DISTRIBUTION OF 8-RANKS OF IMAGINARY QUADRATIC FIELDS, D'APRÉS SMITH

COURSE: PETER KOYMANS
NOTES: ROSS PATERSON

Lecture 1

1. Overview of Arithmetic Statistics

We begin with the following, very influential, conjecture in arithmetic statistics.

Conjecture 1.1 (Cohen–Lenstra, 1984)**.** *Let p be an odd prime and let A be a finite abelian p-group. Then*

$$\lim_{X\to\infty} \frac{\#\{K/\mathbb{Q} \ : \ \text{imag. quad. } |D_K| \leq X, \ \mathrm{Cl}_K[p^\infty] \cong A\}}{\#\{K/\mathbb{Q} \ : \ \text{imag. quad. } |D_K| \leq X\}} = \frac{\prod_{i=1}^\infty (1-p^{-i})}{\#\mathrm{Aut}(A)}.$$

There is very little proved about this conjecture, but a lot of computational evidence has been found and it is strongly believed. Indeed, it is not known for a single pair (p, A)! The main thing that is known is the following average result.

Theorem 1.2 (Davenport–Heilbronn, 1970s)**.** *We have*

$$\sum_{\substack{K/\mathbb{Q} \\ |D_K|\leq X}} \#\mathrm{Cl}_K[3] \sim cX,$$

and the leading constant matches the average predicted by the Cohen–Lenstra distribution.

Conjecture 1.3 (Cohen–Lenstra–Gerth)**.** *Let A be a finite abelian 2-group. Then*

$$\lim_{X\to\infty} \frac{\#\{K/\mathbb{Q} \ : \ \text{imag. quad. } |D_K| \leq X, \ 2\mathrm{Cl}_K[2^\infty] \cong A\}}{\#\{K/\mathbb{Q} \ : \text{imag. quad. } \ |D_K| \leq X\}} = \frac{\prod_{i=1}^\infty (1-2^{-i})}{\#\mathrm{Aut}(A)}.$$

The idea here is that in this setting, $\mathrm{Cl}_K[2]$ is special. It is "predictable" and in fact quite large ($\#\mathrm{Cl}_K[2] = 2^{\omega(D_K)-1}$ where ω is the number of prime divisors). Multiplying by 2 kills this contribution, and the claim is somehow that this is the only predictable behaviour present.

The first result on this is the work of Fouvry–Klüners on the 4-torsion.

Theorem 1.4 (Fouvry–Klüners, 2007)**.** *We have for all $n \geq 0$ that*

$$\lim_{X\to\infty} \frac{\#\{K/\mathbb{Q} \ : \ \text{imag. quad. } |D_K| \leq X, \ 2\mathrm{Cl}_K[4] \cong \mathbb{F}_2^n\}}{\#\{K/\mathbb{Q} \ : \ \text{imag. quad. } |D_K| \leq X\}} = \lim_{r\to\infty} P_{\mathrm{Mat}}(r, n),$$

where $P_{\mathrm{Mat}}(r,n)$ is the probability that a uniformly selected matrix $M \in \mathrm{Mat}(r \times r, \mathbb{F}_2)$ has kernel of dimension n.

Then Alex Smith was able to get from 4 to 8 torsion under GRH.

Theorem 1.5 (Smith, 2016)**.** *Assume GRH. Then for all $m \geq j \geq 0$ we have*

$$\lim_{X\to\infty} \frac{\#\left\{K/\mathbb{Q} \ : \ \text{imag. quad. } |D_K| \leq X,\ 2\mathrm{Cl}_K[8] \cong \mathbb{Z}/2\mathbb{Z}^{m-j} \oplus (\mathbb{Z}/4\mathbb{Z})^j\right\}}{\#\left\{K/\mathbb{Q} \ : \ \text{imag. quad. } |D_K| \leq X,\ 2\mathrm{Cl}_K[4] \cong (\mathbb{Z}/2\mathbb{Z})^m\right\}} = P_{\mathrm{Mat}}(m,j).$$

In 2017, Smith generalised this: there is a filtration of $\mathbb{F}_2$-vector spaces

$$\mathrm{Cl}_K[2] \supseteq 2\mathrm{Cl}_K[4] \supseteq 4\mathrm{Cl}_K[8] \supseteq 8\mathrm{Cl}_K[16] \supseteq \dots,$$

and he studies the behaviour of this filtration. It is a worthwhile exercise to do, to convince yourself that this filtration recovers the isomorphism class of the abelian group. Note also that $2^k\mathrm{Cl}_K[2^{k+1}] = \mathrm{Cl}_K[2] \cap 2^k\mathrm{Cl}_K$.

Definition 1.6. We write $r_{i,K} := \dim_{\mathbb{F}_2} 2^{i-1}\mathrm{Cl}_K[2^i]$.

Smith proves the following.

Theorem 1.7 (Smith, 2017)**.** *Let $k \geq 2$ and let $r_{k+1} \leq r_k \leq \dots \leq r_2$. Then*

$$\lim_{X\to\infty} \frac{\#\left\{K/\mathbb{Q} \ : \ \text{imag. quad. } |D_K| \leq X, \quad r_{i,K} = r_i \quad \forall 2 \leq i \leq k+1\right\}}{\#\left\{K/\mathbb{Q} \ : \ \text{imag. quad. } |D_K| \leq X, \quad r_{i,K} = r_i \quad \forall 2 \leq i \leq k\right\}} = P_{\mathrm{Mat}}(r_{k+1}, r_k).$$

The idea here is that you have some Markov process, so that you don't care about the history of the process beyond the position you're on!

Remark 1.8. This implies Conjecture 1.3 (see K.–Pagano, JEMS, §14). Note that in the setting of Theorem 1.5 we have $r_2 = m$ and $r_3 = j$.

Since this work of Smith, we have seen:

- Generalisation to number fields (under roots of unity assumptions) and quadratic twists of abelian varieties (Smith 2022).
- Counting squarefree $1 \leq d \leq X$ such that $x^2 - dy^2 = -1$ has a solution $x, y \in \mathbb{Z}$ (K.–Pagano 2022).
- Upper and lower bounds (conditional on a converse theorem) on how often as n varies there are solutions of $x^3 + y^3 = n$ for $x, y \in \mathbb{Q}$ (K.–Smith 2024).
- BSD implies Goldfeld's conjecture (Smith 2025).

2. Overview of the lecture series

Our goal in this series is to do the following.

- Sketch a proof of Theorem 1.4.
- Give an unconditional proof of Theorem 1.5 using a new tool from recent work of K.–Smith.
- Time permitting, we might also see how to access Theorem 1.7 in general, but that might be difficult to reach.

The idea is that our proof of Theorem 1.5 will be easier than the original, though it makes use of the many developments since then and is still not terribly easy.

Now for an underlying theme that we will see repeatedly. Accessing the class group directly is very hard, but what we end up being able to do is to *compare* class groups. A promising start for this is as follows. Denote by $H_K(2)$ the maximal abelian unramified extension of K with $\mathrm{Gal}(H_K(2)/K)$ being an $\mathbb{F}_2$-vector space. Then

(1) we can give an explicit description of $H_K(2)$ (next lecture!)
(2) for two different imaginary quadratic fields K, K' there is an interesting intersection $H_K(2) \cap H_{K'}(2)$ (i.e. bigger than $\mathbb{Q}$) which completely fails if 2 is replaced by an odd prime.

The following gives a good mental example to keep in mind when we're talking about comparing class groups, but we will need to do something more general.

Theorem 2.1 (Stevenhagen)**.** *Let $p \equiv 1 \mod 8$ and let χ_p be the quadratic dirichlet character associated to p and let ζ_8 be an element of order 8 in $(\mathbb{Z}/p\mathbb{Z})^\times$. Write $h_d^+ := \#\mathrm{Cl}^+_{\mathbb{Q}(\sqrt{d})}$ for the narrow class number. Then*

$$\begin{aligned} 8 \mid h^+_{-p} &\iff \chi_p(1+\zeta_8^2) = 1 \\ 8 \mid h^+_{-2p} &\iff \chi_p(1+\zeta_8^2)\chi_p(\zeta_8) = 1 \\ 8 \mid h^+_{2p} &\iff \chi_p(1+\zeta_8^2) = \chi(\zeta_8) = 1 \end{aligned}$$

Now suppose that $\chi(1+\zeta_p^2) = \chi(\zeta_8) = 1$. If $\chi(1+\zeta_8) = 1$,

$$16 \mid h^+_{2p} \iff 16 \mid h^+_{-p} \text{ and } 16 \mid h^+_{-2p}$$

Alternatively if $\chi(1+\zeta_8) = -1$ then

$$16 \mid h^+_{2p} \iff 16 \mid h^+_{-2p} \text{ and } 8 \mid\mid h^+_{-p}$$

The slogan here is that for 16-ranks, the relationships between h^+_{2p}, h^+_{-2p}, h^+_{-p} is governed by a splitting condition.

Lecture 2

3. Genus theory and cocycles

Notation 3.1. Let $N := \mathbb{Q}_2/\mathbb{Z}_2 = \varinjlim \mathbb{Z}/2^k\mathbb{Z}$. We view N as a (continuous) $G_\mathbb{Q}$-module via the discrete topology and trivial action. For each $\chi \in \mathrm{Hom}(G_\mathbb{Q}, \{\pm 1\})$ we define $N(\chi)$ to be N as an abelian group but with the $G_\mathbb{Q}$-action for all $\sigma \in G_\mathbb{Q}$ and $n \in N$

$$\sigma(n) = \chi(\sigma)n.$$

Recall that squarefree numbers are in bijection with quadratic extensions of $\mathbb{Q}$ and hence with elements of $\mathrm{Hom}(G_\mathbb{Q}, \{\pm 1\})$. If x is a squarefree number then we define the character χ_x to be the image of x in $\mathrm{Hom}(G_\mathbb{Q}, \{\pm 1\})$, which is precisely the map $\sigma \mapsto \frac{\sigma(\sqrt{x})}{\sqrt{x}}$. We write $N(x) := N(\chi_x)$.

Theorem 3.2. *There is an exact sequence for each $k \geq 1$ and $K = \mathbb{Q}(\sqrt{x})$*

$$0 \to \mathbb{Z}/2^k\mathbb{Z} \to Z^1(\mathrm{Gal}(H_K/\mathbb{Q}), N(x)[2^k]) \to \mathrm{Cl}_K^\vee[2^k] \to 0,$$

where Z^1 denotes the group of 1-cocycles, H_K is the Hilbert class field, and $\vee$ denotes dual.

Proof. We assume the following fact from exercise sheet 2: For every quadratic field $K/\mathbb{Q}$, the Hilbert class field H_K is Galois over $\mathbb{Q}$ and

$$1 \to \mathrm{Gal}(H_K/K) \to \mathrm{Gal}(H_K/\mathbb{Q}) \to \mathrm{Gal}(K/\mathbb{Q}) \to 1$$

is split with $\mathrm{Gal}(H_K/\mathbb{Q}) \cong \mathrm{Gal}(H_K/K) \rtimes \mathrm{Gal}(K/\mathbb{Q})$. Writing $\mathrm{Gal}(K/\mathbb{Q}) = \langle\sigma\rangle$, then the generator σ acts by inversion on $\mathrm{Gal}(H_K/K)$. Moreover by the Artin map $\mathrm{Gal}(H_K/K) \cong \mathrm{Cl}_K$ and this is the natural action of $\mathrm{Gal}(K/\mathbb{Q})$ on Cl_K.

This theorem will essentially follow from inflation restriction, but we need to be quite careful as we must work on the level of cocycles. There is for each $k \geq 1$ and each $K = \mathbb{Q}(\sqrt{x})$ a natural restriction map

$$\begin{aligned} Z^1(\mathrm{Gal}(H_K/\mathbb{Q}), N(x)[2^k]) \xrightarrow{\mathrm{res}} & Z^1(\mathrm{Gal}(H_K/K), N(x)[2^k]) \\ &= \mathrm{Hom}(\mathrm{Gal}(H_K/K), N(x)[2^k]) \\ &= \mathrm{Hom}(\mathrm{Cl}_K, \mathbb{Z}/2^k\mathbb{Z}) \\ &\cong \mathrm{Cl}_K^\vee[2^k]. \end{aligned}$$

Lemma 3.3. res *is surjective.*

Proof. Take $\phi \in \mathrm{Hom}(\mathrm{Gal}(H_K/K), N(x)[2^k])$. By the fact we are assuming above, $\mathrm{Gal}(H_K/\mathbb{Q}) \cong \mathrm{Gal}(H_K/K) \rtimes \mathrm{Gal}(K/\mathbb{Q})$. We define $\psi \in \mathrm{Map}(\mathrm{Gal}(H_K/\mathbb{Q}),\ N(x)[2^k])$ by $\psi = (\phi, 0)$ when written in this semidirect product decomposition. Indeed, ψ is a cocycle, since

$$\begin{aligned} \psi((a_1, g_1)(a_2, g_2)) &= \psi(a_1 + \chi_x(g_1)a_2,\ g_1g_2) \\ &= \phi(a_1) + \chi_x(g_1)\phi(a_2) \\ &= g_1 \cdot_{\chi_x} \psi(a_2, g_2) + \psi(a_1, g_1). \end{aligned}$$

□

Using this, it is now sufficient to see that the kernel of res is a copy of $\mathbb{Z}/2^k\mathbb{Z}$. The kernel is

$$\ker(\mathrm{res}) = \inf(Z^1(\mathrm{Gal}(K/\mathbb{Q}), N(x)[2^k])) \cong \mathbb{Z}/2^k\mathbb{Z},$$

via evaluation at $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$. This we leave as an exercise!

[R:This can be seen as follows: let $\phi \in \ker(\mathrm{res})$. Then representing $\mathrm{Gal}(H_K/\mathbb{Q})$ as the semidirect product in the usual way we have $\phi(\sigma, \tau) = \phi((\sigma, 0) \cdot (0, \tau)) = (\sigma, 0) \cdot \phi(0, \tau) = \phi(0, \tau)$. Hence $\phi \in \inf(Z^1(\mathrm{Gal}(K/\mathbb{Q}), N(x)[2^k])$. The converse is immediate, so the first equality holds. Note that inf is injective on 1-cocycles by definition.

Elements of $Z^1(\mathrm{Gal}(K/\mathbb{Q}), N(x)[2^k])$ are functions $f : \mathrm{Gal}(K/\mathbb{Q}) \to N(x)[2^k]$ which satisfy the cocycle relations, which here (since $\mathrm{Gal}(K/\mathbb{Q})$ is so small) are just given by $0 = f(0) = f(\sigma^2) = \sigma \cdot f(\sigma) + f(\sigma) = (1 - \chi_x(\sigma))f(\sigma)$. But since $1 - \chi_x(\sigma) = 0$ this just means we can choose any image for σ. Hence there are 2^k elements and they are identified by their image under evaluation at σ.] □

Definition 3.4. Let $Z^1_{\mathrm{nr}}(G_\mathbb{Q}, N(x)[2^k]) = \inf\left(Z^1(\mathrm{Gal}(H_K/\mathbb{Q}), N(x)[2^k])\right)$ for $K = \mathbb{Q}(\sqrt{x})$. Note that this unramified decoration really depends on x in general!

These are precisely the cocycles which restrict to $\mathrm{Hom}_{\mathrm{nr}}(G_K, N(x)[2^k]) \cong \mathrm{Cl}_K^\vee[2^k]$, i.e. those that restrict trivially to $\mathrm{Hom}(I_v, N(x)[2^k])$ for all places v where I_v is a choice of inertia subgroup at v.

Theorem 3.5 (Gauss genus theory)**.** *Let $K = \mathbb{Q}(\sqrt{-x})$ with $x \equiv -1 \mod 4$ a positive squarefree integer. Then for each odd prime p let $p^* = \pm p$ with choice of sign so that $p^* \equiv 1 \mod 4$. Then as homomorphism groups we have an equality*

$$Z^1_{\mathrm{nr}}(G_\mathbb{Q}, N(-x)[2]) = \langle \chi_{p^*} \ :\ p \mid x \rangle,$$

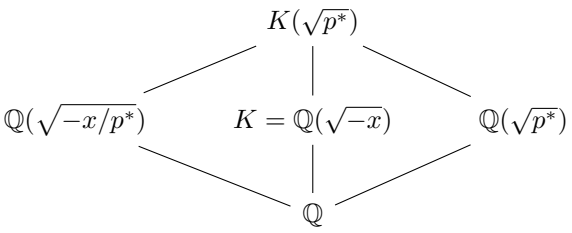
where the χ_{p^*} are all linearly independet over $\mathbb{Q}$ for ramification reasons. In particular

$$\#\mathrm{Cl}_K[2] = 2^{\omega(D_K)-1}$$

via Theorem 3.2.

Proof. We cheat slightly. $\mathrm{Hom}(G_{\mathbb{Q}}, \{\pm 1\})$ has a basis given by the χ_{p^*} for p odd primes, χ_{-1} and χ_2. We will check for each basis element whether it is unramified, which is not sufficient, but the general argument isn't much harder.

Indeed consider the field diagram



If $p \mid x$ then we want to show that χ_{p^*} is unramified on restriction to G_K. Then note that $\mathbb{Q}(\sqrt{p^*})/\mathbb{Q}$ is ramified only at p, and so $K(\sqrt{p^*})/K$ is ramified at most at primes above p. But what about at p? Since p is odd, we get that $\mathbb{Q}(\sqrt{-x/p^*})$ is unramified over p. Hence the total ramification degree is 2 and $K(\sqrt{p^*})/K$ is unramified.

In the second case, if $p \nmid x$ then we want $K(\sqrt{p^*})/K$ to be ramified at p, which is similar. □

Lecture 3

4. The Artin Pairing and the Rédei matrix

Recall that we were planning to understand the behaviour of our class group Cl_K by studying the steps in the filtration

$$\mathrm{Cl}_K[2] \supseteq 2\mathrm{Cl}_K[4] \supseteq 4\mathrm{Cl}_K[8] \supseteq \dots$$

Since Genus theory will describe the group at the top, our job today will be to study the transition between the steps in this filration. This will make use of the Artin pairing.

Definition 4.1 (Artin pairing)**.** For a finite abelian 2-group A we write $A^\vee = \mathrm{Hom}(A, \mathbb{C}^\times)$. We then define pairings

$$\mathrm{Art}_k : 2^{k-1}A[2^k] \times 2^{k-1}A^\vee[2^k] \to \{\pm 1\}$$

given simply by mapping a pair (a, χ) to $\psi(a)$ where $\psi \in A^\vee[2^k]$ is such that $2^{k-1}\psi = \chi$.

Of course, we need to know that this is well defined.

Lemma 4.2. Art_k *is well defined, i.e. it does not depend on the choice of lift* ψ*.*

Proof. Let us take a pair (a, χ) in the domain. Let ψ_1, ψ_2 be two lifts of χ to $A^\vee[2^k]$, so that $2^{k-1}(\psi_1 - \psi_2) = 0$. But if is $b \in A[2^k]$ is a lift of a then

$$\psi_1(a) = \psi_1(a) - \psi_2(a) + \psi_2(a) = (\psi_1 - \psi_2)(2^{k-1}b) + \psi_2(a) = \psi_2(a).$$

□

Why is this pairing useful to us? Well it will tell us precisely which elements lift to the next stage of the filtration.

Lemma 4.3. *The left kernel of* Art_k *is* $2^k A[2^{k+1}]$ *and the right kernel is* $2^k A^\vee[2^{k+1}]$.

Proof. By the independence of lifting in the definition of Art_k, really Art_k arises from the pairing $2^{k-1}A[2^k] \times A^\vee[2^k]$ given by evaluation. In particular, the left kernel pairs trivially with all of $A^\vee[2^k]$ under the evaluation (or canonical) pairing

$$\begin{aligned} A \times A^\vee &\to \mathbb{C}^\times \\ (a, \chi) &\mapsto \chi(a). \end{aligned}$$

It is elementary to show that the orthogonal complement of $A^\vee[2^k]$ under this canonical pairing is $2^k A$. Hence the left kernel of Art is $(2^{k-1}A[2^k]) \cap 2^k A = 2^k A[2^{k+1}]$. □

Now we must relate this finite abelian group theory back to class groups, and explain why we decided to call this the Artin pairing.

Lemma 4.4. *Identifying* $\mathrm{Cl}_K \cong \mathrm{Gal}(H_K/K)$ *via class field theory as before, in the commutative diagram*

$$\begin{array}{ccc} \mathrm{Cl}_K[2] \times \mathrm{Cl}_K^\vee[2] & \xrightarrow{\quad\mathrm{Art}_1\quad} & \{\pm 1\} \\ \sim\uparrow & & \\ \mathrm{Cl}_K[2] \times \mathrm{Hom}(\mathrm{Gal}(H_K/K), \{\pm 1\}) & \xrightarrow{\mathrm{Art}_{K,1}} & \{\pm 1\} \end{array}$$

the bottom row map is given on the class of an unramified prime ideal $\mathfrak{p}$ *and character* χ *by*

$$(\mathfrak{p}, \chi) \mapsto \chi\left(\mathrm{Frob}_{\mathfrak{p}}\right)$$

where $\mathrm{Frob}_{\mathfrak{p}}$ *is Frobenius.*

Now, recall from genus theory that if $-x \equiv 1 \mod 4$ and $x = p_1 \dots p_r$ for distinct odd primes p_i then the map

$$\varphi_x : \mathbb{F}_2^r \to \mathrm{Cl}_K[2]$$

given by $(e_1, \dots, e_r) \mapsto [\prod_{i=1}^r \mathfrak{p}_i^{e_i}]$ where $\mathfrak{p}_i$ are the primes above each p_i. We saw this is surjective with 1-dimensional kernel $(1, 1, \dots, 1)$.

We similarly have a map

$$\psi_x : \mathbb{F}_2^r \to \mathrm{Cl}_K^\vee[2]$$

given by $(e_1, \dots, e_r) \mapsto \prod_{i=1}^r \chi_{p_i^*}^{e_i}$. Again this is surjective with kernel $(1, \dots, 1)$ as before. Note that in fact this ψ naturally factors as the restriction to G_K of a natural map

$$\widetilde{\psi_x} : \mathbb{F}_2^r \to Z^1(\mathrm{Gal}(H_K/\mathbb{Q}), N[2]).$$

We're then going to define a lift of $\mathrm{Art}_{K,1}$ as follows.

Definition 4.5. We define $\widetilde{\mathrm{Art}_{K,1}} : \mathbb{F}_2^r \times \mathbb{F}_2^r \to \{\pm 1\}$ to be the lift of the Artin pairing under φ_x and ψ_x. That is,

$$\widetilde{\mathrm{Art}_{K,1}}(a, b) = \mathrm{Art}_{K,1}(\varphi_x(a), \psi_x(b)).$$

Definition 4.6. Let $-x \equiv 1 \mod 4$ be squarefree, $K = \mathbb{Q}(\sqrt{-x})$, and $r = \omega(x)$. Then the Rédei matrix R_x of K is defined to be the $r \times r$ matrix with entries

$\widetilde{\mathrm{Art}_{K,1}}(\mathbf{e}_i, \mathbf{e}_j)$ where the $\mathbf{e}_i$ are the elementary basis vectors with 1 in the ith component and zeros everywhere else. [R:Note that this is precisely the matrix such that in the standard basis,

$$\widetilde{\mathrm{Art}_{K,1}}(\mathbf{v}, \mathbf{w}) = \mathbf{v}^T R_x \mathbf{w}$$

]

Theorem 4.7. *Let $-x \equiv 1 \mod 4$ be squarefree, $K = \mathbb{Q}(\sqrt{-x})$, and $r = \omega(x)$. Then*

$$\dim_{\mathbb{F}_2} 2\mathrm{Cl}_K[4] = r - 1 - \mathrm{rk}(R_x).$$

Proof. Use that the left kernel of $\mathrm{Art}_{K,1}$ is $2\mathrm{Cl}_K[4]$ and that $\mathbb{F}_2^r \xrightarrow{\varphi_x} \mathrm{Cl}_K[2]$ is surjective with 1 dimensional kernel. □

We now want to determine the entries in the Rédei matrix in terms of symbols we are already comfortable with.

Theorem 4.8. *Write $R_x(i,j)$ for the i,j entry in R_x. Then we have for all indices $1 \le i, j \le r$*

$$R_x(i,j) = \begin{cases} \left(\frac{p_j^*}{p_i}\right) & \text{if } i \neq j \\ \sum_{k \neq i} R_x(i,k) & \text{if } i = j \end{cases}$$

where we interpret Legendre symbols as taking values in $\mathbb{F}_2$.

Proof. In the case $i \neq j$, by Lemma 4.4 and the definition of $\widetilde{\mathrm{Art}_{K,1}}$, we have

$$\begin{aligned} \widetilde{\mathrm{Art}_{K,1}}(\mathbf{e}_i, \mathbf{e}_j) &= \chi_{p_j^*}|_{G_K}(\mathrm{Frob}_{\mathfrak{p}_i}) \\ &= \left(\frac{p_j^*}{p_i}\right). \end{aligned}$$

Note that this second equality is decided by whether the residue degree of p_i is 1 or 2 in the (unramified!) extension $K(\sqrt{p_j^*})/K$ where $K = \mathbb{Q}(\sqrt{-x})$ as before, hence the second equality.

For the case $i = j$, note that $\prod_{j=1}^r \chi_{p_j^*} = \chi_{-x}$ is trivial when restricted to G_K, so all rows have sum equal to 1. □

Hence we can write down R_x as the following matrix (taking Legendre symbols to be valued in $\mathbb{F}_2$)

$$R_x = \begin{pmatrix} * & \left(\frac{p_2^*}{p_1}\right) & \cdots & \left(\frac{p_r^*}{p_1}\right) \\ \left(\frac{p_1^*}{p_2}\right) & * & \cdots & \left(\frac{p_r^*}{p_2}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \left(\frac{p_1^*}{p_r}\right) & \left(\frac{p_2^*}{p_r}\right) & \cdots & * \end{pmatrix} \tag{4.1}$$

where the $*$ are filled in by the sum of the rest of their row.

Example 4.9. Let us consider $x = 3 \times 5 \times 13 = 195$. Then

- $\mathrm{Cl}_K^\vee[2] = \langle \chi_{-3}, \chi_5, \chi_{13} \rangle$
- $\mathrm{Cl}_K[2] = \langle \mathfrak{p}_3, \mathfrak{p}_5, \mathfrak{p}_{13} \rangle$

and hence the Rédei matrix is

$$R = \begin{pmatrix} * & \left(\frac{5}{3}\right) & \left(\frac{13}{3}\right) \\ \left(\frac{-3}{5}\right) & * & \left(\frac{13}{5}\right) \\ \left(\frac{-3}{13}\right) & \left(\frac{5}{13}\right) & * \end{pmatrix} = \begin{pmatrix} * & 1 & 0 \\ 1 & * & 1 \\ 0 & 1 & * \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

This has rank 2 and so we get $\dim_{\mathbb{F}_2} 2\mathrm{Cl}_K[4] = 3 - 1 - 2 = 0$.

Lecture 4

A good heuristic then for $\mathrm{Art}_{K,1}$ will be that it should be a random matrix with entries in $\mathbb{F}_2$ (or ± 1), and that $\widetilde{\mathrm{Art}_{K,1}}$ should be a random matrix in $\mathbb{F}_2$ subject to the constraint that $(1, \dots, 1)$ is in the left and right kernel (i.e. the sums of the rows and columns are zero).

5. Gridding

[R:Our goal is now to let x range among positive squarefree integers and understand $\dim_{\mathbb{F}_2} 2\mathrm{Cl}_{\mathbb{Q}(\sqrt{-x})}[4]$. For convenience, we have restricted to only the odd x with $-x \equiv 1 \mod 4$. By Theorem 4.7, we want to understand the average behaviour of $\omega(x) - \mathrm{rk}(R_x)$ as x varies. A natural way to want to do this is to look at integers with exactly r prime factors for each r individually (freezing the value $r = \omega(x)$, so that R_x represents a pairing on the *fixed* vector space $\mathbb{F}_2^r$). Then we would like to let the prime factors of x vary at random, and use analytic techniques to describe the variation of the Legendre symbols in the entries of R_x to show that these matrices equidistribute among appropriate matrices in $\mathrm{Mat}(r \times r, \mathbb{F}_2)$. Grids are how we make this precise.]

We now allow ourselves a parameter H, which will be large and eventually tend to infinity.

Definition 5.1. Define $\alpha_j := e^{(\log\log H)^{100}}(1 + \frac{1}{\log H})^j$. For $H \geq 100$, a *grid of height* H is a pair

$$(X, i_{\mathrm{small}})$$

where $X = X_1 \times \dots X_r$ is a product such that

(1) Each X_i is a subset of the primes.
(2) For $i \leq i_{\mathrm{small}}$, we ensure $|X_i| = 1$ and moreover the element $p_i \in X_i$ should satisfy $p_i < e^{(\log\log H)^{100}}$.
(3) For $i > i_{\mathrm{small}}$ we insist that X_i is all of the primes in the interval $[\alpha_j, \alpha_{j+1})$ for some j.
(4) For all $1 \leq i < j \leq r$, then for every $p_i \in X_i$ and $p_j \in X_j$ we have $p_i < p_j$.
(5) Every $x = (p_1, \dots, p_r) \in X$ satisfies $p_1 p_2 \dots p_r \leq H$

We view a grid X as a subset of squarefree integers according to their prime factorisations. That is, a tuple $(p_1, \dots, p_r) \in X$ corresponds to the squarefree integer $p_1 \dots p_r$.

Theorem 5.2. *We have $X \cap X' = \emptyset$ for two distinct grids X, X' of height H.*

Proof. Exercise! □

Definition 5.3. We call a grid $X = X_1 \times \dots \times X_r$ good if

(1) $\frac{4}{5}\log\log H \leq r \leq (\log\log H)^2$
(2) $i_{\mathrm{small}} \leq (\log\log\log H)^2$

(3) $\#\left\{i \ : \ X_i \subseteq \left[1, e^{\sqrt{\log H}}\right)\right\} \leq \frac{2}{3}\log\log H.$

What is this saying? Why is this a good notion of good? Well using Erdős–Kac type heuristics: we expect that for a random squarefree integer n we should have about $\log\log y$ prime divisors smaller than y. (1) is concerned with the total number of prime factors, which by the heuristic should be around $\log\log H$, so these bounds describe a 'typical' behaviour. Now (2) is concerned with the number of small prime divisors, those smaller than $e^{(\log\log H)^{100}}$. The number of such primes should be around $\log\log\left(e^{(\log\log H)^{100}}\right) = 100\log\log\log H$. Certainly our bound should then model reality. Finally (3) bounds the number of prime divisors which are of medium size, meaning smaller than $e^{\sqrt{\log H}}$. This should be around $\log\log(e^{\sqrt{\log H}}) = \frac{1}{2}\log\log H$, so we bound by $\frac{2}{3}\log\log H$.

Now we show that one really can restrict to studying integers in good grids.

Theorem 5.4. *We have*

$$\#\left\{n \leq H \ : \ \text{sqf}\,, n \notin \bigcup_{X \text{ good}} X\right\} \ll \frac{H}{\log\log\log H}.$$

Proof. We break this proof into cases. Firstly we will deal with the number of integers which are not in any grid at all, then we deal with those who are not in a *good* grid.

Case 1 (not in a grid): Firstly, we bound those $n \leq H$ which are not in any grid of height H at all. This happens only if one of the following is true.

(1) n has a large neighbour who excludes the grid. That is, for example, say there is an X_i such that the $p_i \mid n$ from that X_i has a neighbour $q \in X_i$ such that $m = qn/p_i > H$. In this case, we would have thrown the grid out. In this case, checking the definitions, we would have

$$n \geq H(1 + \frac{1}{\log H})^{-\omega(n)}.$$

As $\omega(n) \ll \frac{\log H}{\log\log H}$ the last condition implies that $n \geq He^{-c/\log\log H}$ for some $c > 0$. So there are $\ll H/\log\log H$ such integers, which is within our proposed bound.

(2) n has two distinct prime divisors in an interval $[\alpha_j, \alpha_{j+1})$. The number of such n is bounded by

$$\begin{aligned}\sum_{j \ll (\log H)^2} \sum_{p,q \in [\alpha_j, \alpha_{j+1})} \frac{H}{pq} &\ll H \sum_{j \ll (\log H)^2} \sum_{p,q \in [\alpha_j, \alpha_{j+1})} \frac{1}{pq} \\ &\ll H \sum_{j \ll (\log H)^2} \left(\sum_{p \in [\alpha_j, \alpha_{j+1})} \frac{1}{p}\right)^2\end{aligned}$$

Applying Mertens' theorem, $\sum_{p\leq x} p^{-1} = \log\log X + M + O(e^{-\sqrt{\log X}})$ for some constant M. So in particular, our bound becomes

$$\begin{aligned}
&H \sum_{j\ll (\log H)^2} (\log\log \alpha_{j+1} - \log\log \alpha_j)^2 \\
&= H \sum_{j\ll (\log H)^2} \log\left(\frac{\log \alpha_{j+1}}{\log \alpha_j}\right)^2 \\
&= H \sum_{j\ll (\log H)^2} \log\left(\frac{\log \alpha_j + \log(1+\frac{1}{\log H})}{\log \alpha_j}\right)^2 \\
&\ll H \sum_{j\ll (\log H)^2} \left(\frac{1}{\log \alpha_j \log H}\right)^2
\end{aligned}$$

which is smaller than our bound once again.

Case 2 (in a grid, but not a good one): We now need to deal with integers which are in a grid, but not a good one.

Definition 5.5. Now we call a squarefree integer n good if

(1) $\frac{4}{5}\log\log H \leq \omega(n) \leq (\log\log H)^2$
(2) $\#\{p \mid n \ : \ p < \alpha_0\} \leq (\log\log\log H)^2$
(3) $\#\left\{p \mid n \ : \ p \leq 2e^{\sqrt{\log H}}\right\} \leq \frac{2}{3}\log\log H$

In particular, if n is good and in a grid X then it must be in a good grid. It suffices to estimate n failing the conditions (1), (2), or (3). We begin with the first.

Lemma 5.6. *For $X > 3$, we have*

$$\begin{aligned}
\sum_{n\leq X} \omega(n) &= X\log\log X + O(X) \\
\sum_{n\leq X} \omega(n)^2 &= X(\log\log X)^2 + O(X\log\log X)
\end{aligned}$$

In particular, $\sum_{n\leq X} (\omega(n) - \log\log X)^2 \ll X\log\log X$.

Proof. We apply Mertens' theorem

$$\sum_{n\leq X} \omega(n) = \sum_{p\leq X} \sum_{\substack{n\leq X \\ p\mid n}} 1 = \sum_{p\leq X} \left(\frac{X}{p} + O(1)\right) = O\left(\frac{X}{\log X}\right) + X\sum_{p\leq X} \frac{1}{p} = X\log\log X + O(X).$$

Then we compute

$$\begin{aligned}\sum_{n\le X}\omega(n)^2 &= \sum_{p,q\le X}\sum_{\substack{n\le X\\ p\mid n\\ q\mid n}} 1 = O(X\log\log X)+\sum_{p,q\le X}\sum_{\substack{n\le X\\ pq\mid n}} 1 = O(X\log\log X)+\sum_{p,q\le X}\left\lfloor\frac{X}{pq}\right\rfloor\\ &= O(X\log\log X)+\sum_{p,q\le X}\frac{X}{pq}\\ &= O(X\log\log X)+X\left(\sum_{p\le X}\frac{1}{p}\right)^2\\ &= X(\log\log X)^2+O(X\log\log X).\end{aligned}$$

For the final claim, we expand and note that the leading terms cancel and so we are bounded by the error. □

Lemma 5.7. *The number of squarefree $n<H$ which fail to be good because of (1) in Definition 5.5 is $\ll H/\log\log H$.*

Proof. We apply Lemma 5.6, in particular we know a bound on the second moment

$$\sum_{n\le H}(\omega(n)-\log\log H)^2 \ll H\log\log H. \tag{5.1}$$

Note that the contribution of any element of $\left\{n\le X\ :\ \omega(n)\le \frac{4}{5}\log\log H\right\}$ to the sum (5.1) is $\gg(\log\log H)^2$ and so

$$\#\left\{n\le H\ :\ \omega(n)\le\frac{4}{5}\log\log H\right\}\ll\frac{1}{(\log\log H)^2}\sum_{n\le H}(\omega(n)-\log\log H)^2\ll H/\log\log H.$$

Similarly, the contribution of any element of $\left\{n\le X\ :\ \omega(n)>(\log\log H)^2\right\}$ to the second moment sum is $\gg(\log\log H)^4$ and so similarly we obtain

$$\#\left\{n\le X\ :\ \omega(n)>(\log\log H)^2\right\}\ll H/(\log\log H)^3.$$

Then we get that the number of n failing condition (1) is small as required □

The argument for bounding the number who fail (2) or (3) is similar, and left as an exercise. □

Lecture 5

6. The large sieve

The most well-known version of the large sieve is due to Heath-Brown, which we now state below.

Theorem 6.1 (Heath-Brown)**.** *Let $M,N\in\mathbb{Z}_{\ge1}$, and let $\beta_1,\dots,\beta_N\in\mathbb{C}$ with $|\beta_j|\le1$. Then for all $\varepsilon>0$*

$$\sum_{m\le M}\left|\sum_{\substack{n\le N\\ \text{odd}}}\beta_n\left(\frac{m}{n}\right)\right|\ll_\varepsilon\frac{(MN)^{1+\varepsilon}}{\min\left\{M^{1/2},N^{1/2}\right\}}.$$

What really makes this result possible is the double sum, that you have two axes of randomness. Note that on the left there are MN terms, so we're seeing almost square-root saving in one of the variables.

Example 6.2. Take $M = N = X$ and $\beta_n = 1_{\text{prime}}$. Then

$$\sum_{m \leq X} \left| \sum_{\substack{p \leq X \\ \text{prime}}} \left(\frac{m}{p} \right) \right| \ll_\varepsilon X^{3/2+\varepsilon},$$

where the Legendre symbol is really telling us about the splitting of p in $\mathbb{Q}(\sqrt{m})$. From Markov's inequality, for m outside of an exceptional set $\mathcal{E}$ with $|\mathcal{E}| \leq X^{1-\varepsilon/2}$ (i.e. for almost all m)

$$\left| \sum_{\substack{p \leq X \\ \text{prime}}} \left(\frac{m}{p} \right) \right| \ll X^{\frac{1+\varepsilon}{2}}$$

is the GRH bound for $\mathbb{Q}(\sqrt{m})$.

The proof of Heath-Brown's result is quite technical, and we won't have time to get into it in this series, so we'll look at this slightly weaker but more pliable result of Friedlander–Iwaniec.

Theorem 6.3 (Friedlander–Iwaniec)**.** *Let $M, N \geq 1$ and $\beta_1, \dots, \beta_N \in \mathbb{C}$ with $|\beta_i| \leq 1$. Then for all $\varepsilon > 0$*

$$\sum_{\substack{m \leq M \\ \text{odd}}} \left| \sum_{\substack{n \leq N \\ \text{odd}}} \beta_n \left(\frac{m}{n} \right) \right| \ll_\varepsilon \frac{(MN)^{1+\varepsilon}}{\min\left\{ M^{1/6}, N^{1/6} \right\}}.$$

The disadvantage of this result is that the saving is not as good, but the advantage is that it is easier and generalises well to settings like cubic characters, Hecke characters over number fields (see work of Lemke Oliver–Smith), and similar.

Before the theorem, let us prove a lemma.

Lemma 6.4.

$$\sum_{\substack{m \leq M \\ \text{odd}}} \left| \sum_{\substack{n \leq N \\ \text{odd}}} \beta_n \left(\frac{m}{n} \right) \right| \ll_\varepsilon MN^{1/2+\varepsilon} + M^{1/2}N^2.$$

Remark 6.5. This is only nontrivial (meaning beating the trivial bound of MN) when $MN \geq M^{1/2}N^2$. Checking the algebra, this is when $M \geq N^2$.

Remark 6.6. From now on in lectures we will make substantial use of the Cauchy–Schwarz inequality:

$$\left| \sum_{a \leq A} x_a y_a \right| \leq \left(\sum_{a \leq A} x_a^2 \right)^{1/2} \left(\sum_{a \leq A} y_a^2 \right)^{1/2}.$$

We will also make use of Hölder's inequality: for fixed $p, q \in \mathbb{R}_{>1}$ such that $\frac{1}{p}+\frac{1}{q}=1$

$$\sum_{a\le A}|x_a y_a| \le \left(\sum_{a\le A}|x_a|^p\right)^{1/p}\left(\sum_{a\le A}|y_a|^q\right)^{1/q}.$$

Try not to forget these!

Proof. Applying Cauchy–Schwarz with $x_m = 1$ and $y_m = \left|\sum_{\substack{n\le N\\ \text{odd}}}\beta_n\left(\frac{m}{n}\right)\right|$, we have

$$\begin{aligned}\sum_{\substack{m\le M\\ \text{odd}}} 1\cdot\left|\sum_{\substack{n\le N\\ \text{odd}}}\beta_n\left(\frac{m}{n}\right)\right| &\le \left(\sum_{\substack{m\le M\\ \text{odd}}}1\right)^{1/2}\left(\sum_{\substack{m\le M\\ \text{odd}}}\left|\sum_{\substack{n\le N\\ \text{odd}}}\beta_n\left(\frac{m}{n}\right)\right|^2\right)^{1/2}\\ &\le M^{1/2}\left(\sum_{m\le M}\sum_{\substack{n_1,n_2\le N\\ \text{odd}}}\beta_{n_1}\overline{\beta_{n_2}}\left(\frac{m}{n_1n_2}\right)\right)^{1/2}.\end{aligned}$$

Now switching the order of summation and applying the triangle inequality we obtain

$$\begin{aligned}\sum_{\substack{n_1,n_2\le N\\ \text{odd}}}\left|\sum_{m\le M}\left(\frac{m}{n_1n_2}\right)\right| &= \sum_{\substack{n_1,n_2\le N\\ \text{odd}\\ n_1n_2=\square}}\left|\sum_{m\le M}\left(\frac{m}{n_1n_2}\right)\right| + \sum_{\substack{n_1,n_2\le N\\ \text{odd}\\ n_1n_2\ne\square}}\left|\sum_{m\le M}\left(\frac{m}{n_1n_2}\right)\right|\\ &\ll N(\log N)M + N^4.\end{aligned}$$

where for the first term we have applied the trivial bound (since the character is trivial so there is no saving to be had!). For the second term we used that this is a non-principal Dirichlet character of modulus at most N^2 so when we sum over N^2 integers in a row we get zero. Collecting terms, we get the result. □

Proof of Theorem 6.3. We claim that this theorem is equivalent to the following statement: for all α_m, β_n in $\mathbb{C}$ with absolute value at most 1, we have

$$\left|\sum_{\substack{m\le M\\ \text{odd}}}\sum_{\substack{n\le N\\ \text{odd}}}\alpha_m\beta_n\left(\frac{m}{n}\right)\right| \ll_\varepsilon \frac{(MN)^{1+\varepsilon}}{\min\left\{M^{1/6}, N^{1/6}\right\}}. \tag{6.1}$$

The equivalence is as follows.

FI $\implies$ (6.1) Just apply the triangle inequality in (6.1) to get

$$\left|\sum_{\substack{m\le M\\ \text{odd}}}\sum_{\substack{n\le N\\ \text{odd}}}\alpha_m\beta_n\left(\frac{m}{n}\right)\right| \le \sum_{\substack{m\le M\\ \text{odd}}}\left|\sum_{\substack{n\le N\\ \text{odd}}}\alpha_m\beta_n\left(\frac{m}{n}\right)\right|.$$

Then pull the α_m out and conclude.

(6.1) $\implies$ **FI** Take

$$\alpha_m = \begin{cases}\frac{\left|\sum_{n\le N}\beta_n\left(\frac{m}{n}\right)\right|}{\sum_{n\le N}\beta_n\left(\frac{m}{n}\right)} & \text{if the denominator is nonzero,}\\ 0 & \text{else,}\end{cases}$$

and apply (6.1).

Now we prove (6.1), for which a main tool will be Hölder's inequality. WLOG we assume that $M \geq N$ and by the triangle inequality it suffices to bound

$$\sum_{\substack{n\leq N\\ \text{odd}}} 1 \cdot \left| \sum_{\substack{m\leq M\\ \text{odd}}} \alpha_m \left(\frac{m}{n}\right)\right|.$$

By Hölder this is at most

$$\left(\sum_{\substack{n\leq N\\ \text{odd}}} 1\right)^{2/3} \left(\sum_{\substack{n\leq N\\ \text{odd}}} \left| \sum_{\substack{m\leq M\\ \text{odd}}} \alpha_m \left(\frac{m}{n}\right)\right|^3\right)^{1/3} \leq N^{2/3} \left(\sum_{\substack{n\leq N\\ \text{odd}}} \left| \sum_{\substack{m\leq M\\ \text{odd}}} \alpha_m \left(\frac{m}{n}\right)\right|^3\right)^{1/3}.$$

We let S now denote the remaining double sum here, and set

$$\gamma_n = \frac{\left| \sum_{\substack{m\leq M\\ \text{odd}}} \alpha_m \left(\frac{m}{n}\right)\right|^3}{\left(\sum_{\substack{m\leq M\\ \text{odd}}} \alpha_m \left(\frac{m}{n}\right)\right)^3},$$

so that $|\gamma_n| \leq 1$ and

$$S = \sum_{\substack{n\leq N\\ \text{odd}}} \gamma_n \left(\sum_{\substack{m\leq M\\ \text{odd}}} \alpha_m \left(\frac{m}{n}\right)\right)^3.$$

We then expand the cube to obtain

$$\begin{aligned} S &= \sum_{\substack{n\leq N\\ \text{odd}}} \gamma_n \sum_{\substack{m_1,m_2,m_3\leq M\\ \text{odd}}} \alpha_{m_1}\alpha_{m_2}\alpha_{m_3} \left(\frac{m_1m_2m_3}{n}\right) \\ &\leq \sum_{\substack{m_1,m_2,m_3\leq M\\ \text{odd}}} 1 \cdot \left| \sum_{\substack{n\leq N\\ \text{odd}}} \gamma_n \left(\frac{m_1m_2m_3}{n}\right)\right| \\ &\leq \sum_{\substack{\ell\leq M^3\\ \text{odd}}} \tau_3(\ell) \cdot \left| \sum_{\substack{n\leq N\\ \text{odd}}} \gamma_n \left(\frac{\ell}{n}\right)\right| \\ &\ll M^{\varepsilon} \sum_{\substack{\ell\leq M^3\\ \text{odd}}} \left| \sum_{\substack{n\leq N\\ \text{odd}}} \gamma_n \left(\frac{\ell}{n}\right)\right|. \end{aligned}$$

Now we apply Lemma 6.4 and conclude. □

Lecture 6

7. The average 4-torsion

Our goal now is to study the average of the 4-torsion in class groups of imaginary quadratic fields, á la Fouvry–Klüners.

Theorem 7.1 (Fouvry–Klüners 2007)**.** *We have*

$$\lim_{H\to\infty} \frac{\#\left\{K/\mathbb{Q} \text{ imag. quad. } |D_K| \le H,\ 2\mathrm{Cl}_K[4] \cong \mathbb{F}_2^n\right\}}{\#\left\{K/\mathbb{Q} \text{ imag. quad. } |D_K| \le H\right\}} = \lim_{r\to\infty} P_{\mathrm{Mat}}(r,n).$$

Recall that by Theorem 5.4 we simply need to prove this in good grids X of height H. i.e. we want

Theorem 7.2. *For every good grid X of height H we have*

$$\frac{\#\left\{x \in X \ :\ 2\mathrm{Cl}_{\mathbb{Q}(\sqrt{-x})}[4] \cong \mathbb{F}_2^n\right\}}{\#X} = \lim_{r\to\infty} P(r,n) + o(1),$$

as $H \to \infty$.

We take the moment approach here, it is enough to prove that for each $k \ge 1$

$$\sum_{x\in X} \#\left(2\mathrm{Cl}_{\mathbb{Q}(\sqrt{-x})}[4]\right)^k \sim c_k \cdot |X|$$

for c_k the explicit moments of the proposed distribution. Today we prove the $k = 1$ case. Then $k = 2$ is in the exercises, and $k \ge 3$ can be found in the work of Fouvry–Klüners. For $k = 1$ case we will prove

$$\sum_{\substack{x\in X\\ x\equiv 3 \mod 4}} \#2\mathrm{Cl}_{\mathbb{Q}(\sqrt{-x})}[4] \sim 2\#\left\{x \in X \ :\ x \equiv 3 \mod 4\right\}.$$

For convenience, we restrict to the subgrid $X' = X_1' \times \dots X_r'$ with X_i' defined to be $\{p \in X_i \ :\ p \equiv \alpha_i \mod 4\}$ with some fixed $\alpha_i \in \{1,3\}$. Then it suffices to prove for each tuple of α_i such that $\prod_{i=1}^r \alpha_i \equiv 3 \mod 4$ we get

$$\sum_{x\in X'} \#2\mathrm{Cl}_{\mathbb{Q}(\sqrt{-x})}[4] \sim 2\,|X'|.$$

Recall from Theorem 4.7 that we want to study the Rédei matrices since

$$\#2\mathrm{Cl}_{\mathbb{Q}(\sqrt{-x})}[4] = \frac{1}{2}\#\ker(R_x).$$

Let us reshape the Rédei matrices a little bit. Applying quadratic reciprocity, since $\left(\frac{p_i^*}{p_j}\right) = \left(\frac{p_j}{p_i}\right)$

$$R_x = \begin{pmatrix} * & \left(\frac{p_2^*}{p_1}\right) & \dots & \left(\frac{p_r^*}{p_1}\right) \\ \left(\frac{p_1^*}{p_2}\right) & * & \dots & \left(\frac{p_r^*}{p_2}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \left(\frac{p_1^*}{p_r}\right) & \left(\frac{p_2^*}{p_r}\right) & \dots & * \end{pmatrix} = \begin{pmatrix} a_{11} & \left(\frac{p_1}{p_2}\right) & \dots & \left(\frac{p_1}{p_r}\right) \\ \left(\frac{p_2}{p_1}\right) & a_{22} & \dots & \left(\frac{p_2}{p_r}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \left(\frac{p_r}{p_1}\right) & \left(\frac{p_r}{p_2}\right) & \dots & a_{rr} \end{pmatrix} \tag{7.1}$$

where the a_{rr} are such that the column sums are zero. For each vector $\mathbf{v} \in \mathbb{F}_2^r$, we want to understand when $\mathbf{v}$ is in the left kernel of R_x. Note that being in the left

Kernel is equivalent to $\mathbf{v} \cdot R_x(i) = 0$ for all i, where $R_x(i)$ denotes the ith column of the Rédei matrix. For each i, now that we have written down a_{ii}, we can see that

$$\mathbf{v} \cdot R_x(i) = \left(\frac{\prod_{j \neq i} p_j^{v_j + v_i}}{p_i} \right).$$

Hence the indicator function $1_{\mathbf{v} \in \ker_l(R_x)}$ is given explicitly by

$$\begin{aligned} 1_{\mathbf{v} \in \ker_l(R_x)} &= \frac{1}{2^r} \prod_{i=1}^{r} \left(1 + \left(\frac{\prod_{j \neq i} p_j^{v_j + v_i}}{p_i} \right) \right) \\ &= \frac{1}{2^r} \prod_{\substack{i=1 \\ v_i = 0}}^{r} \left(1 + \left(\frac{\prod_{j \neq i} p_j^{v_j}}{p_i} \right) \right) \prod_{\substack{i=1 \\ v_i = 1}}^{r} \left(1 + \left(\frac{\prod_{j \neq i} p_j^{v_j + 1}}{p_i} \right) \right), \end{aligned}$$

where we are considering the Jacobi symbol to be valued in ± 1. Let us now expand these products. For now $\mathbf{v} \in \mathbb{F}_2^r$ and $x = p_1 \dots p_r \in X'$. We let $V_0 = V_0(\mathbf{v})$ denote the set of indices $i \in \{1, \dots, r\}$ such that $v_i = 0$, and similarly V_1, and moreover for any subset $S \subseteq \{1, \dots, r\}$ and element $x = p_1 \dots p_r \in X'$ we write $x(S) = \prod_{i \in S} p_i$. Then we expand the products in the indicator function to obtain

$$\begin{aligned} 2^r \cdot 1_{\mathbf{v} \in \ker_l(R_x)} &= \prod_{\substack{i=1 \\ v_i = 0}}^{r} \left(1 + \left(\frac{x(V_1)}{p_i} \right) \right) \prod_{\substack{i=1 \\ v_i = 1}}^{r} \left(1 + \left(\frac{x(V_0)}{p_i} \right) \right) \\ &= \left(\sum_{S_0 \subseteq V_0} \left(\frac{x(V_1)}{x(S_0)} \right) \right) \left(\sum_{S_1 \subseteq V_1} \left(\frac{x(V_0)}{x(S_1)} \right) \right) \\ &= \sum_{S_0 \subseteq V_0} \sum_{S_1 \subseteq V_1} \left(\frac{x(V_1)}{x(S_0)} \right) \left(\frac{x(V_0)}{x(S_1)} \right) \end{aligned}$$

Partitioning $V_i = T_i \sqcup S_i$ we then have

$$2^r \cdot 1_{\mathbf{v} \in \ker_l(R_x)} = \sum_{S_0 \sqcup T_0 = V_0} \sum_{S_1 \sqcup T_1 = V_1} \left(\frac{x(S_1)x(T_1)}{x(S_0)} \right) \left(\frac{x(S_0)x(T_0)}{x(S_1)} \right).$$

Now, note that the choice of $\mathbf{v} \in \mathbb{F}_2^r$ is equivalent to a choice of V_0 and V_1. Hence when we sum over all such vectors (and relabelling $S_{2+i} = T_i$), we get

$$\sum_{\mathbf{v} \in \mathbb{F}_2^r} 1_{\mathbf{v} \in \ker_l(R_x)} = \frac{1}{2^r} \sum_{S_0 \sqcup S_1 \sqcup S_2 \sqcup S_3 = [r]} \left(\frac{x(S_1)x(S_3)}{x(S_0)} \right) \left(\frac{x(S_0)x(S_2)}{x(S_1)} \right).$$

Feeding this back into our main sum, we then have the following.

$$\begin{aligned} \sum_{x \in X'} \#2\mathrm{Cl}_{\mathbb{Q}(\sqrt{-x})}[4] &= \frac{1}{2} \sum_{x \in X'} \sum_{\mathbf{v} \in \mathbb{F}_2^r} 1_{\mathbf{v} \in \ker_l(R_x)} \\ &= \frac{1}{2^{r+1}} \sum_{x \in X'} \sum_{S_0 \sqcup S_1 \sqcup S_2 \sqcup S_3 = [r]} \left(\frac{x(S_1)x(S_3)}{x(S_0)} \right) \left(\frac{x(S_0)x(S_2)}{x(S_1)} \right) \end{aligned}$$

Observe that sometimes we have $\left(\frac{a}{b}\right)$ and $\left(\frac{b}{a}\right)$ in this product, and sometimes we only have one of the two. When we have exactly one of the pair, we expect that the symbol will oscillate randomly and so we will get savings in most regions. This is codified in the notion of linked indices.

Definition 7.3. We say that 0 is linked with 3 and 1 is linked with 2. All other pairs in $\{1,2,3,4\}$ are declared to be unlinked. We also write i_{medium} for the last index i such that $X_i \subseteq [1, e^{\sqrt{\log(X)}}]$

We will now make use of the large sieve and Siegel–Walfisz results to remove cases.

Case 1: if there are linked i,j such that $S_i \not\subseteq [i_{\text{small}}]$ and $S_j \not\subseteq [i_{\text{small}}]$ then let i_0, j_0 be indices larger that i_{small} in each of these. Then we can freeze the values of the other variables, apply the triangle inequality, and apply the Large sieve over X_{i_0} and X_{j_0} (exercise: what are the coefficients?) to show that these contributions are small.

Case 2: if there are linked i,j such that $S_i \subseteq [i_{\text{small}}]$ and nonempty and $S_j \not\subseteq [i_{\text{medium}}]$ (or vice versa). We apply effective Chebotarev to show that these contributions are small (actually, this only works if there are no Siegel zeroes).

Case 3: In the remaining cases, the following conditions both hold:

- $(S_0, S_3) \subseteq ([i_{\text{small}}], [i_{\text{medium}}]) \cup ([i_{\text{medium}}], [i_{\text{small}}])$ or $\emptyset \in \{S_0, S_3\}$;
- $(S_1, S_2) \subseteq ([i_{\text{small}}], [i_{\text{medium}}]) \cup ([i_{\text{medium}}], [i_{\text{small}}])$ or $\emptyset \in \{S_1, S_2\}$.

Applying Siegel–Walfisz in all cases we have nontrivial characters, we are left with some cases where all variables are at most medium in size (so can be discarded by a trivial bound) and then a main term which correpsonds to a sum of the four cases where our characters are forced to be constant: $(S_0, S_2) = (\emptyset, \emptyset)$, $(S_1, S_3) = (\emptyset, \emptyset)$, $(S_0, S_1) = (\emptyset, \emptyset)$, $(S_2, S_3) = (\emptyset, \emptyset)$. Note that in the first three, the character is trivial, and in the last it is controlled by congruence modulo 4 (which is determined here!).

Lecture 7

8. Repeated Cauchy–Schwarz

Today we'll see a key new ingredient in this area: repeated Cauchy–Schwarz. [R:Our goal for the remaining lectures will be to understand $4\mathrm{Cl}_K[8]$, assuming the structure of $2\mathrm{Cl}_K[4]$. Here is a global sketch of the argument we are going to pursue a subcase of over the coming lectures. We work on a good grid $X = X_1 \times \cdots \times X_r$. For each $x \in X$, writing $K = \mathbb{Q}(\sqrt{-x})$, we have already seen how to determine $2\mathrm{Cl}_K[4]$ via the Rédei matrix R_x which represents the pairing in the commutative diagram]

$$\begin{array}{ccc} \mathbb{F}_2^r \times \mathbb{F}_2^r & \xrightarrow{R_x} & \{\pm 1\} \\ \downarrow{\scriptstyle \varphi_x \times \psi_x} & & \| \\ \mathrm{Cl}_K[2] \times \mathrm{Cl}_K^\vee[2] & \xrightarrow{\mathrm{Art}_{x,1}} & \{\pm 1\}. \end{array}$$

[R: Since we understand the distribution of R_x, we are emboldened to fix the value R for the Rédei matrix, which in turn by Theorem 4.7 fixes $\dim_{\mathbb{F}_2} 2\mathrm{Cl}_K[4]$. This means we are restricting now to working with the subset

$$Y := \{x \in X \ : \ R_x = R\} \subseteq X.$$

A first, rather frustrating, observation is that this is *not* a subgrid of X, so we will need to be able to average over this kind of arithmetically defined subset of a grid. But let us continue.

Having fixed R, lets take $V = \ker_l(R)$ and $W = \ker_r(R)$. Then the pairing $\mathrm{Art}_{x,2}$ for $x \in Y$ extends naturally to a pairing between these two fixed (independent of $x \in Y$!) vector spaces]

$$\begin{array}{ccc} V \times W & \xrightarrow{\widetilde{\mathrm{Art}_{x,2}}} & \{\pm 1\} \\ \downarrow{\scriptstyle \varphi_x \times \psi_x} & & \| \\ 2\mathrm{Cl}_K[4] \times 2\mathrm{Cl}_K^{\vee}[4] & \xrightarrow{\mathrm{Art}_{x,2}} & \{\pm 1\}. \end{array}$$

[R: Where, as usual, $K = \mathbb{Q}(\sqrt{-x})$. Again, it is not hard to see that $4\mathrm{Cl}_K[8]$ is determined by the left kernel of $\widetilde{\mathrm{Art}_{x,2}}$, and again we have turned our problem into studying how a pairing varies randomly on the fixed space $V \times W$. Our goal now is to prove that this pairing, more accurately the corresponding matrix, is uniformly random.

We want to reduce proving such a statement to understanding some kind of character sum as before. Let $(1, \dots, 1), \mathbf{v}_1, \dots, \mathbf{v}_t$ and $(1, \dots, 1), \mathbf{w}_1, \dots, \mathbf{w}_t$ be bases for V and W respectively. Then the matrix which defines $\widetilde{\mathrm{Art}_{x,2}}$ is $(t+1) \times (t+1)$ and (since the all 1 vector, our first basis vector, is still in the kernel) has zeroes in the first row and column, and the remaining submatrix is given by $\widetilde{\mathrm{Art}_{x,2}}(\mathbf{v}_i, \mathbf{w}_j)$. We want to see that this submatrix takes any value with equal probability, how should we do this? Well, it is equivalent to show that the $t \times t$ submatrix is random.

Firstly, we want to know that each entry of this matrix can be ± 1 with equal probability. Hence, for a fixed pair $\mathbf{v}_i, \mathbf{w}_j$, this would be equivalent to showing that

$$\left| \sum_{x \in Y} \widetilde{\mathrm{Art}_{x,2}}(\mathbf{v}_i, \mathbf{w}_j) \right| = o(\#Y),$$

since a saving is equivalent to the densities of $+1$ and -1 have the same leading term. Then we need to check that there isn't some relation between pairs. For example, perhaps two entries always have to be the same. I this case, it's perfectly reasonable for one (and so both) of them to be ± 1 with equal probability but the matrix will still not equidistribute. To avoid such a correlation the values of the pairing on two pairs $(\mathbf{v}_i, \mathbf{w}_j)$ and $(\mathbf{v}_k, \mathbf{w}_\ell)$, we must similarly show that

$$\left| \sum_{x \in Y} \widetilde{\mathrm{Art}_{x,2}}(\mathbf{v}_i, \mathbf{w}_j) \widetilde{\mathrm{Art}_{x,2}}(\mathbf{v}_k, \mathbf{w}_\ell) \right| = o(\#Y).$$

Continuing along this chain of reasoning (triple correlations and so on), we must show that for every choice of $\eta : [t] \times [t] \to \{0, 1\}$ except for the constant zero function, we get

$$\left| \sum_{x \in Y} \prod_{i=1}^{t} \prod_{j=1}^{t} \widetilde{\mathrm{Art}_{x,2}}(\mathbf{v}_i, \mathbf{w}_j)^{\eta(i,j)} \right| = o(\#Y).$$

Over the next few lectures we'll work out the details of a special case of this sketch, which illustrates the key points.]

Our simplified setup for this and the next lecture is as follows: we have three sets of primes

- $X_1 = \{H < p < 2H \ : \ p \equiv 3 \mod 4\}$

- $X_2 = \{2H < q < 3H \ : \ q \equiv 1 \mod 4\}$
- $X_3 = \{3H < r < 4H \ : \ r \equiv 1 \mod 4\}$

and as usual the product $X = X_1 \times X_2 \times X_3$. Note that we've set these up in such a way that for every product $x = pqr$ from X, $-x$ is an imaginary quadratic discriminant.

For each $x = (p, q, r) \in X$, writing $K = \mathbb{Q}(\sqrt{-x})$, we have already seen how to determine $2\mathrm{Cl}_K[4]$ via the Rédei matrix R_x which represents the pairing in the following commutative diagram.

$$\begin{array}{ccc} \mathbb{F}_2^3 \times \mathbb{F}_2^3 & \xrightarrow{R_x} & \{\pm 1\} \\ \downarrow{\scriptstyle \varphi_x \times \psi_x} & & \| \\ \mathrm{Cl}_K[2] \times \mathrm{Cl}_K^\vee[2] & \xrightarrow{\mathrm{Art}_{x,1}} & \{\pm 1\}. \end{array}$$

Our overall goal for the remaining lectures will be to understand $4\mathrm{Cl}_K[8]$, assuming the structure of $2\mathrm{Cl}_K[4]$ or, more accurately, assuming the slightly finer data of a fixed Rédei matrix R_x. Let us first discuss the maximal case where $R_x = 0$, so that via Theorem 4.7 we have

$$\mathrm{Cl}_K[4] \cong (\mathbb{Z}/4\mathbb{Z})^2$$

Then the set of $x \in X$ we are restricting our consideration to is

$$Y := \{x \in X \ : \ R_x = 0\}.$$

Taking the explicit description of the Rédei matrix from (7.1), this can be re-expressed as

$$Y = \left\{(p, q, r) \in X \ : \ \left(\frac{p}{q}\right) = \left(\frac{p}{r}\right) = \left(\frac{q}{r}\right) = +1\right\}.$$

A first, rather frustrating, observation is that this is *not* a subgrid of X, so we will need to be able to average over this kind of arithmetically defined subset of a grid.

Looking on to the 8-torsion, we now want to understand the pairing $\mathrm{Art}_{x,2}$. Having fixed R_x, we will lift $\mathrm{Art}_{x,2}$ to a pairing $\widetilde{\mathrm{Art}_{x,2}}$ which is defined on the fixed (i.e. independent of x) subspace $\ker_l(R_x) \times \ker_r(R_x)$, and again have a problem of a pairing which varies randomly on a fixed vector space. Then we aim to prove that this pairing, more accurately the corresponding matrix, varies uniformly among all possible pairings.

Since we have restricted to the case $R_x = 0$, the left and right kernels are the total space and so we have

$$\begin{array}{ccc} \mathbb{F}_2^3 \times \mathbb{F}_2^3 & \xrightarrow{\widetilde{\mathrm{Art}_{x,2}}} & \{\pm 1\} \\ \downarrow{\scriptstyle \varphi_x \times \psi_x} & & \| \\ 2\mathrm{Cl}_K[4] \times 2\mathrm{Cl}_K^\vee[4] & \xrightarrow{\mathrm{Art}_{x,2}} & \{\pm 1\}. \end{array}$$

Let us work out a construction of the pairing. By definition, $\widetilde{\mathrm{Art}_{x,2}}(\mathbf{v}, \mathbf{w})$ is constructed by choosing an element $f_{x,\mathbf{w}} \in \mathrm{Cl}_K^\vee[4]$ such that $2f_{x,\mathbf{w}} = \psi_x(\mathbf{w})$ and then from this we have $\widetilde{\mathrm{Art}_{x,2}}(\mathbf{v}, \mathbf{w}) = f_{x,\mathbf{w}}(\varphi_x(\mathbf{v}))$. For later convenience, we lift to $Z^1(\mathrm{Gal}(H_K/\mathbb{Q}), N(-x)[4])$ by Theorem 3.2 and choose $\widetilde{f_{x,\mathbf{w}}}$ in this group such that $2\widetilde{f_{x,\mathbf{w}}} = \widetilde{\psi_x}(\mathbf{w})$ exactly. We further, via Definition 3.4, inflate $\widetilde{f_{x,\mathbf{w}}}$ to an element

$\psi_x^4(\mathbf{w}) \in Z^1_{\mathrm{nr}}(G_{\mathbb{Q}}, N(-x)[4])$. In particular, $\psi_x^4(\mathbf{w})$ is well defined on $\mathrm{Gal}(H_K/\mathbb{Q})$ and on $\mathrm{Gal}(H_K/\mathbb{Q})$ we have $2\psi_x^4(\mathbf{w}) = \widetilde{\psi_x}(\mathbf{w})$.[R:One day I will come back and change $\psi_x^4(\mathbf{w})$ to $\psi^4_{x,\mathbf{w}}$ which seems slightly tidier, but that day has not yet come.]

Our goal over the next few lectures is to discuss the following claim:

Theorem 8.1.

$$\left|\sum_{x\in Y} \widetilde{\mathrm{Art}_{x,2}}(\mathbf{v},\mathbf{w})\right| = o(H^3/\log(H)^3).$$

Actually it is possible to reach $o(H^{3-\delta})$, but we'll not work that hard. Why do we want to do this? Obtaining a saving here ensures that the leading terms for the number of x with value $+1$ must be the same as the leading term for those with value -1.

Important: During this and next lecture we will specialise to the setting $\mathbf{v} = (1,0,0)$ and $\mathbf{w} = (0,1,0)$!

Now,

$$\widetilde{\mathrm{Art}_{x,2}}(\mathbf{v},\mathbf{w}) = \psi_x^4(\mathbf{w})(\mathrm{Frob}_{\mathfrak{r}})$$

where $\mathfrak{r} \mid r \in X_3$ is our prime in K, and ψ_x^4 is really just a lift of the character χ_q for $q \in X_1$.

We will begin by looking for relations which constrain ψ_x^4 as an element of $\mathrm{Map}(G_{\mathbb{Q}}, N(-x))$, where we recall that $N(-x)$ is the abelian group $\mathbb{Q}_2/\mathbb{Z}_2$ upon which we act by Galois via the quadratic character χ_{-x}.

We note the following repeated application of Cauchy–Schwarz from exercise sheet 3.

Lemma 8.2.

$$\sum_{x\in Y} \widetilde{\mathrm{Art}_{x,2}}(\mathbf{v},\mathbf{w}) \leq |X|^{3/4} \left| \sum_{\substack{p_1,p_2\in X_1\\ q_1,q_2\in X_2\\ r_1,r_2\in X_3}} \prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k=1}^{2} \widetilde{\mathrm{Art}_{p_iq_jr_k,2}}(\mathbf{v},\mathbf{w}) \right|^{1/8}$$

This will be a key ingredient in proving the following main theorem.

Theorem 8.3. *Let $p_1,p_2 \in X_1$, $q_1,q_2 \in X_2$, $r \in X_3$. Assume that $p_iq_jr \in Y$. Then setting $C := \{p_1,p_2\}\times\{q_1,q_2\}\times\{r\}$, we obtain that*

$$\sum_{x\in C} \psi_x^4(\mathbf{w}) \in \mathrm{Map}(G_{\mathbb{Q}}, N[2]).$$

Moreover, let $d : \mathrm{Map}(G_{\mathbb{Q}}, N) \to \mathrm{Map}(G_{\mathbb{Q}}^2, N)$ be the differential on the chain complexes which define Galois cohomology, we have

$$d\left(\sum_{x\in C} \psi_x^4(\mathbf{w})\right)(\sigma,\tau) = \chi_{p_1p_2}(\sigma)\chi_{q_1q_2}(\tau).$$

Remark 8.4. Recall that $d : \mathrm{Map}(G_{\mathbb{Q}}, N) \to \mathrm{Map}(G_{\mathbb{Q}}^2, N)$ is given by

$$d(f)(\sigma,\tau) = f(\sigma\tau) - f(\sigma) - f(\tau).$$

Note that $(\sigma,\tau) \mapsto \chi_{p_1p_2}(\sigma)\chi_{q_1q_2}(\tau)$ is really the element $\chi_{p_1p_2}\cup\chi_{q_1q_2} = (p_1p_2, q_2q_2) \in H^2(G_{\mathbb{Q}}, \mu_2)$ as a Hilbert symbol, which we already know is trivial everywhere locally

(and hence globally) via our constraints Y. Later we will play off this coboundary against a different nice trivialisation of (i.e. coboundary representing) this class.

Proof. Note that

$$\psi_x(\mathbf{w}) = \begin{cases} \chi_{q_1} & \text{if second coordinate of } x \text{ is } q_1 \\ \chi_{q_2} & \text{if second coordinate of } x \text{ is } q_2 \end{cases}$$

Hence $2\left(\sum_{x\in C}\psi_x^4(\mathbf{w})\right) = \chi_{q_1}+\chi_{q_2}+\chi_{q_1}+\chi_{q_2} = 0$, so this sum must have been valued in 2-torsion.

To see the identity with d note that since $\psi_x^4(\mathbf{w}) \in Z^1(G_{\mathbb{Q}}, N[4])$, we have

$$\psi_x^4(\mathbf{w})(\sigma\tau) = \chi_{-x}(\sigma)\psi_x^4(\mathbf{w})(\tau) + \psi_x^4(\mathbf{w})(\sigma),$$

and hence computing the differential we have

$$\begin{aligned}\left(d\psi_x^4(\mathbf{w})\right)(\sigma,\tau) &= \psi_x^4(\mathbf{w})(\sigma\tau) - \psi_x^4(\mathbf{w})(\sigma) - \psi_x^4(\mathbf{w})(\tau) \\ &= \chi_{-x}(\sigma)\psi_x^4(\mathbf{w})(\tau) + \psi_x^4(\mathbf{w})(\sigma) - \psi_x^4(\mathbf{w})(\sigma) - \psi_x^4(\mathbf{w})(\tau) \\ &= \left(\chi_{-x}(\sigma) - 1\right)\psi_x^4(\mathbf{w})(\tau) \\ &= \begin{cases} 0 & \text{if } \chi_{-x}(\sigma) = +1 \\ -2\psi_x^4(\mathbf{w})(\tau) = \chi_{\pi_2(x)} & \text{if } \chi_{-x}(\sigma) = -1 \end{cases}\end{aligned}$$

where $\pi_2(x) = \pi_2(p,q,r) = q$. Note that this is therefore valued in the 2-torsion! Using the unique non-trivial pairing $N[2]\times N[2] \to N[2]$ we see

$$d\left(\psi_x^4(\mathbf{w})\right)(\sigma,\tau) = \chi_{-x}(\sigma)\chi_{\pi_2(x)}(\tau).$$

Hence

$$\begin{aligned} d\left(\sum_{x\in C}\psi_x^4(\mathbf{w})\right)(\sigma,\tau) &= \chi_{-p_1q_1r}(\sigma)\chi_{q_1}(\tau) + \chi_{-p_1q_2r}(\sigma)\chi_{q_2}(\tau) + \chi_{-p_2q_1r}(\sigma)\chi_{q_1}(\tau) + \chi_{-p_2q_2r}(\sigma)\chi_{q_2}(\tau) \\ &= \chi_{p_1p_2}(\sigma)\chi_{q_1}(\tau) + \chi_{p_1p_2}(\sigma)\chi_{q_2}(\tau) \\ &= \chi_{p_1p_2}(\sigma)\chi_{q_1q_2}(\tau). \end{aligned}$$

□

Lecture 8

9. Comparison of second Artin pairing

Since $\chi_{p_1p_2}, \chi_{q_1,q_2} \in \mathrm{Hom}(G_{\mathbb{Q}}, \mu_2)$, we have $\chi_{p_1p_2} \cup \chi_{q_1q_2} = (p_1p_2, q_1q_2) \in H^2(G_{\mathbb{Q}}, \mu_2)$ which is presented by the global Hilbert symbol or equivalently by the corresponding quaternion algebra. Consider the exact sequence

$$1 \to \mu_2 \to \overline{\mathbb{Q}}^\times \to \overline{\mathbb{Q}}^\times \to 1$$

so after applying Hilbert 90 we have the identification with the 2-torsion in the Brauer group

$$H^2(G_{\mathbb{Q}}, \mu_2) = H^2(G_{\mathbb{Q}}, \overline{\mathbb{Q}}^\times)[2] =: \mathrm{Br}(\mathbb{Q})[2].$$

In particular, by the ABHN theorem we have an inclusion $H^2(G_{\mathbb{Q}}, \mu_2) \hookrightarrow \bigoplus_v H^2(G_{\mathbb{Q}_v}, \overline{\mathbb{Q}_v}^\times)$. Hence $p_iq_jr \in Y$ means that $\left(\frac{p_i}{q_i}\right) = +1$ for all i,j and hence this Hilbert symbol (p_1p_2, q_1q_2) is trivial everywhere locally and so the trivial class.

Hence, there is a continuous 1-cochain $\phi_{p_1p_2,q_1q_2} : G_{\mathbb{Q}} \to \mathbb{F}_2$ such that

$$d\phi_{p_1p_2,q_1q_2}(\sigma,\tau) = \chi_{p_1p_2}(\sigma)\chi_{q_1q_2}(\tau) \tag{9.1}$$

Definition 9.1. We define $\theta \in H^2(\mathbb{F}_2^2, \mathbb{F}_2)$ by $(\sigma,\tau) \mapsto \pi_1(\sigma)\pi_2(\tau)$ where $\pi_i : \mathbb{F}_2^2 \to \mathbb{F}_2$ are the natural coordinate projections. We define

$$G := \mathbb{F}_2 \times_\theta (\mathbb{F}_2^2)$$

as the obvious underlying product of sets with operation

$$(g_1,\sigma_1) *_\theta (g_2,\sigma_2) = (g_1 + g_2 + \theta(\sigma_1,\sigma_2), \sigma_1 + \sigma_2),$$

which constructs G as the central extension of $\mathbb{F}_2^2$ by $\mathbb{F}_2$ corresponding to θ.

Importantly, one can check that by construction the map

$$\psi := (\phi_{p_1p_2,q_1q_2}, (\chi_{p_1p_2}, \chi_{q_1q_2})) : G_{\mathbb{Q}} \to G,$$

is a homomorphism and corresponds to a D_4-extension.

Define $E = \mathbb{Q}(\sqrt{p_1p_2}, \sqrt{q_1q_2})$. Note that $\ker(d) = \mathrm{Hom}(G_{\mathbb{Q}}, \mathbb{F}_2)$, and so the set of possible $\phi_{p_1p_2,q_1q_2}$ differ by a quadratic character: for all $\chi \in \mathrm{Hom}(G_{\mathbb{Q}}, \mathbb{F}_2)$

$$d(\phi_{p_1p_2,q_1q_2} + \chi) = \chi_{p_1p_2} \cup \chi_{q_1q_2}$$

In fact, we may choose $\phi_{p_1p_2,q_1q_2}$ to be unramified over E. Indeed, let σ_p denote a generator of tame inertia at an odd prime p, and write $e_p = \mathrm{ord}(\psi(\sigma_p))$. This e_p is exactly the ramification degree of p in $\overline{\mathbb{Q}}^{\ker(\psi)}/\mathbb{Q}$. Then away from $p_1p_2q_1q_2$ we can twist to ensure that $\psi(\sigma_p) = (*,0,0)$. If $* = 0$ then we get $\psi(\sigma_p) = 0$ so p is unramified, else if $* = 1$ then we add χ_{p^*}.

We now write $\phi^{\mathrm{nr}}_{p_1p_2,q_1q_2}$ for our choice of cochain which is unramified over E.

Theorem 9.2. *Let $p_1,p_2 \in X_1$, $q_1,q_2 \in X_2$, $r \in X_3$. Assume that $p_iq_jr \in Y$. Then setting $C := \{p_1,p_2\} \times \{q_1,q_2\} \times \{r\}$, and making any choice of $\phi^{\mathrm{nr}}_{p_1p_2,q_1q_2}$ we obtain that*

$$\prod_{x \in C} \widetilde{\mathrm{Art}_{x,2}}(\mathbf{v},\mathbf{w}) = (-1)^{\phi^{\mathrm{nr}}_{p_1p_2,q_1q_2}(\mathrm{Frob}_r)},$$

where Frob_r is the genuine Frobenius from $\mathbb{Q}$ in an unramified extension.

In particular, going back to Lemma 8.2, we are now able to bound

$$\sum_{x \in Y} \widetilde{\mathrm{Art}_{x,2}}(\mathbf{v},\mathbf{w}) \leq |X|^{3/4} \left| \sum_{\substack{p_1,p_2 \in X_1 \\ q_1,q_2 \in X_2 \\ r_1,r_2 \in X_3}} (-1)^{\phi^{\mathrm{nr}}_{p_1p_2,q_1q_2}(\mathrm{Frob}_{r_1}) + \phi^{\mathrm{nr}}_{p_1p_2,q_1q_2}(\mathrm{Frob}_{r_2})} \right|^{1/8}$$

which which will be a sum of a quadratic character over $\mathbb{Q}(\sqrt{p_1p_2})$.

Lecture 9: Extra lecture

Continuing in our special case, we now want to prove Theorem 9.2.

Proof of Theorem 9.2. Recall from Theorem 8.3 that

$$d\left(\sum_{x \in C} \psi_x^4(\mathbf{w})\right)(\sigma,\tau) = \chi_{p_1p_2}(\sigma)\chi_{q_1q_2}(\tau)$$

where here $\psi_x^4(\mathbf{w}) \in Z^1_{\mathrm{nr}}(G_{\mathbb{Q}}, N(-x)[4])$ is the element which factors through the Hilbert class field and restricts to $\mathrm{Cl}_K \cong \mathrm{Gal}(H_K/K)$ to satisfy $2\psi_x^4(\mathbf{w}) = \psi_x(\mathbf{w})$. And we have also constructed our element which differentiates to the same thing

$$d\phi^{\mathrm{nr}}_{p_1p_2,q_1q_2} = \chi_{p_1p_2} \cup \chi_{q_1q_2}.$$

Recall that

$$\widetilde{\mathrm{Art}_{x,2}}(\mathbf{v},\mathbf{w}) = (-1)^{\psi_x^4(\mathbf{w})(\mathrm{Frob}_{\mathfrak{r}})}$$

where $\mathfrak{r} \mid r$ in $K = \mathbb{Q}(\sqrt{-x})$ with $x = (p_i, q_j, r) \in C$. The prime $\mathfrak{r}$ splits completely in $K(\sqrt{p_1p_2}, \sqrt{q_1q_2})/K$, and choosing a prime $\mathfrak{r}'$ above $\mathfrak{r}$ in this larger field we have

$$\widetilde{\mathrm{Art}_{x,2}}(\mathbf{v},\mathbf{w}) = (-1)^{\psi_x^4(\mathbf{w})(\mathrm{Frob}_{\mathfrak{r}'})}$$

On this larger field, each $\psi_x^4(\mathbf{w})$ is a homomorphism because we obtain trivial Galois action here, so we can then compute

$$\begin{aligned}\prod_{x\in C} \widetilde{\mathrm{Art}_{x,2}}(\mathbf{v},\mathbf{w}) &= (-1)^{\sum_{x\in C}\psi_x^4(\mathbf{w})(\mathrm{Frob}_{\mathfrak{r}'})} \\ &= (-1)^{\left(\sum_{x\in C}\psi_x^4(\mathbf{w})\right)(\mathrm{Frob}_{\mathfrak{r}'})} \\ &= (-1)^{\phi^{\mathrm{nr}}_{p_1p_2,q_1q_2}(\mathrm{Frob}_{\mathfrak{r}'})+\chi(\mathrm{Frob}_{\mathfrak{r}'})}\end{aligned}$$

for some character $\chi : G_{\mathbb{Q}} \to \mathbb{F}_2$ we discussed above, since

$$\sum_{x\in C}\psi_x^4(\mathbf{w}) - \phi^{\mathrm{nr}}_{p_1p_2,q_1q_2} = \chi \in \ker(d) = \mathrm{Hom}(G_{\mathbb{Q}}, \mathbb{F}_2).$$

What can we say about this character χ? Since we've made our $\phi^{\mathrm{nr}}_{p_1p_2,q_1q_2}$ unramified, and $\psi_x^4(\mathbf{w}) \in Z^1_{\mathrm{nr}}(G_{\mathbb{Q}}, N(-x)[4])$ so the sum ramifies at most at the primes dividing x. Hence this character χ can only ramify at p_1, p_2, q_1, q_2, r. By our choice of Y, note that when we restrict χ to the absolute Galois group of $K(\sqrt{p_1p_2}, \sqrt{q_1, q_2})$, we have that $\chi(\mathrm{Frob}_{\mathfrak{r}'}) = 0$.

Now we wish to lift $\mathrm{Frob}_{\mathfrak{r}'}$ to a choice of Frobenius Frob_r at r such that both $\chi(\mathrm{Frob}_r) = 0$ and $\phi^{\mathrm{nr}}_{p_1p_2,q_1q_2}(\mathrm{Frob}_{\mathfrak{r}'}) = \phi^{\mathrm{nr}}_{p_1p_2,q_1q_2}(\mathrm{Frob}_r)$. The latter constraint is trivial, since $\phi^{\mathrm{nr}}_{p_1p_2,q_1q_2}$ is unramified at r and so any choice of Frobenius lift gives the same value.

Since $\mathfrak{r}$ is totally split in $K(\sqrt{p_1p_2}, \sqrt{q_1q_2})/K$, $\mathrm{Frob}_{\mathfrak{r}} = \mathrm{Frob}_{\mathfrak{r}'}$, so we need only choose $\mathrm{Frob}_r \in G_{\mathbb{Q}}$ to be a lift of $\mathrm{Frob}_{\mathfrak{r}} \in G_K$ such that $\chi(\mathrm{Frob}_r) = \chi(\mathrm{Frob}_{\mathfrak{r}}) = 0$. Since $K/\mathbb{Q}$ is totally ramified, we are able to do this by choosing Frob_r to be the element $\mathrm{Frob}_{\mathfrak{r}} \in G_K$. □

Lecture 10

10. To infinity and beyond

We need a version of Theorem 9.2 for 16-rank and further. Let

$$C = \{p_1, p_2\} \times \{q_1, q_2\} \times \{r_1, r_2\} \times \{d\}$$

and suppose that $\psi_x(\mathbf{w}) \in 4\mathrm{Cl}^{\vee}_{\mathbb{Q}(\sqrt{-x})}[8]$ for all $x \in C$. Fix lifts $\psi_x^8(\mathbf{w}) \in Z^1_{\mathrm{nr}}(G_{\mathbb{Q}}, N(-x)[8])$ for all $x \in C$ (i.e. elements such that $4\psi_x^8(\mathbf{w}) = \psi_x(\mathbf{w})$).

For $T \subseteq \{1, 2, 3\}$, write C_T for the subcube where we fix the ith entry of the triple to be option 2 for all $i \in T$. For example, $C_{\{1,3\}} = \{p_2\} \times \{q_1, q_2\} \times \{r_2\} \times \{d\}$.

Theorem 10.1. *Assume that for all $T \subseteq \{1,2,3\}$*

$$\sum_{x \in C_T} 2^{|T|} \psi_x^8(\mathbf{w}) \in \mathrm{Map}(G_{\mathbb{Q}}, N[2]).$$

Then with $x_0 = p_1 q_1 r_1 d$, the function $d_{x_0}\left(\sum_{x \in C \setminus \{x_0\}} \psi_x^8(\mathbf{w})\right)$ maps a pair (σ, τ) to

$$\begin{aligned}
&\chi_{p_1p_2}(\sigma)\left(\sum_{x \in C_1} 2\psi_x^8(\mathbf{w})(\tau)\right) + \chi_{q_1q_2}(\sigma)\left(\sum_{x \in C_2} 2\psi_x^8(\mathbf{w})(\tau)\right) + \chi_{r_1r_2}(\sigma)\left(\sum_{x \in C_3} 2\psi_x^8(\mathbf{w})(\tau)\right) \\
&\quad + \chi_{p_1p_2}(\sigma)\chi_{q_1q_2}(\sigma)\left(\sum_{x \in C_{\{1,2\}}} 4\psi_x^8(\mathbf{w})(\tau)\right) + \chi_{p_1p_2}(\sigma)\chi_{r_1r_2}(\sigma)\left(\sum_{x \in C_{\{1,3\}}} 4\psi_x^8(\mathbf{w})(\tau)\right) \\
&\quad + \chi_{q_1q_2}(\sigma)\chi_{r_1r_2}(\sigma)\left(\sum_{x \in C_{\{2,3\}}} 4\psi_x^8(\mathbf{w})(\tau)\right).
\end{aligned}$$

Proof. See exercise sheet 4. □

Why do we want this? Well we want to say things like if you give me $\psi_x^8(\mathbf{w})$ for all $x \in C \setminus \{x_0\}$ then I can define

$$\psi_{x_0}^8(\mathbf{w}) = -\sum_{x \in C \setminus \{x_0\}} \psi_x^8(\mathbf{w}) + \text{ [acceptable correction factors]}.$$

Corollary 10.2. *Assume that $\sum_{x \in C_T} 2^{|T|} \psi_x^8(\mathbf{w}) = 0$ for all nonempty T. Then*

$$d_{x_0}\left(\sum_{x \in C \setminus \{x_0\}} \psi_x^8(\mathbf{w})\right) = 0$$

and hence $\sum_{x \in C \setminus \{x_0\}} \psi_x^8(\mathbf{w}) \in Z^1(G_{\mathbb{Q}}, N(-x)[8])$ and $4\sum_{x \in C \setminus \{x_0\}} \psi_x^8(\mathbf{w}) = \psi_x(\mathbf{w})$.

Proof. Immediate from theorem. □

This will give us that the product of the Artin pairing around C is exactly 1. Take $\mathbf{w} = (0,0,1,*)$, we want to find a simple function ϕ which also differentiates to the large equation in Theorem 10.1, so that then this easier function is equal to the sum of our $\psi_x^8(\mathbf{w})$ (up to an acceptable error: the quadratic characters which make up the kernel of the differential). Thinking about the individual terms in this large expression, we can replace $2\psi_x^8(\mathbf{w})$ with $\psi_x^4(\mathbf{w})$

(1) $\chi_{p_1p_2}(\sigma)\left(\sum_{x \in C_1} 2\psi_x^8(\mathbf{w})\right) = \chi_{p_1p_2}(\sigma)\left(\sum_{x \in C_1} \psi_x^4(\mathbf{w})\right)$ is known from the previous layer to be able to be replaced by $\chi_{p_1p_2}(\sigma)\phi^{\mathrm{nr}}_{q_1q_2,r_1r_2}(\tau)$.
(2) Similarly, $\chi_{q_1q_2}(\sigma)\left(\sum_{x \in C_2} 2\psi_x^8(\mathbf{w})\right) = \chi_{q_1q_2}(\sigma)\left(\sum_{x \in C_2} \psi_x^4(\mathbf{w})\right)$ is known from the previous layer to be able to be replaced by $\chi_{q_1q_2}(\sigma)\phi^{\mathrm{nr}}_{p_1p_2,r_1r_2}(\tau)$.
(3) Things now change at the third term: $\chi_{r_1r_2}(\sigma)\left(\sum_{x \in C_3} \psi_x^4(\mathbf{w})\right)$ has the property that $d\left(\sum_{x \in C_3} \psi_x^4(\mathbf{w})\right) = 0$ by Corollary 10.2. Indeed

$$\begin{aligned}
d\psi_x^4(\sigma,\tau) &= \chi_{-x}(\sigma)\psi_x^4(\mathbf{w})(\tau) - \psi_x^4(\mathbf{w})(\tau) \\
&= \begin{cases} 0 & \text{if } \chi_{-x}(\sigma) = +1 \\ \psi_x(\mathbf{w})(\tau) = \chi_{r_2}(\tau) & \text{if } \chi_{-x}(\sigma) = -1 \end{cases}
\end{aligned}$$

but since we are summing around a cube where $r = r_2$ is fixed, we have an even multiple of $\chi_{r_2}(\tau)$ or zero, hence zero. Thus we will let our replacement of this term be zero.

(4) $\chi_{p_1p_2}(\sigma)\chi_{q_1q_2}(\sigma)\sum_{x\in C_{\{1,2\}}}\psi_x(\mathbf{w})(\tau)$. The sum is just literally $\chi_{r_1}+\chi_{r_2}$, and so we can exactly replace with $\chi_{p_1p_2}(\sigma)\chi_{q_1q_2}(\sigma)\chi_{r_1r_2}(\tau)$.

(5) Thinking like before, we now get the sum being $\chi_{r_2}+\chi_{r_2}$ and so we get zero.

(6) Again we get $\chi_{r_2}+\chi_{r_2}$ for the sum which is now zero.

Hence we want to find $\phi : G_{\mathbb{Q}} \to \mathbb{F}_2$ which satisfies

$$d\phi(\sigma,\tau) = \chi_{p_1p_2}(\sigma)\phi^{\mathrm{nr}}_{q_1q_2,r_1r_2}(\tau) + \chi_{q_1q_2}(\sigma)\phi^{\mathrm{nr}}_{p_1p_2,r_1r_2}(\tau) + \chi_{p_1p_2}(\sigma)\chi_{q_1q_2}(\sigma)\chi_{r_1r_2}(\tau).$$

Good properties of such a ϕ (see Higher Genus Theory, Koymans–Pagano) that can be shown are

- $\phi|_{G_{\mathbb{Q}(\sqrt{p_1p_2},\sqrt{q_1q_2})}}$ is a quadratic character
- The Galois group (smallest extension through which it factors) is $\mathbb{F}_2[\mathbb{F}_2^2] \rtimes \mathbb{F}_2^2$.

Moreover, we can choose $\phi = \phi^{\mathrm{nr}}_{p_1p_2,q_1q_2,r_1r_2} \in \mathrm{Cl}^{\vee}_{\mathbb{Q}(\sqrt{p_1p_2},\sqrt{q_1q_2},\sqrt{r_1r_2})}[2]$.

This leads to the goal that

$$\prod_{x\in C}\widetilde{\mathrm{Art}_{x,3}}(\mathbf{v},\mathbf{w}) = (-1)^{\sum_i v_i\phi_{p_1p_2,q_1q_2,r_1r_2}(\mathrm{Frob}_{\pi_i(x)})}.$$

Lecture 11

11. Repeating Repeated Cauchy–Schwarz II

Let's now take ourselves away from the special case from the previous lectures and start working on the proper full case. Let $X = X_1 \times \cdots \times X_r$ be a good grid, and R be a fixed $r \times r$ matrix over $\mathbb{F}_2$ such that

- $\ker_l(R)$ has basis $(1,1,\ldots,1), \mathbf{v}_1, \ldots, \mathbf{v}_s$, and
- $\ker_r(R)$ has basis $(1,1,\ldots,1), \mathbf{w}_1, \ldots, \mathbf{w}_s$.

Our goal, is to show that $\mathrm{Art}_{x,2}$ is a random matrix as x varies in $Y = \{x \in X \ : \ R_x = R\}$. Via orthogonality of characters, this is equivalent to showing that for every tuple $(e_{i,j}) \in \mathbb{F}_2^{s\times s}$ except for the all-zero tuple

$$\left|\sum_{\substack{x\in X\\ R_x=R}}\prod_{i=1}^{s}\prod_{j=1}^{s}\widetilde{\mathrm{Art}_{x,2}}(\mathbf{v}_i,\mathbf{w}_j)^{e_{i,j}}\right| = \text{small}.$$

There are two new steps compared to the special case from before.

- Fix all be three primes in X (i.e. fix indices i_1,i_2,i_3 and vary over the $X_{i_1}\times X_{i_2}\times X_{i_3}$ part). It will be *very important* to choose i_1,i_2,i_3 well.
- Prove some sufficient conditions for

$$\prod_{j=1}^{2}\prod_{k=1}^{2}\widetilde{\mathrm{Art}_{d\pi_j(x)\pi_k(x)}}(\mathbf{v},\mathbf{w}) = 1,$$

where $\pi_j : X \to X_j$ is the projection, and $d \in \prod_{i\neq j,k} X_i$.

Theorem 11.1. *Let $p_1, p_1' \in X_{i_1}$, $p_2, p_2' \in X_{i_2}$, and let $d \in \prod_{i \neq i_1, i_2} X_i$. Assume that*

$$R_{dp_1p_2} = R_{dp_1p_2'} = R_{dp_1'p_2} = R_{dp_1'p_2'} = R.$$

Let $\mathbf{v} \in \ker_l(R)$ and $\mathbf{w} \in \ker_r(R)$. Assume that

$$\pi_{i_1}(\mathbf{v}) = \pi_{i_2}(\mathbf{v}) = \pi_{i_1}(\mathbf{w}) = \pi_{i_2}(\mathbf{w}) = 0.$$

Then

$$\widetilde{\mathrm{Art}_{dp_1p_2,2}}(\mathbf{v},\mathbf{w}) \cdot \widetilde{\mathrm{Art}_{dp_1p_2',2}}(\mathbf{v},\mathbf{w}) \cdot \widetilde{\mathrm{Art}_{dp_1'p_2,2}}(\mathbf{v},\mathbf{w}) \cdot \widetilde{\mathrm{Art}_{dp_1'p_2',2}}(\mathbf{v},\mathbf{w}) = +1.$$

Proof. Recall that

$$\psi_x(\mathbf{w}) = \sum_{i=1}^{r} w_i \chi_{\pi_i(x)^*} \in Z^1(G_{\mathbb{Q}}, N[2]).$$

Fix lifts $\psi_x^4(\mathbf{w}) \in Z^1(G_{\mathbb{Q}}, N(-x)[4])$ with $2\psi_x^4(\mathbf{w}) = \psi_x(\mathbf{w})$. Set $C = \{d\} \times \{p_1, p_1'\} \times \{p_2, p_2'\}$. We make use of the following two claims.

Claim 1: $\sum_{x \in C} \psi_x(\mathbf{w}) = 0$.
Proof: $\sum_{x \in C} \psi_x(\mathbf{w}) = \sum_{x \in C} \sum_{i=1}^{r} w_i \chi_{\pi_i(x)^*} = \sum_{i=1}^{r} \sum_{x \in C} w_i \chi_{\pi_i(x)^*} = 0$. □

Claim 2: $d\left(\sum_{x \in C} \psi_x^4(\mathbf{w})\right) = 0$.
Proof: Same as usual. □

By Claim 1 we know that $\sum_{x \in C} \psi_x^4(\mathbf{w})$ is valued in $N[2]$ and by Claim 2 it satisfies the cocycle relation. Hence it is a quadratic character χ ramified only at $p_1p_1'p_2p_2'd$. Recall that if we write $\underline{\pi_i(x)}$ for a prime above $\pi_i(x)$ in $\mathbb{Q}(\sqrt{-x})$,

$$\widetilde{\mathrm{Art}_{x,2}}(\mathbf{v},\mathbf{w}) = \prod_{i=1}^{r} (-1)^{v_i \psi_x^4(\mathbf{w})(\mathrm{Frob}_{\underline{\pi_i(x)}})}.$$

Hence

$$\prod_{x \in C} \widetilde{\mathrm{Art}_{x,2}}(\mathbf{v},\mathbf{w}) = \prod_{i=1}^{r} (-1)^{v_i \sum_{x \in C} \psi_x^4(\mathbf{w})(\mathrm{Frob}_{\underline{\pi_i(x)}})}.$$

Note that this notation is a bit deceptive, since $\underline{\pi_i(x)}$ are all Frobenii in different fields, which makes it hard to sum these $\psi_{x,\mathbf{w}}$ and replace them with χ. Hence we'd like to put them all over one base field and then sum these up. Note that each $\underline{\pi_i(x)}$ splits completely in the extension $\mathbb{Q}(\sqrt{p_1p_1'}, \sqrt{p_2p_2'}, \sqrt{-x})/\mathbb{Q}(\sqrt{-x})$, and it is an easy exercise to see that this top field is actually independent of $x \in C$. Let us pick a prime $\pi_i(x)'$ in this larger field, and note that $\mathrm{Frob}_{\pi_i(x)'} = \mathrm{Frob}_{\underline{\pi_i(x)}}$ since we are totally split.

Now our Frobenii live in the same field. Moreover, $\chi(\mathrm{Frob}_i(x)')$ is well defined and independent of $x \in C$ away from $i \in \{i_1, i_2\}$. Indeed, this follows from the constant Rédei matrix since p_1p_1' and p_2p_2' are squares locally at primes dividing d. Hence, since $v_{i_1} = v_{i_2} = 0$ we can unambiguously write for every $i \in [r]$

$$v_i \sum_{x \in C} \psi_x^4(\mathbf{w})(\mathrm{Frob}_{\pi_i(x)'}) = v_i \left(\sum_{x \in C} \psi_x^4(\mathbf{w}) \right) (\mathrm{Frob}_{\pi_i(x)'}) = v_i \chi\left(\mathrm{Frob}_{\pi_i(x)'}\right)$$

Thus in our product we have:

$$\begin{aligned}\prod_{i=1}^{r}(-1)^{v_i\sum_{x\in C}\psi_x^4(\mathbf{w})(\underline{\mathrm{Frob}_{\pi_i(x)}})} &= \prod_{i=1}^{r}(-1)^{v_i\sum_{x\in C}\psi_x^4(\mathbf{w})(\mathrm{Frob}_{\pi_i(x)'})}\\ &= \prod_{i=1}^{r}(-1)^{v_i\chi(\mathrm{Frob}_{\pi_i(x)'})}\\ &= 1.\end{aligned}$$

Thus we conclude that the result holds. □

Lecture 12

12. The large sieve and additivity

Pick a topological generator σ_v of I_v^{tame} for each odd v. Pick $\sigma_{-1}, \sigma_2 \in I_2 \subseteq G_{\mathbb{Q}_2}$ dual to χ_{-1} and χ_2. We define

$$\mathbb{Q}^{\mathrm{pro}-2} := \bigcup_{\mathrm{Gal}(K/\mathbb{Q})\ 2\text{-group}} K,$$

with Galois group

$$\mathcal{G}^{\mathrm{pro}-2} := \mathrm{Gal}(\mathbb{Q}^{\mathrm{pro}-2}/\mathbb{Q}),$$

and let

$$\mathfrak{G} := \{\sigma_v \ : \ v \text{ odd}\} \cup \{\sigma_{-1}, \sigma_2\} \subseteq G_{\mathbb{Q}}.$$

Theorem 12.1. *$\mathfrak{G}$ is a minimal set of topological generators for $\mathcal{G}^{\mathrm{pro}-2}$.*

Proof. S is a minimal set of topological generators if and only if S generates $\mathrm{Gal}(K/\mathbb{Q})$ for every Galois $K/\mathbb{Q}$ of degree a power of 2, and for every element $\sigma \in S$ there exists such an extension $K/\mathbb{Q}$ such that $S\setminus\{\sigma\}$ does not generate $\mathrm{Gal}(K/\mathbb{Q})$. For the latter condition, simply take $\mathbb{Q}(\sqrt{p^*})$, $\mathbb{Q}(\sqrt{-1})$ or $\mathbb{Q}(\sqrt{2})$ for $\sigma = \sigma_p, \sigma_{-1}$, or σ_2 respectively. Hence it is sufficient for us to show that S generates the Galois group of every Galois 2-extension.

A general fact about every p-group H is that

$$T \subseteq H \text{ generates } H \iff T \text{ generates } H/\Phi(H),$$

where $\Phi(H)$ is the Frattini subgroup. Another fact in the case $p = 2$ is that $\Phi(H) = H^2[H, H]$.

These facts show that it is enough to check this in the maximal elementary abelian 2-quotient, so hence in multiquadratic fields. This is clear. □

Corollary 12.2. *For each $a, b \in \mathbb{Q}^\times/\mathbb{Q}^{\times 2}$ such that $(a, b)_v = +1$ for all v, there is a unique function $\phi_{a,b}^{\mathfrak{G}} : \mathcal{G}^{\mathrm{pro}-2} \to \mathbb{F}_2$ such that*

- *$d\phi_{a,b}^{\mathfrak{G}} = \chi_a \cup \chi_b$, and*
- *$\phi_{a,b}^{\mathfrak{G}}(\sigma) = 0$ for all $\sigma \in \mathfrak{G}$.*

Proof. Pick some $\phi_{a,b}$ satisfying the first condition, which exists since $\chi_a \cup \chi_b$ represents the quaternion algebra (a, b) which is split everywhere locally and so by the ABHN theorem is trivial in $H^2(G_{\mathbb{Q}}, \mu_2)$.

The coboundary condition has already forced that $\phi_{a,b}(\mathrm{id}) = 0$, and by continuity of $\phi_{a,b}$ (i.e. only finitely many primes can ramify in a finite extension) we see that

there are only finitely many $\sigma \in \mathfrak{G}$ such that $\phi_{a,b}(\sigma) = 1$. Hence we define our element to be the following finite sum

$$\phi_{a,b}^{\mathfrak{G}} := \phi_{a,b} + \sum_{p \text{ odd}} \phi_{a,b}(\sigma_p)\chi_{p^*} + \phi_{a,b}(\sigma_{-1})\chi_{-1} + \phi_{a,b}(\sigma_2)\chi_2.$$

□

Corollary 12.3. *Assume that* $(a, b_1)_v = (a, b_2)_v = +1$ *and* $(a_1, b)_v = (a_2, b)_v = +1$ *for all places* v*, then*

$$\phi_{a,b_1b_2}^{\mathfrak{G}} = \phi_{a,b_1}^{\mathfrak{G}} + \phi_{a,b_2}^{\mathfrak{G}},$$
$$\phi_{a_1a_2,b}^{\mathfrak{G}} = \phi_{a_1,b}^{\mathfrak{G}} + \phi_{a_2,b}^{\mathfrak{G}}.$$

Okay so now we return to our aim from the beginning of the last lecture.

Theorem 12.4.

$$\left| \sum_{\substack{x \in X \\ R_x = R}} \prod_{i=1}^{s} \prod_{j=1}^{s} \widetilde{\mathrm{Art}_{x,2}}(\mathbf{v}_i, \mathbf{w}_j)^{e_{i,j}} \right| = \text{small}.$$

Proof. The left hand side and right hand side have the same d and correct normalisation at $\mathfrak{G}$.

Note that $\phi_{a,b}^{\mathfrak{G}}$ is a valid choice for $\phi_{a,b}^{\mathrm{nr}}$ from our earlier work. Imagine for now, that we can fix indices i_1, i_2, i_3 so that after applying repeated Cauchy–Schwarz, we can use exactly one application of Theorem 9.2 and then the other products go to 1 via Theorem 11.1. In this case, we get

$$\left| \sum_{\substack{x \in X \\ R_x = R}} \prod_{i=1}^{s} \prod_{j=1}^{s} \widetilde{\mathrm{Art}_{x,2}}(\mathbf{v}_i, \mathbf{w}_j)^{e_{i,j}} \right|$$

$$\leq \sum_{d \in \prod_{i \neq i_1, i_2, i_3} X_i} |X_{i_1} \times X_{i_2} \times X_{i_3}|^{3/4} \left| \sum_{\substack{p_1, p_2 \in X_{i_1} \\ q_1, q_2 \in X_{i_2} \\ r_1, r_2 \in X_{i_3} \\ R_{dp_iq_jr_k} = R}} (-1)^{\phi_{p_1p_2,q_1q_2}^{\mathfrak{G}}(\mathrm{Frob}_{r_1}) + \phi_{p_1p_2,q_1q_2}^{\mathfrak{G}}(\mathrm{Frob}_{r_2})} \right|^{1/8}.$$

What do we need from our $i_1, i_2, i_3 > i_{\text{medium}}$ in order to get this?

- Fix i_0, j_0 such that $e_{i_0,j_0} = 1$.
- Pick i_1, i_2, i_3 such that for all $i \neq i_0$ we have
$$\pi_{i_1}(\mathbf{v}_i) = \pi_{i_2}(\mathbf{v}_i) = \pi_{i_3}(\mathbf{v}_i) = 0$$
and such that for all $j \neq j_0$
$$\pi_{i_1}(\mathbf{w}_j) = \pi_{i_2}(\mathbf{w}_j) = \pi_{i_3}(\mathbf{w}_j) = 0.$$
- We have
$$(\pi_{i_1}(\mathbf{v}_{i_0}), \pi_{i_2}(\mathbf{v}_{i_0}), \pi_{i_3}(\mathbf{v}_{i_0})) = (0, 0, 1).$$
- We have
$$(\pi_{i_1}(\mathbf{w}_{j_0}), \pi_{i_2}(\mathbf{w}_{j_0}), \pi_{i_3}(\mathbf{w}_{j_0})) = (0, 1, 0).$$

If we are unable to find such i_1, i_2, i_3 (which should be a rare event!) then the resulting R should be thrown out. Recalling our earlier inner sum

$$S = S(d) = \left| \sum_{\substack{p_1,p_2 \in X_{i_1} \\ q_1,q_2 \in X_{i_2} \\ r_1,r_2 \in X_{i_3} \\ R_{dp_iq_jr_k}=R}} (-1)^{\phi^{\mathfrak{G}}_{p_1p_2,q_1q_2}(\mathrm{Frob}_{r_1})+\phi^{\mathfrak{G}}_{p_1p_2,q_1q_2}(\mathrm{Frob}_{r_2})} \right|,$$

we have

$$S \leq \sum_{\substack{p_1 \in X_{i_1} \\ q_1 \in X_{i_2} \\ r_1 \in X_{i_3}}} \left| \sum_{\substack{p_2 \in X_{i_1} \\ q_2 \in X_{i_2} \\ r_2 \in X_{i_3} \\ R_{dp_iq_jr_k}=R \\ \forall (i,j,k) \neq (2,2,2)}} (-1)^{\phi^{\mathfrak{G}}_{p_1p_2,q_1q_2}(\mathrm{Frob}_{r_1})+\phi^{\mathfrak{G}}_{p_1p_2,q_1q_2}(\mathrm{Frob}_{r_2})} \right|,$$

where the condition that $R_{dp_iq_jr_k} = R$ for all $(i,j,k) \neq (2,2,2)$ can now be absorbed into coefficients $\alpha(p_2,q_2), \beta(p_2,r_2), \gamma(q_2,r_2)$ where these functions depend only on the outer p_1, q_1, r_1.

Miracle: In fact if $R_{dp_iq_jr_k} = R$ for all $(i,j,k) \neq (2,2,2)$ then in fact the final one $R_{dp_1q_2r_2}$ is equal to R for free!

Final idea: Apply bonus repeated Cauchy–Schwarz on the inner sum, applying additivity of our $\phi^{\mathfrak{G}}$ from Corollary 12.3, to bound it by

$$\sum_{\substack{p_1 \in X_{i_1} \\ q_1 \in X_{i_2} \\ r_1 \in X_{i_3}}} \sum_{\substack{p_2,p_3 \in X_{i_1} \\ q_2,q_3 \in X_{i_2} \\ r_2,r_3 \in X_{i_3} \\ (p_2p_3,q_2q_3)=1 \\ \left(\frac{p_2p_3}{r_2}\right)=\left(\frac{p_2p_3}{r_3}\right)=1}} (-1)^{\phi^{\mathfrak{G}}_{p_2p_3,q_2q_3}(\mathrm{Frob}_{r_2})+\phi^{\mathfrak{G}}_{p_2p_3,q_2q_3}(\mathrm{Frob}_{r_3})},$$

to which we can now apply the large sieve and get our saving!

□