F}_p}^\times|$. For any subgroup $H \subset {\mathbb{F}_p}^\times$, let $H^\perp = \{\chi \in X({\mathbb{F}_p}^\times) \mid \chi|_H = 1\}$. Since H is cyclic, one sees easily that $|H| \cdot |H^\perp| = p-1$. It follows that there is a short exact sequence

$$1 \to H^\perp \to X({\mathbb{F}_p}^\times) \to X(H) = Hom(H, \mathbb{C}^\times) \to 1.$$

We apply this to the subgroup $\langle q \rangle \subset {\mathbb{F}_p}^\times$, of order f_q, generated by the residue of q. The set $X(H)$ is then identified with the set of possible f_q-th roots of 1, by $\chi \mapsto \chi(q)$. In other words,

$$\prod_{\chi \neq \chi_0} (1 - \chi(q)T) = \prod_{\zeta^{f_q}=1} (1 - \zeta^{f_q} T)^{|H^\perp|},$$

each value counted $|H^\perp| = [{\mathbb{F}_p}^\times : H] = g_q$ times. But $\prod_{\zeta^{f_q}=1}(1 - \zeta^{f_q}T) = (1 - T^{f_q})$. So this completes the proof for $q \neq p$. For $q = p$, we have seen that there is exactly one $\mathfrak{p}$ dividing p, of norm p, and on the other hand for $\chi \neq \chi_0$ one has $\chi(p) = 0$. So the two sides match for p as well. □

Theorem 9.14. *(i) Let $\chi \neq \chi_0$; then $L(1, \chi) \neq 0$. (ii) Let $K = \mathbb{Q}(\zeta_p)$. Then the Dedekind zeta function $\zeta_K(s)$ has a pole at $s = 1$.*

Proof: Follow Serre, Cours d'arithmétique. Since $\zeta_K(s) = \zeta(s) \times \prod_{\chi \neq \chi_0} L(s, \chi)$, (ii) is a consequence of (i) and of the existence of a pole of $\zeta(s)$. So we prove (i). There are many proofs, and we use one specific to this situation. Recall that $\zeta(s) = \frac{1}{s-1} + \phi(s)$ where $\phi(s)$ is holomorphic for $Re(s) > 0$. Thus if $L(1, \chi) = 0$ for one χ then ζ_K would be holomorphic up to $Re(s) > 0$. On the other hand, ζ_K is a Dirichlet series with positive coefficients. By Landau's lemma, it suffices to show that its region of absolute convergence is for $Re(s) > \sigma_0$ with σ_0 strictly greater than 0, in order to derive a contradiction.

Now for $q \neq p$, the q factor of ζ_K is

$$\frac{1}{(1 - q^{-f_q s})^{g_q}} = (1 + q^{-f_q s} + p^{-2f_q s} + \dots)^{g_q}$$

which (for s real) dominates the series

$$1 + q^{-f_q g_q s} + q^{-2 f_q g_q s} + \cdots = \sum_{i=1}^{\infty} q^{-i(p-1)s}.$$

It follows that the n-th coefficient of ζ_K is greater than or equal to the n-th coefficient of

$$\sum_{(p,n)=1} \frac{1}{n^{(p-1)s}}$$

which diverges for $s = \frac{1}{p-1}$. Thus ζ_K has a pole on $Re(s) > 0$, and thus all the $L(1, \chi) \neq 0$.

10. Day 10: Dirichlet's theorem on primes in an arithmetic progression

- (a) Dirichlet density
- (b) Primes in an arithmetic progression

First, some unfinished business with $\zeta(s)$.

Proposition 10.1. *The series $\sum_p p^{-s}$, where p runs over all rational primes, is asympotic to $\log(\frac{1}{s-1})$ when $s \to 1$. The series $\sum_p \sum_{k \geq 2} p^{-ks}$ is bounded when $s \to 1$.*

Proof. The variable s is real and greater than 1. We consider

$$\log(\zeta(s)) = -\sum_p \log(1 - p^{-s}) = \sum_p \sum_{k \geq 1} \frac{1}{k p^{ks}} = \sum_p p^{-s} + \sum_p \sum_{k \geq 2} \frac{1}{k p^{ks}}.$$

Now the second sum is majorized by $\sum_p \sum_{k \geq 2} p^{-ks}$ which is the sum of the geometric series

$$\sum_p \frac{1}{p^s(p^s - 1)} \leq \sum_p \frac{1}{p(p-1)} \leq \sum_n \frac{1}{n(n-1)} = 1.$$

So the second term is bounded in a neighborhood of 1, and so $\sum_p p^{-s}$ is asymptotic to $\log(\zeta(s)$ which is asymptotic to $\log(\frac{1}{s-1})$ when $s \to 1$. □

This justifies the following definition.

Definition 10.2. Let A be a subset of the set P of prime numbers. The Dirichlet density of A, if it exists, is the number k such that

$$\lim_{s \to 1} \frac{\sum_{p \in A} p^{-s}}{log(\frac{1}{s-1})} = k.$$

In particular, the set of all prime numbers has density 1, a finite set has density 0, and k is necessarily in $[0, 1]$ if it exists.

Theorem 10.3. *(Dirichlet's theorem). Let a and m be relatively prime integers, $m > 0$. Let P_a be the set of primes q such that $q \equiv a \pmod{m}$. Then P_a has density $\frac{1}{\varphi(m)}$.*

We will prove this for $m = p$ prime. The proof in general is the same but requires the factorization of the Dedekind zeta function of the zeta function of the field of mth roots of 1. In particular, note that P_a is an infinite set!

The proof uses the results on Dirichlet L-functions proved in the previous session. For χ a Dirichlet character mod p, let

$$f_\chi(s) = \sum_{q \neq p} \frac{\chi(q)}{q^s}.$$

This series converges absolutely for $Re(s) > 1$.

Lemma 10.4. *If $\chi = \chi_0$, then f_χ is asympotic to $\log(\frac{1}{s-1})$ when $s \to 1$. If $\chi \neq \chi_0$, then f_χ is bounded when $s \to 1$.*

Proof. For $\chi = \chi_0$, f_χ is just the sum studied before, with the term $q = p$ removed, so the first statement is clear. For the second term, we consider $\log L(s, \chi)$ in a neighborhood of $s = 1$. This is well-defined since each $\chi(q)q^{-s}$ is of absolute value < 1 as

$$\sum_q \log(\frac{1}{1 - \chi(q)q^{-s}}) = \sum_q \sum_k \frac{\chi(q)^k}{kq^{sk}}.$$

As before, we have

$$\log L(s, \chi) = f_\chi(s) + \sum_q \sum_{k \geq 2} \frac{\chi(q)^k}{kq^{sk}}.$$

But since $|\chi(q)| = 1$, the second term is bounded in a neighborhood of 1, by our previous calculation. Thus $f_\chi s$ is asymptotic to $\log L(s, \chi)$ as $s \to 1$, and since $L(s, \chi)$ is holomorphic in a neighborhood of $s = 1$ and doesn't vanish at $s = 1$, this is finite.

□

This is the main application of the result $L(1, \chi) \neq 0$. We can now prove Dirichlet's theorem. Let a be an integer prime to p, or a residue class modulo p. Define

$$g_a(s) = \sum_{q \in P_a} \frac{1}{q^s}.$$

Lemma 10.5. $g_a(s) = \frac{1}{[K:\mathbb{Q}]} \sum_\chi \chi(a)^{-1} f_\chi(s)$.

Proof. This is a simple consequence of the orthogonality relations: if a is an integer, then

$$T(a) = \sum_\chi \chi(a) = 0 \ a \neq 1 \pmod p;$$

$$\sum_\chi \chi(a) = p - 1 = [K : \mathbb{Q}] \ a = 1 \pmod p.$$

This is proved by a familiar argument: if $a \neq 1$ then there is a Dirichlet character χ' with $\chi'(a) \neq 1$. Indeed, if $a = \alpha^b$ for a cyclic generator α of ${\mathbb{F}_p}^\times$, with $b < p-1$, we can define χ' by $\chi'(\alpha) = e^{\frac{2\pi i}{p-1}}$; then $\chi'(a) = e^{\frac{2\pi i b}{p-1}} \neq 1$. Now

$$T(a) \sum_\chi \chi(a) = \sum_\chi \chi'(a)\chi(a) = \chi'(a)T(a)$$

and we conclude as before.

Now consider the right-hand side of the lemma:

$$\sum_\chi \chi(a)^{-1} f_\chi(s) = \sum_\chi \sum_{q \neq p} \frac{\chi(a)^{-1}\chi(q)}{q^s} = \sum_{q \neq p} [\sum_\chi \chi(a)^{-1}\chi(q)] \frac{1}{q^s} = \sum_{q \neq p} T(a^{-1}q)\frac{1}{q^s}.$$

By the orthogonality relation, this is

$$\sum_{q \in P_a} \frac{[K : \mathbb{Q}]}{q^s}.$$

□

Now we complete the proof. This lemma states that $g_a(s)$ is asymptotic to $\frac{1}{[K:\mathbb{Q}]} \sum_\chi \chi(a)^{-1} f_\chi(s)$ when $s \to 1$. But the f_χ are bounded for $\chi \neq \chi_0$, and f_{χ_0} is asymptotic to $\log(\frac{1}{s-1})$. So $g_a(s)$ is asymptotic to $\frac{1}{[K:\mathbb{Q}]} \log(\frac{1}{s-1})$ when $s \to 1$, which is exactly the statement that P_a has density $\frac{1}{p-1}$.

11. Day 11: Prime number theorem

(Hindry/my notes)

- (a) The Riemann zeta function (done)
- (b) The Dedekind zeta function of a number field (done)
- (c) Analytic properties for $s > 1$ (done)
- (d) The formula for the residue (done)

12. Day 12: Newman's proof

- (a) Newman's proof of the prime number theorem
- (b) Non vanishing along the line $Re(s) = 1$.
- (c) Newman's analytic theorem

For any real number x, denote by $\pi(x)$ the number of primes less than x, or equivalently

$$\pi(x) = \sum_{p \leq x} 1.$$

Theorem 12.1. *(Prime Number Theorem)* $\pi(x) \sim \frac{x}{\log(x)}$.

The theorem was conjectured by Gauss and was proved by Hadamard and de la Vallée Poussin in the 1890s. Introduce the notation

$$\theta(x) = \sum_{p \leq x} log(p)$$

where p runs through prime numbers.

Proposition 12.2. *The Prime Number Theorem is equivalent to the assertion*

$$\theta(x) \sim x.$$

This proposition follows from Abel summation and we will prove it if we have time. Meanwhile, here is another statement

Theorem 12.3. *The integral*

$$\int_1^\infty (\theta(t) - t)t^{-2}dt$$

is convergent.

Here is the proof that this convergence implies $\theta(x) \sim x$. Suppose $\limsup \theta(x)/x > 1$. Thus there exists $\epsilon > 0$ and a sequence of x_n tending to ∞ such that $\theta(x_n)/x_n \geq 1+\epsilon$. Consider $t \in [x_n, (1+\epsilon/2)x_n]$. For such t we have (since θ is increasing)

$$(\theta(t) - t)t^{-2} \geq (\theta(x_n) - (1 + \epsilon/2)x_n)/t^2 \geq \frac{\epsilon x_n/2}{t^2} \geq \frac{\epsilon}{2((1 + \epsilon/2)^2 x_n}$$

Thus

$$\int_{x_n}^{(1+\epsilon/2)x_n} (\theta(t) - t)t^{-2}dt \geq [\epsilon/2x_n]\frac{\epsilon}{2((1 + \epsilon/2)^2 x_n} \geq \frac{\epsilon^2}{4(1 + \epsilon/2)^2}$$

which is a constant independent of n; so the integral diverges. If $\liminf \theta(x)/x < 1$ then there exists $\epsilon > 0$ and a sequence of x_n tending

to ∞ such that $\theta(x_n)/x_n \leq 1-\epsilon$. On the interval $[(1-\epsilon/2)x_n, x_n]$ we then have

$$(\theta(t)-t)t^{-2} \leq (\theta(x_n)-(1-\epsilon/2)x_n)/t^2 \leq -\frac{\epsilon x_n/2}{t^2}$$

and in the same way we find an infinite sum of negative terms. So it follows that $\theta(x) \sim x$.

We use Newman's theorem from complex analysis. First note that if $h(t)$ be a bounded, piecewise continuous function, then the integral

$$F(s) = \int_0^\infty h(u)e^{-su}du$$

is convergent and defines a holomorphic function on the half-plane $Re(s) > 0$. (For this we can replace h by 1 and then the result follows by simple integration.

Theorem 12.4. *Let $h(t)$ be a bounded, piecewise continuous function. Suppose the holomorphic function*

$$F(s) = \int_0^\infty h(u)e^{-su}du$$

on the half-plane $Re(s) > 0$ can be analytically continued to $Re(s) \geq 0$. Then the integral converges for $s = 0$ and

$$F(0) = \int_0^\infty h(u)du$$

The theorem includes the claim that $F(s)$ is convergent and holomorphic for $Re(s) > 0$; this is a consequence of the boundedness and piecewise continuity of h, because e^{-su} is holomorphic and rapidly decreasing. The proof is not obvious but uses only classic complex analysis (the residue theorem). We prove it after deriving the convergence of the integral. We apply it to the function

$$\begin{aligned} F(s) = \int_1^\infty (\theta(t)-t)t^{-s-2}dt &= \int_0^\infty \frac{\theta(e^u)-e^u}{e^{u(s+2)}}e^u du \\ &= \int_0^\infty [\theta(e^u)e^{-u}-1]e^{-us}du \end{aligned}$$

We want to apply the analytic theorem to $h(u) = \theta(e^u)e^{-u} - 1$. It is piecewise continuous because $\theta(x)$ is. It is also bounded. Indeed, I claim

Lemma 12.5.

$$\theta(x) \leq (2log(4))x.$$

Here is the (elementary) argument.

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

Hence

$$nlog(4) \geq log\binom{2n}{n} \geq \sum_{n<p\leq 2n} log(p) = \theta(2n) - \theta(n),$$

where the first inequality follows from the previous one, the second one comes from the factorization of the binomial coefficient, and the final equality is the definition. Applying this when $n = 2^m$ we find

$$\theta(2^m) = \sum_{k=0}^{m-1} \theta(2^{k+1}) - \theta(2^k) \leq \sum_{k=0}^{n-1} 2^k log(4) = (2^m - 1)log(4).$$

So if $2^m \leq x < 2^{m+1}$ we have

$$\theta(x) \leq \theta(2^{m+1}) \leq (2^{m+1})log(4) \leq 2log(4)x.$$

So it remains to show that $F(s)$ can be analytically continued to a holomorphic function to the closed half-plane in order to show that $\theta(x) \sim x$. Rewrite $F(s)$ for $Re(s) > 0$ as

$$\begin{aligned} F(s) = \int_1^\infty (\theta(t) - t)t^{-s-2}dt &= \sum_{n=1}^\infty \int_n^{n+1} \theta(t)t^{-s-2}dt - \int_1^\infty t^{-s-1}dt \\ &= \sum_{n=1}^\infty \theta(n)\frac{n^{-s-1} - (n+1)^{-s-1}}{s+1} - \frac{1}{s} \\ &= \frac{1}{s+1}\sum_{n=1}^\infty n^{-s-1}[\theta(n) - \theta(n-1)] - \frac{1}{s} \end{aligned}$$

and since $\theta(n) - \theta(n-1) = 0$ if n is not prime and equals $log(p)$ if $n = p$ is prime, this is

$$\frac{1}{s+1}\sum_p p^{-s-1}log(p) - \frac{1}{s} = \frac{1}{s+1}D(s+1) - \frac{1}{s},$$

where $D(s) = \sum_p p^{-s}log(p)$

On the other hand, for $Re(s) > 1$, $-\zeta'(s)/\zeta(s)$ can be written

$$\begin{aligned} -\frac{d}{ds}log(\zeta(s)) = \sum_p \frac{d}{ds}log(1-p^{-s}) &= \sum_p log(p)p^{-s}\sum_{m\geq 0} p^{-ms} \\ &= \sum_{p,m\geq 1} log(p)p^{-ms} = \sum_p log(p)p^{-s} + \sum_{p,m\geq 2} p^{-ms}. \end{aligned}$$

We have seen that the second term is absolutely convergent and holomorphic for $Re(s) > \frac{1}{2}$, so

$$D(s) = -\zeta'(s)/\zeta(s) + r(s)$$

where $r(s)$ is holomorphic for $Re(s) > \frac{1}{2}$ Thus $F(s) = \frac{1}{s+1}D(s+1) - \frac{1}{s} = -\frac{\zeta'(s+1)}{(s+1)\zeta(s+1)} - \frac{1}{s} + r(s+1)$ where $h(s) = r(s+1)$ is holomorphic for $Re(s) > -\frac{1}{2}$. We need to show that the pole at $s = 0$ of $-\frac{1}{s}$ is precisely compensated by a pole of $-\frac{\zeta'(s+1)}{(s+1)\zeta(s+1)}$ The key theorem is then

Theorem 12.6. *(Hadamard, de la Vallée-Poussin), The function $\zeta(s)$ has no zeroes on the line $Re(s) = 1$.*

Proof. This is based on a simple inequality:

$$4cos(x) + cos(2x) + 3 = 2(1 + cos(x)^2) \geq 0.$$

Now using the Taylor series for $\log(1 - p^{-s})$ for $Re(s) > 1$,

$$\log(\zeta(\sigma + i\tau)) = \sum_{p,m} p^{-m(\sigma+i\tau)}/m$$

and so its real part is

$$\log|\zeta(\sigma + i\tau)| = \sum_{p,m} p^{-m\sigma} cos(m\tau \log p)).$$

Apply this to the product $\zeta(\sigma + i\tau)^4\zeta(\sigma + i2\tau)\zeta(\sigma)^3$:

$$\log|\zeta(\sigma+i\tau)^4\zeta(\sigma+i2\tau)\zeta(\sigma)^3| = \sum_{p,m} p^{-m\sigma}[4cos(m\tau \log(p))+cos(2m\tau \log(p))+3] \geq 0$$

by the simple inequality.

Now if $\sigma > 1$, this implies that

$$|\zeta(\sigma + i\tau)^4\zeta(\sigma + i2\tau)\zeta(\sigma)^3| \geq 1.$$

Suppose $\zeta(s)$ had a zero of order k at $1 + i\tau$ and order ℓ at $1 + 2i\tau$. Then (as $\sigma \to 1^+$)

$$|\zeta(\sigma + i\tau)|^4 \sim a(\sigma - 1)^{4k};\ |\zeta(\sigma + 2i\tau)| \sim b(\sigma - 1)^{\ell};\ |\zeta(\sigma)| \sim (\sigma - 1)^{-1}.$$

Thus $|\zeta(\sigma + i\tau)^4\zeta(\sigma + i2\tau)\zeta(\sigma)^3| \sim c(\sigma - 1)^{4k+\ell-3}$; but since we know it is ≥ 1 for $\sigma > 1$ it follows that $4k + \ell - 3 \leq 0$ which is impossible if $k > 0$. □

Corollary 12.7. *The function $G(s) = F(s-1) - r(s) = -\frac{\zeta'(s)}{(s)\zeta(s)} - \frac{1}{s-1}$ on $Re(s) > 1$ extends holomorphically to $Re(s) \geq 1$.*

This completes the proof that $\theta(x) \sim x$, assuming Newman's analytic theorem, and thus the Prime Number Theorem, assuming the relation between $\theta(x)$ and $\pi(x)$.

Proof. The previous theorem shows that ζ'/ζ is holomorphic on the line $Re(s) = 1$ except at $s = 1$ where ζ has a pole. So it remains to understand $G(s)$ in a neighborhood of $s = 1$. Now $\zeta(s)$ has a simple pole at $s = 1$, $\zeta(s) = \frac{1}{s-1}f(s)$ (for f is holomorphic and non-vanishing near $s = 1$) so

$$\zeta'(s)/\zeta(s) = f'/f - dlog(s-1) = f'/f - 1/(s-1) = -1/(s-1) + g(s)$$

where g is holomorphic in a neighborhood of 1. Thus $G(s) = -g(s)$ in a neighborhood of $s = 1$ and we are done. □

Proof of the analytic theorem. For any $T >> 0$ let $F_T(s) = \int_0^T h(t)e^{-st}dt$. Since h is bounded, say $|h(t)| \leq M$, and piecewise continuous, this is a (finite) sum of entire functions. We need to show

- $\lim_{T \to \infty} F_T(0)$ exists
- $\lim_{T \to \infty} F_T(0) = F(0)$.

Consider for $R >> 0$ the contour $\gamma = \gamma(R, \delta)$ which bounds the region

$$S = \{s \in \mathbb{C} \mid Re(z) > -\delta, |s| < R\}.$$

The hypothesis that F extends holomorphically to $Re(s) \geq 0$ implies that, for fixed R, there exists $\delta > 0$ such that $F(s)$ is analytic on (the closure of) S.

Now let

$$G_T(s) = (F(s) - F_T(s))e^{sT}(1 + s^2/R^2).$$

Then $G_T(0) = F(0) - F_T(0)$. It therefore suffices to prove

- $\lim_{T \to \infty} G_T(0)$ exists
- $\lim_{T \to \infty} G_T(0) = 0$.

We use the residue theorem:

$$G_T(0) = F(0) - F_T(0) = \frac{1}{2\pi i}\int_\gamma G_T(s)\frac{ds}{s} = \frac{1}{2\pi i}\int_\gamma (F(s) - F_T(s))e^{sT}(1 + s^2/R^2)\frac{ds}{s}.$$

Write $\gamma = \gamma_1 \cup \gamma_2$ where $\gamma_1 \subset \{Re(s) > 0\}$ and $\gamma_2 \subset \{Re(s) < 0\}$. The integral over γ_1 is for $|s| = R$, i.e. $s = Re^{i\theta}$. Then

$$|e^{sT}(1 + s^2/R^2)\frac{1}{s}| = e^{Re(s)T}|\frac{Re^{-i\theta}}{R^2} + \frac{Re^{i\theta}}{R^2}| = e^{Re(s)T}\frac{2Re(s)}{R^2}.$$

Moreover

$$|F(s) - F_T(s)| = |\int_T^\infty h(t)e^{-st}dt| \leq M\int_T^\infty e^{-Re(s)t}dt = \frac{Me^{-Re(s)T}}{Re(s)}.$$

So the integral over γ_1 is bounded: since the length of γ_1 is πR we find

$$|\frac{1}{2\pi i}\int_{\gamma_1} (F(s) - F_T(s))e^{sT}(1 + s^2/R^2)\frac{ds}{s}| \leq \frac{Me^{-Re(s)T}}{Re(s)} \cdot e^{Re(s)T}\frac{2Re(s)}{R^2} \cdot (\pi R) \leq \frac{M}{R}.$$

So as $R \to \infty$, this part of the integral is arbitrarily small.

Now we bound the integral over γ_2. Write

$$\frac{1}{2\pi i}\int_{\gamma_2} G_T(s)\frac{ds}{s} = I_1 - I_2$$

where

$$I_1 = I_1(T) = \frac{1}{2\pi i}\int_{\gamma_2} F(s)e^{sT}(1+s^2/R^2)\frac{ds}{s},\ I_2 = \frac{1}{2\pi i}\int_{\gamma_2} F_T(s)e^{sT}(1+s^2/R^2)\frac{ds}{s}.$$

First note that $F_T(s)$ is entire, so we can use the Cauchy integral theorem to replace the contour γ_2 by the arc of the circle of radius R in $Re(s) < 0$. There is the bound

$$|F_T(s)| = |\int_0^T h(t)e^{-st}dt| \leq M\int_0^T e^{-Re(s)t}dt = \frac{M[1-e^{-Re(s)T}]}{Re(s)}.$$

Thus

$$|I_2| \leq \frac{1}{2\pi}(\pi RM)\cdot\frac{M|1-e^{-Re(s)T}|}{Re(s)}e^{Re(s)T}\frac{2Re(s)}{R^2} = \frac{M}{R}\cdot|e^{Re(s)T}-1|$$

but the term $|e^{Re(s)T}-1|$ is ≤ 1 on $Re(s) < 0$. So $|I_2| \leq \frac{M}{R}$ is negligeable when $R \to \infty$. As for I_1, let K be a compact subset $K \subset \{Re(s) < 0\}$ where F is holomorphic, and say $|F(s)| \leq M_K$ is bounded on K. Say $K \subset \{Re(s) \leq -\eta\}$. Then for $Re(s) \in K$, $\lim_{T\to\infty}|F(s)e^{sT}(1+\frac{s^2}{R^2})\frac{1}{s}|$ is dominated by $e^{-\eta T}$ which tends uniformly to 0. Thus

$$\lim_{T\to\infty} I_1(T) = 0.$$

It follows that

$$|F(0) - F_T(0)| \leq \frac{2M}{R} + \epsilon(T)$$

for some T that tends to 0 (depending on R). This suffices to show that $\lim F_T(0) = F(0)$.

13. Odds and ends

13.1. **The function θ.** Here are some elementary considerations about the function θ.

$$\theta(x) = \sum_{p\leq x}\log p \leq \log(x)\sum_{p\leq x} 1 = \log(x)\pi(x),$$

so

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

On the other hand, if $2 \leq y < x$,

$$\pi(x) - \pi(y) = \sum_{y<p\leq x} 1 \leq \frac{1}{\log y} \sum_{y<p\leq x} \log p = \frac{1}{\log y}[\theta(x) - \theta(y)].$$

Thus

$$\pi(x) \leq \frac{\theta(x)}{\log y} + \pi(y) \leq \frac{\theta(x)}{\log y} + y.$$

Choose $y = \frac{x}{(\log x)^2}$. Then

$$\frac{\theta(x)}{\log x} \leq \pi(x) \leq \frac{\theta(x)}{\log x - 2\log\log x} + \frac{x}{(\log x)^2}.$$

Thus since we know that $\theta x \sim x$ we have

$$\frac{\theta(x)}{\log x} \sim \frac{x}{\log x};$$

$$\frac{\theta(x)}{\log x - 2\log\log x} + \frac{x}{(\log x)^2} \sim \frac{x}{\log x - 2\log\log x} + \frac{x}{(\log x)^2} \sim \frac{x}{\log x}.$$

which together imply

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

13.2. **Finiteness of the class number.** This is an argument in *geometry of numbers*.

Note first that

Theorem 13.1. *Let K be a number field. Let $S = \sigma_1, \ldots, \sigma_{r_1}$ be the set of distinct real embeddings, $T = \tau_1, \ldots, \tau_{r_2}$ (half) the set of distinct complex embeddings of K, so that $\tau_i \neq \bar{\tau}_j$ for all i, j. Thus $|S| = r_1$, $|T| = r_2$, and $r_1 + 2r_2 = n = [K : \mathbb{Q}]$. We embed $\mathcal{O}_K$ in $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ by*

$$\Phi : x \mapsto (\sigma_1(x), \ldots, \sigma_{r_1}(x), \tau_1(x), \ldots, \tau_{r_2}(x)).$$

Then the image of $\mathcal{O}_K$ is discrete in the n-dimensional real vector space $V = \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$.

Proof. Let $K(t) \subset \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ be the (euclidean) set defined by $|x_i| \leq t$, $|z_j| \leq t$ for $i = 1, \ldots, r_1$, $j = 1, \ldots, r_2$. We show that $\mathcal{O}_K \cap K(t)$ is finite. Indeed, every element $x \in \mathcal{O}_K \cap K(t)$ has the property that $|\sigma_i(x)|, |\tau_j(x)| \leq t$. This means the coefficients of its minimal polynomial are all bounded in terms of t, but the coefficients are integers, so (as we have already seen) this means they all belong to a finite set. □

It follows that $\mathcal{O}_K \subset V$ is a lattice, in other words $\mathcal{O}_K$ contains an $\mathbb{R}$-basis of V. The parallelipiped defined by such a basis is thus a fundamental domain for $\mathcal{O}_K \subset V$; we let v_K denote its volume, which is $|\det(e_1, \ldots, e_n)|$ where $\{e_1, \ldots, e_n\}$ is a basis for the lattice (it's actually the square root of the absolute value of the discriminant).

Theorem 13.2. *(Minkowski) Let $\Omega \subset \mathbb{R}^n$ be a compact, convex, and symmetric set (symmetric means that if $x \in \Omega$ then $-x \in \Omega$). Let $\Lambda \subset V$ be a lattice. Assume $vol(\Omega) > 2^n |\det(\Lambda)|$. Then $\Omega \cap \Lambda$ has a non-zero element.*

Proof. Suppose $\Lambda = A(\mathbb{Z}^n)$ for some matrix A, so that $|\det(\Lambda)| = |\det(A)|$. The image of Ω under A^{-1} is again compact, convex, and symmetric; so we may as well assume $\Lambda = \mathbb{Z}^n$. Let $C = [0, 1[^n$. Let $T \subset \mathbb{R}^n$ and suppose $(T + \lambda) \cap (T + \mu) = \emptyset$ for $\lambda \neq \mu \in \mathbb{Z}^n$. Now

$$T = \cup_{\lambda \in \mathbb{Z}^n} (T \cap (C + \lambda))$$

(disjoint union). So

$$vol(T) = \sum_{\lambda \in \mathbb{Z}^n} vol(T \cap (C + \lambda)) = \sum_{\lambda \in \mathbb{Z}^n} vol((T - \lambda) \cap C);$$

but by the hypothesis on T, this is

$$vol((\cup_{\lambda \in \mathbb{Z}^n} vol((T - \lambda) \cap C) \leq vol(C) = 1.$$

Thus if $vol(T) > 1$ there exists $x \in T \cap (T + \lambda)$ with $0 \neq \lambda \in \mathbb{Z}^n$ and we have $x = t_1 = t_2 + \lambda$ so $\lambda = t_1 - t_2 \in T - T$.

Now we let $T = 1/2\Omega = \{\frac{x}{2}, x \in \Omega\}$. So $K = T - T$ and $vol(T) = 2^{-n}vol(\Omega) > 1$ by hypothesis. Then we have seen that $\Omega \cap \mathbb{Z}^n - \{0\} \neq \emptyset$. □

Lemma 13.3. *There exists $c_1 = c_1(K) > 0$ such that, if $I \subset \mathcal{O}$ is a non-zero ideal, then there is a non-zero $\alpha \in I$ such that $|N_{K/\mathbb{Q}}(\alpha)| \leq c_1 N(I)$.*

Proof. Let $\Omega_t \subset V$ be the compact, convex, symmetric subset defined above, a product of r_1 intervals and r_2 circular disks, of volume $2^{r_1}\pi^{r_2}t^n$. The lattice $\Phi(I) \subset V$ has volume $v_K N(I)$. Thus, if $2^{r_1}\pi^{r_2}t^n > 2^n v_K N(I)$ we have $\Phi(I) \cap \Omega_t \neq \{0\}$. We take any t satisfying this inequality (actually we can take t defined by the *equality* using compactness); there is $0 \neq \alpha \in I$ with $\Phi(\alpha) \in \Omega_t$, so $|N_{K/\mathbb{Q}}(\alpha)| \leq t^n \leq c_1 N(I)$ where c_1 is any constant larger than $(4/\pi)^{r_2}v_K$. □

Theorem 13.4. *The set of ideal classes of $\mathcal{O}_K$ is finite.*

Proof. Let c_1 be as above and let $m = [c_1]!$. For I as in the lemma, let $0 \neq \alpha \in I$ as in the lemma; so $|I/(\alpha)| \leq c_1$. It follows that $mI \subset \alpha\mathcal{O}_K$. Let $J = \frac{m}{\alpha}I$. Then $J \subset \mathcal{O}_K$ and J and I are in the same ideal class; indeed $\alpha J = mI$. Moreover, $\alpha \in I$, so $m\alpha \in \alpha J$ which means that $m \in J$. So any ideal is equivalent to one that contains m. But the set of ideals containing m is finite. □

We can do better than that. Let c_1 be the constant in the Lemma.

Corollary 13.5. *Every ideal class in $\mathcal{O}_K$ contains an element of norm $< c_1$.*

Proof. Let κ be an ideal class, and let I be an ideal in the inverse class. The lemma states that there is a non-zero $\alpha \in I$ such that $|N_{K/\mathbb{Q}}(\alpha)| \leq c_1 N(I)$. Now $(\alpha) \subset I$ implies that there is an ideal $J \subset \mathcal{O}$ such that $(\alpha) = IJ$; in other words, J is in the class κ and is integral. But now $N(IJ) = N_{K/\mathbb{Q}}(\alpha) < c_1 N(I)$, thus $N(J) < c_1$. □

Recall that $c_1 = (4/\pi)^{r_2}v_K$, where v_K is the volume of the lattice $\Phi(\mathcal{O}_K) \subset \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$. If $(\alpha_1, \ldots, \alpha_n)$ is a $\mathbb{Z}$-basis of $\mathcal{O}_K$, this volume is $|\det(A)|$ where A is the real matrix with rows (or columns)

$$(\sigma_1(\alpha_j) \ldots \sigma_{r_1}(\alpha_j) Re(\tau_1(\alpha_j)) Im(\tau_1(\alpha_j)) \ldots Re(\tau_{r_2}(\alpha_j)) Im(\tau_r(\alpha_j))$$

for $j = 1, \ldots, n$. If we consider instead the complex matrix C where the σ's are unchanged but the $Re(\tau_i)$ and $Im(\tau_i)$ are replaced by τ_i, $\bar{\tau}_i$ (complex conjugate), this corresponds to multiplication of A by a block matrix $\begin{pmatrix} I_{r_1} & 0 \\ 0 & B \end{pmatrix}$ where B consists of diagonal 2×2 blocks all of which

are given by $\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}$. So $|\det(C)| = |\det(B)||\det(A)| = |-2i|^{r_2} v_K$; $v_K = 2^{-r_2}|\det(C)|$.

However, if we let $(s_i) = (\sigma_i, \tau_j, \bar{\tau}_j)$ be uniform notation, then $|\det(C)| = \sqrt{|(\det(s_i(\alpha_j)))^2|} = \sqrt{|\Delta_K|}$ where Δ_K is the *discriminant* of $\mathcal{O}_K$. Indeed,

Lemma 13.6. *Let $L \subset K$ be any lattice, and define*

$$\Delta_L = (\det(s_i(\alpha_j)))^2.$$

Then
(i) $\Delta_L \in \mathbb{Q}^\times$;
(ii) If $L \subset \mathcal{O}_K$, then $\Delta_L \in \mathbb{Z}$.
(iii) If $L' \subset L$, then $|\Delta_{L'}| = [L : L']^2 |\Delta_L|$.

We admit the lemma for the moment, and write Δ_K instead of $\Delta_{\mathcal{O}_K}$.

Corollary 13.7. *Every ideal class in $\mathcal{O}_K$ contains an integral ideal of norm at most $(2/\pi)^{r_2}\sqrt{|\Delta_K|}$.*

There is a better bound due to Minkowski:

Corollary 13.8. *Every ideal class in $\mathcal{O}_K$ contains an integral ideal of norm at most $(\frac{4}{\pi})^{r_2}\frac{n!}{n^n}\sqrt{|\Delta_K|}$.*

This involves paying closer attention to the calculation above, but the one we have defined will suffice. However, since no ideal has norm less than 1, Minkowski's bound implies in particular that $(4\pi)^{r_2}\frac{n!}{n^n}\sqrt{|\Delta_K|} \geq 1$, in other words that $\sqrt{|\Delta_K|} \geq (\pi/4)^{r_2} \cdot \frac{n^n}{n!} > 1$ (the other bound is > 1 provided $r_2 > 0$).

Now we prove the lemma about Δ_L. First, we let $K' \supset K$ be a Galois extension of $\mathbb{Q}$ containing K and contained in $\mathbb{C}$, so that K' contains the images of all complex embeddings of K. We check that Δ_L is invariant under $Gal(K'/\mathbb{Q})$. If $\sigma \in Gal(K'/\mathbb{Q})$, then the n-tuple $(\sigma \circ s_i)$ is a permutation of (s_i). (If $\sigma \circ s_i = \sigma \circ s_i$ then $s_i = s_j$.) So σ permutes the rows (or columns) of the matrix C and therefore multiplies its determinant by ± 1; but it multiplies the square of the determinant by 1. This proves (i). If $L \subset \mathcal{O}_K$, then $\Delta_L \in \mathbb{Q} \cap \mathcal{O}_K = \mathbb{Z}$. Finally, the (absolute value of) determinant of a basis of a sublattice is the index of the sublattice multiplied by the covolume of the lattice, and the index is squared in the square of the determinant. (We should work with $\mathbb{R}^n$ and replace the τ_i by their real and imaginary part, but this just amounts to multiplying by 2^{-r_2} again.)

If $K = \mathbb{Q}(\sqrt{d})$ then Δ_K is given by the square of the determinant of the matrix $\begin{pmatrix} 1 & 1 \\ \sqrt{d} & -\sqrt{d} \end{pmatrix}$ if $d \equiv 2, 3 \pmod 4$ $\begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{d}}{2} & \frac{1-\sqrt{d}}{2} \end{pmatrix}$ if $d \equiv 1 \pmod 4$. Thus $\Delta_K = 4d$ or $\Delta_K = d$ in the two cases. Note that the divisors of Δ_K are exactly the primes that ramify in K!

We note the following addendum to quadratic reciprocity.

Lemma 13.9. *The prime 2 is ramified in $\mathbb{Q}(\sqrt{d})$ if $d \equiv 2, 3 \pmod 4$, it is split if $d \equiv 1 \pmod 8$, and stays prime if $d \equiv 5 \pmod 8$.*

Proof. Consider the ring $\mathcal{O}_K/(2) = \mathbb{F}_2[X]/(f)$, $f(X) = X^2 - X - \frac{d-1}{4}$. We need to know when f has a root in $\mathbb{F}$. Obviously it has no double root, so this is only possible if $f(0) = 0 \pmod 2$, i.e. if and only if $\frac{d-1}{4} \equiv 0 \pmod 2$. □

Examples. Let $K = \mathbb{Q}(\sqrt{-19})$. Then $|\Delta_K| = 19$, and every ideal class contains an ideal of norm $< 2\sqrt{19}/3 = 2.906\cdots < 3$. Now 2 stays prime in $\mathcal{O}_K$ by the lemma, so any prime dividing 2 has norm 4, and so any ideal has norm at least 3. Thus $\mathbb{Q}(\sqrt{-19})$ has class number 1 and its ring of integers is principal.

14. The discriminant

We need to prove the following fact.

Proposition 14.1. *Let $\Gamma \subset \mathbb{R}^n$ be a discrete subgroup. Then Γ has a basis over $\mathbb{Z}$ formed of r linearly independent vectors over $\mathbb{R}$, for some $r \leq n$, with equality if and only if Γ has a compact fundamental domain in $\mathbb{R}^n$.*

Proof. Let $e_1, \ldots, e_r \subset \Gamma$ be a maximal set of $\mathbb{R}$-linearly independent elements. If we show that the subgroup Λ generated by the e_i is of finite index in Γ then we know that the rank of Γ is r, which is less than or equal to $\dim \mathbb{R}^n = n$.

In any case, every element of Γ is a linear combination, with $\mathbb{R}$-coefficients, of the e_i. Say $y \in \Gamma$, $y = \sum_{i=1}^r x_i e_i$ for some $x_i \in \mathbb{R}$. We want to show there is a finite set of possibilities for y modulo Λ. Subtracting an element of Λ we may assume

$$y \in C := \{\sum_{i=1}^r a_i e_i, 0 \leq a_i \leq 1\}.$$

Now C is compact, and a discrete subset of a compact space C is closed (because it is equal to all its limit points) and therefore compact; but it can be covered by a collection of open sets with one member in each

set, and this has a finite subcover, so the set is itself finite. This implies that Γ/Λ is finite, and completes the proof. Obviously if $r < n$ we can find an independent e_{r+1} and then the fundamental domain will have to include all multiples of e_{r+1}, hence is not compact. On the other hand, if $r = n$ we have seen that the fundamental domain is compact. □

It follows from Lemma 13.6 that

Corollary 14.2. *Let K be a number field of degree n and let $R \subset \mathcal{O}_K$ be a subring of rank n over $\mathbb{Z}$. Suppose $|\Delta_R| = |\Delta_K|$. Then $R = \mathcal{O}_K$. Moreover, if Δ_R is square-free then $R = \mathcal{O}_K$.*

Proof. Let $d = [\mathcal{O}_K : R]$. Then we have seen that $|\Delta_R| = d^2|\Delta_K|$, which implies the corollary. □

We can use the discriminant to determine the ring $\mathcal{O}_K$ even when the corollary does not apply directly. First we note that the discriminant can be defined when the inclusion $\mathbb{Z} \subset \mathcal{O}$ is replaced by $\mathbb{Z}_S \subset \mathcal{O}_S$ for any multiplicative subset $S \subset \mathbb{Z}$. So if p is a prime, we can consider $\mathbb{Z}_{(p)} \subset \mathcal{O}_{(p)}$. Since $\mathcal{O}_{(p)}$ has only finitely many prime ideals, it is a PID. Notation is self-explanatory.

Lemma 14.3. *The ideal $\Delta_{\mathcal{O}_{(p)}/\mathbb{Z}_{(p)}}$ is the localization at p of Δ_K.*

Proof. If $\{\alpha_1, \ldots, \alpha_n\}$ is a $\mathbb{Z}$-basis for $\mathcal{O}$ it is also a $\mathbb{Z}_{(p)}$ basis for $\mathcal{O}_{(p)}$. So the assertion is clear. □

The prime decomposition of p in $\mathcal{O}$ can be determined after localization at p: if $p\mathcal{O} = \prod_i \mathfrak{p}_i^{e_i}$ then

$$p\mathcal{O}_{(p)} = \prod_i (\mathfrak{p}_i \mathcal{O}_{(p)})^{e_i} = \prod_i (\mathfrak{p}_{i,(p)})^{e_i}$$

(the second is alternative notation).

Lemma 14.4. *Let $\alpha \in \mathcal{O}$ be an element of degree n over $\mathbb{Q}$ and let p be a prime such that $\mathbb{Z}[\alpha]_{(p)} = \mathcal{O}_{(p)}$. Then p ramifies in K if and only if p divides Δ_K. (See Hindry, Exercise 6.14)*

Proof. We have seen that p ramifies in K if and only if $p\mathcal{O}_{(p)}$ is divisible by a prime of $\mathcal{O}_{(p)}$ with multiplicity greater than 1. This is true if and only if $\mathcal{O}_{(p)}/(p) = \mathbb{Z}_{(p)}[\alpha]/(p)$ has nilpotents, as we have already seen; and this is true if and only if the minimal polynomial f of α has a multiple root mod p. But this holds if and only if f and f' have a common root, say α', mod p. More precisely, we have

$$\prod_i \mathcal{O}_{(p)}/(\mathfrak{p}_{i,(p)})^{e_i} = \mathcal{O}_{(p)}/(p) = \mathbb{Z}_{(p)}[\alpha]/(p) = \mathbb{F}_p[X]/(f) = \prod_i \mathbb{F}_p[X]/((\phi_i)^{e_i})$$

and we can match $\mathfrak{p}_{i,(p)}$ with the polynomial ϕ_i. Say $\phi = \phi_i$ for some i with ramification index $e = e_i > 1$. In the Galois closure K^+ of K, f has a root β that reduces mod p to a root of ϕ, and there is an element $s \in Gal(K^+/\mathbb{Q})$ such that $s(\alpha) = \beta$, and $f'(\beta)$ is divisible by a prime of $s(\mathcal{O}_{(p)} = \mathbb{Z}_{(p)}[\beta]$ dividing p. Thus p ramifies in K if and only if

$$N_{s(K)/\mathbb{Q}}(f'(\beta)) = N_{K/\mathbb{Q}}(f'(\alpha)) \in p\mathbb{Z}_{(p)}.$$

which is true if and only if p divides $(N_{K/\mathbb{Q}}(f'(\alpha))) = (\Delta_{\mathcal{O}_{(p)}/\mathbb{Z}_{(p)}})$. $\square$

Now for any $\alpha \in \mathcal{O}$ of degree n, $[\mathcal{O} : \mathbb{Z}[\alpha]] = N$ is finite, so for any prime not dividing n, $\mathbb{Z}[\alpha]_{(p)} = \mathcal{O}_{(p)}$. In this way we see that the set of ramified primes is finite.

Examples. 1. Let p be a prime, $r \geq 1$, and consider $\zeta_{p^r} = e^{2\pi i/p^r}$, a primitive p^r-th root of 1. Let $K = \mathbb{Q}(\zeta_{p^r})$, the cyclotomic field of degree $\phi(p^r)$ over $\mathbb{Q}$. Let $R = \mathbb{Z}[\zeta_{p^r}]$. We have seen in the homework that

$$\Delta_R = (p^{p^{r-1}(pr-r-1)})$$

is divisible by no prime other than p. Since there is an integer d such that $\Delta_R = d^2\Delta_K$, it follows that p is the only prime that ramifies in K.

2. On the other hand, suppose $[K : \mathbb{Q}] = n$, p is a prime, and $\alpha \in \mathcal{O}_K$ is an element such that $(\alpha^n) = p\mathcal{O}_K$ as ideals. Then p is ramified in K, and indeed (α) is a prime ideal. The proof is obvious: let $(\alpha) = \prod_i \mathfrak{p}_i^{e_i}$ be the prime factorization of the principal ideal (α). Then

$$(p\mathcal{O}_K) = \prod_{i=1}^{g} \mathfrak{p}_i^{ne_i}$$

and if f_i is the residue degree of $\mathfrak{p}_i$ we have

$$\sum ne_if_i = n; \sum_{i=1}^{g} e_if_i = 1$$

which means that $g = 1$ and $e_1 = f_1 = 1$, $(\alpha) = \mathfrak{p}_1$. For example, suppose $K = \mathbb{Q}(\sqrt[5]{3})$. Take $p = 3$, $\alpha = \sqrt[5]{3}$. Then $n = 5$ and $(\alpha^5) = (3)$, so 3 is ramified in K and (α) is a prime ideal. Moreover, letting $f = X^5 - 3$ be the minimal polynomial of α, we see that

$$N_{K/\mathbb{Q}}(f'(\alpha)) = 5^5 \cdot \prod_{i=0}^{4} (\zeta_5^i)(\alpha)^4 = 5^5 \cdot 3^4.$$

Thus if $R = \mathbb{Z}[\alpha]$, $\Delta_R = \pm 5^5 \cdot 3^4$ and the only primes that can ramify in K are 3 and 5. We have seen that 3 ramifies, and 5 has to divide Δ_K because Δ_K/Δ_R is a square. We have not yet shown that every

prime that divides Δ_K necessarily ramifies. For this we would have to determine the localization of $\mathcal{O}$ at 5.

Integers in a cyclotomic field. We consider the field $K = \mathbb{Q}(\zeta_{p^r})$ for $r \geq 1$. Write $\zeta = \zeta_{p^r}$ (any primitive root of $\Phi_{p^r}(X)$. We show that $[K : \mathbb{Q}] = \phi(p^r)$. In any case, it is generated by a root of a polynomial of degree $\phi(p^r)$, so the degree n is at most $\phi(p^r)$.

Proposition 14.5. *Let ζ' be another primitive p^rth root of 1. Then $\frac{1-\zeta}{1-\zeta'}$ is a unit in $\mathcal{O}_K$.*

Proof. It's the same as for the case $r = 1$. □

It follows that

$$\prod_{\zeta'}(1-\zeta') = u(1-\zeta)^{\phi(p^r)}$$

for some unit u But the left-hand side is $\Phi_{p^r}(1) = p$. So $(p)\mathcal{O}_K = (1-\zeta)^{\phi(p^r)}$. It follows that $(1-\zeta)$ is a prime ideal, say $\mathfrak{p}$, and that p is totally ramified; in particular, $\mathcal{O}/\mathfrak{p} = \mathbb{F}_p$.

Lemma 14.6 (Nakayama's Lemma)**.** *Let R be a local ring with maximal ideal $\mathfrak{p}$, M a finitely generated R module. Suppose $\mathfrak{p}M = M$. Then $M = 0$. In particular, if $N \subset M$ is a pair of finitely generated R-modules such that $N/\mathfrak{p}N = M/\mathfrak{p}M$ then $N = M$.*

Proof. Say M is generated by m elements $e_1, \ldots, e_m$, and assume the statement for modules with fewer generators. There exists an expression

$$e_1 = \sum_i a_i e_i$$

with $a_i \in \mathfrak{p}$. Thus

$$(1-a_1)e_1 = \sum_{i>1} a_i e_i$$

But $1 - a_1 \notin \mathfrak{p}$, so it is a unit; thus M has fewer than m generators, and we conclude. For the final statement, we apply the first statement to M/N. Every element $m \in M$ can be written $am + n$ for some $a \in \mathfrak{p}$ and $n \in N$, in other words $M/N = \mathfrak{p}M/N$. □

Proposition 14.7. *Let R denote the localization of $\mathcal{O}$ at p. Then $R = \mathbb{Z}_{(p)}[\zeta]$.*

Proof. Let $n = \phi(p^r)$, $\alpha = 1 - \zeta$. The ring R has a unique prime ideal (α) dividing p. Consider $S = \mathbb{Z}_{(p)}[\zeta]$. It is a $\mathbb{Z}_{(p)}$-submodule of R. We are going to apply Nakayama's Lemma. It suffices to show that

the map $S \to R/pR$ is surjective. But $pR = \mathfrak{p}^n R$. Thus R/pR has a composition series

$$R/pR \supset \mathfrak{p}/pR \supset \mathfrak{p}^2/pR \cdots \supset \mathfrak{p}^{n-1}/pR.$$

As in the proof that $\sum e_i f_i = n$, we know that each $\mathfrak{p}^i/\mathfrak{p}^{i+1}$ is isomorphic to $\mathbb{F}_p$. It follows that as a vector space,

$$R/pR = \mathbb{F}_p + \mathbb{F}_p\alpha + \mathbb{F}_p\alpha^2 + \cdots + \mathbb{F}_p\alpha^{n-1}$$

which is the image of S. □

Now p is the only prime that ramifies in K. So to show that $\mathcal{O}_K = \mathbb{Z}[\zeta]$ it suffices to show that they have the same discriminant after localization at p. But we have just seen that the rings become equal after localization at p. Thus

Proposition 14.8. $\mathcal{O}_K = \mathbb{Z}[\zeta]$.

Theorem 14.9. *Let n be any integer, $K = \mathbb{Q}(\zeta_n)$. Then $[K : \mathbb{Q}] = \phi(n)$.*

Proof. This is an induction on the number of distinct primes dividing n. Suppose $n = n'p^r$ where $(p, n') = 1$. So we have $K \supset K' = \mathbb{Q}(\zeta_{n'})$ and $K \supset L = \mathbb{Q}(\zeta_{p^r})$. Because p is totally ramified in L, it is also totally ramified in any intermediate extension. But p is unramified in K'. So $K' \cap L = \mathbb{Q}$. On the other hand, $K = K'L$, and one knows by Galois theory that the map $Gal(K/\mathbb{Q}) \to Gal(K'/\mathbb{Q}) \times Gal(L/\mathbb{Q})$ has image the set of $\{x, y\} \in Gal(K'/\mathbb{Q}) \times Gal(L/\mathbb{Q})$ such that x and y have the same restriction to $K' \cap L$. In other words, the map is surjective. It follows that $[K : \mathbb{Q}] = [K' : \mathbb{Q}][L : \mathbb{Q}]$. □

One also proves that $\mathcal{O}_K = \mathbb{Z}[\zeta_n]$, but for this one needs the full theory of the discriminant.