Knots and Primes

Summer 2012 Tutorial

Chao Li
chaoli@math.harvard.edu

Charmaine Sia
sia@math.harvard.edu

Course notes for an undergraduate tutorial taught at Harvard University during summer 2012. Figures borrowed from *Knots and Primes: An Introduction to Arithmetic Topology* by Masanori Morishita (course reference text), *Algebraic Topology* by Allen Hatcher, *The Knot Book* by Colin Adams, the Knot Atlas (http://katlas.math.toronto.edu/) and Wikipedia.

Lecture 1. (July 2, 2012)

1. Analogy between knots and primes

Knots and primes are the basic objects of study in knot theory and number theory respectively. Surprisingly, these two seemingly unrelated concepts have a deep analogy discovered by Barry Mazur in the 1960s while studying the Alexander polynomial, which initiated the study of what is now known as arithmetic topology. As motivation for this analogy, we first consider the correspondence between commutative rings and spaces in algebraic geometry.

1.1. **Commutative rings and spaces.**

Example 1.1. Consider the polynomial ring $\mathbb{C}[t]$: it has transcendence degree one over the field $\mathbb{C}$, which we think of as one degree of freedom. We represent it by a complex line, denoted by $\operatorname{Spec}\mathbb{C}[t]$. Hilbert's Nullstellensatz tells us that there is a bijective correspondence between elements $a \in \mathbb{C}$ and maximal ideals $(t-a)$ of functions that vanish at a. Since every nonzero prime ideal of $\mathbb{C}[t]$ is a maximal ideal, this justifies us labeling the complex line as $\operatorname{Spec}\mathbb{C}[t]$, the set of prime ideals of the ring $\mathbb{C}[t]$. (The zero ideal corresponds to the generic point, which one should think of as the entire line.) The inclusion of the point representing $(t-a)$ into the complex line corresponds to the quotient map $\mathbb{C}[t] \twoheadrightarrow \mathbb{C}[t]/(t-a) \cong \mathbb{C}$ in the opposite direction and is denoted by a map $\operatorname{Spec}\mathbb{C} \hookrightarrow \operatorname{Spec}\mathbb{C}[t]$.

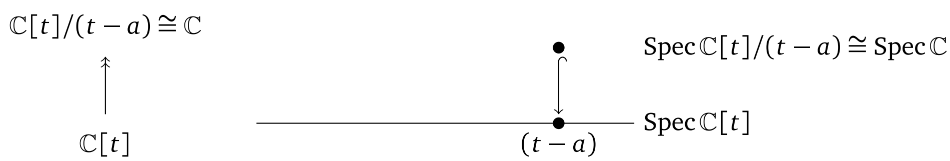


Figure 1. Inclusion $\operatorname{Spec}\mathbb{C} \hookrightarrow \operatorname{Spec}\mathbb{C}[t]$

Example 1.2. There is a similar story for the ring of integers $\mathbb{Z}$. Above, we used transcendence degree as a measure of dimension; however, we could equally well have used Krull dimension, that is, the supremum of all integers n such that there is a strict chain of prime ideals $\mathfrak{p}_0 \subset \mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_n$, as the Krull dimension of a domain finitely generated over a field is equal to its transcendence degree. Krull dimension turns out to be the "correct" notion of dimension in algebraic geometry, as it is defined for all commutative rings. The Krull dimension of $\mathbb{Z}$ is one, so once again we represent it by a line, denoted by $\operatorname{Spec}\mathbb{Z}$; its points are prime ideals (p) where p is a prime number. As before, the inclusion of the point representing (p) into the complex line corresponds to the quotient map $\mathbb{Z} \twoheadrightarrow \mathbb{Z}/(p) \cong \mathbb{F}_p$ in the opposite direction and is denoted by a map $\operatorname{Spec}\mathbb{F}_p \hookrightarrow \operatorname{Spec}\mathbb{Z}$.

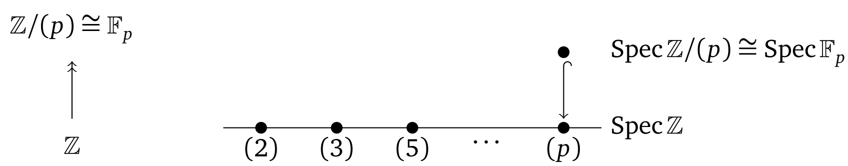


Figure 2. Inclusion $\operatorname{Spec}\mathbb{F}_p \hookrightarrow \operatorname{Spec}\mathbb{Z}$

1.2. **Knots and primes.** The key idea behind the analogy between knots and primes is to use a different notion of dimension, namely étale cohomological dimension. The space $\operatorname{Spec}\mathbb{F}_p$ has étale homotopy groups

$$\pi_1^{\text{ét}}(\operatorname{Spec}\mathbb{F}_p) = \operatorname{Gal}(\overline{\mathbb{F}_p}/\mathbb{F}_p) = \hat{\mathbb{Z}}, \qquad \pi_i^{\text{ét}}(\operatorname{Spec}\mathbb{F}_p) = 0 \ (i \geq 2)$$

(here $\hat{\mathbb{Z}}$ is the profinite completion of $\mathbb{Z}$). Since the circle S^1 has homotopy groups

$$\pi_1(S^1) = \operatorname{Gal}(\mathbb{R}/S^1) = \mathbb{Z}, \qquad \pi_i(S^1) = 0 \ (i \geq 2),$$

this suggests that $\operatorname{Spec}\mathbb{F}_p$ should be regarded as an arithmetic analogue of S^1. (It is a classical theorem in algebraic topology that a space with only one nonzero homotopy group, called an *Eilenberg-MacLane space*, is unique up to homotopy equivalence.) On the other hand, the space $\operatorname{Spec}\mathbb{Z}$ (or in fact $\operatorname{Spec}\mathcal{O}_k$, where $\mathcal{O}_k$ is the ring of integers of a number field k) satisfies Artin-Verdier duality, which one can think of as some sort of Poincaré duality for 3-manifolds, and $\pi_1^{\text{ét}}(\operatorname{Spec}\mathbb{Z}) = 1$. Hence it makes sense to regard $\operatorname{Spec}\mathbb{Z}$ as an analogue of $\mathbb{R}^3$. (The reader may wonder why we regard $\operatorname{Spec}\mathbb{Z}$ as an analogue of $\mathbb{R}^3$ instead of S^3. It turns out that the correct analogue of S^3 is $\operatorname{Spec}\mathbb{Z} \cup \{\infty\}$ (the prime at infinity), just as $S^3 = \mathbb{R}^3 \cup \{\infty\}$.) Thus, the embedding

$$\operatorname{Spec}\mathbb{F}_p \hookrightarrow \operatorname{Spec}\mathbb{Z}$$

is viewed as the analogue of an embedding

$$S^1 \hookrightarrow \mathbb{R}^3.$$

This yields an analogy between knots and primes.

This analogy can be extended to many concepts in knot theory and number theory. We list some of these analogies in Table 1.

Knots	Primes
Fundamental/Galois groups	
$\pi_1(S^1) = \mathrm{Gal}(\mathbb{R}/S^1)$ $= \langle [l] \rangle$ $= \mathbb{Z}$ Circle $S^1 = K(\mathbb{Z}, 1)$	$\pi_1(\mathrm{Spec}(\mathbb{F}_q)) = \mathrm{Gal}(\overline{\mathbb{F}_q}/\mathbb{F}_q)$, $q = p^n$ $= \langle [\sigma] \rangle$ $= \hat{\mathbb{Z}}$ Finite field $\mathrm{Spec}(\mathbb{F}_q) = K(\hat{\mathbb{Z}}, 1)$
Loop l	Frobenius automorphism σ
Universal covering $\mathbb{R}$	Separable closure $\overline{\mathbb{F}_q}$
Cyclic covering $\mathbb{R}/n\mathbb{Z}$	Cyclic extension $\mathbb{F}_{q^n}/\mathbb{F}_q$
Manifolds	**Spec of a ring**
$V \simeq S^1$ $V \setminus S^1 \simeq \partial V$ ($\simeq$ denotes homotopy equivalence)	$\mathrm{Spec}(\mathcal{O}_p) \simeq \mathrm{Spec}(\mathbb{F}_q)$ $\mathrm{Spec}(\mathcal{O}_\mathfrak{p}) \setminus \mathrm{Spec}(\mathbb{F}_q) \simeq \mathrm{Spec}(k_\mathfrak{p})$ ($\simeq$ denotes étale homotopy equivalence; $\mathcal{O}_\mathfrak{p}$ is a $\mathfrak{p}$-adic integer ring whose residue field is $\mathbb{F}_q$ and whose quotient field is $k_\mathfrak{p}$)
Tubular neighborhood V	$\mathfrak{p}$-adic integer ring $\mathrm{Spec}(\mathcal{O}_\mathfrak{p})$
Boundary ∂V	$\mathfrak{p}$-adic field $\mathrm{Spec}(k_\mathfrak{p})$
3-manifold M	Number ring $\mathrm{Spec}(\mathcal{O}_k)$
Knot $S^1 \hookrightarrow \mathbb{R}^3 \cup \{\infty\} = S^3$	Rational prime $\mathrm{Spec}(\mathbb{F}_p) \hookrightarrow \mathrm{Spec}(\mathbb{Z}) \cup \{\infty\}$
Any connected oriented 3-manifold is a finite covering of S^3 branched along a link (Alexander's theorem)	Any number field is a finite extension of $\mathbb{Q}$ ramified over a finite set of primes
Knot group	**Prime group**
$G_K = \pi_1(M \setminus K)$	$G_{\{\mathfrak{p}\}} = \pi_1^{\text{ét}}(\mathrm{Spec}(\mathcal{O}_k \setminus \{\mathfrak{p}\}))$
$G_K \cong G_L \iff K \sim L$ for prime knots K, L	$G_{\{(p)\}} \cong G_{\{(q)\}} \iff p = q$ for primes p, q
Linking number	**Legendre symbol**
Linking number $\mathrm{lk}(L, K)$	Legendre symbol $(\frac{q^*}{p})$, $q^* := (-1)^{\frac{q-1}{2}} q$
Symmetry of linking number $\mathrm{lk}(L, K) = \mathrm{lk}(K, L)$	Quadratic reciprocity law $(\frac{q}{p}) = (\frac{p}{q})$ ($p, q \equiv 1 \bmod 4$)
Alexander-Fox theory	**Iwasawa theory**
Infinite cyclic covering $X_\infty \to X_K$ $\mathrm{Gal}(X_\infty/X_K) = \langle \tau \rangle \cong \mathbb{Z}$	Cyclotomic $\mathbb{Z}_p$-extension k_∞/k $\mathrm{Gal}(k_\infty/k) = \langle \gamma \rangle \cong \mathbb{Z}_p$
Knot module $H_1(X_\infty)$	Iwasawa module H_∞
Alexander polynomial $\det(t \cdot \mathrm{id} - \tau \mid H_1(X_\infty) \otimes_\mathbb{Z} \mathbb{Q})$	Iwasawa polynomial $\det(T \cdot \mathrm{id} - (\gamma - 1) \mid H_\infty \otimes_{\mathbb{Z}_p} Q_p)$

TABLE 1. Analogies between knots and primes

2. Preliminaries on Knot Theory

Definition 2.1. A *knot* is the image of an embedding of S^1 into S^3 (or more generally, into an orientable connected closed 3-manifold M). A *knot type* is the equivalence class of embeddings that can be obtained from a particular one under ambient isotopy. (However, following common parlance, we shall often refer to a knot type simply as a knot when there is no danger of confusion.)

We shall be concerned only with *tame* knots, that is, knots which possess a tubular neighborhood. A knot is tame if and only if it is ambient isotopic to a piecewise-linear knot, or equivalently, to a smooth knot.

2.1. **Knot diagrams.** Let K be a knot. By removing a point in S^3 not contained in K (call it ∞), we may assume that $K \subset \mathbb{R}^3$.

Definition 2.2. A projection of K onto a plane in $\mathbb{R}^3$ is called *regular* if it has only a finite number of multiple points, all of which are double points.

Clearly, any knot projection can be transformed into a regular projection by a slight perturbation of the knot. All the knot projections we consider will be regular, with the over- and undercrossings marked.

Definition 2.3. The *crossing number* of a knot (type) is the least number of crossings in any projection of a knot of that type.

Definition 2.4. Given two knots J and K, the *connected sum* or *composition* of J and K, denoted $J\#K$, is the knot obtained by removing a small arc from each knot projection and connecting the endpoints by two new arcs, as in Figure 3.

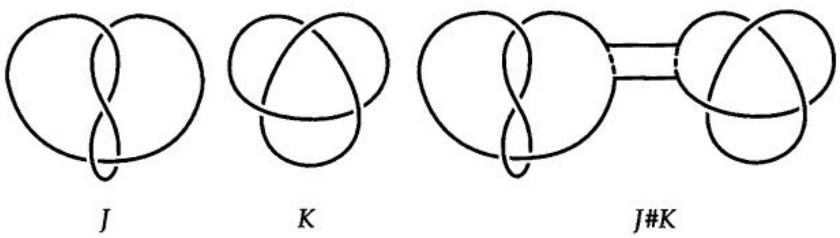
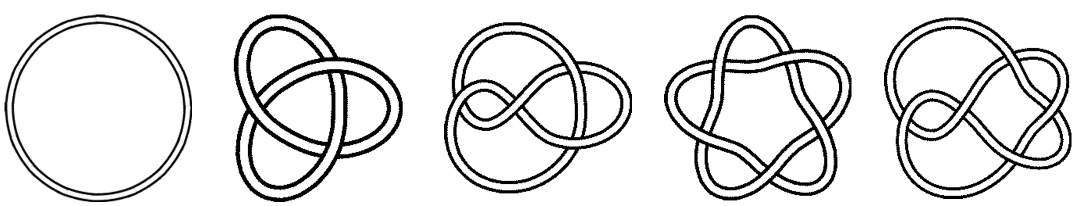


FIGURE 3. Connected sum of two knots

Note that in general, the connected sum of unoriented knots is not well-defined—more than one knot may arise as the connected sum of two unoriented knots. However, the connected sum is well-defined if we put an orientation on each knot and insist that the orientation of the connected sum matches the orientation of each of the factor knots. A knot is called *prime* if it cannot be written as the connected sum of two non-trivial knots, and *composite* otherwise.

Example 2.5.



(A) 0_1 Unknot

(B) 3_1 $(3,2)$-torus knot Trefoil

(C) 4_1 Figure eight

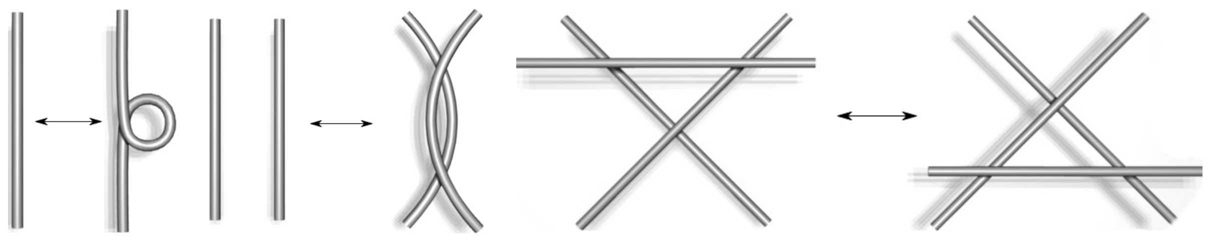
(D) 5_1 $(5,2)$-torus knot

(E) 5_2

FIGURE 4. Prime knots (i.e., knots that cannot be expressed as the connected sum of two knots, neither of which is the trivial knot) with crossing number at most 5. The knots are labelled using Alexander-Briggs notation: the regularly-sized number indicates the crossing number, while the subscript indicates the order of that knot among all knots with that crossing number in the Rolfson classification.

Remark 2.6. A knot is called *alternating* if it has a projection in which the crossings alternate between over- and undercrossings as one travels along the knot. All prime knots with crossing number less than 8 are alternating (there are three non-alternating knots with crossing number 8); moreover, it is a theorem of Thistlewaite, Kauffman and Murasugi (one of the Tait conjectures) that any minimal crossing projection of an alternating knot is an alternating projection. This provides a useful way to check if one has drawn a projection of a low-crossing knot correctly.

Two knot projections represent the same knot if and only if, up to planar isotopy, one can be obtained from the other via a sequence of *Reidemeister moves*, moves repesenting ambient isotopies that change the relations between the crossings. The Reidemeister moves are shown in Figure 5.



(A) Type I
Twist/untwist

(B) Type II
Move over/under a strand

(C) Type III
Move over/under a crossing

FIGURE 5. Reidemeister moves

2.2. **The knot group.** Let K be a knot. We fix the following notation and terminology.

Definition 2.7. Denote by V_K a tubular neighborhood of K. The complement $X_K := S^3 \backslash \mathrm{int}(V_K)$ of an open tubular neighborhood $\mathrm{int}(V_K)$ in S^3 is called the *knot exterior*. (Note that X_K is a compact 3-manifold with boundary a torus.) A *meridian* of K is a closed (oriented) curve on ∂X_K which is the boundary of a disk D^2 in V_K. A *longitude* of K is a closed curve on ∂X_K which intersects with a meridian at one point and is null-homologous in X_K. (See Figure 6.)

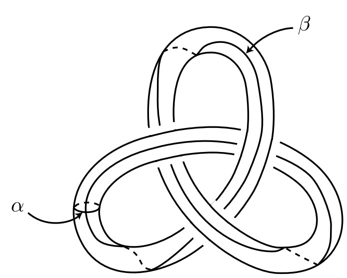


FIGURE 6. Tubular neighborhood of a knot with a meridian α and a longitude β.

The most obvious invariant of a knot K is the *knot group* G_K, which is defined to be the fundamental group of the knot exterior $\pi_1(X_K) = \pi_1(S^3 \setminus K)$. Given a regular presentation of a knot, one can obtain a presentation of the knot group, known as a *Wirtinger presentation*.

Theorem 2.8. *Given a regular presentation of a knot K, give the knot an orientation and divide it into arcs c_1, c_2, ..., c_n such that c_i is connected to c_{i+1} at a double point (with the convention that $c_{n+1} = c_1$), as in Figure 7. The knot group G_K has a Wirtinger presentation*

$$G_K = \langle x_1, \ldots, x_n \mid R_1, \ldots, R_n \rangle,$$

where the relation R_i has the form $x_i x_k x_{i+1}^{-1} x_k^{-1}$ or $x_i x_k^{-1} x_{i+1}^{-1} x_k$ depending on whether the crossing at a double point is a positive or negative crossing, as specified by Figure 8.

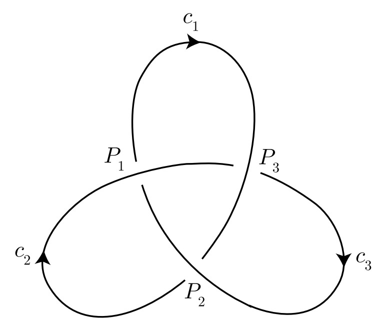
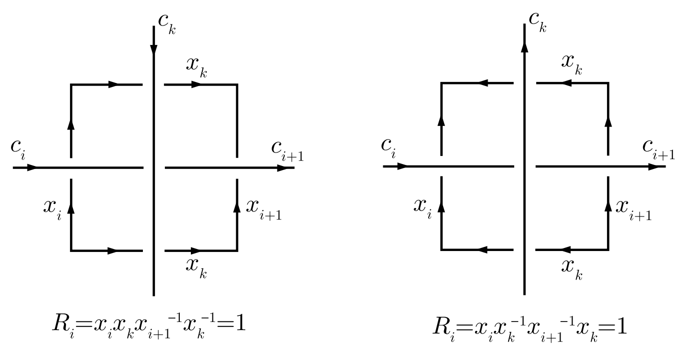


FIGURE 7. Oriented knot K, divided into arcs $c_1, c_2, \ldots, c_n$.



(A) Positive crossing (B) Negative crossing

FIGURE 8. Relation in knot group depending on the type of crossing

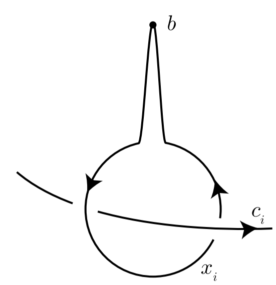


FIGURE 9. Loop x_i passing through the point at infinity and going once under c_i from the right to the left.

Proof. For $1 \leq i \leq n$, let x_i be a loop passing through ∞ and which goes once under c_i from the right to the left, as shown in Figure 9.

It is clear that the loops x_i generate the group G_K. Suppose that the arcs c_i and c_{i+1} are separated by c_k at the i-th crossing. If the crossing is positive (respectively negative), one can concatenate the loops x_i, x_k, x_{i+1}^{-1}, x_k^{-1} (respectively x_i, x_k^{-1}, x_{i+1}^{-1}, x_k) to obtain a null-homologous loop. Hence the relations R_i, $1 \leq i \leq n$, hold in G_K. (Note that the relation R_i implies any cyclic permutation of it by conjugation.) Moreover, the generators x_i and relations R_i form a presentation for G_K: by considering the projection of a loop ℓ in X_K onto the plane of the knot projection, one can write ℓ in terms of the x_i's. When a homotopy is performed on ℓ, the word representing ℓ changes only when the projection of ℓ passes through the crossings of K. □

Fact 2.9. One of the relations among the R_i is redundant, that is, we can derive any one of the relations R_i from the others.

Corollary 2.10. *G_K has a presentation with deficiency 1, that is, a presentation where the number of relations is one fewer than the number of generators.*

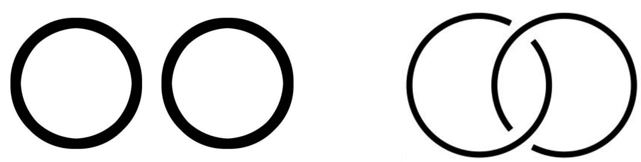
Definition 2.11. A r-component *link* L is the image of an embedding of a disjoint union of r copies of S^1 into S^3 (or more generally, into an orientable connected closed 3-manifold M). Thus one may write $L = K_1 \cup \cdots \cup K_r$ where the K_i are mutually disjoint knots. (As before, we shall often refer to an equivalence class of links under ambient isotopy simply as a link.)

The *link group* G_L is defined to be $\pi_1(S^3 \setminus L)$. Similarly to the case of knots, G_L has a Wirtinger presentation of deficiency 1. In general, for a knot K or link L in an orientable connected closed 3-manifold M, the knot group $G_K(M) := \pi_1(M \setminus K)$ or link group $G_L(M) := \pi_1(M \setminus L)$ also has deficiency 1, but may not have a Wirtinger presentation.

Example 2.12 (Knot group of trefoil)**.** Consider the trefoil knot from Figure 7. Its knot group has a Wirtinger presentation $\langle x_1, x_2, x_3 \mid x_2x_1x_3^{-1}x_1^{-1}, x_3x_2x_1^{-1}x_2^{-1}, x_1x_3x_2^{-1}x_3^{-1}\rangle$. The product of the three relations in reverse order is 1, hence any one of the relations is redundant. From the second relation, we obtain $x_3 = x_2x_1x_2^{-1}$, and substituting this into the first relation, we see that the knot group of the trefoil is the braid group $B_3 = \langle x_1, x_2 | x_1x_2x_1 = x_2x_1x_2\rangle$.

Exercise 2.13. Show that the above knot group is isomorphic to the group $\langle a, b \mid a^3 = b^2\rangle$. (In general, a (p,q)-torus knot has fundamental group $\langle a, b \mid a^p = b^q\rangle$, but this is harder to show.)

Exercise 2.14. Show that two unlinked circles (Figure 10a) and the Hopf link (Figure 10b) are not equivalent.



(A) Two unlinked circles (B) Hopf link

FIGURE 10. Two non-equivalent links

Remark 2.15. A knot is said to be *chiral* if it is not equivalent to its mirror image, and *achiral* or *amphichiral* otherwise. Clearly, the knot group cannot detect whether a knot is chiral. The other knot invariant that we shall introduce in this tutorial, the Alexander polynomial, is also unable to detect chirality since it is defined in terms of a homology group. However, other knot invariants such as the Jones polynomial are able to detect the chirality of some knots.

LECTURE 2. (JULY 6, 2012)

3. QUADRATIC RECIPROCITY

The story starts with the French "amateur" mathematician Fermat in the 17th century. Fermat was once interested in representing integers as the sum of two squares. He was amazed when he found an elegant criterion as to whether a prime number can be written in the form $x^2 + y^2$ and could not wait to communicate his result to another French mathematician, Mersenne, on Christmas day of 1640.

Example 3.1. $5 = 1^2 + 2^2$, $13 = 2^2 + 3^2$, $17 = 1^2 + 4^2$, $29 = 2^2 + 5^2$. But other primes like $7, 11, 19, 23$ cannot be represented in this way. The difference seems to depend on p mod 4.

Theorem 3.2 (Fermat)**.** *An odd prime p can be written as $p = x^2 + y^2$ if and only if $p \equiv 1 \pmod 4$.*

But why? The "only if" is obvious, but the other direction is far from trivial. In his letter, Fermat claimed that he had a "solid proof." But nobody was able to find the proof among his work—apparently the margin was never big enough for Fermat. The only clue is that he used a "descent argument": if such a prime p is not of the required form, then one can construct another smaller prime and so on, until a contradiction occurs when one encounters 5, the smallest such prime.

Euler later gave the first rigorous proof, which consists of two steps:

(1) If $p \mid x^2+y^2$ with $(x,y)=1$, then $p=x^2+y^2$. This assertion uses a descent argument.
(2) If $p \equiv 1 \pmod 4$, then $p \mid x^2+y^2$ with $(x,y)=1$.

We will not go into details of the first step but concentrate on the second. Historically, it took more time for Euler to figure out the proof of the second step. From the modern point of view, the key observation is that it has to do with whether or not -1 is a quadratic residue mod p.

Definition 3.3. Let p be a prime and a be an integer. We say a is a *quadratic residue mod p* if $x^2 \equiv a \pmod p$ has a solution, i.e., a mod p is a square in $\mathbb{F}_p$.

The following lemma then follows easily.

Lemma 3.4. *Let p be an odd prime, then $p \mid x^2+y^2$ with $(x,y)=1$ if and only if -1 is a quadratic residue.*

Proof. Because $x^2+y^2 \equiv 0 \pmod p$ if and only if $(xy^{-1})^2 \equiv -1 \pmod p$. □

So the remaining question is to find a way to determine all the quadratic residues.

Example 3.5. One naive way is to enumerate all the squares. For example when $p=11$,

x	1	2	3	4	5	6	7	8	9	10
x^2	1	4	9	5	3	3	5	9	4	1

we find that $1,3,4,5,9$ are quadratic residues mod 11 and $2,6,7,8,10$ are not quadratic residues mod 11.

We now introduce a more powerful tool in dealing with quadratic residues—the Legendre symbol.

Definition 3.6 (Legendre Symbol)**.** Let p be an odd prime and a be an integer coprime to p. We define the *Legendre symbol*

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

Remark 3.7. This definition seems a bit arbitrary. It is a good time for us to translate things in a more conceptual manner. Consider the inclusion $(\mathbb{F}_p^\times)^2 \hookrightarrow \mathbb{F}_p^\times$. Since $\mathbb{F}_p^\times$ is a cyclic group of order $p-1$, we know that the subgroup $(\mathbb{F}_p^\times)^2$ consisting of squares has index 2. So we have an exact sequence

$$1 \to (\mathbb{F}_p^\times)^2 \to \mathbb{F}_p^\times \to \{\pm 1\} \to 1.$$

Therefore the Legendre symbol $\left(\frac{\cdot}{p}\right)$ is nothing but the quotient map $\mathbb{F}_p^\times \to \{\pm 1\}$. In particular, this map is a group homomorphism, so we have

Proposition 3.8. *The Legendre symbol is* multiplicative*, namely,*

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

The multiplicativity pins down the above seemingly arbitrary definition.
To actually compute the Legendre symbol, one way is to find out an explicit expression of the map $\mathbb{F}_p^\times \to \{\pm 1\}$.

Proposition 3.9. *Suppose $a \in \mathbb{F}_p^\times$, then*

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

Proof. Since $\mathbb{F}_p^\times$ is a cyclic group of order $p-1$, we know that $a^{p-1}=1$ (Fermat's little theorem), thus $a^{\frac{p-1}{2}} \in \{\pm 1\}$. It is clear that the kernel consists of $(\mathbb{F}_p^\times)^2$. □

This proposition allows us to compute the Legendre symbol without enumerating all squares in $\mathbb{F}_p^\times$.

Example 3.10. Let us compute $\left(\frac{3}{11}\right)$. By the previous proposition,

$$\left(\frac{3}{11}\right) \equiv 3^5 \equiv (-2)^2 \cdot 3 \equiv 1 \pmod{11}.$$

This coincides with the fact that 3 is a quadratic residue mod 11: $5^2 \equiv 3 \pmod{11}$.

Obviously this could become tedious when p is bigger. We can do better using Proposition 3.11 together with the famous quadratic reciprocity law we will introduce in a moment.

Proposition 3.11. *The following formulas hold:*

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

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

Proof. The first formula follows from Proposition 3.9. For the second formula, we need to compute $2^{\frac{p-1}{2}}$ in $\mathbb{F}_p$. Let α be an 8th root of unity in $\overline{\mathbb{F}_p}$ such that $\alpha^8 = 1$ and $\alpha^4 = -1$. Then $x = \alpha + \alpha^{-1}$ satisfies $x^2 = 2$ and $2^{\frac{p-1}{2}} = x^{p-1}$. Notice that $x^p = \alpha^p + \alpha^{-p}$, we know that

$$x^p = \begin{cases} \alpha + \alpha^{-1} = x, & p \equiv \pm 1 \pmod 8, \\ \alpha^3 + \alpha^{-3} = -x, & p \equiv \pm 3 \pmod 8. \end{cases}$$

Therefore $x^{p-1} = \pm 1$ according to the residue of p mod 8. □

Theorem 3.12 (Quadratic Reciprocity)**.** *Let p and q be distinct odd primes. Then*

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

Before giving its proof, some examples are in order to demonstrate how the quadratic reciprocity can help us to simplify the computation of Legendre symbols.

Example 3.13. Let us compute $\left(\frac{3}{11}\right)$ in the previous example again. By quadratic reciprocity,

$$\left(\frac{3}{11}\right) = -\left(\frac{11}{3}\right) = -\left(\frac{2}{3}\right) = -(-1) = 1.$$

Observe that quadratic reciprocity helps us to decrease the size of the modulus quite effectively!

Example 3.14. Let us compute $\left(\frac{137}{227}\right)$. Notice that

$$\left(\frac{137}{227}\right) = \left(\frac{-90}{227}\right) = \left(\frac{-1}{227}\right)\left(\frac{2}{227}\right)\left(\frac{3^2}{227}\right)\left(\frac{5}{227}\right).$$

By definition, $\left(\frac{3^2}{227}\right) = 1$. By Proposition 3.11, we know that

$$\left(\frac{-1}{227}\right) = -1, \quad \left(\frac{2}{227}\right) = -1.$$

The quadratic reciprocity gives

$$\left(\frac{5}{227}\right) = \left(\frac{227}{5}\right) = \left(\frac{2}{5}\right) = -1.$$

So

$$\left(\frac{137}{227}\right) = -1.$$

Now let us come back to the proof of the quadratic reciprocity law. Gauss discovered the quadratic reciprocity law in his youth. Like many fundamental results in mathematics (e.g., the fundamental theorem of algebra), tons of different proofs of the quadratic reciprocity law have been found (six of them are due to Gauss), varying from counting lattice points to the infinite product expansion of sine functions. In the book *Reciprocity Laws: From Euler to Eisenstein* by Franz Lemmermeyer, 233 different proofs are collected[1] with bibliography! Here we give a simple lowbrow group-theoretic proof due to Rousseau. The only thing we really need is the Chinese remainder theorem.

[1]An online list: `http://www.rzuser.uni-heidelberg.de/~hb3/fchrono.html`.

Proof of Quadratic Reciprocity. By the Chinese remainder theorem, we have $\mathbb{Z}/pq \cong \mathbb{Z}/p \times \mathbb{Z}/q$, thus

$$(\mathbb{Z}/pq)^\times \cong \mathbb{F}_p^\times \times \mathbb{F}_q^\times.$$

Now we choose a set of coset representatives of $\{\pm 1\}$ of $(\mathbb{Z}/pq)^\times$ in two ways.

(1) We choose the coset representatives

$$S = \left\{(a,b) \in \mathbb{F}_p^\times \times \mathbb{F}_q^\times : a = 1, \dots, p-1; b = 1, \dots, \frac{q-1}{2}\right\}.$$

(2) We choose the coset representatives

$$T = \left\{(c \bmod p, c \bmod q) \in \mathbb{F}_p^\times \times \mathbb{F}_q^\times : c = 1, \dots, \frac{pq-1}{2}.\right\}.$$

Now let us compute the product of elements of S and the product of elements of T and compare the results. The products should differ at most by a sign since S can be obtained by replacing elements of T by their opposites. For the first case,

$$\prod_{(a,b)\in S} (a,b) = \left((p-1)!^{\frac{q-1}{2}}, ((q-1)/2)!^{p-1}\right).$$

For the second case, the first factor $(\text{mod} p)$ is equal to

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

by Proposition 3.9. By symmetry, we know that

$$\prod_{(c,c)\in T} (c,c) = \left((p-1)!^{\frac{q-1}{2}} \left(\frac{q}{p}\right), (q-1)!^{\frac{p-1}{2}} \left(\frac{p}{q}\right)\right).$$

Notice that

$$((q-1)/2)!^2 = (-1)^{\frac{q-1}{2}}(q-1)! \mod q,$$

so comparing the two products we conclude that

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

as desired. ☐

Remark 3.15. Quadratic reciprocity has been vastly generalized to *Artin reciprocity*, in the framework of class field theory. Hopefully we will be able to give another highbrow proof after introducing some elements of class field theory, one of the greatest mathematical achievements of the 20th century.

Remark 3.16. Next time we will introduce the notion of number fields and number rings and reformulate Fermat's result on $p = x^2 + y^2$ as prime decomposition in the number ring $\mathbb{Z}[i]$.

Exercise 3.17. Does the equation

$$2x^2 \equiv -21 \pmod{79}$$

have a solution?

Exercise 3.18 (Optional)**.** Show that there are infinitely many primes $p \equiv 1 \pmod 4$.

LECTURE 3. (JULY 9, 2012)

4. COVERING SPACES

In Lecture 1, we saw how one could obtain a presentation of the knot group from a knot projection. However, this is not the best way to study the knot group topologically, as the presentation obtained depends on the choice of knot projection. A more topological way of studying the knot group is provided by covering spaces.

Covering spaces enable one to study fundamental groups via their action on topological spaces, similarly to how group representations enable one to study abstract groups through their action on vector spaces. As we shall see, the fundamental group $\pi_1(X)$ of a space X can be thought of as a Galois (automorphism) group $\mathrm{Gal}(\tilde{X}/X)$, where $\tilde{X}$ is the universal covering space of X. This perspective gives rise to a fundamental analogy between topological and arithmetic fundamental/Galois groups in arithmetic topology.

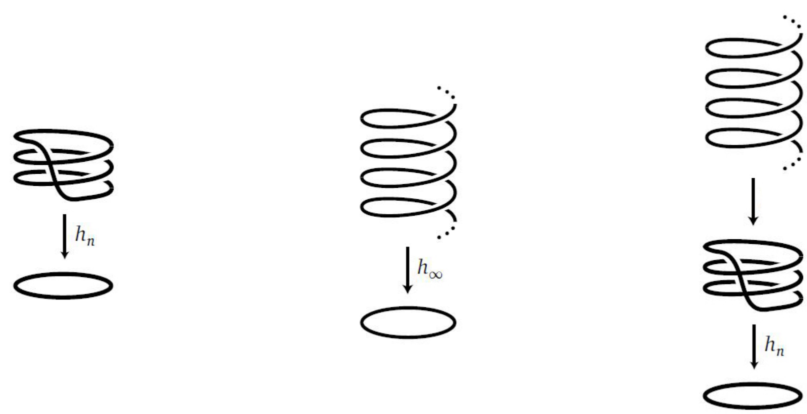
4.1. **Unramified coverings.** We recall the basic definitions and results regarding covering spaces. In what follows, we shall assume that the base space X is a connected topological manifold. (This will allow us to ignore technicalities about local path-connectedness and semi-locally simply-connectedness, which one can tell from the names will cause headaches.)

Definition 4.1. Let X be a space. A continuous map $h : Y \to X$ is called an *(unramified) covering* if for any $x \in X$, there is an open neighborhood U of x such that $h^{-1}(U)$ is a disjoint union of open sets in Y, each of which is mapped homeomorphically onto U by h. (Note that we do not require $h^{-1}(U_\alpha)$ to be non-empty, so h need not be surjective.) The set of automorphisms $Y \xrightarrow{\cong} Y$ over X forms a group, called the group of *covering transformations* of $h : Y \to X$, denoted by $\mathrm{Aut}(Y/X)$.

Example 4.2 (Coverings of S^1).

- $h_n : S^1 \to S^1$, $h(z) = z^n$, where n is a positive integer and we view z as a complex number with $|z| = 1$. (Figure 11a shows $n = 3$.)
- $h_\infty : \mathbb{R}^1 \to S^1$, $h(t) = (\cos 2\pi t, \sin 2\pi t)$. (Figure 11b.)

In fact, the covering $h_\infty : \mathbb{R}^1 \to S^1$ actually covers the coverings $h_n : S^1 \to S^1$, as shown in Figure 11c. As we shall see later, $\mathbb{R}^1$ is the universal covering space of S^1, that is, it is a covering of any other (connected) covering space of S^1 (which turn out to be the finite coverings h_n).



(A) $h_n : S^1 \to S^1$ (B) $h_\infty : \mathbb{R}^1 \to S^1$ (C) h_n is a subcovering of h_∞

FIGURE 11. Coverings of S^1

Example 4.3. We consider a higher-dimensional example. Let Y be the closed orientable surface of genus 11, the "11-hole torus," as in Figure 12. This has 5-fold rotational symmetry generated by a rotation of angle $2\pi/5$, and hence an action of the cyclic group $\mathbb{Z}/5\mathbb{Z}$. The quotient space $X = Y/(\mathbb{Z}/5\mathbb{Z})$ is a surface of genus 3, obtained from one of the five subsurfaces by identifying two boundary circles C_i and C_{i+1}. Thus we have a covering space $M_{11} \to M_3$, where M_g denotes the closed orientable surface of genus g. This example clearly generalizes by replacing the 2 holes in each "arm" of M_{11} by m holes and the 5-fold symmetry by n-fold symmetry to give covering spaces $M_{mn+1} \to M_{m+1}$.

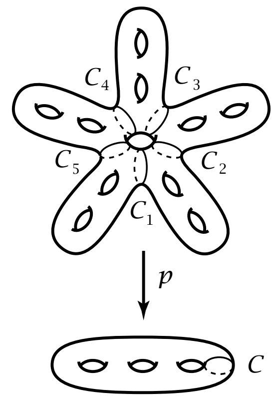


FIGURE 12. Covering of the 3-hole torus by the 11-hole torus.

In what follows, we shall restrict our attention to connected covering spaces, since a general covering space is just a disjoint union of connected ones.

Definition 4.4. A covering $h : Y \to X$ is called *Galois* or *normal* if for each $x \in X$ and each pair of lifts $\tilde{x}'$, $\tilde{x}'$ of x, there is a covering transformation taking $\tilde{x}$ to $\tilde{x}'$. For a Galois covering $h : Y \to X$, we call $\mathrm{Aut}(Y/X)$ the *Galois group* of Y over X and denote it by $\mathrm{Gal}(Y/X)$.

Intuitively, a Galois covering is one with maximal symmetry, in analogy with Galois extensions, which as splitting fields of polynomials, can be considered "maximally symmetric."

Recall that the main theorem of Galois theory gives a bijective correspondence between intermediate field extensions and subgroups of the Galois group. There is a similar version of the main theorem for coverings, which relates connected coverings of a given space X and subgroups of $\pi_1(X)$.

Theorem 4.5. *The induced map $h_* : \pi_1(Y, y) \to \pi_1(X, x)$ is injective, and there is a bijection*

$$\{\textit{connected coverings } h : Y \to X\}/\textit{isom.} \xrightarrow{\cong} \{\textit{subgroups of } \pi_1(X, x)\}/\textit{conj.}$$

$$(h : Y \to X) \mapsto h_*(\pi_1(Y, y)) \qquad\qquad (y \in h^{-1}(x))$$

with the property that $h : Y \to X$ is a Galois covering if and only if $h_*(\pi_1(Y, y))$ is a normal subgroup of $\pi_1(X, x)$. In this case $\mathrm{Gal}(Y/X) \cong \pi_1(X, x)/h_*(\pi_1(Y, y))$.

The covering $h : \tilde{X} \to X$ (up to isomorphism over X) which corresponds to the identity subgroup of $\pi_1(X, x)$ is called the *universal covering* of X; it is a covering space of any other covering space of X. Since the map $h_* : \pi_1(Y, y) \to \pi_1(X, x)$ is injective, $\pi_1(\tilde{X}) = 1$, i.e. the universal covering is simply-connected, and $\mathrm{Gal}(\tilde{X}/X) \cong \pi_1(X)$. The two most important types of covering spaces we shall consider are the universal covering space and the cyclic covering spaces.

Example 4.6. We return to Example 4.2, the coverings of S^1. We have seen that $\mathbb{R}^1 \to S^1$ is a covering. Since $\mathbb{R}^1$ is simply-connected, this tells us that this is in fact the universal covering of $\mathbb{R}^1 \to S^1$. Indeed, $\mathrm{Gal}(\mathbb{R}^1/S^1) \cong \mathbb{Z} \cong \pi_1(S^1)$, with a generator τ of $\mathrm{Gal}(\mathbb{R}^1/S^1)$ acting by a shift of the helix. Moreover, since the only quotient groups of $\mathbb{Z}$ are the finite cyclic groups $\mathbb{Z}/n\mathbb{Z}$ and we have found coverings $h_n : S^1 \to S^1$ with exactly these Galois groups, this tells us that these are the only other coverings of S^1.

Theorem 4.5 tells us that instead of using the Wirtinger presentation, one can also study a knot group $G_K = \pi_1(X_K)$ by studying the covering spaces of the knot exterior X_K. While the universal covering space of a knot exterior depends on the particular knot, there is a uniform procedure for constructing the cyclic coverings of a knot exterior. The idea is to reverse the reasoning in Example 4.3: instead of taking one of several copies of a manifold with boundary and gluing the boundaries together, we want to slice the base space open and glue several copies of the resulting space together along the boundaries appropriately.

How can we slice the knot complement in a natural way? If K is the unknot, then there is an obvious way to slice the knot complement such that the knot figures prominently: we slice along the intersection of the knot complement with the disk bounded by the unknot. Is it possible to do this for a general knot K? The answer is in the affirmative: Seifert showed in 1934 that for every knot or link, there exists an compact oriented connected surface, called a *Seifert surface*, whose boundary is that knot or link.

Example 4.7.

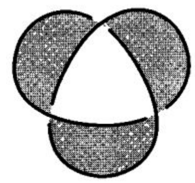


FIGURE 13. Möbius band with three half-twists, with boundary a trefoil. Note that the Möbius band is not considered to be a Seifert surface for the trefoil as it is not orientable.

Exercise 4.8. Figure 14 shows a Seifert surface for a knot, since this surface is compact, oriented and connected. What knot is this a Seifert surface of? What type of surface is this? (Hint: use Euler characteristic and the classification theorem for surfaces by genus, orientability and number of boundary components.)

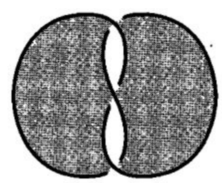


FIGURE 14

Example 4.9 (Infinite cyclic covering). Let $K \subseteq S^3$ be a knot. From a Wirtinger presentation for G_K, one sees that $x_1, \ldots, x_n$ are mutually conjugate. Hence the abelianization $G_K/[G_K, G_K]$ is an infinite cyclic group generated by the class of a meridian α of K. Let $\psi_\infty : G_K \to \mathbb{Z}$ be the surjective homomorphism sending α to 1, and let $h_\infty : X_\infty \to X_K$ be the covering corresponding to $\ker(\psi_\infty)$ in Theorem 4.5. The covering space X_∞ is independent of the choice of α and is called the *infinite cyclic covering* of X_K. It is constructed as follows. Let Σ_K be a Seifert surface of K. Let Y be the space obtained by cutting X_K along $X_K \cap \Sigma_K$, and let Σ^+ and Σ^- be the two surfaces homeomorphic to $X_K \cap \Sigma_K$ obtained from the cut, as in Figure 15.

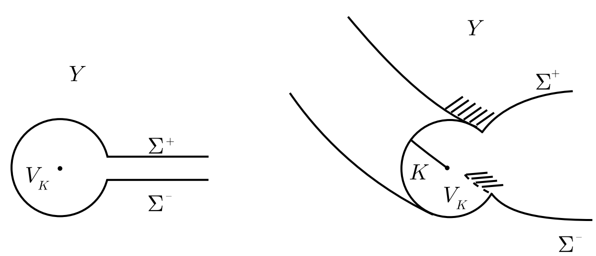


FIGURE 15. Copies Σ^+ and Σ^- of $X_K \cap \Sigma_K$ obtained by cutting along $X_K \cap \Sigma_K$.

Let Y_i $(i \in \mathbb{Z})$ be copies of Y. The space X_∞ is obtained from the disjoint union of all the Y_i's by identifying Σ_i^+ with Σ_{i+1}^- $(i \in \mathbb{Z})$, as in Figure 16, and a generator τ of $\mathrm{Gal}(X_\infty/X_K)$ is given by the shift sending Y_i to Y_{i+1} $(i \in \mathbb{Z})$.

Example 4.10 (Finite cyclic covering). For each $n \in \mathbb{N}$, let $\psi_n : G_K \to \mathbb{Z}/n\mathbb{Z}$ be the composite of ψ_∞ with the surjection $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$, and let $h_n : X_n \to X_K$ be the covering corresponding to $\ker(\psi_n)$. Then $\mathrm{Gal}(X_n/X_K) \cong \mathbb{Z}/n\mathbb{Z}$. The covering spaces X_n are constructed similarly to X_∞, except that we now take n copies $Y_0, \ldots, Y_{n-1}$ of Y, and identify Σ_{n-1}^+ with Σ_0^- instead, as in Figure 17. A generator τ of $\mathrm{Gal}(X_n/X_K)$ corresponding to 1 mod $n\mathbb{Z}$ is given by the shift sending Y_i to Y_{i+1} $(i \in \mathbb{Z}/n\mathbb{Z})$.

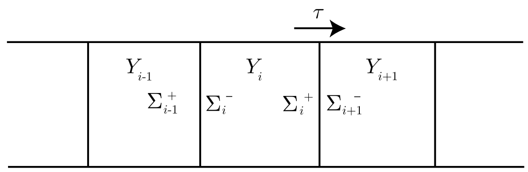


FIGURE 16. Space X_∞ obtained from the disjoint union of the Y_i's by identifying Σ_i^+ with Σ_{i+1}^- $(i \in \mathbb{Z})$

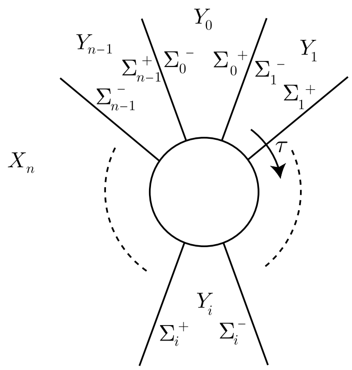


FIGURE 17. Space X_∞ obtained from the disjoint union of $Y_0, \ldots, Y_{n-1}$ by identifying Σ_i^+ with Σ_{i+1}^- $(i \in \mathbb{Z}/n\mathbb{Z})$

4.2. **Ramified coverings.** Above, we considered (unramified) coverings of the knot exterior X_K. However, we may wish to consider a covering of the entire space S^3 extending the above coverings. This is accomplished by ramified or branched coverings.

Let M and N be n-manifolds $(n \geq 2)$ and let $f : N \to M$ be a continuous map. Define $S_N := \{y \in N \mid f \text{ is not a homeomorphism in a neighborhood of } y\}$ and $S_M := f(S_N)$.

Definition 4.11. The map $f : N \to M$ is called a *covering ramified over* S_M if the following conditions are satisfied:

(1) $f|_{N\setminus S_N} : N \setminus S_N \to M \setminus S_M$ is an (unramified) covering, and
(2) for any $y \in S_N$, there exist neighborhoods V of y and U of $f(y)$, and homeomorphisms $\varphi : V \xrightarrow{\approx} D^2 \times D^{n-2}$ and $\psi : U \xrightarrow{\approx} D^2 \times D^{n-2}$, such that $(g_e \times \mathrm{id}_{D^{n-2}}) \circ \varphi = \psi \circ f$ for some positive integer $e = e(y)$, where $g_e(z) := z^e$ for $z \in D^2 = \{z \in \mathbb{C} \mid |z| \leq 1\}$. (That is, f acts locally like the eth-power map.)

One can "complete" an unramified covering to obtain a ramified covering, as follows.

Example 4.12 (Fox completion). Consider the n-fold cyclic covering $h_n : X_n \to X_K$ from Example 4.10. The restriction $h_n|_{\partial X_n} : \partial X_n \to \partial X_K$ is an n-fold covering of tori and $n\alpha$ is a meridian of ∂X_n. Attach $V = D^2 \times S^1$ to X_n by gluing ∂V and ∂X_n in such a way that a meridian of ∂V coincides with $n\alpha$. Denote by M_n the closed 3-manifold obtained in this manner.

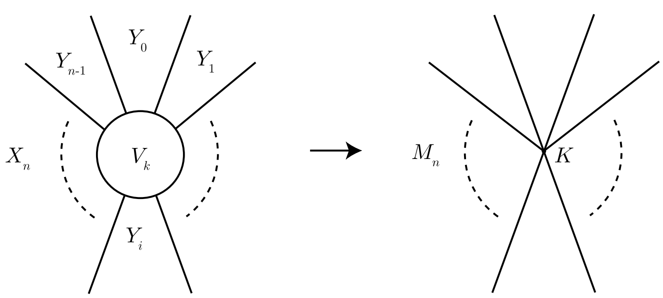


FIGURE 18. Fox completion M_n of X_n

Define $f_n : M_n \to S^3$ by $f_n|_{X_n} := h_n$ and $f_n|_V := f_n \times \mathrm{id}_{S^1}$. Then f_n is a covering ramified over K, which we call the *Fox completion* of $h_n : X_n \to X_K$.

LECTURE 4. (JULY 11, 2012)

5. NUMBER RINGS AND ÉTALE COVERINGS

Recall Fermat's theorem: an odd prime p is of the form $p = x^2 + y^2$ if and only if $p \equiv 1 \pmod 4$. We proved this using the point of view of finite field arithmetic and quadratic residues. Now we are going to shift our view once again (is this called capricious, flighty, mercurial, or fickle?) and see how it can do us a favor. Using the imaginary number $i = \sqrt{-1}$, we know that in this case p can be decomposed into the product $p = (x+yi)(x-yi)$, where

$$x \pm yi \in \mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$$

are *Gaussian integers*.

Notice that the Gaussian integer ring $\mathbb{Z}[i]$, like $\mathbb{Z}$, is a *unique factorization domain* (UFD), i.e., every element in $\mathbb{Z}[i]$ can be uniquely decomposed into the product of prime elements. More precisely, if $a \in \mathbb{Z}[i]$ and there are two decompositions $a = \prod_{i=1}^n \alpha_i$ and $a = \prod_{j=1}^m \alpha'_i$, then $n = m$, and after a possible permutation, we have $(\alpha_i) = (\alpha'_i)$, namely α_i and α'_i are the same up to a unit.

Now the above classical result of Fermat can be reformulated as the basic rule of prime decomposition in the bigger ring $\mathbb{Z}[i]$ (rather than the usual integer ring $\mathbb{Z}$) as follows.

Proposition 5.1. *Let p be a prime number.*

(1) *If $p \equiv 1 \pmod 4$, then $p = \alpha\bar{\alpha}$, where $\alpha, \bar{\alpha} \in \mathbb{Z}[i]$ are prime elements, $\bar{\alpha}$ is the conjugate of α and $(\alpha) \neq (\bar{\alpha})$.*
(2) *If $p \equiv 3 \pmod 4$, then p is a prime element.*
(3) *If $p = 2$, then $2 = (1+i)^2 \times (-i)$, where $1+i$ is a prime element and $-i$ is a unit.*

Proof. (1) By Fermat's theorem, we can find integers $x, y \in \mathbb{Z}$ such that $p = x^2 + y^2$. Set $\alpha = x + yi$, then it suffices to show that $x + yi$ is a prime element. Define the *norm map*

$$\mathbb{N} : \mathbb{Z}[i] \to \mathbb{Z}_+, \quad a + bi \mapsto (a+bi)(a-bi) = a^2 + b^2,$$

then $\mathbb{N}$ is clearly multiplicative. Assume $x + yi = \alpha_1\alpha_2$, then taking norms gives $p = \mathbb{N}(x+yi) = \mathbb{N}(\alpha_1)\mathbb{N}(\alpha_2)$. Hence one of the $\mathbb{N}(\alpha_i)$'s is equal to 1, so it must be a unit and $x + yi$ is a prime element.

(2) Suppose $p = \alpha_1\alpha_2$ is not a prime element, then taking norms gives that $\mathbb{N}(\alpha_1) = \mathbb{N}(\alpha_2) = p$. This contradicts the fact that p is not of the form $x^2 + y^2$ by Fermat's theorem.

(3) This follows from the fact that $\mathbb{N}(1+i) = 2$ is a prime number.

□

Exercise 5.2. Show that $(\alpha) \neq (\bar{\alpha})$ in the first case to complete the proof.

So the arithmetic problem of the sum of two squares is essentially equivalent to finding the prime decomposition of p in the ring $\mathbb{Z}[i]$. This elegant point of view helps us to vastly and systematically generalize the arithmetic objects we study.

Definition 5.3. A *number field K* is a finite extension of the field $\mathbb{Q}$ of rational numbers. The elements of K are called *algebraic numbers*. The *number ring* (or *ring of integers*) $\mathcal{O}_K$ of K is the integral closure of $\mathbb{Z}$ in K. In concrete terms, $\mathcal{O}_K$ consists of *algebraic integers*, namely roots of monic polynomials in $\mathbb{Z}[x]$.

Example 5.4. The simplest number field other than $\mathbb{Q}$ is $K = \mathbb{Q}(i)$, an imaginary quadratic extension of $\mathbb{Q}$. Let us compute the number ring of $K = \mathbb{Q}(i)$. Suppose $x = a + bi$ with $a, b \in \mathbb{Q}$. Then $x \in \mathcal{O}_K$ if and only if x satisfies a quadratic monic equation $X^2 - sX + t = 0$ with $s, t \in \mathbb{Z}$. We know that $2a = s$ and $a^2 + b^2 = t$. So $s^2 + 4b^2 = 4t$. Set $n = 2b$, then n is an integer and $s^2 + n^2 = 4t$. Therefore s and n are multiples of 2. We conclude that $a, b \in \mathbb{Z}$, so $\mathcal{O}_K = \mathbb{Z}[i]$, which is exactly the Gaussian integer ring.

Example 5.5. An important class of number fields are the *cyclotomic fields* $K = \mathbb{Q}(\zeta_n)$, generated by a primitive nth root of unity ζ_n. It has Galois group $(\mathbb{Z}/n\mathbb{Z})^\times$. It can be shown in general that $\mathcal{O}_K = \mathbb{Z}[\zeta_n]$. In particular, taking $n = 4$ gives us again the Gaussian integer ring.

Exercise 5.6. Determine the ring of integers of the field $K = \mathbb{Q}(\sqrt{-7})$.

Remark 5.7. The cyclotomic fields were studied by Kummer in order to attack Fermat's last theorem (Fermat, our old friend). Kummer factorized the equation $z^n = x^n + y^n$ as

$$z^n = (x+y)(x+\zeta_n y)\cdots(x+\zeta_n^{n-1}y).$$

To match the factors, he was forced to consider prime decomposition in the ring $\mathbb{Z}[\zeta_n]$. However, a crucial caveat is that $\mathbb{Z}[\zeta_n]$ is not always a UFD, so the rule of unique decomposition into prime elements is not always possible.

Fortunately, Kummer considered a generalized notion of "ideals" and the decomposition of ideals into prime ideals is still available for all number rings. We state this version of the fundamental theorem of the arithmetic of number rings without proof.

Theorem 5.8. *Let $\mathcal{O}_K$ be a number ring and $\mathfrak{a}$ be a nontrivial ideal of $\mathcal{O}_K$. Then $\mathfrak{a}$ can be uniquely (up to permutation) decomposed into a product of prime ideals*

$$\mathfrak{a} = \mathfrak{p}_1^{e_1}\mathfrak{p}_2^{e_2}\cdots\mathfrak{p}_m^{e_m}, \quad e_i \geq 1.$$

Example 5.9. In $\mathbb{Z}[\sqrt{-5}]$, the element 6 has two decompositions $6 = 2\cdot 3 = (1+\sqrt{-5})(1-\sqrt{-5})$ where none of $2, 3, 1+\sqrt{-5}, 1-\sqrt{-5}$ can be further decomposed. The problem occurring here is exactly that none of them generate prime ideals. The prime decomposition promised by the previous theorem is given by

$$(6) = (2, 1+\sqrt{-5})\cdot(2, 1-\sqrt{-5})\cdot(3, 1+\sqrt{-5})\cdot(3, 1-\sqrt{-5}).$$

As we have seen, in order to study the problem concerning rational numbers and integers, we need to work in a new world of extensions of $\mathbb{Q}$ and $\mathbb{Z}$, the number fields and number rings. It is amusing to compare this with the topological setting: in order to study the topology of a space, one way is to work with its unramified covering spaces instead. We now carry this key idea further, leading to the notion of *finite étale coverings* and *étale fundamental groups* in this algebraic setting.

space X	scheme Spec$A$
unramified covering	finite étale covering
fundamental group	étale fundamental group

Let us look at several examples to motivate.

Example 5.10. One can define the fundamental group using loops. Though Spec$A$ can be endowed with a topology (the Zariski topology), it is too coarse to contain any loop in the usual sense. Alternatively, we will define the étale fundamental group as the automorphism group of its "universal covering."

Example 5.11. However, another difference in the algebraic setting is that we cannot always expect the existence of the (usually *infinite*) universal covering. For example, the universal covering $\mathbb{R}^1 \to S^1$ is given by a transcendental function $t \mapsto e^{it}$, which does not make sense in the algebraic world. So we are going to step back and find an object which approximates all the *finite* étale coverings best.

Example 5.12. Remember that the ring $\mathbb{Z}$ corresponds to a space (an affine scheme) Spec $\mathbb{Z}$, which can be geometrically represented by a line. A maximal ideal (p) corresponds to a closed point Spec $\mathbb{F}_p \hookrightarrow$ Spec $\mathbb{Z}$. Unlike Spec $\mathbb{C}[t]$, where all the residue fields are the same field $\mathbb{C}$, the points on Spec $\mathbb{Z}$ have different residue fields $\mathbb{F}_p$, which are not algebraically closed. In other words, there are many finite extensions of $\mathbb{F}_p$ (one for each degree n). So we have many finite "covering spaces," although each of these covering spaces is also a point geometrically. Intuitively, we will draw a slightly bigger point to stand for these finite extensions Spec $\mathbb{F}_{p^n}$. From this point of view, the space Spec $\mathbb{F}_p$ is not "simply connected" because it has nontrivial finite covering spaces. It is now very natural to define the "fundamental group" of Spec $\mathbb{F}_p$ as the automorphism group $\mathrm{Gal}(\overline{\mathbb{F}_p}/\mathbb{F}_p)$, viewing Spec $\overline{\mathbb{F}_p}$ as the "universal covering" since $\overline{\mathbb{F}_p}$ is the union of all finite extensions of $\mathbb{F}_p$.

Example 5.13. The inclusion $\mathbb{Z} \hookrightarrow \mathbb{Z}[i]$ gives a map π : Spec $\mathbb{Z}[i] \to$ Spec $\mathbb{Z}$. The fiber of a prime $(p) \in$ Spec $\mathbb{Z}$ is given by Spec $\mathbb{Z}[i] \otimes_{\mathbb{Z}} \mathbb{F}_p =$ Spec $\mathbb{Z}[i]/(p\mathbb{Z}[i])$. From the prime decomposition in $\mathbb{Z}[i]$ we have the following situation:

(5)	$(1+2i)(1-2i)$	$\mathbb{Z}[i]/(5) \cong \mathbb{Z}[i]/(1+2i) \times \mathbb{Z}[i]/(1-2i) \cong \mathbb{F}_5 \times \mathbb{F}_5$	two points
(3)	(3)	$\mathbb{Z}[i]/(3) \cong \mathbb{F}_9$	one bigger point
(2)	$(1+i)^2$	$\mathbb{Z}[i]/(2) \cong \mathbb{F}_2[i] = \{0, 1, i, 1+i\}$	one double point

This matches the geometry:

(1) For primes $p \equiv 1 \pmod 4$, $(p) = \mathfrak{p}_1\mathfrak{p}_2$. The $\mathfrak{p}_1$ and $\mathfrak{p}_2$ correspond to two points lying above $(p) \in$ Spec $\mathbb{Z}$.

(2) For primes $p \equiv 3 \pmod 4$, $(p) = \mathfrak{p}$ remains prime, which corresponds to a slightly bigger point lying above $(p) \in \operatorname{Spec}\mathbb{Z}$.
(3) For $p = 2$, $(2) = (1+i)^2$ is a power of prime, which corresponds to a double point geometrically. In this case the tensor product is no longer a field ($1+i$ is a nilpotent).

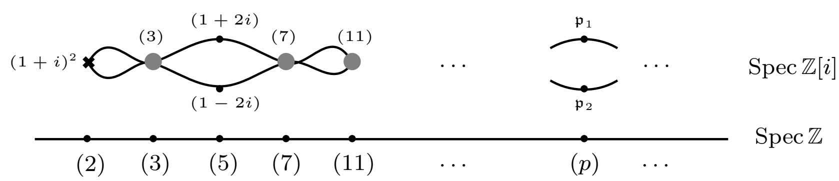


FIGURE 19. Geometry of the map $\operatorname{Spec}\mathbb{Z}[i] \to \operatorname{Spec}\mathbb{Z}$.

So only in the last case is the geometrical picture not an unramified covering: two points are somehow collapsing together. This is characterized by the fact that $\mathbb{Z}[i] \otimes_{\mathbb{Z}} \mathbb{F}_p$ is not a field extension of $\mathbb{F}_p$.

With the above said, it is not at all absurd to introduce the following definition, with the mind that "étale" is intended to mean an unramified covering in the algebraic setting.

Definition 5.14. Let k be a field. An k-algebra is called *finite étale* if it is a finite product of finite separable extensions of k.

Definition 5.15. A map $\operatorname{Spec} B \to \operatorname{Spec} A$ (or equivalently, a ring homomorphism $A \to B$) is called a *finite étale* map if B is a finitely generated flat A-module and for any prime $\mathfrak{p} \in \operatorname{Spec} A$, $B \otimes_A \kappa(\mathfrak{p})$ is a finite étale $\kappa(\mathfrak{p})$-algebra, where $\kappa(\mathfrak{p}) = \operatorname{Frac}(A/\mathfrak{p})$ is the residue field of $\mathfrak{p}$. In this case we say that B is a *finite étale A-algebra* or $\operatorname{Spec} B$ is a *finite étale covering* of $\operatorname{Spec} A$.

Next time we will define the étale fundamental group in terms of finite étale coverings and make more concrete sense with examples.

LECTURE 5. (JULY 13, 2012)

6. MONODROMY PERMUTATION REPRESENTATION

In Lecture 3, we saw how $\operatorname{Aut}(Y/X)$ acts on a covering $h : Y \to X$. Restricting to a fiber $h^{-1}(x)$, this gives us an action $\operatorname{Aut}(Y/X)$ on $h^{-1}(x)$. One can also define an action of $\pi_1(X,x)$ on a fiber $h^{-1}(x)$ in terms of liftings of maps, which is closely related to the former action when the covering $h : Y \to X$ is Galois. For this, we need the following proposition.

Proposition 6.1 (Homotopy lifting property)**.** *Given a covering space $h : Y \to X$, a homotopy $f_t : Z \to X$ ($t \in [0,1]$) and a map $\widetilde{f_0} : Z \to Y$ lifting f_0, there exists a unique homotopy $\widetilde{f_t} : Z \to Y$ of $\widetilde{f_0}$ lifting f_t.*

In particular, taking Z to be a point, we obtain the *path lifting property* of a covering space $h : Y \to X$: for any path $f : [0,1] \to X$ and any lift y_0 of the starting point $f(0) = x_0$, there exists a unique path $\tilde{f} : [0,1] \to Y$ lifting f starting at y_0. Furthermore, for any homotopy f_t of f, there exists a unique lift $\widetilde{f_t}$ of f_t such that $\widetilde{f_t}$ is a homotopy of $\tilde{f}$.

The path lifting property now enables us to define an action of $\pi_1(X,x)$ on a fiber $h^{-1}(x)$ as follows. For a loop $[l] \in \pi_1(X,x)$ and a point $y \in h^{-1}(x)$, we define $y \cdot [l]$ to be the ending point $\tilde{l}(1)$, where $\tilde{l}$ is the lift of l with starting point $\tilde{l}(0) = y$. The induced representation $\rho_x : \pi_1(X,x) \to \operatorname{Aut}(h^{-1}(x))$ is called the *monodromy permutation representation* of $\pi_1(X,x)$. Moreover, one can show that if $h : Y \to X$ is a Galois covering, then the composite

$$\pi_1(X,x) \twoheadrightarrow h_*(\pi_1(Y,y))\backslash\pi_1(X,x) \cong \operatorname{Gal}(Y/X) \xrightarrow{\text{restrict to a fiber } h^{-1}(x)} \operatorname{Aut}(h^{-1}(x))$$

is precisely the monodromy permutation representation. Conversely, from the monodromy permutation representation, one can recover the covering $h : Y \to X$ up to isomorphism.

Example 6.2. Suppose our base space X is the circle S^1. If $h : Y \to X$ is one of the finite cyclic coverings $h_n : S^1 \to S^1$ as in Example 4.2, then $\mathrm{Gal}(Y/X)$ is a finite cyclic group $\mathbb{Z}/n\mathbb{Z}$, $h^{-1}(x)$ consists of n points, $\mathrm{Aut}(h^{-1}(x))$ is the symmetric group S_n and the image of the inclusion $\mathrm{Gal}(Y/X) \hookrightarrow S_n$ is the subgroup generated by the cyclic permutation $(1\ 2\ \cdots\ n)$. If $h : Y \to X$ is the universal covering $h_\infty : \mathbb{R}^1 \to S^1$, then $\mathrm{Gal}(\mathbb{R}^1/S^1)$ is the infinite cyclic group $\mathbb{Z}$, $\mathrm{Aut}(h^{-1}(x))$ is the infinite symmetric group S_∞ and the image of the inclusion $\mathrm{Gal}(\mathbb{R}^1/S^1) \hookrightarrow S_\infty$ is the subgroup generated by the shift $m \mapsto m+1$ ($m \in \mathbb{Z}$).

7. Linking numbers and Legendre symbols

To get a better feeling for working not only with knots, but also with links, and not only at a single prime, but with multiple primes, the first analogy between knot theory and number theory that we shall study is that between linking numbers and Legendre symbols. The linking number and Legendre symbol are the first invariants that come to mind when one considers a 2-component link and a pair of primes respectively, and surprisingly, there is an analogy between them.

7.1. **Linking numbers.**

Definition 7.1. Let K and L be disjoint oriented simple closed curves K and L in S^3 (i.e. a 2-component link). The *linking number of K and L*, denoted by $\mathrm{lk}(L,K)$, is defined as follows. Let Σ_L be a Seifert surface of L; by perturbing Σ_L suitably, we may assume that K intersects Σ_L transversely. Let $P_1, \ldots, P_m$ be the set of intersection points of K and Σ_L. According as the tangent vector of K at P_i has the same or opposite direction as the normal vector of Σ_L at P_i, assign a number $\varepsilon(P_i) := 1$ or -1 to each P_i, as in Figure 20. The linking number $\mathrm{lk}(L,K)$ is defined by

$$\mathrm{lk}(L,K) := \sum_{i=1}^{m} \varepsilon(P_i).$$

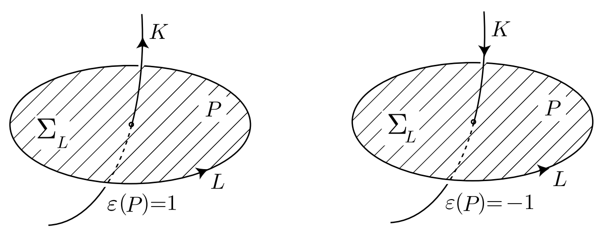


Figure 20. Calculation of the linking number.

Remark 7.2. The linking number can also be computed from a link diagram by the formula

$$\mathrm{lk}(L,K) = \frac{1}{2}(\#\text{ positive crossings between } K \text{ and } L - \#\text{ negative crossings between } K \text{ and } L),$$

from which we see that $\mathrm{lk}(L,K)$ is symmetric:

$$\mathrm{lk}(L,K) = \mathrm{lk}(K,L).$$

Example 7.3.

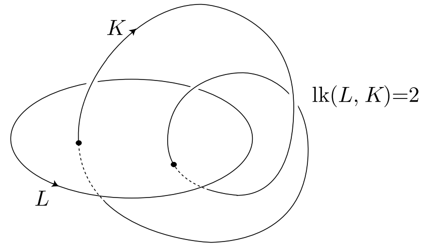


Figure 21. 2-component link with linking number 2.

We have seen that covering spaces provide a link between topological and arithmetic fundamental groups. Thus, it is natural to try to formulate the linking number in terms of covering spaces as a first step in establishing the analogy between linking numbers and Legendre symbols. We use the notation in Example 4.9: for a meridian α of L, let $\psi_\infty : G_L \to \mathbb{Z}$ be the surjective homomorphism sending α to 1, let X_∞ be the infinite cyclic covering of X_L corresponding to $\ker(\psi_\infty)$ and let τ denote the generator of $\mathrm{Gal}(X_\infty/X_L)$ corresponding to $1 \in \mathbb{Z}$. Let $\rho_\infty : G_L \twoheadrightarrow \mathrm{Gal}(X_\infty/X_L)$ be the natural surjection. We shall want to think of this as a monodromy permutation reprsentation.

Proposition 7.4. $\rho_\infty([K]) = \tau^{\mathrm{lk}(L,K)}$.

Proof. Recall (Example 4.9) that X_∞ is constructed by gluing together copies Y_i $(i \in \mathbb{Z})$ of the space Y obtained by cutting X_L along the Seifert surface Σ_L of L, as in Figure 22.

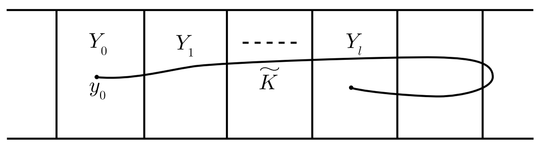


FIGURE 22. Lift $\tilde{K}$ of K to X_∞.

Let $\tilde{K}$ be a lift of K to X_∞. Then, when K crosses Σ_L with intersection number $\varepsilon = 1$ (respectively -1), $\tilde{K}$ crosses from Y_i to Y_{i+1} (respectively from Y_{i+1} to Y_i) for some i since Σ_i^+ is identified with Σ_{i+1}^- in X_∞. Therefore, if the starting point y_0 of $\tilde{K}$ is in Y_0, then its ending point lies in Y_l, $l = \mathrm{lk}(L,K)$, that is, $\rho_\infty([K])(y_0) \in Y_l$. Since τ maps Y_i onto Y_{i+1}, it follows that $\rho_\infty([K]) = \tau^{\mathrm{lk}(L,K)}$. □

Let $\psi_2 : G_L \to \mathbb{Z}/2\mathbb{Z}$ be the composite of ψ_∞ with the surjection $\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$, and let $h_2 : X_2 \to X_L$ be the double covering corresponding to $\ker(\psi_2)$. Let $\rho_2 : G_L \twoheadrightarrow \mathrm{Gal}(X_2/X_L)$ be the natural surjection, then by Proposition 7.4, the image of $[K]$ in $\mathrm{Gal}(X_2/X_L) \cong \mathbb{Z}/2\mathbb{Z}$ under ρ_2 is given by $\mathrm{lk}(L,K) \bmod 2$. A similar argument to that in the proof of Proposition 7.4 tells us that

$$\rho_2([K])(y) = \text{ending point of a lift of } K \text{ with starting point } y.$$

We conclude that

$$\rho_2([K]) = \mathrm{id}_{X_2} \iff h_2^{-1}(K) = K_1 \cup K_2 \quad \text{(2-component link)};$$
$$\rho_2([K]) = \tau \iff h_2^{-1}(K) = \mathfrak{K} \quad \text{(knot in } X_2\text{)}.$$

We thus obtain the following result:

Proposition 7.5.

$$h_2^{-1}(K) = \begin{cases} K_1 \cup K_2 & \text{if } \mathrm{lk}(L,K) \equiv 0 \pmod 2, \\ \mathfrak{K} & \text{if } \mathrm{lk}(L,K) \equiv 1 \pmod 2. \end{cases}$$

Note that this result has the same form as the decomposition of prime numbers in the ring of Gaussian integers $\mathbb{Z}[i]$ that was established in Lecture 4: a prime $p \in \mathbb{Z}$ decomposes in $\mathbb{Z}[i]$ as

$$p = \begin{cases} \alpha\bar{\alpha} & \text{if } \left(\frac{-1}{p}\right) = 1, \\ p & \text{if } \left(\frac{-1}{p}\right) = -1. \end{cases}$$

We shall see that this result can be extended more generally to the ring of integers $\mathcal{O}_k$ of any quadratic field $\mathbb{Q}[\sqrt{q}]$.

Example 7.6.

Let $K \cup L$ be the two-component link in Figure 23. Since $\mathrm{lk}(L,K) = 2$, K decomposes in the two-sheeted cover X_2 of X_L as $h^{-1}(K) = K_1 \cup K_2$. We can see this pictorially as follows. The knot complement X_L is homeomorphic to a solid torus. (Imagine expanding the two linked solid tori in Figure 24 via a homeomorphism to fill up the ambient space.) The two-sheeted cover X_2 is obtained by slicing X_L along the disc bounded by L and gluing together two sliced copies of X_L, hence it is also a solid torus, with each of the copies of X_L "stretched out" to

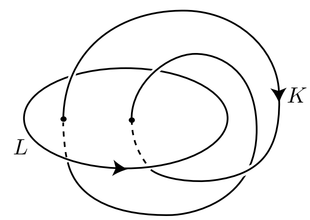


FIGURE 23. 2-component link $K \cup L$ with $\mathrm{lk}(L, K) = 2$.

form half of it, as in Figure 25. (As a guide, the left intersection point of K with Σ_L in Figure 23 lifts to the left intersection point along the gluing boundary on the left in Figure 25, and the right intersection point along the gluing boundary on the right.) Thus, $h^{-1}(K)$ is a Hopf link in X_2.

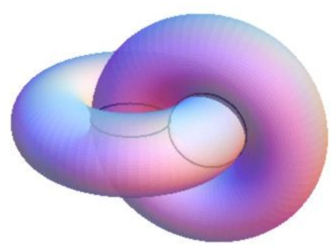


FIGURE 24. The complement of a solid torus in S^3 is homeomorphic to another solid torus.

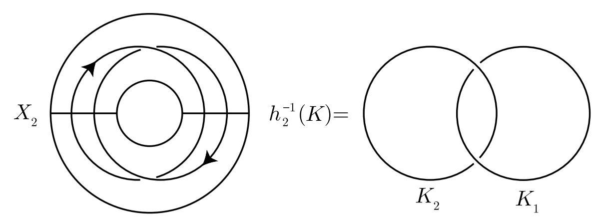


FIGURE 25. h^{-1} is a Hopf link in X_2.

Exercise 7.7. Let $K \cup L$ be the two-component link in Figure 26. What knot or link does K lift to in the two-sheeted cover X_2 of X_L?

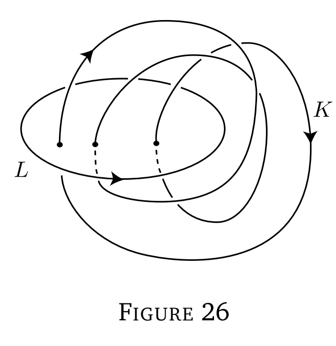


FIGURE 26

Exercise 7.8 (Optional)**.** Find a two-component link $K \cup L$ such that K lifts to a figure eight knot in the two-sheeted cover X_2 of X_L.

Lecture 6. (July 16, 2012)

8. Etale fundamental groups

Now we are in a position to define the (no longer mysterious) notion of étale fundamental groups using all finite étale coverings. In the topological setting, the universal covering $\tilde{X}$ "covers" all other coverings $h : Y \to X$. Moreover, for a fixed $x \in X$, giving a cover $\tilde{X} \to Y$ boils down to a choice of a point in $h^{-1}(x)$. Precisely speaking, we have

Proposition 8.1. *Let X be a topological space and $\tilde{h} : \tilde{X} \to X$ be a universal covering (unique up to isomorphism over X). Fix a point $x \in X$, then for any covering $h : Y \to X$, there is a bijection*

$$\mathrm{Hom}_X(\tilde{X}, Y) \cong h^{-1}(x) = \mathrm{Hom}_X(x, Y).$$

Namely, giving a map from $\tilde{X}$ to Y is the same as choosing a preimage of x on Y.

In other words, the functor $F_x : \mathrm{Hom}_X(x, -)$ is represented by the universal covering $\tilde{X}$. This universal property can be used similarly to define the étale universal covering in the algebraic setting.

Theorem 8.2. *Let $X = \mathrm{Spec} A$ and $x = \mathrm{Spec}\,\Omega \hookrightarrow X$ be a geometric point (i.e., Ω is a separably algebraically closed field). Define the functor*

$$F_x : \{\text{finite étale coverings of } X\} \to \mathbf{Sets}, \quad (h : Y \to X) \mapsto h^{-1}(x) := \mathrm{Hom}_X(x, Y).$$

Then the functor F_x is pro-represented, i.e., there exists an inverse system $(X_i)_{i\in I}$ of finite étale coverings such that

$$\varinjlim_{i\in I} \mathrm{Hom}_X(X_i, Y) \cong \mathrm{Hom}_X(x, Y).$$

Remark 8.3. As discussed before, the functor F_x is not genuinely representable for the reason that the universal covering usually does not exist in the algebraic category.

Definition 8.4. We define $\tilde{X} := \varprojlim_{i\in I} X_i$ and the *étale fundamental group*

$$\pi_1(X, x) = \mathrm{Gal}(\tilde{X}/X) := \varprojlim_{i\in I} \mathrm{Gal}(X_i/X).$$

These are all very nice except that we have not computed a single example of étale fundamental groups. Now we will introduce more algebraic number theory, get into the computation and make more concrete sense of the inverse limit constructions if you are not entirely comfortable with them.

Example 8.5. Let us compute the étale fundamental group of $X = \mathrm{Spec}\,\mathbb{F}_q$ (it is nontrivial even though geometrically $\mathrm{Spec}\,\mathbb{F}_q$ is a single point!). Giving a geometric point x of X is the same as fixing an algebraic closure $\overline{\mathbb{F}_q}$ of $\mathbb{F}_q$. By definition, a finite étale covering of X is a finite disjoint union of $Y = \coprod_{j=1}^m \mathrm{Spec}\, k_j$, where the k_j's are finite extensions of $\mathbb{F}_p$. So

$$\mathrm{Hom}_X(x, Y) = \coprod_j \mathrm{Hom}_X(\mathrm{Spec}\,\overline{\mathbb{F}_q}, \mathrm{Spec}\, k_j).$$

Notice that for each j,

$$\mathrm{Hom}_X(\mathrm{Spec}\,\overline{\mathbb{F}_q}, \mathrm{Spec}\, k_j) = \varinjlim_n \mathrm{Hom}_X(\mathrm{Spec}\,\mathbb{F}_{q^n}, \mathrm{Spec}\, k_j).$$

So the universal covering is

$$\tilde{X} = \varprojlim_n \mathrm{Spec}\,\mathbb{F}_{q^n} = \mathrm{Spec}\,\overline{\mathbb{F}_q}.$$

Recalling that $\mathrm{Gal}(\mathbb{F}_{q^n}/\mathbb{F}_q)$ is the cyclic group $\mathbb{Z}/n\mathbb{Z}$ generated by the Frobenius automorphism $x \mapsto x^q$, we know that the étale fundamental group is

$$\pi_1(\mathrm{Spec}\,\mathbb{F}_q, x) = \mathrm{Gal}(\overline{\mathbb{F}_q}/\mathbb{F}_q) = \varprojlim_n \mathrm{Gal}(\mathbb{F}_{q^n}/\mathbb{F}_q) \cong \varprojlim_n \mathbb{Z}/n\mathbb{Z} =: \hat{\mathbb{Z}}.$$

Concretely,

$$\varprojlim_n \mathbb{Z}/n\mathbb{Z} = \{(a_n)_{n\geq 1} : \phi_{n,m}(a_n) = a_m \text{ for } m \mid n\},$$

where $\phi_{n,m} : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ is the natural quotient map for $m \mid n$. Therefore $\hat{\mathbb{Z}}$ is a gigantic group compared to $\mathbb{Z}$ (e.g., it is uncountable and contains infinitely many copies of $\mathbb{Z}$). The natural map $\mathbb{Z} \to \hat{\mathbb{Z}}$ carries $1 \in \mathbb{Z}$ to the Frobenius automorphism $(\sigma : x \mapsto x^q) \in \mathrm{Gal}(\overline{\mathbb{F}_q}/\mathbb{F}_q)$.

Remark 8.6. $\hat{\mathbb{Z}}$ is an example of a *profinite group*, i.e., an inverse limit of finite groups. By definition, every étale fundamental group is a profinite group. We endow a profinite group with the topology induced from the product topology. So every profinite group is compact by Tychonoff's theorem. The following fact matches our intuition about "completion."

Exercise 8.7. Show that the natural map $\mathbb{Z} \to \hat{\mathbb{Z}}$ is injective and has dense image.

Remark 8.8. In general, for any group G, its normal subgroups of finite index naturally form an inverse system (N_i) and the profinite group

$$\hat{G} := \varprojlim_i G/N_i$$

is called the *profinite completion* of G. For example, the profinite completion of $\mathbb{Z}$ is $\hat{\mathbb{Z}}$. Since the group $\mathbb{Z}$ is not profinite (it is not compact), it can never be an étale fundamental group. Therefore, as the arithmetic counterpart to $\pi_1(S^1) = \mathbb{Z}$, the profinite completion $\hat{\mathbb{Z}}$ may be the best possible candidate we can hope for. Luckily, we already know that $\pi_1(\operatorname{Spec}\mathbb{F}_p) = \hat{\mathbb{Z}}$.

Remark 8.9. A fundamental comparison theorem asserts that for any varieties over $\mathbb{C}$, the profinite completion of its fundamental group (under the complex analytic topology) is the same as its étale fundamental group.

Next time we will use beautiful facts from algebraic number theory to compute the étale fundamental group of $\operatorname{Spec}\mathbb{Z}$ and use our computation to keep exploring the analogy between knots and primes.

Lecture 7. (July 18, 2012)

Last time we computed the étale fundamental group of $\operatorname{Spec}\mathbb{F}_p$ to be $\operatorname{Gal}(\overline{\mathbb{F}_p}/\mathbb{F}_p) = \hat{\mathbb{Z}}$. The same arugment goes for the étale fundamental group of the spectrum of any field.

Example 8.10. In general, for a field F, the fundamental group of $\operatorname{Spec}F$ is the *absolute Galois group* $\operatorname{Gal}(F^s/F)$, where F^s is the separable closure of F (by the same argument). The absolute Galois group of $\mathbb{Q}$ contains monstrous secrets about all number fields and thus is of great interest to study in number theory. People are far from understanding the whole absolute Galois group $\operatorname{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ as of today.

Our next goal is to compute the étale fundamental group $\pi_1(\operatorname{Spec}\mathbb{Z})$. The following algebraic fact allows us to consider only the étale coverings arising from number rings.

Theorem 8.11. *Suppose* $\operatorname{Spec}A$ *is normal (i.e., A is an integrally closed domain). Let K_i be a finite extension of* $\operatorname{Frac}(A)$ *and B_i be the integral closure of A in K_i. Then the universal covering is $\tilde{X} = \varprojlim_i X_i$, where X_i runs over the* $\operatorname{Spec}B_i$ *such that* $\operatorname{Spec}B_i \to \operatorname{Spec}A$ *is finite étale.*

Remark 8.12. Indeed, the same thing is true for an arbitrary normal scheme.

Now applying the theorem to the normal ring $A = \mathbb{Z}$, to determine $\pi_1(\operatorname{Spec}\mathbb{Z})$ it suffices to find all the number rings $\mathcal{O}_K$ that are finite étale over $\mathbb{Z}$. From the point of view of prime decomposition, this means that in the decomposition of each prime

$$(p) = \prod_{i=1}^{m} \mathfrak{p}_i^{e_i}$$

in $\mathcal{O}_K$, all the e_i's are equal to 1.

Definition 8.13. A prime $p \in \mathbb{Z}$ is called *unramified* in $\mathcal{O}_K$ if all the e_i's are equal to 1, and *ramified* otherwise. The number e_i is called the *ramification index* of $\mathfrak{p}_i$.

So the K_i's in the previous theorem are exactly the number fields unramified at each p. If we define the *maximal unramified extension* $\mathbb{Q}^{ur}$ of $\mathbb{Q}$ as the union of all such $K_i's$, then $\pi_1(\operatorname{Spec}\mathbb{Z}) = \operatorname{Gal}(\mathbb{Q}^{ur}/\mathbb{Q})$. To compute $\mathbb{Q}^{ur}$, the key input is to relate the ramification behavior of primes to a numerical invariant of number rings—the discriminant.

Definition 8.14. Let K be a number field of degree n and $\{\alpha_j\}_{j=1}^n$ be a $\mathbb{Z}$-basis of the number ring $\mathcal{O}_K$. Then we define the (absolute) *discriminant* $d_K = \left(\det(\sigma_i(\alpha_j))_{ij}\right)^2$, where σ_i runs over the embeddings $K \hookrightarrow \mathbb{C}$. The discriminant is an integer independent of the choice of $\mathbb{Z}$-basis.

We omit the proof of the following important theorem.

Theorem 8.15. *A prime p is ramified in $\mathcal{O}_K$ if and only if $p \mid d_K$.*

Example 8.16. Let $K = \mathbb{Q}(i)$. We know that $\{1, i\}$ is a $\mathbb{Z}$-basis of $\mathbb{Z}[i]$. So

$$d_K = \left(\det \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix}\right)^2 = (-2i)^2 = -4.$$

The previous theorem tells us that only the prime 2 is ramified in $\mathbb{Z}[i]$, which coincides with what we discovered before.

Exercise 8.17. Let p be an odd prime. Determine which primes are ramified in the cyclotomic field $K = \mathbb{Q}(\zeta_p)$.

So showing that a number field K is unramified everywhere is equivalent to showing that $|d_K| = 1$. Are there any such number fields? Minkowski used his method of geometry of numbers to bound $|d_k|$ from below.

Theorem 8.18 (Minkowski)**.** *Let K be a number field of degree n . Then*

$$|d_K|^{1/2} \geq \frac{n^n}{n!}\left(\frac{\pi}{4}\right)^{n/2}.$$

Corollary 8.19. *If K is a number field other than $\mathbb{Q}$, then $|d_K| > 1$.*

Proof. Minkowski's theorem gives that $|d_K| \geq a_n := \left(\frac{n^n}{n!}\right)^2 \left(\frac{\pi}{4}\right)^n$. The result then follows since $a_2 = \pi^2/4 > 1$ and $a_{n+1} \geq a_n$. □

Combining Theorem 8.15 with the previous corollary, we know that the only number ring unramified everywhere over $\mathbb{Z}$ is $\mathbb{Z}$ itself (thus there is no nontrivial finite étale $\mathbb{Z}$-algebras except $\mathbb{Z}^n$). So the maximal unramified extension $\mathbb{Q}^{ur} = \mathbb{Q}$ and we have

Corollary 8.20. $\pi_1(\operatorname{Spec}\mathbb{Z}) = 1$.

We collect our computation so far as follows. If my (or Charmaine's) word that $\operatorname{Spec}\mathbb{Z}$ is "3-dimensional" can be trusted, then it is natural to believe that $\operatorname{Spec}\mathbb{Z}$ should behave like a "simply connected 3-manifold."

$K : S^1 \hookrightarrow \mathbb{R}^3$	$\operatorname{Spec}\mathbb{F}_p \hookrightarrow \operatorname{Spec}\mathbb{Z}$
$\pi_1(S^1) = \mathbb{Z}$	$\pi_1(\operatorname{Spec}\mathbb{F}_p) = \hat{\mathbb{Z}}$
$\pi_1(\mathbb{R}^3) = 1$	$\pi_1(\operatorname{Spec}\mathbb{Z}) = 1$

We can easily add a new row concerning the knot group $G_K = \pi_1(\mathbb{R}^3 \setminus K)$. The knot group corresponds to the unramified coverings of $\mathbb{R}^3 \setminus K$, so the arithmetic counterpart should correspond to the finite étale coverings of $\operatorname{Spec}\mathbb{Z} \setminus \{p\}$, or equivalently $\operatorname{Spec}\mathbb{Z}[1/p]$ (we can kill the prime ideal (p) by inverting p).

Definition 8.21. We define the *prime group* to be the étale fundamental group

$$G_{\{p\}} := \pi_1(\operatorname{Spec}\mathbb{Z} \setminus \{p\}) = \pi_1(\operatorname{Spec}\mathbb{Z}[1/p]).$$

Even though there are no nontrivial finite étale coverings of the whole space $\operatorname{Spec}\mathbb{Z}$, there do exist finite coverings of $\operatorname{Spec}\mathbb{Z}$ that are étale outside a prime p (e.g., our favorite example $\operatorname{Spec}\mathbb{Z}[i] \to \operatorname{Spec}\mathbb{Z}$ is étale outside 2), so the prime group may be nontrivial.

What is the right analogy for the tubular neighborhood V_K of a knot K? This is not that easily seen and leads to the beautiful idea of completion. It is already quite surprising that we have gone so far away without even mentioning p-adic numbers.

Example 8.22. Consider the complex line $\mathbb{A}^1 = \operatorname{Spec}\mathbb{C}[t]$. The point $\{0\} \hookrightarrow \mathbb{A}^1$ corresponds to the quotient map $\mathbb{C}[t] \to \mathbb{C}[t]/(t) \cong \mathbb{C}$. How do we describe a neighborhood of $\{0\}$ algebraically? Notice that $\mathbb{C}[t] \to \mathbb{C}[t]/(t)$ is nothing but the evaluation map $f \mapsto f(0)$, which only gives the information about the point $\{0\}$. If we would like to remember the first derivative of f, then the quotient map $\mathbb{C}[t] \to \mathbb{C}[t]/(t^2)$ is better. Geometrically, the nilpotent element t adds a bit of "fuzz" to the point along the t direction (we have seen a similar example $\mathbb{Z}[i]/(1+i)^2$ before), so $\operatorname{Spec}\mathbb{C}[t]$ stands for a double point (as the intersection of a parabola $y = t^2$ and the line $y = 0$). In general, the quotient map $\mathbb{C}[t] \to \mathbb{C}[t]/(t^{n+1})$ remembers all derivatives of f up to order n and provides us an order n "fuzz" around the point. Of course there is no reason to stop us at any specific n. So we can take the inverse limit of the inverse system

$$\cdots \to \mathbb{C}[t]/(t^n) \to \mathbb{C}[t]/(t^{n-1}) \to \cdots \to \mathbb{C}[t]/(t),$$

which is exactly the power series ring

$$\varprojlim_n \mathbb{C}[t]/(t^n) = \mathbb{C}[[t]].$$

We call $\mathbb{C}[[t]]$ the *completion* of $\mathbb{C}[t]$ at the prime ideal (t). It provides geometrically an infinitesimal neighborhood of the point $t=0$ to help us read the local information about that point.

Example 8.23. We now mimic the completion process of $\mathbb{C}[t]$ at $t=0$ to give an infinitesimal neighborhood of $\operatorname{Spec}\mathbb{F}_p \hookrightarrow \operatorname{Spec}\mathbb{Z}$. We take the inverse limit of the inverse system

$$\cdots \to \mathbb{Z}/(p^n) \to \mathbb{Z}/(p^{n-1}) \to \cdots \to \mathbb{Z}/(p),$$

and define

$$\mathbb{Z}_p := \varprojlim_n \mathbb{Z}/(p^n).$$

As a group, $\mathbb{Z}_p$ is profinite. Even more, since each of the finite group $\mathbb{Z}/(p^n)$ is of p-power order, $\mathbb{Z}_p$ is a *pro-p group* and it is the *pro-p completion* of $\mathbb{Z}$. Geometrically, $\operatorname{Spec}\mathbb{Z}_p$ should be thought of as an infinitesimal neighborhood of $\operatorname{Spec}\mathbb{F}_p$ encoding all the local information of $\operatorname{Spec}\mathbb{Z}$ at p.

$K: S^1 \hookrightarrow \mathbb{R}^3$	$\operatorname{Spec}\mathbb{F}_p \hookrightarrow \operatorname{Spec}\mathbb{Z}$
$\pi_1(S^1) = \mathbb{Z}$	$\pi_1(\operatorname{Spec}\mathbb{F}_p) = \hat{\mathbb{Z}}$
$\pi_1(\mathbb{R}^3) = 1$	$\pi_1(\operatorname{Spec}\mathbb{Z}) = 1$
$G_K = \pi_1(\mathbb{R}^3 \setminus K)$	$G_{\{p\}} = \pi_1(\operatorname{Spec}\mathbb{Z}[1/p])$
V_K	$\operatorname{Spec}\mathbb{Z}_p$

Definition 8.24. The elements of the ring $\mathbb{Z}_p$ are called *p-adic integers*.

We do not know much about the p-adic integers but the following exercise can be tackled right now.

Exercise 8.25. Show that $\hat{\mathbb{Z}} \cong \prod_p \mathbb{Z}_p$, where p runs over all prime numbers.

In the sequel, we will investigate more basic properties of $\mathbb{Z}_p$, study the prime group $G_{\{p\}}$ using class field theory and reinterpret the Legendre symbol to draw the connection to linking numbers at last.

LECTURE 8. (JULY 20, 2012)

9. p-ADIC INTEGERS AND PRIME GROUPS

As an arithmetic analogue of the tubular neighborhood, last time we defined the ring of p-adic integers to be the inverse limit $\varprojlim_n \mathbb{Z}/(p^n)$. Here is a concrete way to understand the p-adic integers.

Example 9.1. Instead of arranging the integers on a line in the usual order, let us classify them by their residues modulo $p, p^2, \dots, p^n, \dots$. We think that two integers x, y are closer if $x-y$ is more divisible by p. Because a p-adic integer is a sequence (a_n) where $a_{n+1} - a_n$ is divisible by p^n, the a_n's are getting closer and closer when n gets bigger and thus has a "limit" under this p-adic sense of distance. The p-adic integer ring $\mathbb{Z}_p$ simply consists of all these limits and $\mathbb{Z}$ is dense in $\mathbb{Z}_p$. In sum, a p-adic integer is a limit of integers under the p-adic distance.

To make more precise the meaning of the p-adic distance, we shall discuss some basic algebraic properties of $\mathbb{Z}_p$. The following will not surprise you if you are familiar with the power series ring $\mathbb{C}[[t]]$.

Theorem 9.2. *A nonzero element $x \in \mathbb{Z}_p$ can be uniquely written as $x = u \cdot p^k$, where $u \in \mathbb{Z}_p^\times$ is a unit and $k \geq 0$ is an integer.*

Proof. Write $x = (x_n)$ and let k be the largest power of p dividing x (equivalently, $x_k = 0$ but $x_{k+1} \neq 0$). Then we can find a unique element $u = (u_n) \in \mathbb{Z}_p$ such that $x = p^k u$. Moreover, $p \nmid u_n$, so $u_n \in \mathbb{Z}/(p^n)$ is a unit. □

So the arithmetic in $\mathbb{Z}_p$ is much easier than that in $\mathbb{Z}$. In particular,

Corollary 9.3. *$\mathbb{Z}_p$ is a UFD and all its nonzero ideals are of the form (p^k). (0) and (p) are the only two prime ideals.*

Exercise 9.4. The natural map $\mathbb{Z}_p \to \mathbb{Z}/(p^k)$ induces an isomorphism $\mathbb{Z}_p/p^k\mathbb{Z}_p \cong \mathbb{Z}/p^k\mathbb{Z}$.

The ring of p-adic integers $\mathbb{Z}_p$ is like a local version of $\mathbb{Z}$ obtained via throwing away all primes other than p. The usage of p-adic numbers is ubiquitous in modern number theory, reflecting the importance of the local-global point of view. $\mathbb{Z}_p$ has many properties similar to $\mathbb{Z}$ (integrally closed, prime decomposition, Krull dimension 1, etc.), but it is much simpler than $\mathbb{Z}$ (local, complete, discrete valuation, etc.). It allows us to study arithmetic problems by studying one prime at a time and then tying the local information thus obtained together.

The following Hensel's lemma is the crucial property of $\mathbb{Z}_p$ as a result of the completion process.

Theorem 9.5 (Hensel's Lemma)**.** *Let $f(x) \in \mathbb{Z}[x]$ and $\bar{a} \in \mathbb{F}_p$ be a simple root of the reduction $\bar{f}(x) = f(x) \bmod p \in \mathbb{F}_p[x]$. Then $\bar{a}$ lifts to a root $a \in \mathbb{Z}_p$ of $f(x)$.*

Sketch of the proof. The key idea is to produce the solutions modulo p^k inductively from $\bar{a}$. Then taking the limit of these solutions gives a solution in $\mathbb{Z}_p$. □

Corollary 9.6. $\mathbb{Z}_p^\times \cong \mathbb{F}_p^\times \times (1 + p\mathbb{Z}_p)$.

Proof. Consider the quotient map: $\mathbb{Z}_p^\times \to \mathbb{F}_p^\times, a \mapsto a \bmod p$. Clearly it is surjective and has kernel $1 + p\mathbb{Z}_p$. So it suffices to show that the exact sequence

$$1 \to 1 + p\mathbb{Z}_p \to \mathbb{Z}_p^\times \to \mathbb{F}_p^\times \to 1$$

splits, which follows from Hensel's lemma since we can lift each solution of $x^{p-1} = 1$ in $\mathbb{F}_p$ to a solution in $\mathbb{Z}_p$. □

Remark 9.7. The proof shows that all $(p-1)$th roots of unity exist in $\mathbb{Z}_p^\times$.

Now let us turn back to the analogy between knot groups and prime groups, and the linking number and the Legendre symbol. Recall that we can interpret the linking number using covering spaces. Let $L \cup K$ be a link and X_2 be the double covering of the knot complement X_L corresponding to the map $\rho_2 : G_L \to \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$. Then $\rho_2([K]) = \mathrm{lk}(L, K) \bmod 2$.

Let p and q be two odd primes. Let us first work out the arithmetic analogue $\rho_2 : G_{\{q\}} \to \mathbb{Z}/2\mathbb{Z}$. Class field theory classifies all number fields with abelian Galois groups. The following can be derived easily using class field theory.

Theorem 9.8. *The maximal abelian extension of $\mathbb{Q}$ unramified outside p is $\mathbb{Q}(\zeta_{p^\infty}) = \bigcup_n \mathbb{Q}(\zeta_{p^n})$.*

From Exercise 8.17, we know that $\mathbb{Q}(\zeta_{p^\infty})$ satisfies the unramified-outside-p condition and class field theory furthermore ensures that it is the maximal one. It follows that

$$G_{\{q\}}^{\mathrm{ab}} \cong \mathrm{Gal}(\mathbb{Q}(\zeta_{p^\infty}/\mathbb{Q}) = \varprojlim_n (\mathbb{Z}/(p^n))^\times = \mathbb{Z}_p^\times.$$

Using the structure of $\mathbb{Z}_p^\times$, we construct the natural quotient map

$$\rho_2 : G_{\{q\}} \to G_{\{q\}}^{\mathrm{ab}} \to \mathbb{F}_q^\times \to \mathbb{Z}/2\mathbb{Z}.$$

ρ_2 should correspond to a quadratic extension of $\mathbb{Q}$ unramified outside q. What is it? Let us assume for simplicity that q is congruent to 1 modulo 4. Then the natural option $K = \mathbb{Q}(\sqrt{q})$ is in fact a quadratic extension unramified outside q. In fact, its number ring is $\mathcal{O}_K = \mathbb{Z}[\frac{1+\sqrt{q}}{2}]$ with discriminant $d_K = q$. Now we can similarly define the *mod 2 linking number* to be $\mathrm{lk}_2(q, p) := \rho_2(\sigma_p)$, where $\sigma_p \in \mathrm{Gal}(\mathbb{Q}(\sqrt{q})/\mathbb{Q})$ is the Frobenius automorphism associated to p.

Theorem 9.9. $(-1)^{\mathrm{lk}_2(q,p)} = \left(\frac{q}{p}\right)$.

Proof. Notice that $\mathrm{lk}_2(q, p) = 0$ is equivalent to $\rho_2(\sigma_p) = \mathrm{id}$, or $\sigma_p(\sqrt{q}) = \sqrt{q}$. By the definition of the Frobenius automorphism, this is equivalent to $\sqrt{q} \in \mathbb{F}_p^\times$, or $q \in (\mathbb{F}_p^\times)^2$, which happens exactly when $\left(\frac{q}{p}\right) = 1$. □

So the Legendre symbol tells us how primes are "linked" together. This extra structure shows the advantage of viewing primes as knots in a 3-dimensional space rather than plain points on the line. Figure 27 shows five primes “linked” as the Olympic rings.

We summarize the analogy obtained as follows.

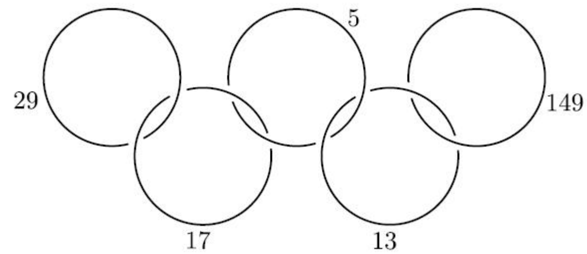


FIGURE 27. Primes 29, 17, 5, 13 and 149 "linked" as the Olympic rings.

G_L	$G_{\{q\}}$
$G_L^{\text{ab}} \cong \mathbb{Z}$	$G_{\{q\}}^{\text{ab}} \cong \mathbb{Z}_q^\times \cong \mathbb{F}_q^\times \times (1+q\mathbb{Z}_p)$
$\rho_2 : G_L \to \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$	$\rho_2 : G_{\{q\}} \to \mathbb{F}_q^\times \to \mathbb{Z}/2\mathbb{Z}$
$\rho_2([K]) = \text{lk}_2(L,K)$	$\rho_2(\sigma_p) = \text{lk}_2(q,p)$
$\text{lk}(L,K) = \text{lk}(K,L)$	$\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right)$
$h_2^{-1}(K) = \begin{cases} K_1 \cup K_2, & \text{lk}_2(L,K) = 0, \\ \mathfrak{K}, & \text{lk}_2(L,K) = 1 \end{cases}$	$(p) = \begin{cases} \mathfrak{p}_1\mathfrak{p}_2, & \left(\frac{q}{p}\right) = 1 \\ \mathfrak{p}, & \left(\frac{q}{p}\right) = -1 \end{cases}$

We finally tie up the beautiful story of the seemingly unrelated linking number and the Legendre symbol. And it took only three weeks. This is so amazing, isn't it? Next time we will start a new story line: the analogy between the Alexander polynomial and Iwasawa theory.

LECTURE 9. (JULY 23, 2012)

We have seen how the knot group, that is, the fundamental group of the knot exterior, gives us a knot invariant, whereas its abelianization, which by the Hurewicz theorem is also the first homology group of the knot exterior, does not help us in distinguishing knots because it is always $\mathbb{Z}$. Although the knot group is an obvious knot invariant, it is not the most useful invariant when studying knots because it is not that easy to work with nonabelian groups. However, we can obtain a polynomial (strictly speaking, Laurent polynomial) knot invariant, the Alexander polynomial, from the first homology group of the infinite cyclic cover of the knot exterior. This knot invariant turns out to have a close analogy with the Iwasawa polynomial in number theory.

10. HOMOLOGY GROUPS

By the Hurewicz theorem, which states that the first homology group of a space is isomorphic to the abelianization of its fundamental group, one could conceivably give a definition of the Alexander polynomial in terms of fundamental groups and abelianizations. However, in order to provide more intuition regarding the Alexander polynomial, we shall first introduce homology groups.

As we shall see, homology is directly related to the decomposition of a space into cells, and may be regarded as an algebraization of the most obvious geometry in a cell structure: how cells of dimension n attach to cells of dimension $n-1$. It provides information about the number of "holes of each dimension" in a space.

Example 10.1. Let us do an example to get an idea of what homology measures. Consider the graph X_1 shown in Figure 28, which consists of two vertices joined by four edges; for the purpose of computing homology, we shall make the edges directed. When studying the fundamental group of X_1, we consider loops formed by sequences of edges, starting and ending at a fixed basepoint. For example, we could first travel along the edge a, then backward along the edge b, to obtain a loop ab^{-1}. A more complicated loop would be $ab^{-1}ac^{-1}ba^{-1}cb^{-1}$. Now, a key feature of the fundamental group is that it is nonabelian in the sense that ab^{-1} is regarded as a different element from $b^{-1}a$ and $ab^{-1}ac^{-1}ba^{-1}cb^{-1}$, which enriches but simultaneously complicates the theory.

The idea of homology is to try to simplify matters by abelianizing. For example, the loops ab^{-1}, $b^{-1}a$ and $ab^{-1}ac^{-1}ba^{-1}cb^{-1}$ are regarded as equal if we make a, b, c and d all commute with each other. Thus, one consequence of abelianizing is that loops are no longer required to have a fixed basepoint; rather, they become *cycles*, without a chosen basepoint.

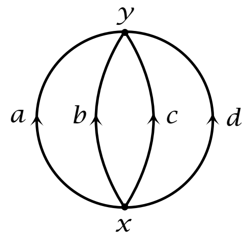


FIGURE 28. Directed graph X_1.

Having abelianized, let us switch to additive notation instead. Then cycles become linear combinations of edges with integer coefficients, for example $a - b + c - b = a - 2b + c$. We call such linear combinations *chains* of edges. For example, $a - b + c$ is a chain that cannot arise from a cycle, whereas the chain $a - 2b + c$ can arise from the loop $ab^{-1}cb^{-1}$. We will call a *cycle* (in the algebraic sense) any chain that can be decomposed into one or more cycles in the previous geometric sense. (Thus, for example, we can also think of the chain $a - 2b + c$ as arising from a disjoint union of the (geometric) cycles $a - b$ and $c - b$.)

A natural question then arises: when is a chain a cycle (in the algebraic sense)? A disjoint union of geometric cycles is characterised by the property that the number of ingoing edges is equal to the number of outgoing edges at every vertex. For an arbitrary chain $ka + lb + mc + nd$, the number of ingoing edges at y is $k + l + m + n$, while the number of outgoing edges is $-k - l - m - n$ (and vice versa at x). Hence the condition for $ka + lb + mc + nd$ to be a cycle is that $k + l + m + n = 0$.

Let us try to describe this result in a way that generalizes to all directed graphs. Let C_1 be the free abelian group on the edges (in this case a, b, c and d) and let C_0 be the free abelian group on the vertices (in this case x and y). The elements of C_1 are chains of edges (1-dimensional chains), while the elements of C_0 are chains of vertices (0-dimensional chains). Define a homomorphism $\partial_1 : C_1 \to C_0$ by

$$\partial_1(\text{edge}) = (\text{vertex at head of edge}) - (\text{vertex at tail of edge})$$

(in this case, ∂_1 sends all four of a, b, c and d to $y - x$). Then the (algebraic) cycles are precisely the kernel of ∂_1. In our example, one can easily check that the kernel of ∂_1 is generated by the three cycles $a - b$, $b - c$ and $c - d$. This conveys the information that the graph X_1 has three visible "holes" bounded by these cycles.

Let us now enlarge the graph X_1 by attaching a 2-cell, that is, an open disc A along the cycle $a - b$ to produce a 2-dimensional *CW complex* or *cell complex*, as in Figure 29. (We shall not define a CW complex rigorously, but it should be clear from this construction what is allowed.) If we think of the 2-cell A as being oriented clockwise, then we can think of its boundary as the cycle $a - b$. This cycle is now homotopically trivial since we have filled in the "hole" bounded by the cycle $a - b$. This suggests that we should form a quotient group of cycles by modding out by the subgroup generated by $a - b$. In this quotient the cycles $a - c$ and $b - c$ are equivalent, which is consistent with the fact that they are homotopic in X_2.

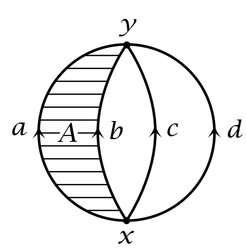


FIGURE 29. Directed graph X_2 obtained by attaching a 2-cell to X_1.

Once again, we can describe this result in a way that generalizes to all 2-dimensional cell complexes as follows. Let C_2 be the free abelian group generated by the 2-cells, and define ∂_2 to be the homomorphism taking a cell to its boundary (in this case, $\partial_2(A) = a - b$). We thus have a sequence of homomorphisms $C_2 \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0$, and the quotient group that we are interested in is $\ker \partial_1 / \operatorname{im} \partial_2$, the 1-dimensional cycles modulo boundaries. This quotient group is the first *homology group* $H_1(X_2)$.

We can continue this procedure by considering CW complexes obtained by adding cells in higher dimensions. It is clear what the general pattern should be: for a CW complex X, one has chain groups $C_n(X)$ which are the free abelian groups generated by the n-cells (i.e. open n-discs) of X, and homomorphisms $\partial_n : C_n(X) \to C_{n-1}(X)$ that are defined by linearity and the property that a cell is sent to its boundary.

Definition 10.2. The nth *homology group* $H_n(X)$ is the quotient group $\ker \partial_n / \operatorname{im} \partial_{n+1}$.

We call chains in the kernel of ∂_n cycles and chains in the image of ∂_{n+1} boundaries, so one should think of homology as the quotient group of cycles modulo boundaries; as in the 1-dimensional example, it should be the case that boundaries are always cycles. The major difficulty lies in defining ∂_n in general. We have seen how ∂_1 and ∂_2 should be defined, but it is less clear how to define ∂_n for $n \geq 3$. One solution to this problem is to consider CW complexes built from simplices (the interval, triangle and tetrahedron are the 1-, 2- and 3-dimensional instances of simplices respectively); these are called Δ-*complexes*. In this case, there is a easy formula for the boundary map ∂_n, and the homology groups thus defined are called *simplicial homology groups*. The drawback of this approach, however, is that this is a rather restrictive class of spaces. (For instance, the CW complex structure in Example 10.1 is not a Δ-complex structure. The type of homology we have computed in Example 10.1 is called *cellular homology*, which turns out to be equivalent to simplicial homology.) Moreover, an obvious question arises: given two different CW or Δ-complex structures on X, do they give rise to isomorphic homology groups? To address this problem, one introduces *singular homology groups*, which are defined for all spaces X, not just CW or Δ-complexes, and then shows that the cellular, simplicial and singular homology groups coincide whenever defined. We shall not define singular homology groups, since we are more interested in the *properties* of homology groups instead of their various equivalent definitions. (It turns out that singular homology can be characterized by a set of axioms consisting of the main properties it satisfies.) Here we simply state some of the main properties of homology groups that one can deduce using singular homology:

- The homology groups $H_n(X)$ only depend on the space X and not on the CW or Δ-complex structure.
- A map of spaces $f : X \to Y$ induces a homomorphism of homology groups $f_* : H_n(X) \to H_n(Y)$ such that $(fg)_* = f_* g_*$ and $\mathrm{id}_* = \mathrm{id}$; in particular, homeomorphic spaces have isomorphic homology groups. In fact, homotopy equivalent spaces have isomorphic homology groups.
- If $f \simeq g : X \to Y$, then $f_* = g_* : H_n(X) \to H_n(Y)$.
- A pair of topological spaces (X, A), $A \subseteq X$ induces a long exact sequence in homology via the inclusions $i : A \to X$ and $j : (X, \emptyset) \to (X, A)$ (we can think of a space X as a pair $(X, \emptyset)$):

$$\cdots \to H_1(X) \xrightarrow{j_*} H_1(X, A) \to H_0(A) \xrightarrow{i_*} H_0(X) \xrightarrow{j_*} H_0(X, A) \to 0.$$

  (Here $H_n(X, A)$ is a *relative homology group*. This is defined as follows: the *relative chains* $C_n(X, A)$ are the quotient group $C_n(X)/C_n(A)$ with relative boundary map $\partial_n : C_n(X, A) \to C_{n-1}(X, A)$ induced from $\partial_n : C_n(X) \to C_{n-1}(X)$. Unraveling $\ker \partial_n$ and $\operatorname{im} \partial_{n+1}$, we see that the *relative cycles* are (equivalence classes of) n-chains $\alpha \in C_n(X)$ such that $\partial \alpha \in C_{n-1}(A)$, and the *relative boundaries* are (equivalence classes of) n-chains $\alpha = \partial \beta + \gamma$ for some $\beta \in C_{n+1}(X)$ and $\gamma \in C_n(A)$. If A and X are CW complexes with $A \neq \emptyset$ a subcomplex of X, one can show that $H_n(X, A) = H_n(X/A)$ for $n \geq 1$ and $H_0(X, A) \oplus \mathbb{Z} \cong H_0(X/A)$.)

11. Skein relation definition of the Alexander polynomial

There are several equivalent ways of defining the Alexander polynomial. The most elementary way is in terms of a skein relation, that is, a relationship between three knot diagrams that are identical except at the same one crossing. Specifically, the Alexander polynomial $\Delta(t)$ is defined by the following rules:

- $\Delta_{\text{unknot}}(t) = 1$.
- Given three (oriented) links L_+, L_- and L_0 that are identical except as depicted in Figure 30 at one particular crossing, the Alexander polynomials of the links L_+, L_- and L_0 satisfy the relation

$$\Delta_{L_+}(t) - \Delta_{L_-}(t) + (t^{1/2} - t^{-1/2})\Delta_{L_0}(t) = 0.$$

Example 11.1 (Alexander polynomial of trefoil). Treating the trefoil as L_+, with the circled crossing as the one that appears in Figure 30, we have

Δ ( ) $-\, \Delta$ ( ) $+ (t^{1/2} - t^{-1/2})\Delta$ ( ) $= 0,$

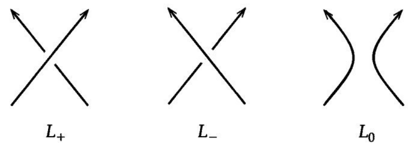
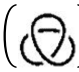
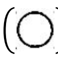


FIGURE 30. Variations in crossing of L_+, L_- and L_0.

where

Δ () $= \Delta$ () $= 1$

and

Δ () $- \Delta$ () $+ (t^{1/2} - t^{-1/2})\Delta$ () $= 0.$

Exercise 11.2. Show that Δ () $= 0$.

Hence

Δ () $= -t^{1/2} + t^{-1/2}$

and thus

Δ () $= t - 1 + t^{-1}.$

This definition in terms of a skein relation allows one to express both the Alexander polynomial (discovered surprisingly early in 1928) and the Jones polynomial (discovered much later in 1984) as special cases of a more general polynomial knot invariant, the HOMFLY polynomial (discovered almost immediately after in 1985). In general, knot invariants that can be defined in terms of a skein relation are usually quantum invariants of interest to quantum topologists. However, to establish the analogy between the Alexander polynomial and the Iwasawa polynomial, we shall use a different definition of the Alexander polynomial, in terms of homology. Historically, the Alexander polynomial was first formulated by Alexander in terms of homology, although he also showed that the Alexander polynomial satisfies a similar skein relation; Conway later rediscovered the skein relation in a different form and proved that the skein relation together with a choice of value for the unknot suffice to determine the polynomial.

LECTURE 10. (JULY 25, 2012)

12. THE ALEXANDER POLYNOMIAL IN TERMS OF HOMOLOGY

Let K be a knot in S^3. We have noted that the first homology group $H_1(X_K)$ of the knot exterior X_K is not useful in distinguishing knots, since it is always isomorphic to $\mathbb{Z}$. However, the situation is different if we consider the infinite cyclic covering $h_\infty : X_\infty \to X_K$: since $\mathrm{Gal}(X_\infty/X_K) \cong \mathbb{Z}$, we see that h_∞ corresponds to the abelianization $\psi : G_K \to G_K^{\mathrm{ab}} \cong \mathbb{Z} \cong \mathrm{Gal}(X_\infty/X_K)$, so $\pi_1(X_\infty) \cong \ker\psi \cong [G_K, G_K]$ and $H_1(X_\infty) \cong [G_K, G_K]^{\mathrm{ab}}$ is in general not a constant group.

We can attempt to understand the group $H_1(X_\infty)$ by considering the long exact sequence in homology associated to the pair $(X_\infty, h_\infty^{-1}(x_0))$, where x_0 is a fixed basepoint in X_K:

$$\cdots \to H_1(h_\infty^{-1}(x_0)) \to H_1(X_\infty) \to H_1(X_\infty, h_\infty^{-1}(x_0)) \to H_0(h_\infty^{-1}(x_0)) \to H_0(X_\infty) \to H_0(X_\infty, h_\infty^{-1}(x_0)).$$

- $\underline{H_1(h_\infty^{-1}(x_0))}$: since $h_\infty^{-1}(x_0)$ consists of a discrete set of points, $H_1(h_\infty^{-1}(x_0)) = 0$.
- $\underline{H_0(h_\infty^{-1}(x_0))}$: $H_0(h_\infty^{-1}(x_0))$ is the free abelian group on the set of points $h_\infty^{-1}(x_0)$, which are indexed by elements of $G_K^{\mathrm{ab}} \cong \mathbb{Z}$, so $H_0(h_\infty^{-1}(x_0)) \cong \mathbb{Z}[G_K^{\mathrm{ab}}] \cong \mathbb{Z}[t, t^{-1}] =: \Lambda$. (Here t corresponds to the class of a meridian α that generates G_K^{ab}.)
- $\underline{H_0(X_\infty)}$: since X_∞ is connected, $H_0(X_\infty) \cong \mathbb{Z}$.
- $\underline{H_0(X_\infty, h_\infty^{-1}(x_0))}$: by the CW approximation theorem, we can approximate the pair $H_0(X_\infty, h_\infty^{-1}(x_0))$ by a CW pair (X, A) that is weakly homotopy equivalent and which will thus have the same homology groups. But $H_0(X_\infty, h_\infty^{-1}(x_0)) \oplus \mathbb{Z} \cong H_0(X, A) \oplus \mathbb{Z} \cong H_0(X/A) \cong Z$ since X/A is connected, so $H_0(X_\infty, h_\infty^{-1}(x_0)) = 0$.

- $\underline{H_1(X_\infty, h_\infty^{-1}(x_0))}$: here we use the fact that relative cycles are (equivalence classes of) chains in $C_n(X_\infty)$ whose boundary lies in $h_\infty^{-1}(x_0)$. For $g = [l] \in G_K$, let $\tilde{l}$ denote the lift of l with starting point $y_0 \in h_\infty^{-1}(x_0)$. Then $\tilde{l} \in C_1(X_\infty, h_\infty^{-1}(x_0))$ and we have a map
$$d : G_K \to H_1(X_\infty, h_\infty^{-1}(x_0)), \quad d(g) := [\tilde{l}].$$
For $g_1 = [l_1]$, $g_2 = [l_2] \in G_K$, let $\tilde{l_1}$, $\tilde{l}_2$ and $\widetilde{l_1 \cdot l_2}$ be the lifts of l_1, l_2 and $l_1 \cdot l_2$ respectively with starting point y_0, and let $\tilde{l}_2'$ be the lift of l_2 whose starting point is the ending point of l_1. Then $d(g_1 g_2) = [\widetilde{l_1 \cdot l_2}] = [\tilde{l}_1] + [\tilde{l}_2']$ and $d(g_1) + \psi(g_1) d(g_2) = [\tilde{l}_1] + \psi(g_1)[\tilde{l}_2] = [\tilde{l}_1] + [\tilde{l}_2']$. Hence $d(g_1 g_2) = d(g_1) + \psi(g_1) d(g_2)$ in $H_1(X_\infty, h_\infty^{-1}(x_0))$.

Moreover, each of the above are (left) $\mathbb{Z}[G_K^{\mathrm{ab}}]$-modules. This is an appropriate time to pause and introduce the following definition.

Definition 12.1. Let A_K be the quotient module of the (left) free $\mathbb{Z}[G_K^{\mathrm{ab}}]$-module $\bigoplus_{g \in G_K} \mathbb{Z}[G_K^{\mathrm{ab}}] dg$ on the symbols dg $(g \in G_K)$ by the (left) $\mathbb{Z}[G_K^{\mathrm{ab}}]$-submodule generated by elements of the form $d(g_1 g_2) - dg_1 - \psi(g_1) d(g_2)$ for g_1, $g_2 \in G_K$:
$$A_K := \left(\bigoplus_{g \in G_K} \mathbb{Z}[G_K^{\mathrm{ab}}] dg \right) / \langle d(g_1 g_2) - dg_1 - \psi(g_1) d(g_2) \, (g_1, g_2 \in G_K) \rangle_{\mathbb{Z}[G_K^{\mathrm{ab}}]}.$$

By definition, the map $d : G \to A_K$ defined by the correspondence $g \mapsto dg$ is a ψ-differential, namely for g_1, $g_2 \in G$, one has
$$d(g_1 g_2) = d(g_1) + \psi(g_1) d(g_2),$$
and A_K is universal for this property in the sense that for any (left) $\mathbb{Z}[G_K^{\mathrm{ab}}]$-module A and any ψ-differential $\partial : G \to A$, there exists a unique $\mathbb{Z}[G_K^{\mathrm{ab}}]$-homomorphism $\varphi : A_K \to A$ such that $\varphi \circ d = \partial$.

Fact 12.2. There is an exact sequence of (left) $\mathbb{Z}[G_K^{\mathrm{ab}}]$-modules
$$0 \to [G_K, G_K]^{\mathrm{ab}} \xrightarrow{\theta_1} A_K \xrightarrow{\theta_2} \mathbb{Z}[G_K^{\mathrm{ab}}] \xrightarrow{\epsilon_{\mathbb{Z}[G_K^{\mathrm{ab}}]}} \mathbb{Z} \to 0$$
called the *Crowell exact sequence*, where θ_1 is the homomorphism induced by $n \mapsto dn$ $(n \in [G_K, G_K])$, θ_2 is the homomorphism induced by $dg \mapsto \psi(g) - 1$ $(g \in G_K)$ and $\epsilon_{\mathbb{Z}[G_K^{\mathrm{ab}}]}(\sum a_g g) := \sum a_g$.

One can check that the isomorphisms above commute with the maps in the Crowell exact sequence and the homology exact sequence, and hence that the Crowell exact sequence is simply the homology exact sequence in another guise.

Definition 12.3. The $\mathbb{Z}[G_K^{\mathrm{ab}}]$-module module A_K is called the *Alexander module* of the knot K.

Using the Fox free differential calculus, one can give an explicit resolution for the Alexander module A_K.

Definition 12.4. Let F be the free group on the generators $x_1, x_2, \ldots, x_m$. The *Fox free derivative* $\frac{\partial}{\partial_{x_i}} : \mathbb{Z}[F] \to \mathbb{Z}[F]$ is defined by the following axioms:

- $\dfrac{\partial}{\partial x_i}(u + v) = \dfrac{\partial}{\partial x_i} u + \dfrac{\partial}{\partial x_i} v$ for any $u, v \in \mathbb{Z}[F]$,
- $\dfrac{\partial}{\partial x_i} e = 0$,
- $\dfrac{\partial}{\partial x_i} x_j = \partial_{ij}$ where ∂_{ij} is the Kronecker delta,
- $\dfrac{\partial}{\partial x_i}(uv) = \dfrac{\partial}{\partial x_i} u + u \dfrac{\partial}{\partial x_i} v$ for any $u, v \in F$.

One can check that this system of axioms is consistent. As a consequence of the axioms, we also have the following formula for inverses:
$$\frac{\partial}{\partial x_i} u^{-1} = -u^{-1} \frac{\partial}{\partial x_i} u \quad \text{for any } u \in F.$$

Theorem 12.5. *Let $G_K = \langle x_1, \dots, x_m \mid R_1, R_2, \dots, R_{m-1}\rangle$ be a presentation of the knot group G_K (e.g. a Wirtinger presentation). Let F be the free group on x_1, …, x_m, and let $\pi : F \to G_K$ be the natural homomorphism. (We shall also denote by π the induced map of group rings $\mathbb{Z}[F] \to \mathbb{Z}[G_K]$.) The Alexander module A_K has a free resolution over $\mathbb{Z}[G_K^{ab}]$:*

$$\mathbb{Z}[G_K^{ab}]^{m-1} \xrightarrow{Q} \mathbb{Z}[G_K^{ab}]^m \to A_K \to 0.$$

Here the $(m-1) \times m$ presentation matrix Q_K, called the Alexander matrix *of K, is given by*

$$Q_K = \left((\psi \circ \pi) \left(\frac{\partial R_i}{\partial x_j} \right) \right)_{ij} \in \mathbb{Z}[G_K^{ab}]^{(m-1)\times m} \cong \Lambda^{(m-1)\times m}.$$

Note that the Alexander matrix depends on a choice of presentation for G_K!

Example 12.6 (Alexander matrix for the trefoil)**.** Let us compute Alexander matrices for the trefoil using two different presentations of the knot group: $\langle a, b \mid aba - bab\rangle$ and $\langle x, y \mid x^3 - y^2\rangle$.

- $\langle a, b \mid aba - bab\rangle$:
$$\frac{\partial}{\partial a} = 1 + ab - b, \quad \frac{\partial}{\partial b} = -1 - ba + a,$$
so the associated Alexander matrix is $[\psi(1 + ab - b) \quad \psi(-1 - ba - a)] = [1 + t^2 - t \quad -1 - t^2 + t]$. (Abelianizing sends both a and b to t.)
- $\langle x, y \mid x^3 - y^2\rangle$:
$$\frac{\partial}{\partial x} = 1 + x + x^2, \quad \frac{\partial}{\partial y} = -1 - y,$$
so the associated Alexander matrix is $[\psi(1+x+x^2) \quad \psi(-1-y)] = [1+t^2+t^4 \quad -1-t^3]$. (Here one needs to be careful: abelianizing sends x to t^3 and y to t^2!)

For a commutative ring R and a finitely presented R-module M, let

$$R^s \xrightarrow{Q} R^r \to M \to 0$$

be a free resolution of M over R with presentation matrix Q, and define $E(M)$ to be the ideal of R generated by the maximal minors of Q.

Definition 12.7. The *Alexander ideal* is the ideal $E(A_K)$ of $\mathbb{Z}[G_K^{\mathrm{ab}}] \cong \Lambda$ generated by the $(m-1)$-minors of Q_K.

Note that the Alexander ideal can be defined because G_K^{ab} is abelian by definition. It is a theorem of Crowell and Fox that the Alexander ideal is independent of a choice of a free resolution for A_K, unlike the Alexander matrix. Moreover, Alexander proved that the Alexander ideal is always a principal ideal (although $\mathbb{Z}[t, t^{-1}]$ is *not* a PID!). Thus a generator of the Alexander ideal is defined up to multiplication by a unit of Λ, namely $\pm t^n$ for some integer n.

Example 12.8 (Alexander ideal for the trefoil)**.** It is clear from the first part of Example 12.6 that the Alexander ideal of the trefoil is $\langle 1 - t + t^2\rangle$. We can also see this from the second part of Example 12.6 by the calculations $1 + t^2 + t^4 = (1 - t + t^2)(1 + t + t^2)$, $(1 + t^3) = (1 - t + t^2)(1 + t)$.

Definition 12.9. The *Alexander polynomial* $\Delta_K(t)$ of a knot K is a generator of $E(A_K)$ (and hence is defined up to multiplication by $\pm t^n$ for some integer n).

Example 12.10. The Alexander polynomial of the trefoil is $\Delta_K(t) = 1 - t + t^2$. Note that this matches our computation using skein relations up to a factor of t.

Exercise 12.11. Compute the Alexander polynomial of the figure eight knot, using both skein relations and the Fox free derivative.

Working through the definition of the Alexander ideal using a Wirtinger presentation for K, one can show that the effect of taking the mirror image of a knot K is to make the substitution $t \longleftrightarrow t^{-1}$. On the other hand, by the theorem of Crowell and Fox and the definition in terms of the first homology group of X_∞, the Alexander ideal does not depend on the chirality of K. Hence we conclude that the Alexander polynomial is symmetric, that is, up to multiplication by t^n, it has the form $a_r t^{-r} + \cdots + a_1 t^{-1} + a_0 + a_1 t + \cdots + a_r t^r$.

Lecture 11. (July 27, 2012)

13. Class field theory

In Lecture 8, we learned that $\mathbb{Z}_p$ consists of limits of integers under the "p-adic" distance. Let us make this more precise.

Definition 13.1. Let $\mathbb{Q}_p$ be the fraction field of $\mathbb{Z}_p$. The elements of the field $\mathbb{Q}_p$ are called *p-adic numbers*. Every nonzero p-adic number is of the form $u \cdot p^k$, where $k \in \mathbb{Z}$ and $u \in \mathbb{Z}_p^\times$.

Definition 13.2. We define the *p-adic valuation*

$$v_p : \mathbb{Q}_p \to \mathbb{Z} \cup \infty$$

by $v_p(x) = k$ if $x = u \cdot p^k$ and $v_p(0) = \infty$, and the *p-adic metric* $|\cdot|_p : \mathbb{Q}_p \to \mathbb{R}$ by $|x|_p = p^{-v_p(x)}$.

One can easily show that $|\cdot|_p$ satisfies the axioms for a metric, thus the p-adic distance $d(x,y) := |x-y|_p$ makes sense. This distance matches our previous intuition: the higher the power of p dividing $x-y$, the "closer" x and y are.

Remark 13.3. $\mathbb{Q}_p$ is nothing but the completion of $\mathbb{Q}$ under the p-adic metric, just as $\mathbb{R}$ is the completion of $\mathbb{Q}$ under the usual Euclidean metric. But unlike $\mathbb{R}$, the topology of $\mathbb{Q}_p$ is totally disconnected: its connected components are one-point sets.

This completion process works in general for any number field.

Definition 13.4. Let K be a number field and $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$ lying above p. We denote by $\mathcal{O}_\mathfrak{p}$ the *$\mathfrak{p}$-adic completion* of $\mathcal{O}_K$ at $\mathfrak{p}$ and $K_\mathfrak{p} = \mathrm{Frac}(\mathcal{O}_\mathfrak{p})$ its fraction field. $K_\mathfrak{p}$ is then a p-adic field, i.e., a finite extension of $\mathbb{Q}_p$.

Analogous to $\mathbb{Q}_p$, each element of the field $K_\mathfrak{p}$ can be written as $x = u \cdot \pi^n$, where π is called a *uniformizer* and $u \in \mathcal{O}_\mathfrak{p}^\times$. The p-adic fields are surprisingly useful in modern number theory. It turns out that all inequivalent metrics on a number field K are divided into two classes: either a $\mathfrak{p}$-adic metric coming from $K \hookrightarrow K_\mathfrak{p}$, or a usual metric coming from an embedding $K \hookrightarrow \mathbb{R}$ or $K \hookrightarrow \mathbb{C}$. So unsurprisingly we decide to give them a name.

Definition 13.5. Let K be a number field. An prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$ is called a *finite place* (or *non-archimedean place*) of K. An embedding $K \hookrightarrow \mathbb{R}$ or $K \hookrightarrow \mathbb{C}$ is called an *infinite place* (or *archimedean place*) of K. A complex embedding $K \hookrightarrow \mathbb{C}$ and its complex conjugate are viewed as the same complex place.

By definition, the finite places of $\mathbb{Q}$ are the prime numbers $v = 2, 3, 5, \ldots$ and the only infinite place (usually denoted by $v = \infty$) of $\mathbb{Q}$ is the embedding $\mathbb{Q} \hookrightarrow \mathbb{R}$. Just as we study a space locally by studying the neighborhoods of its points, the fields $K_v = \mathbb{R}$, $\mathbb{C}$ or $K_\mathfrak{p}$ encode all the local information of the global object—the number field K. For this reason, the terminology *local fields* and *global fields* are commonly used.

Our next goal is to state the main results of class field theory and draw some important consequences out of it. We will proceed in two steps. The first step is local class field theory, which classifies all the abelian extensions of a p-adic field. Then binding the local information at all finite and infinite places appropriately gives the structure of all abelian extensions of a number field, which is the content of global class field theory.

Recall that $\mathrm{Spec}\,\mathbb{Z}_p$ is constructed as a tubular neighborhood of $\mathrm{Spec}\,\mathbb{F}_p$, so we would expect that they have the same étale fundamental groups.

Theorem 13.6. $\pi_1(\mathrm{Spec}\,\mathbb{Z}_p) \cong \pi_1(\mathrm{Spec}\,\mathbb{F}_p) = \hat{\mathbb{Z}}$.

In terms of field theory, this can be rephrased as $\mathrm{Gal}(\mathbb{Q}_p^{\mathrm{ur}}/\mathbb{Q}_p) \cong \mathrm{Gal}(\overline{\mathbb{F}_p}/\mathbb{F}_p)$, namely there is a bijection between unramified finite extensions of $\mathbb{Q}_p$ and finite extensions of the residue field $\mathbb{F}_p$. The same thing is true for any p-adic field K, thus at least we understand the structure of the subfield K^{ur} of the maximal abelian extension K^{ab} of K: $\mathrm{Gal}(K^{\mathrm{ur}}/K) \cong \hat{\mathbb{Z}}$. Now we are in a position to state the main theorem of local class field theory.

Theorem 13.7 (Local class field theory)**.** *Let K be a p-adic field.*

- *(Local reciprocity) There exists a unique homomorphism (called the* local reciprocity map*)* $\rho_K : K^\times \to \mathrm{Gal}(K^{\mathrm{ab}}/K)$ *satisfying:*

(1) *The following diagram commutes:*

$$\begin{array}{ccc} K^\times & \xrightarrow{\rho_K} & \mathrm{Gal}(K^{\mathrm{ab}}/K) \\ \downarrow{\scriptstyle v} & & \downarrow \\ \mathbb{Z} & \longrightarrow & \mathrm{Gal}(K^{\mathrm{ur}}/K) \cong \hat{\mathbb{Z}}, \end{array}$$

where v is the valuation map.

(2) *For any finite abelian extension L/K, ρ_K induces an isomorphism*

$$K^\times/\mathbb{N}_{L/K}L^\times \cong \mathrm{Gal}(L/K),$$

where $\mathbb{N}_{L/K} : L^\times \to K^\times$ is the norm map.

- *(Existence) There is a bijection*

$$\{\text{subgroups of finite index of } K^\times\} \longleftrightarrow \{\text{finite abelian extensions of } K\}$$

given by $\mathbb{N}_{L/K}L^\times \leftarrow\!\shortmid L$.

Remark 13.8. Local class field theory tells us that the finite abelian extensions of K are essentially classified by the intrinsic group structure of $K^\times$! Using the local reciprocity law, we know that

$$\mathrm{Gal}(K^{\mathrm{ab}}/K) \cong \varprojlim_L \mathrm{Gal}(L/K) \cong \varprojlim_L K^\times/\mathbb{N}_{L/K}L^\times.$$

This is the same as the profinite completion of $K^\times$ by the existence theorem. Since

$$K^\times \cong \mathcal{O}^\times \times \pi^{\mathbb{Z}} \cong \mathcal{O}^\times \times \mathbb{Z},$$

we are able to conclude that

Corollary 13.9. $\mathrm{Gal}(K^{ab}/K) \cong \mathcal{O}^\times \times \hat{\mathbb{Z}}$.

Example 13.10. $\mathrm{Gal}(\mathbb{Q}_p^{\mathrm{ab}}/\mathbb{Q}_p) \cong \mathbb{Z}_p^\times \times \hat{\mathbb{Z}}$. Moreover, an extension $L/\mathbb{Q}_p$ is unramified if and only if $\rho_{\mathbb{Q}_p}(\mathbb{Z}_p^\times) = 1$ in $\mathrm{Gal}(L/\mathbb{Q}_p)$.

Example 13.11. Let p be an odd prime. We can further determine all the quadratic extensions of $\mathbb{Q}_p$. By local class field theory, a quadratic extension corresponds to a subgroup of $\mathbb{Q}_p^\times$ of index 2. Since such a subgroup must contain $(\mathbb{Q}_p^\times)^2$, it is equivalent to finding all subgroups of $\mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2$ of index 2. Notice that

$$\mathbb{Q}_p^\times \cong \mathbb{Z}_p^\times \times \mathbb{Z} \cong \mathbb{F}_p^\times \times (1 + p\mathbb{Z}_p) \times \mathbb{Z} \cong \mathbb{Z}/(p-1) \times \mathbb{Z}_p \times \mathbb{Z},$$

where in the second equality the exponential map induces an isomorphism between $\mathbb{Z}_p$ and $(1 + p\mathbb{Z}_p)$. We have $\mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2 \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, which has exactly three subgroups of index 2. So there are exactly three quadratic extensions of $\mathbb{Q}_p$.

Exercise 13.12. Show that $\mathbb{Q}_p(\sqrt{c})$, $\mathbb{Q}_p(\sqrt{p})$ and $\mathbb{Q}_p(\sqrt{cp})$ are three non-isomorphic quadratic extensions, where c is any quadratic non-residue modulo p.

Now we intend to bind all local fields associated to a number field K. A natural option is to take the direct product $\prod_v K_v^\times$, where v runs over all places. But it turns out to be too huge to deal with arithmetic problems. For example, any element of K only has nonzero v-valuation at finitely many places v.

Definition 13.13. We define the *idèle group* J_K to be the subgroup of $\prod_v K_v^\times$ consisting of elements (x_v) such that x_v has nonzero v-adic valuation for only finitely many finite places v (in other words, all but finitely many x_v lie in $\mathcal{O}_v^\times$). So $K^\times$ naturally sits inside J_K. We define the *idèle class group* to be the quotient group $C_K := J_K/K^\times$.

Exercise 13.14. Show that $C_{\mathbb{Q}} \cong \mathbb{R}_{>0} \prod_p \mathbb{Z}_p^\times$.

The idele class group C_K is the group encoding both local and global information and turns out to be the right object characterizing the abelian extensions of K.

Theorem 13.15 (Global class field theory)**.** *Let K be a number field.*

- *(Global reciprocity) There exists a unique continuous homomorphism (called the* global reciprocity map*) $\rho_K : C_K \to \mathrm{Gal}(K^{\mathrm{ab}}/K)$ satisfying:*

(1) *The following diagram commutes (compatibility with local reciprocity):*

$$\begin{array}{ccc} K_v & \xrightarrow{\rho_{K_v}} & \mathrm{Gal}(K_v^{\mathrm{ab}}/K_v) \\ \downarrow & & \downarrow \\ C_K & \xrightarrow{\rho_K} & \mathrm{Gal}(K^{\mathrm{ab}}/K) \end{array}$$

(2) *For any finite abelian extension L/K, ρ_K induces an isomorphism*

$$C_K/\mathbb{N}_{L/K}(C_L) \cong \mathrm{Gal}(L/K),$$

where $\mathbb{N}_{L/K}(x)_v = \prod_{w \text{ above } v} \mathbb{N}_{L_w/K_v}(x_w)$.

- *(Existence) There is a bijection*

$$\{\textit{open subgroups of finite index of } C_K\} \longleftrightarrow \{\textit{finite abelian extensions of } K\}$$

given by $\mathbb{N}_{L/K} C_L \leftarrow\!\shortmid L$.
- *ρ_K is surjective and the kernel is the connected component of 1 in C_K.*

Example 13.16. The connected component of 1 in $C_{\mathbb{Q}}$ is $\mathbb{R}_{>0}$, therefore global class field theory gives $\mathrm{Gal}(\mathbb{Q}^{\mathrm{ab}}/\mathbb{Q}) \cong \prod_p \mathbb{Z}_p^\times$. Moreover, an extension $L/\mathbb{Q}$ is unramified at p if and only $\rho_K(\mathbb{Z}_p^\times) = 1$ in $\mathrm{Gal}(L/\mathbb{Q})$. In particular, we know that the maximal unramified extension of $\mathbb{Q}$ unramified outside p has Galois group $\mathbb{Z}_p^\times$ (the prime group as promised last time).

The name of reciprocity maps hints at a possible relation with quadratic reciprocity. This is the case and quadratic reciprocity can be easily deduced from global class field theory. Indeed, generalizing quadratic reciprocity and other higher reciprocity laws is one of the major motivations for developing class field theory historically. The name of the idèle class group hints at a possible relation with the ideal class group. This is also the case. We will justify this point and then introduce the basics of Iwasawa theory next time.

Lecture 12. (July 30, 2012)

14. The Alexander polynomial in terms of homology, continued

Last time, we defined the Alexander ideal $E(A_K)$ to be the (principal) ideal generated by the maximal minors of a presentation matrix for the Alexander module A_K, and the Alexander polynomial to be a generator of this ideal. Here is another way to understand the Alexander polynomial. Recall that we have a Crowell exact sequence

$$0 \to H_1(X_\infty) \to A_K \xrightarrow{\theta_2} \mathbb{Z}[G_K^{\mathrm{ab}}] \xrightarrow{\epsilon_{\mathbb{Z}[G_K^{\mathrm{ab}}]}} \mathbb{Z} \to 0$$

of $\mathbb{Z}[G_K^{\mathrm{ab}}]$-modules, where θ_2 is the homomorphism induced by $dg \mapsto \psi(g) - 1$ ($g \in G_K$, ψ is the abelianization map) and $\epsilon_{\mathbb{Z}[G_K^{\mathrm{ab}}]}(\sum a_g g) := \sum a_g$. We can separate this exact sequence into two short exact sequences, one of which is

$$(\dagger) \qquad 0 \to H_1(X_\infty) \to A_K \to \ker \epsilon_{\mathbb{Z}[G_K^{\mathrm{ab}}]} \to 0.$$

Exercise 14.1. Show that $\ker \epsilon_{\mathbb{Z}[G_K^{\mathrm{ab}}]} \cong \mathbb{Z}[G_K^{\mathrm{ab}}]$ as $\mathbb{Z}[G_K^{\mathrm{ab}}]$-modules.

It follows that $\ker \epsilon_{\mathbb{Z}[G_K^{\mathrm{ab}}]}$ is a free $\mathbb{Z}[G_K^{\mathrm{ab}}]$-module. We can define a section η of θ_2 by setting $\eta(\psi(g)-1) = dg$ for some lift g of $\psi(g)$ on a basis for $\ker \epsilon_{\mathbb{Z}[G_K^{\mathrm{ab}}]}$, and extending by linearity. Thus the short exact sequence (†) splits, that is, $A_K \cong H_1(X_\infty) \oplus \mathbb{Z}[G_K^{\mathrm{ab}}]$ as $\mathbb{Z}[G_K^{\mathrm{ab}}]$-modules.

From now on, we make use of the isomorphism $\Lambda := \mathbb{Z}[t, t^{-1}] \cong \mathbb{Z}[G_K^{\mathrm{ab}}]$. From the direct sum $A_K \cong H_1(X_\infty) \oplus \Lambda$, we may assume that $Q_K = (Q_1 \mid 0)$ by some elementary operations if necessary, that is, Q_1 is a square presentation matrix for the Λ-module $H_1(X_\infty)$. Since this does not change the ideal generated by the maximal minors of Q_K, we have the following proposition.

Proposition 14.2. *The Alexander ideal and Alexander polynomial are also given by $E(H_1(X_\infty)) = \langle \det(Q_1) \rangle$ and $\Delta_K(t) = \det(Q_1)$ respectively.*

Moreover, since $\Lambda_{\mathbb{Q}} := \Lambda \otimes_{\mathbb{Z}} \mathbb{Q} = \mathbb{Q}[t, t^{-1}]$ is a PID, we have a $\Lambda_{\mathbb{Q}}$-isomorphism

$$H_1(X_\infty) \otimes_{\mathbb{Z}} \mathbb{Q} \cong \bigoplus_{i=1}^{s} \Lambda_{\mathbb{Q}}/(f_i), \quad f_i \in \Lambda_{\mathbb{Q}}.$$

Since a generator τ of $\mathrm{Gal}(X_\infty/X_K)$ acts on the right hand side as multiplication by t, one thus obtains

$$\Delta_K(t) = f_1 \cdots f_s = \det(t \cdot \mathrm{id} - \tau \mid H_1(X_\infty) \otimes_{\mathbb{Z}} \mathbb{Q}) \bmod \Lambda_{\mathbb{Q}}^{\times}.$$

15. Asymptotic formulas on the orders of the first homology groups of cyclic ramified coverings

The infinite cyclic covering X_∞ and its first homology group are our main objects of study; however, the size of X_∞, which allows for its richness of information, also makes it more difficult to study. Taking a leaf from the study of field extensions, we can attempt to simplify the problem by studying its finite subcoverings X_n instead. The group $H_1(X_n)$ is still an infinite group since it has an infinite subgroup with generator given by the homology class of ∂X_n; however, we can remove this subgroup by filling in the tube enclosed by ∂X_n. This naturally leads us to consider the Fox completion M_n of X_n. Recall that M_n is constructed from X_n by gluing a tube $V = D^2 \times S^1$ to X_n along ∂V_n and ∂X_n in such a way that a meridian of ∂V coincides with $n\alpha$.

We can summarize this discussion in the following diagram:

$$\begin{array}{ccc} X_\infty & & \\ \downarrow & & \\ X_n & \subset & M_n \\ \downarrow & & \downarrow \\ X_K & \subset & S^3 \end{array}$$

The main result we shall prove is that $\#H_1(M_n)$ is finite and grows exponentially asymptotically; in fact, we shall have an explicit formula for the constant involved in terms of the Alexander polynomial of K. We can see an inkling of this in the following proposition.

Proposition 15.1. *$H_1(M_n) \cong H_1(X_\infty)/(t^n - 1)H_1(X_\infty)$ for $n \geq 1$.*

Proof. There is an exact sequence

$$H_1(X_\infty) \xrightarrow{t^n - 1} H_1(X_\infty) \to H_1(X_n) \to \mathbb{Z} \to 0$$

that arises from the homology exact sequence associated to a short exact sequence of chain complexes

$$0 \to C_*(X_\infty) \xrightarrow{t^n - 1} C_*(X_\infty) \to C_*(X_n) \to 0$$

(the map $(t^n - 1)_* : H_0(X_\infty) \to H_0(X_\infty)$ is zero). Hence

$$H_1(X_n) \cong H_1(X_\infty)/(t^n - 1)H_1(X_\infty) \oplus \mathbb{Z}.$$

Here $1 \in \mathbb{Z}$ corresponds to a lift $[\tilde{\alpha}^n]$ of $[\alpha^n]$ to X_n. (Since the image of α^n in $\mathrm{Gal}(X_n/X_K) \cong \mathbb{Z}/n\mathbb{Z}$ is 0, α^n can be lifted to X_n.) But $H_1(X_n) \cong H_1(M_n) \oplus \langle [\tilde{\alpha}^n] \rangle$, so we obtain the assertion. □

Note that by taking $n = 1$ in Proposition 15.1, we get

$$H_1(X_\infty)/(t-1)H_1(X_\infty) \cong H_1(M_1) = H_1(S^3) = 0,$$

that is, $H_1(X_\infty)$ is a torsion Λ-module.

The following lemma, whose proof we omit, guarantees that the groups $H_1(M_n)$ are finite under nice circumstances and gives us a way to compute them.

Lemma 15.2. *Let N be a finitely generated, torsion Λ-module and suppose that $E(N) = (\Delta)$. Then, for any $f(t) \in \mathbb{Z}[t]$, $N/f(t)N$ is a finite abelian group if and only if $\Delta(\xi) \neq 0$ for all nonzero roots $\xi \in \overline{\mathbb{Q}}$ of $f(t) = 0$. Moreover, if $f(t)$ can be decomposed as $\pm \prod_{j=1}^{s}(t - \xi_j)$, then*

$$\left| N/f(t)N \right| = \prod_{j=1}^{s} \left| \Delta(\xi_j) \right|.$$

Taking $N = H_1(X_n)$, $f(t) = t^n - 1$ and considering the Alexander polynomial as an integer polynomial with nonzero constant term, we see from Propositions 14.2 and 15.1 and Lemma 15.2 that if the equation $\Delta_K(t) = 0$ does not have a root that is a root of unity, then all the first homology groups $H_1(M_n)$ are finite and

$$\#H_1(M_n) = \prod_{j=0}^{n-1} \left|\Delta_K\left(\zeta_n^j\right)\right|, \tag{‡}$$

where ζ_n is a primitive nth root of unity. In this case, it makes sense to talk about the rate of growth of $\#H_1(M_n)$. This turns out to be a function of the Mahler measure of the Alexander polynomial.

Definition 15.3. For a nonconstant polynomial $g(t) \in \mathbb{R}[t]$, define the *Mahler measure* $m(g)$ of $g(t)$ by

$$m(g) := \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \log\left|g\left(e^{i\theta}\right)\right| \, d\theta\right).$$

A question arises: how can one compute $m(g)$? A method is given by Jensen's formula in complex analysis.

Exercise 15.4. Show that if $g(t)$ splits over $\mathbb{C}$ as $g(t) = c\prod_{i=1}^{d}(t-\xi_i)$, then $m(g) = |c|\prod_{i=1}^{d}\max(\left|\xi_i\right|, 1)$. (Hint: Jensen's formula states that if f is a holomorphic function with no zeroes on the circle $|z| = r$, zeroes $a_1, \ldots, a_k$ in the open disk $|z| < r$ (and possibly other zeroes elsewhere), and $f(0) \neq 0$, then

$$\frac{1}{2\pi}\int_0^{2\pi} \log\left|f(re^{i\theta})\right| \, d\theta = \log\left|f(0)\right| + \sum_{j=1}^{k}(\log r - \log\left|a_k\right|).)$$

We are now ready to state the main theorem of this section.

Theorem 15.5. *Assume that there is no root of $\Delta_K(t) = 0$ that is a root of unity. Then*

$$\lim_{n\to\infty} \frac{1}{n} \log \#H_1(M_n) = \log m(\Delta_K).$$

That is, $\#H_1(M_n)$ grows like $m(\Delta_K)^n$.

Proof. From Equation (‡), we have

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} \log \#H_1(M_n) &= \lim_{n\to\infty} \frac{1}{n} \prod_{j=0}^{n-1} \left|\Delta_K\left(\zeta_n^j\right)\right| \\
&= \int_0^1 \log\left|\Delta_K\left(e^{2\pi i x}\right)\right| \, dx \\
&= \frac{1}{2\pi}\int_0^{2\pi} \log\left|\Delta_K\left(e^{i\theta}\right)\right| \, d\theta \\
&= \log m(\Delta_K).
\end{aligned}$$

□

Example 15.6. Let K be the figure eight knot. In one of the exercises, we computed the Alexander polynomial of the figure eight knot to be

$$\Delta_K(t) = t^2 - 3t + 1 = \left(t - \frac{3+\sqrt{5}}{2}\right)\left(t - \frac{3-\sqrt{5}}{2}\right).$$

Hence, by Exercise 15.4, we have

$$\lim_{n\to\infty} \frac{1}{n} \log \#H_1(M_n) = \log m(\Delta_K) = \log \frac{3+\sqrt{5}}{2}.$$

Lecture 13. (August 1, 2012)

16. Ideal class groups

Recall that class field theory classifies finite abelian extensions of a number field K in terms of the idèle class group C_K. There is a reciprocity map $\rho_K : C_K \to \mathrm{Gal}(K^{\mathrm{ab}}/K)$ inducing the isomorphism $C_K/\mathbb{N}_{L/K}C_L \cong \mathrm{Gal}(L/K)$. We found out that $\mathrm{Gal}(\mathbb{Q}^{\mathrm{ab}}/\mathbb{Q}) \cong \prod_p \mathbb{Z}_p^\times$ and now can easily recover the classical Kronecker-Weber theorem.

Theorem 16.1 (Kronecker-Weber). *Each finite abelian extension of $\mathbb{Q}$ is a subfield of the cyclotomic field $\mathbb{Q}(\zeta_N)$ for some N.*

Proof. Notice that $\prod_p \mathbb{Z}_p^\times = \hat{\mathbb{Z}}^\times$ and $\varprojlim_N \mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q}) \cong \varprojlim(\mathbb{Z}/N\mathbb{Z})^\times \cong \hat{\mathbb{Z}}^\times$. So $\mathbb{Q}^{\mathrm{ab}}$ is explictly the union of all cyclotomic fields $\mathbb{Q}(\zeta_N)$. □

Remark 16.2. This theorem marks the significance of cyclotomic fields for studying abelian number fields.

This global reciprocity map comprises local reciprocity maps $\rho_{K_v} : K_v^\times \to \mathrm{Gal}(K_v^{\mathrm{ab}}/K_v)$ for each place v of K. When v is a finite place, ρ_{K_v} is provided by local class field theory and gives a bijection between finite abelian extensions of K_v and subgroups of $K_v^\times$ of finite index.

In view of class field theory, the infinite places should play an equal role as the finite places. It is also amusing to compare the picture at finite places with the two other local fields $\mathbb{C}$ and $\mathbb{R}$, though their extensions are rather simple. In the case of $\mathbb{C}$, it has no nontrivial finite extension and $\mathbb{C}^\times$ also has no nontrivial subgroup of finite index. In the case of $\mathbb{R}$, it has only one nontrivial finite extension $\mathbb{C}/\mathbb{R}$ of order 2 and $\mathbb{R}^\times$ has only one nontrivial subgroup $(\mathbb{R}^\times)^2 = \mathbb{R}_{>0}$ of index 2. Moreover, $\mathbb{R}_{>0}$ is exactly the image of the norm map $\mathbb{N}_{\mathbb{C}/\mathbb{R}}\mathbb{C}^\times$!

Now we can make sense of ramification of infinite places.

Definition 16.3. Let L/K be an extension of number fields. Let v be an infinite place. We say that an infinite place w of v *lies above* v (denoted by $w \mid v$) if w extends v. We say v is *ramified in* L if v is real and w is complex for some w lying above v, and *unramified in* L otherwise.

Example 16.4. The infinite place $v = \infty$ of $\mathbb{Q}$ is unramified in $\mathbb{Q}(\sqrt{2})$ but is ramified in $\mathbb{Q}(i)$.

By local class field theory, a finite place v is unramified in L if and only if $\rho_K(\mathcal{O}_v^\times) = 1$ in $\mathrm{Gal}(L/K)$. For the infinite place, the above definition gives an analogous result: let v be an infinite place, then v is unramified in L if and only if $\rho_K(K_v^\times) = 1$, since v being unramified simply means there is no appearance of the nontrivial element in $\mathrm{Gal}(\mathbb{C}/\mathbb{R})$ at any infinite place.

We already know that the maximal unramified abelian extension of $\mathbb{Q}$ is $\mathbb{Q}$ itself. What is the maximal unramified abelian extension of a general number field K? The question is a bit more subtle for fields other than $\mathbb{Q}$. We need to distinguish the following two definitions.

Definition 16.5. Let K be a number field. The maximal abelian extension of K unramified at all places (denoted by H) is called the *Hilbert class field* of K. The maximal abelian extension of K unramified at all finite places (denoted by H^+) is called the *narrow Hilbert class field* of K. So $\pi_1^{\mathrm{ab}}(\mathrm{Spec}\,\mathcal{O}_K) \cong \mathrm{Gal}(H^+/K)$.

Example 16.6. For $K = \mathbb{Q}$, we have $H = H^+ = \mathbb{Q}$. But for general number fields, the inclusion $K \subseteq H \subseteq H^+$ may be strict.

In general, the Hilbert class field H is closely related to an intrinsic invariant of the number field K—its ideal class group. The *fractional ideals* of $\mathcal{O}_K$ are of the form $\prod_i \mathfrak{p}_i^{e_i}$, where $e_i \in \mathbb{Z}$. So the fractional ideals form a group under multiplication. The ideal class group measures how far these fractional ideals are from being principal ones.

Definition 16.7. Denote the *group of fractional ideals* of K by $I(K)$. Denote by $P(K)$ the subgroup of principal fractional ideals (i.e., those generated by an element in $K^\times$). We define the *(ideal) class group* of K to be the quotient group $H(K) := I(K)/P(K)$. A fundamental theorem in algebraic number theory asserts that $H(K)$ is always a finite group. The order of $H(K)$ is called the *class number* of K.

Remark 16.8. By definition, K has class number one if and only if $\mathcal{O}_K$ is a PID, if and only if $\mathcal{O}_K$ is a UFD.

Let $\phi : K^\times \to I(K)$ be the natural map $a \mapsto (a)$. Then $P(K) = \mathrm{Im}\,\phi$ and $H(K) = \mathrm{Coker}\phi$. On the other hand, as an abstract group, $I(K) \cong \bigoplus_{\mathfrak{p}} \mathbb{Z}$. So ϕ is nothing but the valuation map $K^\times \to \bigoplus_{\mathfrak{p}} \mathbb{Z}$. This valuation map natually extends to the idèle group $J_K \to \bigoplus_{\mathfrak{p}} \mathbb{Z}$ and hence induces a map $C_K \to H(K)$. This is a surjection and the kernel is exactly $\prod_{v\nmid\infty} \mathcal{O}_v^\times \prod_{v\mid\infty} K_v^\times$. By class field theory, $C_K/\mathcal{O}_v^\times \prod_{v\mid\infty} K_v^\times \cong H(K)$ corresponds to the Hilbert class field H. So we have the following isomorphism between the class group and the Hilbert class field.

Theorem 16.9. *The reciprocity map induces an isomorphism* $H(K) \cong \mathrm{Gal}(H/K)$.

Remark 16.10. This important result justifies several terminologies: the “idèles” are a generalization of “ideals” and the Hilbert “class” field is the field with Galois group canonically isomorphic to the “class” group. So class field theory is a vast generalization of this correspondence between abelian number fields and idèle class groups.

Similarly, the narrow Hilbert class field theory H^+ corresponds to $C_K/\prod_{v\nmid\infty} \mathcal{O}_v^\times \prod_{v\mid\infty} (K_v^\times)^2$. Instead of removing all elements in $K^\times$, we should remove those elements lying in $(K_v^\times)^2$ for all infinite places. This motivates the following definition.

Definition 16.11. An element $a \in K^\times$ is called *totally positive* if $\sigma(a) > 0$ for every real embedding $\sigma : K \hookrightarrow \mathbb{R}$. Let $P^+(K)$ be the subgroup of $P(K)$ generated by all totally positive elements. We define the *narrow class group* to be $H^+(K) := I(K)/P^+(K)$.

Theorem 16.12. *The reciprocity map induces an isomorphism* $H^+(K) \cong \mathrm{Gal}(H^+/K)$.

Notice that by definition $\mathrm{Gal}(H^+/K) = \pi_1^{\mathrm{ab}}(\mathrm{Spec}\,\mathcal{O}_K)$. So miraculously we can read off information about the étale fundamental group by computing the narrow class group!

Example 16.13. $K = \mathbb{Q}$. Each fractional ideal of $\mathbb{Q}$ can be generated by a positive rational number. So $H^+(\mathbb{Q}) = 1$, which corresponds to the familiar fact that the narrow class field $H^+ = \mathbb{Q}$ and $\pi_1^{\mathrm{ab}}(\mathrm{Spec}\,\mathbb{Z}) = 1$.

Example 16.14. $K = \mathbb{Q}(\sqrt{i})$ has class number 1 and has no real places. So $H(K) = H^+(K) = 1$ and $H = H^+ = K$.

When K has class number greater than 1, the Hilbert class field may not that easy to find. The following proposition about ramification in general number fields is quite handy.

Proposition 16.15. *Let K be a number field and $L = K(\sqrt[n_1]{a_1}, \dots, \sqrt[n_m]{a_m})$. For a prime $\mathfrak{p}$ of K, if $a_i \notin \mathfrak{p}$ and $n_i \notin \mathfrak{p}$, then $\mathfrak{p}$ is unramified in L.*

Example 16.16. $K = \mathbb{Q}(\sqrt{-5})$ has class number 2 and has no real places. $H(K) = H^+(K)$ is represented by (1) and $(2, 1+\sqrt{-5})$. The Hilbert class field should be a quadratic extension of K. We claim that $H = H^+ = K(i)$. To prove it, it suffices to show that $K(i)/K$ is unramified at every prime. By the proposition, we know that only the primes of K above (2) can be ramified in $K(i)$. Since $K(i) = K(\sqrt{5}) \hookrightarrow K(\zeta_5)$ ($\cos\frac{2\pi}{5} = \frac{\sqrt{5}-1}{4}$), the proposition also tells us that only the primes of K over (5) can be ramified in $K(i)$. Therefore $K(i)/K$ is unramified everywhere.

Exercise 16.17. For $K = \mathbb{Q}(\sqrt{-6})$, show that $H = H^+ = K(\sqrt{-3})$. (Hint: use the fact that K has class number 2.)

Remark 16.18. Here is a beautiful connection between the solution of the $p = x^2 + ny^2$ problem and the Hilbert class field of $\mathbb{Q}(\sqrt{-n})$.

$p = x^2 + y^2$	$p \equiv 1 \pmod 4$	$\mathbb{Q}(i) = \mathbb{Q}(\zeta_4)$
$p = x^2 + 5y^2$	$p \equiv 1, 9 \pmod{20}$	$\mathbb{Q}(\sqrt{-5}, i) \hookrightarrow \mathbb{Q}(\zeta_{20})$
$p = x^2 + 6y^2$	$p \equiv 1, 7 \pmod{24}$	$\mathbb{Q}(\sqrt{-6}, \sqrt{-3}) \hookrightarrow \mathbb{Q}(\zeta_{24})$

In general, we have a criterion of the form $p \equiv \cdots \pmod N$ if and only if the Hilbert class field H of $\mathbb{Q}(\sqrt{-n})$ is an abelian extension of $\mathbb{Q}$ and in that case N is the the smallest number such that $H \hookrightarrow \mathbb{Q}(\zeta_N)$ (ensured by the Kronecker-Weber theorem). Moreover, the correct residues modulo N is exactly the subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$ corresponding to the subfield $H \hookrightarrow \mathbb{Q}(\zeta_N)$ via Galois theory.

Here we also supply an example for which the narrow class group and the class group are different.

Example 16.19. For $K = \mathbb{Q}(\sqrt{3})$, $H = K$ and $H^+ = K(i)$.

Finally, we summarize our analogy between the first homology group and class groups as follows. The class group is one of the most important arithmetic invariants of a number field. The analogy suggests that we can study $H(K)$ by studying an infinite tower of extensions of K, in the way we obtained formulas for $\#H_1(M_n)$ via going to the infinite cyclic covering $X_\infty \to X_K$ for a knot K. We will talk more about this key idea in Iwasawa theory next time.

homology group	ideal class group
$H_1(M) = \pi_1^{\mathrm{ab}}(M)$	$H^+(K) = \pi_1^{\mathrm{ab}}(\operatorname{Spec} \mathscr{O}_K)$ $H(K) = \pi_1^{\mathrm{ab}}(\operatorname{Spec} \mathscr{O}_K \cup \{\infty\text{-places}\})$

Lecture 14. (August 3, 2012)

17. Iwasawa theory

Last time we found the relationship between the class group and the Hilbert class field via class field theory. The class group measures the failure of unique factorization and is one of the most important arithmetic invariants of a number field.

Example 17.1. When trying to solve the Fermat equation

$$x^p + y^p = z^p, \quad p \text{ an odd prime},$$

we factorize it as

$$\prod_i (x + \zeta_p^i y) = z^p$$

and hope to conclude that $x + \zeta_p^i y$ is a pth power. In fact, one can show that the ideals $(x + \zeta_p^i y)$ are mutually coprime, so by unique factorization of prime ideals, we know that $(x + \zeta_p^i y)$ is a pth power of an ideal $\mathfrak{a}$. Now, if we require $\mathfrak{a}$ to be a principal ideal, it suffices to assume that the class group $H(\mathbb{Q}(\zeta_p))$ has no element of p-power order. This is the way Kummer found the following famous criterion.

Theorem 17.2. *If $p \nmid \#H(\mathbb{Q}(\zeta_p))$, then $x^p + y^p = z^p$ has no nontrivial integer solutions.*

The primes satisfying this condition are called *regular primes*. Kummer computed the class group for $p < 100$ and showed that there are only three irregular primes $p = 37, 59, 67$ in this range. This was the best result on Fermat's last theorem for a long period.

So understanding the p-part of the class group $H(\mathbb{Q}(\zeta_p))$ is of great arithmetic interest.

Example 17.3. Let us consider the first irregular prime $p = 37$. The 37-part of the cyclotomic fields $\mathbb{Q}(\zeta_{37^n})$ turns out to be $\mathbb{Z}/37^n\mathbb{Z}$.

Example 17.4. 691 is also an irregular prime. The 691-part of of the cyclotomic fields $\mathbb{Q}(\zeta_{691^n})$ turns out to be $(\mathbb{Z}/691^n\mathbb{Z})^2$.

You may smell something. We now introduce a general definition and state Iwasawa's class number formula in the more general setting.

Definition 17.5. Let p be an odd prime. We know that $\mathrm{Gal}(\mathbb{Q}(\zeta_{p^\infty})/\mathbb{Q}) \cong \mathbb{Z}_p^\times \cong \mathbb{F}_p^\times \times \mathbb{Z}_p$. We define $\mathbb{Q}_\infty \subseteq \mathbb{Q}(\zeta_{p^\infty})$ to be the fixed field of $\mathbb{F}_p^\times$ (so $\mathrm{Gal}(\mathbb{Q}_\infty/\mathbb{Q}) \cong \mathbb{Z}_p$). Let K be a number field. We define $K_\infty := K\mathbb{Q}_\infty$. Then $\mathrm{Gal}(K_\infty/K) \cong \mathbb{Z}_p$ and we call K_∞ the *cyclotomic $\mathbb{Z}_p$-extension* of K. We denote by K_n the finite extension of K corresponding to the group $\mathbb{Z}_p/p^n\mathbb{Z}_p \cong \mathbb{Z}/p^n\mathbb{Z}$ and $H_n = H(K_n)$.

Example 17.6. For $K = \mathbb{Q}(\zeta_p)$, $K_\infty = \mathbb{Q}(\zeta_{p^\infty})$ is the cyclotomic $\mathbb{Z}_p$-extension of K and $K_n = \mathbb{Q}(\zeta_{p^{n+1}})$.

Theorem 17.7 (Iwasawa's class number formula)**.** *There exist constants $\mu, \lambda \geq 0$ and ν such that when n is sufficiently large,*

$$\log_p \#H_n = \mu p^n + \lambda n + \nu.$$

Example 17.8. We have $\mu = 0$, $\lambda = \nu = 1$ for $K = \mathbb{Q}(\zeta_{37})$ and $\mu = 0$, $\lambda = 2$, $\nu = 1$ for $K = \mathbb{Q}(\zeta_{691})$.

In view of the analogy between 3-manifolds and number fields, we obtain the following table.

infinite cyclic covering $X_\infty \to X_K$	Cyclotomic $\mathbb{Z}_p$ extension K_∞/K
homology group $H_1(M_n)$	class group $H_n := H(K_n)[p]$
asymptotic formula on homology groups	asymptotic formula on class groups

Our goal today is to sketch the main ingredients of the proof and see how it is amazingly similar to the case of knots and Alexander polynomials.

Recall that the Alexander polynomial is defined via the action of $\Lambda = \mathbb{Z}[G_K^{\mathrm{ab}}] \cong \mathbb{Z}[\mathrm{Gal}(X_\infty/X_K)]$ on $H_1(X_\infty)$. Since $\Lambda \cong \mathbb{Z}[t^{\pm 1}]$ and $\Lambda_\mathbb{Q} = \mathbb{Z}[t^{\pm 1}]$ is a PID, we know that

$$H_1(X_\infty) \otimes \mathbb{Q} \cong \bigoplus_i \Lambda_\mathbb{Q}/(f_i)$$

as $\Lambda_\mathbb{Q}$-modules. Since the presentation matrix is simply the diagonal matrix (f_i), the Alexander polynomial is nothing but the product $\prod_i f_i$ up to $\Lambda_\mathbb{Q}^\times$.

What is the arithmetic analogue of Λ acting on $H_n = H(K_n)[p]$? By class field theory, $H_n \cong \mathrm{Gal}(L_n/K_n)$, where L_n is the maximal unramified abelian p-extension of K_n. Write $L_\infty := \bigcup L_n$ and

$$H_\infty := \varprojlim_n H_n = \varprojlim_n \mathrm{Gal}(L_n/K_n) = \mathrm{Gal}(L_\infty/K_\infty).$$

Since $\mathbb{Z}_p[\mathrm{Gal}(K_n/K)]$ acts on $H_n \cong \mathrm{Gal}(L_n/K_n)$, we know that $\varprojlim_n \mathbb{Z}_p[\mathrm{Gal}(K_n/K)]$ acts on H_∞.

Definition 17.9. We define the *Iwasawa algebra* to be $\mathbb{Z}_p[[\mathrm{Gal}(K_\infty/K)]] := \varprojlim_n \mathbb{Z}_p[\mathrm{Gal}(K_n/K)]$.

The following theorem tells us that the Iwasawa algebra has a neat description and thus plays a similar role as $\mathbb{Z}[t^{\pm 1}]$.

Theorem 17.10. *Let γ be the topological generator of* $\mathrm{Gal}(K_\infty/K) \cong \mathbb{Z}_p$*. Then $\gamma \mapsto 1+T$ induces an isomorphism $\mathbb{Z}_p[[\mathrm{Gal}(K_\infty/K)]] \cong \mathbb{Z}_p[[T]]$ of $\mathbb{Z}_p$-algebras.*

This is one major reason why we want to consider K_∞ and H_∞: the action of the nice algebra $\mathbb{Z}_p[[T]]$ on the class groups cannot be seen at finite levels. Assuming for simplicity that there is only one prime of K ramified in K_∞ and is totally ramified (this is the case for $K = \mathbb{Q}(\zeta_p)$), then we can recover H_n from H_∞ analogously to the knot situation, which is another major reason we would like to consider the infinite tower of fields rather than K itself.

Theorem 17.11. *We have an isomorphism $H_n \cong H_\infty/((1+T)^{p^n}-1)H_\infty$.*

Morally the Iwasawa algebra $\hat{\Lambda} = \mathbb{Z}_p[[T]]$ behaves as the power series ring $\mathbb{C}[[X,Y]]$ and we have the following p-adic version of the Weierstrass preparation theorem.

Theorem 17.12. *Suppose $f(T) \in \hat{\Lambda}$ is nonzero. Then $f(T)$ can be uniquely written as*

$$f(T) = p^\mu g(T) u(T),$$

where $\mu \geq 0$, $u(T) \in \hat{\Lambda}^\times$ and $g(T)$ is a Weierstrass polynomial*, namely $g(T) = T^\lambda + c_1 T^{\lambda-1} + \cdots + c_\lambda$ with $c_i \equiv 0$* (mod p)*. We call μ and λ the μ*-invariant *and λ*-invariant *of $f(T)$.*

$\hat{\Lambda} = \mathbb{Z}_p[[T]]$ is not a PID but fortunately the following structure theorem for finitely generated $\hat{\Lambda}$-modules is still valid.

Theorem 17.13. *Let N be a finitely generated $\hat{\Lambda}$-module, then we have an pseudo-isomorphism (i.e., a homomorphism with finite kernel and cokernel)*

$$N \sim \hat{\Lambda}^r \bigoplus_i \hat{\Lambda}/(p^{m_i}) \bigoplus_j \hat{\Lambda}/(f_j^{e_j}),$$

where $r, m_i, e_i \geq 0$ and f_i's are irreducible Weierstrass polynomials. The polynomial

$$f := \prod_i p^{m_i} \prod_j f_j^{e_j}$$

is called the Iwasawa polynomial *of N, well-defined up to $\hat{\Lambda}^\times$. We define the μ*-invariant *and λ*-invariant of *N to be $\mu = \sum_i m_i$ and $\lambda = \sum_j (\deg f_j)^{e_j}$ (the sum of individual the μ- and λ-invariants).*

Using the finiteness of the class group H_n and a Nakayama lemma argument, one can show that H_∞ is a finitely generated torsion $\hat{\Lambda}$-module. So the Iwasawa polynomial, μ-invariant and λ-invariant of H_∞ are all well defined due to the previous structure theorem. Now we can restate Theorem 17.7 more precisely.

Theorem 17.7. Let μ and λ be the μ-invariant and λ-invariant of H_∞. Then there exists a constant ν such that when n is sufficiently large,

$$\log_p \# H_n = \mu p^n + \lambda n + \nu.$$

Proof sketch. Suppose $H_\infty \sim \hat{\Lambda}/(p^{m_i}) \bigoplus_j \hat{\Lambda}/(f_j^{e_j})$. By Theorem 17.11, we need to compute $\# N/((1+T)^{p^n}-1)N$ for $N = \hat{\Lambda}/(p^m)$ or $N = \hat{\Lambda}/(g)$ for g a Weierstrass polynomial. The first case contributes p^{mp^n} and the second case contributes $\prod_{\zeta^{p^n}=1} |g(\zeta-1)|_p^{-1}$. The latter one is dominated by the leading term since g is a Weierstrass polynomial and becomes $n \deg g + c$ for some constant c when n is sufficiently large. □

Remark 17.14. One can further show that $\mu = 0$ for K any abelian extension of $\mathbb{Q}$. Moreover when $K = \mathbb{Q}(\zeta_p)$, the formula $\# H_n = \prod |f(\zeta-1)_p^{-1}|$ is true for any $n \geq 1$, where $f(T)$ is the Iwasawa polynomial of H_∞.

Remark 17.15. The Iwasawa polynomial is harder to compute by hand compared to Alexander polynomial, since we do not have a nice arithmetic analogue of the Wirtinger presentation and also the ring $\mathbb{Z}_p$ is much larger than $\mathbb{Z}$. Computer algorithms have been developed for the computation.

Remark 17.16. When is p an irregular prime, i.e., when does $H(\mathbb{Q}(\zeta_p))$ have nontrivial p-part? Miraculously, Kummer proved that it happens exactly when p appears in one of the numerators of the Riemann zeta values $\zeta(-1), \zeta(-3), \zeta(-5), \ldots$! For example,

$$\zeta(-11) = \frac{691}{32760}$$

and

$$\zeta(-31) = \frac{37 \cdot 683 \cdot 305062827}{2^6 \cdot 3 \cdot 5 \cdot 17}$$

shows that 691 and 37 are irregular. The connection between Iwasawa polynomials and Riemann zeta functions can be made rigorous and is the content of the Iwasawa main conjecture. This conjecture is much deeper and was first proved by Mazur-Wiles in 1984 and reproved by Rubin in 1994 using Euler systems. It is not so surprising that Wiles' proof of Fermat's last theorem used Iwasawa theory (and more generally Galois representations of tower of fields and p-adic L-functions) in an essential way.

We end the lectures with the following summary.

infinite cyclic covering $X_\infty \to X_K$	Cyclotomic $\mathbb{Z}_p$ extension K_∞/K
homology group $H_1(M_n)$	class group $H_n := H(K_n)$
asymptotic formula on homology groups	asymptotic formula on class groups
$\Lambda = \mathbb{Z}[G_K^{ab}] \cong \mathbb{Z}[t^{\pm 1}]$	$\hat{\Lambda} = \mathbb{Z}_p[[\mathrm{Gal}(K_\infty/K)]] \cong \mathbb{Z}_p[[T]]$
$H_1(M_n) \cong H_1(X_\infty)/(t^n-1)H_1(X_\infty)$	$H_n \cong H_\infty/((1+T)^{p^n}-1)H_\infty$
Alexander polynomial	Iwasawa polynomial
$\# H_1(M_n) = \prod |\Delta(\zeta)|$	$\# H_n = \prod |f(\zeta-1)|_p^{-1}$

There are many more beautiful stories to tell and discover. Now—it's your turn!