COHEN–LENSTRA HEURISTICS: INFORMAL NOTES

VICTOR WANG

Abstract. Following Smith [8], except in some of the definitions, details, and appendices. I'm not sure yet where things break down for real quadratic fields.

Contents

1. Introduction

Roughly speaking, the algebraic input relies on three principles:

(1) Linear algebra and combinatorics over $\mathbb{F}_2$, including (but not limited to) the torsion class pairing below, and the notion of minimality (Definition 4.12 and Appendix D).[1]
(2) Class field theory, including (but not limited to) representing class group characters of K using Galois subextensions of H_K/K, which are actually Galois over $\mathbb{Q}$; and also calculating local Artin symbols. (See Propositions 2.1 and 2.5.)
(3) "Dihedral-like" Galois extensions (with restricted ramification) over $\mathbb{Q}$ can be "parameterized" using suitable $\mathbb{Q}$-cocycles. (See Propositions 3.3 and 3.7.)

Date: April 1, 2018.

[1] As Bjorn Poonen once said, "Your success in life is determined by how much linear algebra you know."

The third might be the most significant, because $\mathbb{Q}$-cocycles can be easily manipulated as we vary the quadratic field data encoded in "dihedral-like" groups. It can be motivated in at least two ways: extending characters over K to cocycles over $\mathbb{Q}$ (surjectivity of inflation-restriction map), or the class-Selmer analogy (where Selmer groups are already defined over $\mathbb{Q}$). The challenge is then finding nontrivial relations in families: Theorems 4.15 and TBD.

2. Computing the torsion class pairing

The first principle extracts 2^{k+1}-ranks from the torsion class pairing. The left kernel of

$$\mathrm{Cl}[2] \times \widehat{\mathrm{Cl}}[2] \xrightarrow{(x,\psi)\mapsto\psi(x)} \mu_2 = \pm 1$$

is $2\,\mathrm{Cl}[4]$ (giving the 4-rank of Cl), because $\psi(x)^2 = \psi(x^2)$ (left kernel detects whether x is a square). Generally, to find the 2^{k+1}-rank using $\mathbb{F}_2$-linear algebra, consider the pairing

$$2^{k-1}\,\mathrm{Cl}[2^k] \times 2^{k-1}\widehat{\mathrm{Cl}}[2^k] \xrightarrow{(2^{k-1}x,2^{k-1}\psi)\mapsto 2^{k-1}\psi(x)} \pm 1$$

(easily check well-defined) with left kernel $2^k\,\mathrm{Cl}[2^{k+1}]$. This should all be classical.

2.1. Global computation: general case.

Fix $k \geq 1$ and let $K/\mathbb{Q}$ be *any* number field.

Proposition 2.1. *For $u \in 2^{k-1}\widehat{\mathrm{Cl}}_K[2^k]$ and $v \in 2^{k-1}\,\mathrm{Cl}_K[2^k]$, the above natural pairing $2^{k-1}\widehat{\mathrm{Cl}}_K[2^k] \times 2^{k-1}\,\mathrm{Cl}_K[2^k] \to \mu_2$ is given by the Artin symbol formula*

$$\langle u, v\rangle := \psi_k(v) = \mathrm{rec}_{L/K}(v) \in \mathrm{Gal}(L/K)[2] \hookrightarrow \mu_2,$$

where we have chosen $\psi_k \in \widehat{\mathrm{Cl}}_K[2^k]$ with $u = 2^{k-1}\psi_k$, and where $L = L(\psi_k) := H_K^{\mathrm{rec}(\ker\psi_k)}$ is the fixed field of $\ker\psi_k \leq \mathrm{Cl}_K$ acting on the Hilbert class field H_K/K.

Remark 2.2. Implicit in the identification $\mathrm{Gal}(L/K)[2] \hookrightarrow \mathbb{F}_2$ (usually an isomorphism, unless ψ_k is the trivial character) is the fact that L/K is cyclic of order $\#\,\mathrm{im}\,\psi_k \mid 2^k$.

Proof. Use global Hilbert class field theory. The composite map

$$\phi_k\colon\ \mathrm{Gal}(H_K/K) \xrightarrow{\mathrm{rec}^{-1}\colon\cong} \mathrm{Cl}_K \xrightarrow{\psi_k} \mu_{2^k}$$

has kernel $\mathrm{Gal}(H_K/L)$ (by definition of L/K), so it induces an injection of quotients:

$$\phi_k\colon\ \mathrm{Gal}(L/K) \xrightarrow{\mathrm{rec}_{L/K}^{-1}\colon\cong} \mathrm{Cl}_K/\ker\psi_k \hookrightarrow \mu_{2^k}.$$

In particular, L/K is cyclic, and $\phi_k(\mathrm{rec}_{L/K}(v)) = \psi_k(v) \in 2^{k-1}\mu_{2^k} = \mu_2$, so $\langle u, v\rangle := \psi_k(v) \in \mu_2$ has the same order as $\mathrm{rec}_{L/K}(v) \in \mathrm{Gal}(L/K)[2] \hookrightarrow \mu_2$. Now $\mathrm{Aut}(\mu_2 \cong \mathbb{Z}/2) = 1$ allows the desired identification $\psi_k(v) = \mathrm{rec}_{L/K}(v) \in \mu_2$. $\square$

Remark 2.3. The permissible fields L/K are precisely the degree $2^{\leq k}$ cyclic unramified extensions over K containing $H_K^{\ker u}$ (an unramified quadratic extension of K; these have been studied more explicitly in classical genus theory and subsequent work).

Proposition 2.4. *Every unramified*[2] *abelian extension of a quadratic field $K/\mathbb{Q}$ is Galois.*

Proof. $H_K^+/\mathbb{Q}$ is Galois by maximality. To prove that every subgroup of $\mathrm{Gal}(H_K^+/K)$ is normal in $\mathrm{Gal}(H_K^+/\mathbb{Q})$, use Artin reciprocity and the fact that $\sigma(I)I \in P_K^+$ for $I \in I_K$. $\square$

[2] unramified at finite places (allowing ramification at ∞)

2.2. **Local computation: quadratic case.** Now specialize and simplify locally.

Proposition 2.5. *In the notation of Proposition 2.1, if $K/\mathbb{Q}$ is quadratic, then*

(1) $L/\mathbb{Q}$ is dihedral Galois, and
(2) for any prime $p \mid \Delta_K$ with $p\mathcal{O}_K = \mathfrak{p}^2$, every decomposition field over $\mathbb{Q}_p$ is abelian *with Galois group C_2 or C_2^2. The character*

$$\psi_k \operatorname{rec}_{L/K}^{-1} \colon \operatorname{Gal}(L/K) \to \mu_{2^k}$$

restricts on the *abelian decomposition group $D_\mathfrak{p}$ to*

$$\chi|_{G_{K_\mathfrak{p}}} \colon D_\mathfrak{p} \to \mu_2$$

for some local unramified quadratic or trivial character $\chi \colon G_{\mathbb{Q}_p} \to \mu_2$ over $\mathbb{Q}_p$.

Furthermore, $\operatorname{rec}_{L/K}(\mathfrak{p}) = (\chi, b)_p = \operatorname{inv}_p(\chi \cup \chi_b)$ for any uniformizer b of $\mathbb{Q}_p$.

Remark 2.6. Here $(\chi, b)_p$ is understood to mean $(\operatorname{Disc} \overline{\mathbb{Q}_p}^{\ker \chi}/\mathbb{Q}_p, b)_p$, where $\overline{\mathbb{Q}_p}^{\ker \chi} = F$ below is at most quadratic over $\mathbb{Q}_p$, and $\operatorname{Disc} F$ is defined up to a square in $\mathbb{Q}_p$.

Proof that $L/\mathbb{Q}$ is dihedral Galois. Since L/K is a cyclic unramified extension of K, it is Galois over $\mathbb{Q}$ by Proposition 2.4. To prove $L/\mathbb{Q}$ dihedral, use Artin reciprocity together with the fact that $\operatorname{Gal}(K/\mathbb{Q})$ inverts ideal classes in Cl_K. □

Proof of local restriction. Given $L/K/\mathbb{Q}$, choose primes $\mathfrak{q}/\mathfrak{p}/p$ with $p \mid \Delta_K$ (i.e. p ramified in K, so $p\mathcal{O}_K = \mathfrak{p}^2$). Note that $L_\mathfrak{q}/K_\mathfrak{p}$ is (cyclic and) unramified. On the other hand, if F denotes the maximal unramified sub-extension of $L_\mathfrak{q}$ over $\mathbb{Q}_p$ (so $F/\mathbb{Q}_p$ is cyclic) then $L_\mathfrak{q}/F$ is totally ramified by local structure theory. So $L_\mathfrak{q}/FK_\mathfrak{p}$ is both unramified and totally ramified, hence trivial. Furthermore, F and $K_\mathfrak{p}$ must be linearly disjoint over $\mathbb{Q}_p$, so $L_\mathfrak{q}/\mathbb{Q}_p$ is abelian with Galois group $\operatorname{Gal}(F/\mathbb{Q}_p) \times \operatorname{Gal}(K_\mathfrak{p}/\mathbb{Q}_p)$. But F must then be at most quadratic, because C_2^2 (Klein four group) is the only non-cyclic abelian subgroup of $\operatorname{Gal}(L/\mathbb{Q})$, the dihedral group of size $2[L : K]$.[3] Now

$$D_\mathfrak{p} := \operatorname{Gal}(L_\mathfrak{q}/K_\mathfrak{p}) \xrightarrow{\text{res}\colon \cong} \operatorname{Gal}(F/\mathbb{Q}_p)$$

is at most order two. Yet $F/\mathbb{Q}_p$ is unramified by definition, so $\psi_k \operatorname{rec}_{L/K}^{-1}$ indeed restricts on $D_\mathfrak{p}$ to a local character $\chi|_{G_{K_\mathfrak{p}}}$, with χ defined over $F/\mathbb{Q}_p$ with the desired properties. □

Proof of Artin symbol calculation. Fix $b \in \mathbb{Q}_p$ with $v_p(b) = 1$, so b/p, a unit, must be a norm in $F/\mathbb{Q}_p$ (an unramified local extension). Then $\operatorname{rec}_{L/K}(\mathfrak{p})$ is trivial if and only if $L_\mathfrak{q} = K_\mathfrak{p}$ if and only if $F = \mathbb{Q}_p$ if and only if $p \in N_{F/\mathbb{Q}_p}(F^\times)$ if and only if $b \in N_{F/\mathbb{Q}_p}(F^\times)$ if and only if $\operatorname{inv}_p(\chi \cup \chi_b) = 0$. But $\operatorname{rec}_{L/K}(\mathfrak{p})$ (killed by squaring) and $\operatorname{inv}_p(\chi \cup \chi_b) = (\chi, b)_p$ (a quadratic Hilbert symbol[4]) are both $\mathbb{Z}/2$-valued, so they must coincide. □

3. Relating characters and cocycles

Definition 3.1. Let $K/\mathbb{Q}$ be quadratic with character

$$\delta_K \colon G_\mathbb{Q} \to \operatorname{Gal}(K/\mathbb{Q}) = \pm 1 \in \operatorname{End}_\mathbb{Z}(\mathbb{Q}_2/\mathbb{Z}_2),$$

and set $M_K := \mathbb{Q}_2/\mathbb{Z}_2$ with Galois action $gn := \delta_K(g)n$ for $g \in G_\mathbb{Q}$. Let $\iota_K \colon M_K \to \mathbb{Q}_2/\mathbb{Z}_2$ be the forgetful map (an identity of abelian groups).

[3] Better proof of $[L_\mathfrak{q} : K_\mathfrak{p}] \le 2$ using Artin reciprocity: $\mathfrak{p}$ must split into at most $[L : K]/2$ primes (so each/the decomposition group has size at most 2) because it has order at most 2 in $I_K/P_K N_{L/K}(L^\times)$.

[4] see Serre, Local Fields, p. 207, Proposition 5, for the invariant map interpretation

Remark 3.2. Why define ι_K? When K varies later on, we will want to think of the underlying abelian group $\mathbb{Q}_2/\mathbb{Z}_2$ as being fixed, with only the action $\delta_K\colon G_{\mathbb{Q}} \to \operatorname{End}_{\mathbb{Z}}(\mathbb{Q}_2/\mathbb{Z}_2)$ varying.

3.1. **Global extension of characters.** We now use *additive* notation for characters.

Proposition 3.3 ([8, Cf. Proposition 2.7])**.** *Define $M = M_K$ as above. The cocycle group*

$$\overline{\mathrm{Cl}}_K^{\vee}[2^k] := Z^1_{\mathrm{cts}}(\operatorname{Gal}(K^{\mathrm{ur}}/\mathbb{Q}), M[2^k])$$

surjects onto $\widehat{\mathrm{Cl}}_K[2^k] = \operatorname{Hom}_{\mathrm{cts}}(\operatorname{Gal}(K^{\mathrm{ur}}/K), 2^{-k}\mathbb{Z}/\mathbb{Z})$, via restriction of cocycles. Consequently, for $k \geq 1$, the image of $2^{k-1}\overline{\mathrm{Cl}}_K^{\vee}[2^k]$ under $\overline{\mathrm{Cl}}_K^{\vee}[2] \to \widehat{\mathrm{Cl}}_K[2]$ is $2^{k-1}\widehat{\mathrm{Cl}}_K[2^k]$.

Remark 3.4. Smith explicitly extends K-characters to $\mathbb{Q}$-objects. It may be instructive to work this out later. See crossed homomorphism or MSE: motivating inhomogeneous cochains (esp. Mariano answer about section interpretation) for inspiration.

Here is another perspective.

Proof. For $1 \leq k \leq \infty$, consider the inflation-restriction exact sequence

$$\begin{aligned} 0 &\to H^1(G/N, M[2^k]^N) \xrightarrow{\text{inf}} H^1(G, M[2^k]) \xrightarrow{\text{res}} H^1(N, M[2^k])^{G/N} \\ &\to H^2(G/N, M[2^k]^N) \xrightarrow{\text{inf}} H^2(G, M[2^k]) \end{aligned}$$

with $G = \operatorname{Gal}(H_K/\mathbb{Q})$ and $N = \operatorname{Gal}(H_K/K) \cong \mathrm{Cl}_K$. Since N and G/N act trivially on M and $H^1(N, M[2^k])$, resp., and $G/N = \operatorname{Gal}(K/\mathbb{Q}) = \pm 1$ is cyclic, the sequence simplifies to

$$0 \to M[2^k]/2M[2^k] \to H^1(G, M[2^k]) \to \operatorname{Hom}(N, M[2^k]) \to M[2] \to H^2(G, M[2^k]).$$

One can abstractly conclude $2^{k-1}H^1(G, M[2^k]) \xrightarrow{\sim} 2^{k-1}\operatorname{Hom}(N, M[2^k])$ for $k \geq 2$ (this is also true for $k = 1$: any quadratic character on N lifts to a Klein four character on G), but in fact, Smith explicitly proves that the restriction map is surjective for $M[2^k]$.[5] □

Remark 3.5. To see why G/N acts trivially on $H^1(N, M[2^k])$, recall that G acts on $Z^1(N, -)$ by sending $n \mapsto a_n$ to $gng^{-1} \mapsto ga_n$. For $a \in H^1(N, M[2^k]) = \operatorname{Hom}(N, M[2^k])$, Artin reciprocity over K gives $a_{gng^{-1}} = ga_n$, since $a_{n^{-1}} = -a_n$, and g acts on Cl_K by $\delta_K(g)$. Of course, Smith's proof crucially relies on this "dihedral-like" structure as well.

Remark 3.6. Consider the class-Selmer analogy (which Smith says Fouvry–Klüners used earlier): the Selmer groups involve $H^1(G_{\mathbb{Q}}, -)$ *by definition*, perhaps motivating the above $H^1(G_{\mathbb{Q}}, -)$ extension of the dual class group. Alternative motivation: $\mathbb{Q}$-cocycles can be added over varying ground fields $K/\mathbb{Q}$, while K-characters maybe cannot (I'm not sure yet).

3.2. **Local restriction of cocycles.** We want to express Proposition 2.5 using cocycles.

Proposition 3.7. *In Proposition 2.5, suppose the character $\psi_k\colon \operatorname{Gal}(L/K) \to 2^{-k}\mathbb{Z}/\mathbb{Z}$ extends to a cocycle $\phi_k\colon \operatorname{Gal}(L/\mathbb{Q}) \to M_K[2^k]$. Then the local restriction $\phi_k|_{\operatorname{Gal}(L_{\mathfrak{q}}/\mathbb{Q}_p)}$ is*

(1) a quadratic character extending $\psi_k|_{D_{\mathfrak{p}}} = \chi|_{G_{K_{\mathfrak{p}}}}$, where $D_{\mathfrak{p}} = \operatorname{Gal}(L_{\mathfrak{q}}/K_{\mathfrak{p}})$;

(2) the sum of χ with one of the two characters of $\operatorname{Gal}(K_{\mathfrak{p}}/\mathbb{Q}_p)$, say χ'.

Furthermore, $(\chi', b)_p = 0$ and

$$\operatorname{rec}_{L/K}(\mathfrak{p}) = (\chi, b)_p = (\phi_k, b)_p = \operatorname{inv}_p(\phi_k \cup \chi_b)$$

for any uniformizer b of $\mathbb{Q}_p$, as long as $b \in N_{K_{\mathfrak{p}}/\mathbb{Q}_p}(K_{\mathfrak{p}}^{\times})$.

[5] It should also be possible to show (through a computation likely boiling down to Smith's argument) that the transgression (boundary) map [6, Proposition 1.6.6, p. 65] is zero.

Remark 3.8. The appearance of quadratic "η" in $(\eta, b)_p$ is shorthand for the discriminant of the at most quadratic field of definition of η. In particular, $(\phi_k, b)_p = (\chi, b)_p + (\chi', b)_p$.

Remark 3.9. Later on, b will be the norm of an ideal $w(K)$ depending on K, such that $w(K) \in 2\operatorname{Cl}_K[4]$. In particular, $w(K) = \beta I^2$ for some element $\beta \in K^\times$ and fractional ideal I. But $N(I^2) = N(I)^2$ is the norm of $N(I) \in K^\times$, so b is the norm of the element $\beta N(I) \in K^\times$. So $w(K) \in 2\operatorname{Cl}_K[4]$ will give us $b \in N_{K/\mathbb{Q}}(K^\times)$ for free, even as K varies.

Proof. The global inflation-restriction sequence (see Proposition 3.3) restricts down to

$$0 \to H^1(\operatorname{Gal}(K_\mathfrak{p}/\mathbb{Q}_p), M[2^k]) \xrightarrow{\text{inf}} H^1(\operatorname{Gal}(L_\mathfrak{q}/\mathbb{Q}_p), M[2^k]) \xrightarrow{\text{res}} H^1(\operatorname{Gal}(L_\mathfrak{q}/K_\mathfrak{p}), M[2^k]) \\ \to H^2(\operatorname{Gal}(K_\mathfrak{p}/\mathbb{Q}_p), M[2^k]).$$

Claim: everything is defined over $M[2]$, i.e. the inclusion $M[2] \to M[2^k]$ defines an isomorphism of inflation-restriction sequences. Proof: compute for the left and right H^1 terms and the H^2 term, perhaps using cyclic Tate cohomology. Then use the 5-lemma.

Now, over $M[2] = 2^{-1}\mathbb{Z}/\mathbb{Z}$, all Galois actions are trivial, so $H^1 = Z^1 = \operatorname{Hom}$. By Proposition 2.5, $\operatorname{Gal}(L_\mathfrak{q}/\mathbb{Q}_p) = \operatorname{Gal}(F/\mathbb{Q}_p) \times \operatorname{Gal}(K_\mathfrak{p}/\mathbb{Q}_p)$, so the H^1's must form a *split* short exact sequence of character groups. The splitting expresses $\phi_k|_{\operatorname{Gal}(L_\mathfrak{q}/\mathbb{Q}_p)}$ as the desired sum $\chi + \chi'$. Finally, $b \in N_{K_\mathfrak{p}/\mathbb{Q}_p}(K_\mathfrak{p}^\times)$ implies $(\chi', b)_p = 0$, even if χ' is nontrivial. □

4. Relating different ground fields

4.1. Defining families of objects.

Definition 4.1. Fix a quadratic field $K/\mathbb{Q}$ of discriminant $\Delta_K < 0$. Let $X_1, \dots, X_d$ be pairwise disjoint sets of odd primes $p \nmid \Delta_K$. Let $\mathbf{X} = \mathbf{X}_{[d]}(K)$ denote the product $X_1 \times \cdots \times X_d$, with ith projection π_i to X_i. As $\mathbf{x} = \mathbf{x}_{[d]} \in \mathbf{X}$ varies, define the *family of quadratic fields*

$$K(\mathbf{x}) := \mathbb{Q}(\sqrt{\Delta_K \pi_1(\mathbf{x}) \dots \pi_d(\mathbf{x})}).$$

Call this family *simple* if $p \pmod 4$ is constant for $p \in X_i$.

Remark 4.2. Simplicity requires the sign of the prime discriminant $p^* = (-1)^{(p-1)/2}p$ to be constant on each set X_i. This is natural when applying genus theory in families.

Definition 4.3. Let $\mathbf{X}$ represent a simple family. Call w_b a *constant family* of 2-torsion elements if there exists a constant discriminant $\Delta_b \mid \Delta_K$ such that $w_b(\mathbf{x})$ is the image of

$$(\Delta_b)^{\frac{1}{2}} := \prod_{\wp \in \operatorname{Spec} \mathcal{O}_{K(\mathbf{x})}} \wp^{\frac{1}{2} v_\wp(\Delta_b)} \in I_{K(\mathbf{x})}^\delta \le I_{K(\mathbf{x})}$$

in $\overline{\operatorname{Cl}}_{K(\mathbf{x})}[2] := I^\delta_{K(\mathbf{x})}/I_\mathbb{Q}$, for all $\mathbf{x} \in \mathbf{X}$. Let the *level* be the largest integer $k \ge 1$ such that $w_b(\mathbf{x}) \in 2^{k-1}\overline{\operatorname{Cl}}_{K(\mathbf{x})}[2^k]$ for all $\mathbf{x} \in \mathbf{X}$.

Remark 4.4. Appendix A relates $\overline{\operatorname{Cl}}_{K(\mathbf{x})}[2]$ to the actual 2-torsion group $\operatorname{Cl}_{K(\mathbf{x})}[2]$.

Definition 4.5. Let $\mathbf{X}$ represent a family. Call w_a a *constant family* of characters if there exists a constant discriminant $\Delta_a \mid \Delta_K$ such that $w_a(\mathbf{x})$ is the image of

$$\chi_{\Delta_a} \colon G_\mathbb{Q} \to 2^{-1}\mathbb{Z}/\mathbb{Z} = M_K[2] = M_{K(\mathbf{x})}[2]$$

in $\overline{\operatorname{Cl}}^\vee_{K(\mathbf{x})}[2] = \operatorname{Hom}_{\text{cts}}(\operatorname{Gal}(K(\mathbf{x})^{\text{ur}}/\mathbb{Q}), M_{K(\mathbf{x})}[2])$, for all $\mathbf{x} \in \mathbf{X}$. Let the *level* be the largest integer $k \ge 1$ such that $w_a(\mathbf{x}) \in 2^{k-1}\overline{\operatorname{Cl}}^\vee_{K(\mathbf{x})}[2^k]$ for all $\mathbf{x} \in \mathbf{X}$.

Remark 4.6. Proposition 3.3 relates $\overline{\mathrm{Cl}}^{\vee}_{K(\mathbf{x})}[2]$ to the actual dual 2-torsion group $\widehat{\mathrm{Cl}}_{K(\mathbf{x})}[2]$.

To use Proposition 2.1, we need to *witness the level* of w_a using elements of $\overline{\mathrm{Cl}}^{\vee}_{K(\mathbf{x})}[2^k]$.

Definition 4.7. Call

$$\mathfrak{R}(\mathbf{X}) = (\psi_1(\mathbf{x}), \dots, \psi_{k(\mathbf{x})}(\mathbf{x}))_{\mathbf{x}\in\mathbf{X}}$$

a *set of raw cocycles* (resp. *cochains*) if $\psi_j(\mathbf{x})$ is a $G_{\mathbb{Q}}$-cocycle in $Z^1_{\mathrm{cts}}(G_{\mathbb{Q}}, M_{K(\mathbf{x})}[2^j])$ (resp. $G_{\mathbb{Q}}$-cochain in $C^1_{\mathrm{cts}}(G_{\mathbb{Q}}, M_{K(\mathbf{x})}[2^j])$) for each $j \in [k(\mathbf{x})]$ and $\mathbf{x} \in \mathbf{X}$, such that $\psi_j(\mathbf{x}) = 2\psi_{j+1}(\mathbf{x})$ for $j = 1, \dots, k(\mathbf{x}) - 1$. Let the *level* be the largest integer $k \geq 1$ such that $k(\mathbf{x}) \geq k$ for all $\mathbf{x} \in \mathbf{X}$. If w_a is a family of characters, say that $\mathfrak{R}(\mathbf{X})$ *witnesses* w_a *to level* ℓ if it is a set of raw cocycles such that $\psi_1(\mathbf{x}) = w_a(\mathbf{x})$ and $k \geq \ell$.[6]

Remark 4.8. We do not require $\psi_j(\mathbf{x})$ to be defined over $\mathrm{Gal}(K(\mathbf{x})^{\mathrm{ur}}/\mathbb{Q})$. That is OK for Proposition D.3, a key combinatorial result. But ramification considerations will play a big role in the setup and proof of Theorem 4.15, due to the use of Proposition 3.3.

Definition 4.9. Call a $G_{\mathbb{Q}}$-cocycle *unramified over* L if it is defined over $\mathrm{Gal}(L^{\mathrm{ur}}/\mathbb{Q})$.

Remark 4.10. A cocycle in $Z^1_{\mathrm{cts}}(G_{\mathbb{Q}}, M_K[2^k])$ is unramified over K if and only if it lies in $\overline{\mathrm{Cl}}^{\vee}_{K}[2^k]$. See Proposition C.1 for how to think about fields of definition more precisely.

4.2. **Raw cocycles: consistency and minimality.**

Definition 4.11. If $\mathfrak{R}(\mathbf{X})$ is a set of raw cochains of level $k \geq d$, define the set map

$$\psi_d(\mathbf{X}) := \sum_{\mathbf{x}\in\mathbf{X}} \iota_{\mathbf{x}}\psi_d(\mathbf{x}) \colon G_{\mathbb{Q}} \to \mathbb{Q}_2/\mathbb{Z}_2,$$

where $\iota_{\mathbf{x}}$ is the forgetful map $\iota_{K(\mathbf{x})} \colon M_{K(\mathbf{x})} \to \mathbb{Q}_2/\mathbb{Z}_2$.

Definition 4.12. Let $\mathbf{X} = \mathbf{X}_{[d]}(K)$ represent a family. Call $\mathfrak{R}(\mathbf{X})$ *minimal* or *oscillatory* if it is a set of raw cocycles of level $k \geq d$, and the set map $\psi_d(\mathbf{X})$ is 0.

Remark 4.13. The subtlest part of the definition is $k \geq d$. Cf. Smith's notion of *consistency*, which makes sense at level $k = 1$ for any d. When $k = d = 1$, the notions agree.

To appreciate minimality, and to formulate Theorem 4.15 below, we need to understand the combinatorics of restricted variation.

Definition 4.14. Let $\mathbf{X} = \mathbf{X}_{[d]}(K)$ represent a family, $S \subseteq [d]$ a set of *variation indices*, and $T = [d] - S$ the complementary set of *fixed indices*, with a choice of primes $\mathbf{y} = (q_i)_{i\in T} \in \prod_{i\in T} X_i$. Let $\Delta_{\mathbf{y}}$ denote the discriminant of the quadratic $K_{\mathbf{y}} := \mathbb{Q}(\sqrt{\Delta_K \prod_{i\in T} q_i}) = K(\mathbf{y})$, and $\mathbf{X}_S$ the product set $\prod_{i\in S} X_i$, representing the *restricted family* of fields

$$K_{\mathbf{y}}(\mathbf{x}_S) := \mathbb{Q}(\Delta_{\mathbf{y}}^{1/2} \prod_{i\in S} p_i^{1/2}) = K(\mathbf{y} \sqcup \mathbf{x}_S)$$

for $\mathbf{x}_S = (p_i)_{i\in S} \in \prod_{i\in S} X_i$. One can then define constant families, *restricted* levels, minimality, and so on *with respect to* the data $\mathbf{y}, S$.

Constancy of families is stable under restriction, while level is nondecreasing, witnessing (of w_a by raw cocycles) is stable, and **minimality is stable** (see Appendix D for details).

[6] This means w_a has level at least ℓ, but possibly greater.

4.3. **First half of main theorem.** Let $\mathbf{X} = \mathbf{X}_{[d]}(K)$ represent a simple family of fields. Let w_b denote a constant family of 2-torsion elements, and w_a a constant family of characters. Assume the following conditions:

(1) w_b is of level at least d, where $d \geq 2$.
(2) $|X_i| = 2$ for all $i \in [d]$, with a distinguished point $\mathbf{x}_0 = (p_i)_{i\in[d]}$.
(3) $\mathfrak{R}(\mathbf{X})$ is a set of raw cocycles with $k(\mathbf{x}) \geq d$ and $\psi_1(\mathbf{x}) = w_a(\mathbf{x})$ for all $\mathbf{x} \neq \mathbf{x}_0$, such that $\psi_j(\mathbf{x})$ is unramified over $K(\mathbf{x})$ for all $j \in [d]$. (No condition at $\mathbf{x}_0$.)
(4) For every index $i \in [d]$ and complementary variation set $S = [d] - i$, the set $\mathfrak{R}(\mathbf{X})$ is minimal with respect to the data q_i, S for all $q_i \in X_i \setminus p_i$.

Theorem 4.15 ([8, Theorem 2.8(1)])**.** *Above, $\mathfrak{R}(\mathbf{X})$ can be modified at $\mathbf{x}_0$ so that*

(1) $\mathfrak{R}(\mathbf{X})$ witnesses w_a to level d;
(2) $\psi_j(\mathbf{x}_0)$ is unramified over $K(\mathbf{x}_0)$ for all $j \in [d]$; and
(3) $\psi_d(\mathbf{X})$ is a quadratic $G_{\mathbb{Q}}$-character defined over $\prod_{\mathbf{x}\in\mathbf{X}} K(\mathbf{x})^{\mathrm{ur}}$.

Furthermore,

$$\sum_{\mathbf{x}\in\mathbf{X}_{[d]}} \langle w_a(\mathbf{x}), w_b(\mathbf{x})\rangle = 0,$$

where the pairing $\langle -, -\rangle$ is induced by the torsion class pairing computed in Proposition 2.1.

Remark 4.16. The sum is independent of the witness $\mathfrak{R}(\mathbf{X})$. Can the theorem be strengthened (e.g. smaller sums)? Or can it be weakened (e.g. larger sums) with an easier proof?

Proof. Whenever $k(\mathbf{x}) \geq d$, Proposition 2.1 says

$$\langle w_a(\mathbf{x}), w_b(\mathbf{x})\rangle = \psi_d(\mathbf{x})(w_b(\mathbf{x})) = \mathrm{rec}_{L(\psi_d(\mathbf{x}))/K(\mathbf{x})}(w_b(\mathbf{x})) \in 2^{-1}\mathbb{Z}/\mathbb{Z},$$

where $L(\psi_d(\mathbf{x}))$ is the fixed field of $\ker \psi_d(\mathbf{x})|_{G_{K(\mathbf{x})}}$ acting on the Hilbert class field $H_{K(\mathbf{x})}/K(\mathbf{x})$.

Since w_b is a **constant family** of 2-torsion elements, there is a constant discriminant $\Delta_b \mid \Delta_K$ such that $w_b(\mathbf{x}) = (\Delta_b)^{\frac{1}{2}} \pmod{I_{\mathbb{Q}}}$, the ideal square root taking place in $I_{K(\mathbf{x})}$. As the relevant Artin symbol at $w_b(\mathbf{x})$ is $\mathbb{F}_2$-valued, we can ignore any squares in Δ_b. In other words, let $b = \Delta_b$ if Δ_b is odd, and $b = \Delta_b/4$ otherwise. Then b is squarefree, and equal to the norm of $w_b(\mathbf{x})$, up to a rational square. **Since w_b is of level $d \geq 2$, the ideal class of $w_b(\mathbf{x})$ is a square**, so $b \in N_{K(\mathbf{x})/\mathbb{Q}}(K(\mathbf{x})^\times)$ by the remark following Proposition 3.7. By Propositions 2.5 and 3.7 applied to primes $p \mid b$ of the form $p\mathcal{O}_{K(\mathbf{x})} = \mathfrak{p}(\mathbf{x})^2$, we find

$$\langle w_a(\mathbf{x}), w_b(\mathbf{x})\rangle = \mathrm{rec}_{L(\psi_d(\mathbf{x}))/K(\mathbf{x})}(w_b(\mathbf{x})) = \sum_{p\mid b} \mathrm{rec}_{L(\psi_d(\mathbf{x}))/K(\mathbf{x})}(\mathfrak{p}(\mathbf{x})) = \sum_{p\mid b} (\psi_d(\mathbf{x}), b)_p.$$

We can at last modify $\mathfrak{R}(\mathbf{X})$ at $\mathbf{x}_0$. Provisionally define a 1-cochain

$$\psi_d(\mathbf{x}_0) = -\iota_{\mathbf{x}_0}^{-1} \sum_{\mathbf{x}\neq\mathbf{x}_0} \iota_{\mathbf{x}}\psi_d(\mathbf{x})\colon G_{\mathbb{Q}} \to M_{K(\mathbf{x}_0)},$$

which is in fact a cocycle by Proposition D.3(2). Although this is a continuous 1-cocycle $G_{\mathbb{Q}} \to M_{K(\mathbf{x}_0)}[2^d]$, it may be ramified. For now, multiplying by 2^{d-1} gives $\psi_1(\mathbf{x}_0) = w_a(\mathbf{x}_0)$ by **constancy** of w_a and **oddness** of the number of summation indices $\mathbf{x} \neq \mathbf{x}_0$.

We now study ramification. Minimality with respect to $q_i, [d] - i$ for $q_i \neq p_i$ implies

$$\psi_{d-1}(\mathbf{x}_0) = 2\psi_d(\mathbf{x}_0) = \sum_{\mathbf{x}\neq\mathbf{x}_0 : \pi_i(\mathbf{x}) = p_i} \iota_{\mathbf{x}_0}^{-1}\iota_{\mathbf{x}} \underbrace{\psi_{d-1}(\mathbf{x})}_{\text{defined over } K(\mathbf{x})^{\mathrm{ur}}},$$

for each index $i \in [d]$. If $q_i \in X_i \backslash p_i$, then $\psi_{d-1}(\mathbf{x}_0)$ is unramified at[7] q_i, since $K(\mathbf{x})^{\mathrm{ur}}/K(\mathbf{x})/\mathbb{Q}$ is unramified at q_i for $\mathbf{x} \in \mathbf{X}$ such that $\pi_i(\mathbf{x}) = p_i$. By Lemma B.3, it follows that $\psi_{d-1}(\mathbf{x}_0)$ is defined over $K(\mathbf{x}_0)^{\mathrm{ur}}$. Let $L_0/K(\mathbf{x}_0)/\mathbb{Q}$ be the smallest Galois extension $E/\mathbb{Q}$ containing $K(\mathbf{x}_0)$ such that $\psi_{d-1}(\mathbf{x}_0)$ can be defined over $\mathrm{Gal}(E/\mathbb{Q})$ (see Proposition C.1).

Letting ψ denote the provisional choice of $\psi_d(\mathbf{x}_0)$, Lemma C.3 furnishes $c \in \mathbb{Q}^\times$ such that the cocycle $\psi + \chi_c$ is defined over a Galois tower $L/L_0/K(\mathbf{x}_0)/\mathbb{Q}$ with L/L_0 quadratic and $L/K(\mathbf{x}_0)$ unramified. **Redefine** $\psi_d(\mathbf{x}_0) := \psi + \chi_c$, now inside $\overline{\mathrm{Cl}}^\vee_{K(\mathbf{x}_0)}[2^d]$; since $2\chi_c = 0$, this definition preserves $\psi_{d-1}(\mathbf{x}_0)$ and lower, including $\psi_1(\mathbf{x}_0) = w_a(\mathbf{x}_0)$.

With this new definition,

$$\sum_{\mathbf{x}\in\mathbf{X}} \psi_d(\mathbf{x}) = \chi_c + 0 = \chi_c.$$

So

$$\sum_{\mathbf{x}\in\mathbf{X}_{[d]}} \langle w_a(\mathbf{x}), w_b(\mathbf{x})\rangle = \sum_{\mathbf{x}\in\mathbf{X}_{[d]}} \sum_{p\mid b} (\psi_d(\mathbf{x}), b)_p = \sum_{p\mid b} (\chi_c, b)_p = \sum_{p\mid b} (c, b)_p.$$

By **Hilbert reciprocity**, $\sum_{p\in\operatorname{Spec}\mathbb{Z}} (c,b)_p = 0$, so the previous sum vanishes if and only if

$$\sum_{p\nmid b} (c, b)_p = 0.$$

In fact, each term vanishes! Fix $p \nmid b$. For convenience, replace c with the discriminant of $\mathbb{Q}(\sqrt{c})/\mathbb{Q}$. If $p \nmid c$, then b, a unit, must be a norm in the unramified local extension $\mathbb{Q}_p(\sqrt{c})/\mathbb{Q}_p$, so $(c,b)_p = 0$.

Now suppose $p \mid c$; we will uniformly treat odd and even p. Recall from earlier that $b \in N_{K(\mathbf{x})/\mathbb{Q}}(K(\mathbf{x})^\times)$ for all $\mathbf{x} \in \mathbf{X}$, so $(\Delta_{K(\mathbf{x})}, b)_p = 0$. Since $\mathbf{X}$ is a **simple family**, the 2-part $\Delta_2 \in \{-4, \pm 8\}$ of the discriminant of $K(\mathbf{x})$ is constant as $\mathbf{x} \in \mathbf{X}$ varies. Since χ_c is defined over $\prod K(\mathbf{x})^{\mathrm{ur}}$, Proposition B.2 says the prime discriminant Δ_p of c must lie in the prime discriminant factorization of $K(\mathbf{x})$ for some $\mathbf{x} \in \mathbf{X}$, **even if** $p = 2$. So $\mathbb{Q}(\sqrt{c\Delta_{K(\mathbf{x})}})$ is unramified at p, even if $p = 2$![8] As in the $p \nmid c$ case, we get $(c\Delta_{K(\mathbf{x})}, b)_p = 0$. Finally,

$$(c, b)_p = (\Delta_{K(\mathbf{x})}, b)_p + (c\Delta_{K(\mathbf{x})}, b)_p = 0 + 0 = 0$$

by bilinearity of the quadratic Hilbert symbol, as desired. □

Remark 4.17. On the Selmer side, Smith's proof of [8, Theorem 2.9] seems easier, without need for anything like Lemma C.3. If we weakened Theorem 4.15 by doubling the sizes of the sums, I imagine we would have a correspondingly easier proof here, but I may be missing the bigger picture (either in terms of analytic input, or class-Selmer analogy).

Remark 4.18. We can say more about $L_0 \le K(\mathbf{x}_0)^{\mathrm{ur}}$. Since $\psi_{d-1}(\mathbf{x}_0)$ kills G_{L_0}, the restricted character kernel $G_{F_0} := \ker \psi_{d-1}(\mathbf{x}_0)|_{G_{K(\mathbf{x}_0)}}$ contains G_{L_0}, so $F_0 \le L_0$, so $F_0 \le K(\mathbf{x}_0)^{\mathrm{ur}}$. But $M_{K(\mathbf{x}_0)}[2^{d-1}]$ cyclic implies $F_0/K(\mathbf{x}_0)$ cyclic Galois, so $F_0 \le H^+_{K(\mathbf{x}_0)}$. By Proposition 2.4, $F_0/\mathbb{Q}$ is Galois. Now Corollary C.2 says $\psi_{d-1}(\mathbf{x}_0)$ is defined over F_0, so $F_0 = L_0$.

Remark 4.19. We can say more about L as well. Tracing through Lemma C.3, one sees $G_L := \ker \psi_d(\mathbf{x}_0)|_{G_{L_0}}$. On the other hand, the character kernel $G_F := \ker \psi_d(\mathbf{x}_0)|_{G_{K(\mathbf{x}_0)}}$ lies in G_{F_0}, so $G_F = \ker \psi_d(\mathbf{x}_0)|_{G_{F_0}}$. But $G_{F_0} = G_{L_0}$ from the previous remark, so $G_F = G_L$ and $F = L$. As before, $M_{K(\mathbf{x}_0)}[2^d]$ cyclic implies $F \le H^+_{K(\mathbf{x}_0)}$, so $L = F \le H^+_{K(\mathbf{x}_0)}$.

[7] i.e. "defined over a field unramified at"

[8] For $p = 2$, the point is that $c\Delta_{K(\mathbf{x})}$ is Δ_2^2 times a product of odd prime discriminants.

Appendix A. Genus theory and the 2-class group

For $K/\mathbb{Q}$ quadratic, let $\overline{\mathrm{Cl}}_K[2]$ be the $\mathbb{F}_2$-vector subspace of $I_K/I_\mathbb{Q}$ generated by the finite primes of K ramified over $\mathbb{Q}$. An easy computation gives a short exact sequence $I_\mathbb{Q} \hookrightarrow I_K^\sigma \twoheadrightarrow \overline{\mathrm{Cl}}_K[2]$, so $\overline{\mathrm{Cl}}_K[2]$ can also be described as $I_K^\sigma/I_\mathbb{Q}$. Define the map $\iota\colon \overline{\mathrm{Cl}}_K[2] \to \mathrm{Cl}_K^+[2]$, where $\mathrm{Cl}_K^+ := I_K/P_K^+$ denotes the narrow class group.[9] For convenience, let $K_\infty^\times$ denote the group of totally positive elements of $K^\times$. Write $2^{k-1}\overline{\mathrm{Cl}}_K[2^k] := \iota^{-1}(2^{k-1}\,\mathrm{Cl}_K^+[2^k])$.

Proposition A.1. *The map ι is surjective. Its kernel is isomorphic to $\mathbb{Z}/2$, generated by $(x)I_\mathbb{Q}$, where x is given uniquely up to unique $\mathcal{O}_K^\times\mathbb{Q}^\times$-scalar by*

- *$\sqrt{\Delta_K}$ if $K/\mathbb{Q}$ is imaginary;*
- *$\epsilon + \epsilon^{-1} \in \mathbb{Q}^\times\sqrt{\Delta_K}$ if $K/\mathbb{Q}$ is real with fundamental unit ϵ of norm -1; and*
- *$1+\epsilon$ otherwise, if $K/\mathbb{Q}$ is real with $N\epsilon = +1$, where ϵ is chosen to lie in $K_\infty^\times$.*

Remark A.2. For a "dual" perspective, see Milovic's (master?) thesis on (and slightly generalizing) the work of Fouvry–Klüners. Early on it has a description mapping out of $\mathrm{Cl}^+/2\,\mathrm{Cl}^+$ (instead of mapping into $\mathrm{Cl}^+[2]$) using Hilbert symbols and reciprocity.

Proof. The ideal norm $N = 1+\sigma$ maps into $I_\mathbb{Q} \le P_K^+$, so an ideal $I \in I_K$ satisfies $I^2 \sim (1)$ if and only if $(1-\sigma)I = (x)$ for some $x \in K_\infty^\times$. In this case, $(Nx) = N(1-\sigma)I = (1)$, so $Nx = \pm 1$; total positivity forces $Nx = +1$. By **Hilbert 90**, $x = (1-\sigma)y$ for some $y \in K^\times$, so $(1-\sigma)(Iy^{-1}) = (1)$, i.e. $Iy^{-1} \in I_K^\sigma$. Since $x = y/\sigma y$ is totally positive, y must be either totally positive or negative, so (y) admits a totally positive generator. Hence $[I] = [Iy^{-1}] \in [I_K^\sigma] = \operatorname{im}\iota$, **establishing surjectivity** of ι. □

Remark A.3. We started with the equivalence $I^2 \sim (1) \iff \sigma(I) \sim I$. The latter is natural for generalization to cyclic extensions $K/\mathbb{Q}$: see Klys [5] or Emerton's notes.

Remark A.4. For examples of $1+\epsilon$ in the third case, see fundamental unit tables. For $d = 21$, we have $\epsilon = (5+\sqrt{21})/2$, so $1+\epsilon = (7+\sqrt{21})/2$. For $d = 33$, we have $\epsilon = 23 + 4\sqrt{33}$, so $1+\epsilon = 24 + 4\sqrt{33}$. In general, $N(1+\epsilon) = 2 + a$ if $2\epsilon = a + b\sqrt{d}$ (where $a^2 - db^2 = 4$).

Appendix B. Results on ramification

Proposition B.1. *Suppose M/K is generated by two subextensions E, F. If $K = E \cap F$ and either E or F is finite Galois over K, then E, F are linearly disjoint over K.*

Proof. Say $E = K(\alpha)$ is finite Galois over K. The minimal polynomial f of α over K remains irreducible over F, because $K = E \cap F$. See MSE for further discussion. □

The following results are used to control the ramification of fields and objects of interest.

Proposition B.2. *For $\mathbf{X}$ a family, $\prod_{\mathbf{x}\in\mathbf{X}} K(\mathbf{x})^{\mathrm{ur}}/E$ is unramified, where $E := \prod K(\mathbf{x})/\mathbb{Q}$. If $F/\mathbb{Q}$ is a* quadratic *subfield of $\prod K(\mathbf{x})^{\mathrm{ur}}$, then Δ_F is, up to a square, a product of prime discriminants in $P(\mathbf{X})$, the union of the prime discriminants of $\Delta_{K(\mathbf{x})}$ for $\mathbf{x} \in \mathbf{X}$.*

Proof. C/A and D/B unramified implies CD/AB unramified, so $\prod K(\mathbf{x})^{\mathrm{ur}}/E$ is unramified. Now use the structure of multiquadratic fields: E lies in a linearly disjoint compositum (see Proposition B.1) of "prime discriminant fields" $\mathbb{Q}(\sqrt{\Delta_p})/\mathbb{Q}$, where $\Delta_2 \in \{-4, \pm 8\}$ (any two of which disjointly generate the third), and $\Delta_p = (-1)^{(p-1)/2}p$ if p is odd.

[9] For all imaginary quadratics, and most real quadratics, these coincide. For the purposes of Cohen–Lenstra, see Gerth [4, p. 490–491].

- Let E_{gen} be the smallest compositum of prime discriminant fields such that $E \le E_{\mathrm{gen}}$. Then the odd Δ_p's all lie in $P(\mathbf{X})$, while the Δ_2's in E_{gen} either arise from $P(\mathbf{X})$ or a product of Δ_2's from $P(\mathbf{X})$, up to a square (e.g. -4 is $(+8)(-8)$ up to a square).
- E is ramified precisely at primes dividing $\prod \Delta_{K(\mathbf{x})}$, i.e. the underlying primes of $P(\mathbf{X})$. Easily check that $E_{\mathrm{gen}} \le \prod K(\mathbf{x})^{\mathrm{ur}}$.

The quadratic $F/\mathbb{Q}$ lies in $\prod K(\mathbf{x})^{\mathrm{ur}}$, so every prime discriminant Δ_q of Δ_F must either be in $P(\mathbf{X})$ or a product of Δ_2's from $P(\mathbf{X})$, up to a square. Otherwise, F and E would be linearly disjoint over $\mathbb{Q}$ (again, see Proposition B.1), and FE/E would be ramified at q. □

Lemma B.3. *For* $\mathbf{X}$ *a* simple *family with* $\mathbf{x}_0 \in \mathbf{X}$ *distinguished,* $\prod K(\mathbf{x})^{\mathrm{ur}}/K(\mathbf{x}_0)$ *is* unramified outside of $R := \bigcup_{i\in[d]}(X_i \setminus \pi_i(\mathbf{x}_0))$*. Consequently,* $K(\mathbf{x}_0)^{\mathrm{ur}}$ *is the maximal subextension of* $\prod_{\mathbf{x}\in\mathbf{X}} K(\mathbf{x})^{\mathrm{ur}}/\mathbb{Q}$ unramified at *every prime in* R*.*

Proof. The first part of Proposition B.2 says $\prod K(\mathbf{x})^{\mathrm{ur}}/E$ is unramified. Since $\mathbf{X}$ is a **simple family**, the 2-part $\Delta_2 \in \{-4, \pm 8\}$ of $\Delta_{K(\mathbf{x})}$ is constant as $\mathbf{x} \in \mathbf{X}$ varies. Thus $K(\mathbf{x})K(\mathbf{x}_0)/K(\mathbf{x}_0)$ can only be ramified over *odd* primes $p \mid \Delta_{K(\mathbf{x})}$ with $p \nmid \Delta_{K(\mathbf{x}_0)}$. This automatically excludes the primes $p \mid \Delta_K$. We are left with precisely the primes $p \in R$ as possibilities. In other words, $E/K(\mathbf{x}_0)$ is unramified outside of R. Thus the whole tower $K(\mathbf{x})^{\mathrm{ur}}/E/K(\mathbf{x}_0)$ is unramified outside of R, proving the first part of the lemma.

We then immediately get that $\prod K(\mathbf{x})^{\mathrm{ur}}/K(\mathbf{x}_0)^{\mathrm{ur}}$ is unramified outside of R. Yet by definition of $K(\mathbf{x}_0)^{\mathrm{ur}}$, every subextension $E/K(\mathbf{x}_0)^{\mathrm{ur}}$ of $\prod K(\mathbf{x})^{\mathrm{ur}}/K(\mathbf{x}_0)^{\mathrm{ur}}$ is ramified, hence ramified somewhere over R. So $K(\mathbf{x}_0)^{\mathrm{ur}}$ has the desired maximality property. □

Remark B.4. Similarly, if $p^* \in P(\mathbf{X})$, then $\prod K(\mathbf{x})^{\mathrm{ur}}/E/\mathbb{Q}(\sqrt{p^*})$ is unramified at p.

Lemma B.5. *Let* $K = F(\sqrt{a})$ *and* $L = F(\sqrt{b})$ *be two* ramified *quadratic extensions of local fields over* $\mathbb{Q}_p$*. If* KL/L *is* unramified*, then so is* $F(\sqrt{ab})/F$*.*

Proof. KL/F has $e = 2 \ge f$, so $F(\sqrt{ab})/F$ must be the maximal unramified extension. □

Appendix C. Fields of definition of cocycles

Proposition C.1. *Let* N *be a* $G_{\mathbb{Q}}$*-module, and let* $\psi\colon G_{\mathbb{Q}} \to N$ *be a continuous 1-cocycle. Let* G_L *be a* normal *open subgroup in the kernel of set map* ψ*. Then* ψ *is defined over* L*.*

Proof. Take $h \in G_L$ in $\psi(gh) = g\psi(h) + \psi(g)$ to get $\psi(gh) = \psi(g)$ for all $g \in G_{\mathbb{Q}}$. Since G_L is normal, ψ induces a set map $\overline{\psi}\colon \mathrm{Gal}(L/\mathbb{Q}) = G_{\mathbb{Q}}/G_L \to N$. Now $\overline{\psi}(\overline{g}\overline{h}) = g\overline{\psi}(\overline{h}) + \overline{\psi}(\overline{g})$ for any $g, h \in G_{\mathbb{Q}}$, so $g\overline{\psi}(\overline{h})$ is independent of the coset representative $g \in \overline{g}$. Thus $\overline{\psi}\colon \mathrm{Gal}(L/\mathbb{Q}) \to N^{G_L}$ is a finite cocycle with $G_{\mathbb{Q}}$-inflation ψ, as desired. □

Corollary C.2. *Take* $K/\mathbb{Q}$ *Galois, and take* N *on which* G_K *acts trivially. Let* G_L *be the kernel of the homomorphism* $\psi|_{G_K}\colon G_K \to N$*. If* $L/\mathbb{Q}$ *is Galois, then* ψ *is defined over* L*.*

The following "quadratic twist" result is used in proving Theorem 4.15. For $\mathbf{X}$ a simple family with $\mathbf{x}_0 \in \mathbf{X}$ distinguished, let ψ be a cocycle $G_{\mathbb{Q}} \to M_{K(\mathbf{x}_0)}$ such that

(1) ψ is defined over $\prod_{\mathbf{x}\in\mathbf{X}} K(\mathbf{x})^{\mathrm{ur}}$, and
(2) 2ψ is defined over L_0, where $K(\mathbf{x}_0) \le L_0 \le K(\mathbf{x}_0)^{\mathrm{ur}}$ and $L_0/\mathbb{Q}$ is finite Galois.

Lemma C.3. *In the setting above, if* χ_c *denotes the quadratic character of* $\mathbb{Q}(\sqrt{c})/\mathbb{Q}$*, then for any* $c \in \mathbb{Q}^\times$*, the twist* $\psi + \chi_c$ *is a cocycle defined over a Galois tower* $L^c/L_0/\mathbb{Q}$*, with* L^c/L_0 *at most quadratic. Furthermore, there exists* c *such that* $L^c/K(\mathbf{x}_0)$ *is unramified, i.e.*

$$\psi + \chi_c \in \overline{\mathrm{Cl}}^{\vee}_{K(\mathbf{x}_0)}.$$

Proof of field of definition. Take $g \in G_{\mathbb{Q}}$ and $n \in G_{L_0}$. Then $2\psi(n) = 0$ by definition of L_0, so g acts trivially on $\psi(n) \in 2^{-1}\mathbb{Z}/\mathbb{Z}$. Also, $n \in G_{L_0} \le G_{K(\mathbf{x}_0)}$, so $\delta_{K(\mathbf{x}_0)}(n) = +1$. Thus

$$\begin{aligned}\psi(gng^{-1}) &= \psi(g) + g\psi(n) + gn\psi(g^{-1}) \\ &= \psi(g) + \psi(n) + g\psi(g^{-1}) = \psi(n) + \psi(gg^{-1}) = \psi(n).\end{aligned}$$

In particular, G_L, the subgroup of G_{L_0} killed by ψ, is normal in $G_{\mathbb{Q}}$, so $L/\mathbb{Q}$ is Galois ψ is defined over L by Proposition C.1. The group G_L is actually the kernel of $\psi|_{G_{L_0}} \colon G_{L_0} \to M_{K(\mathbf{x}_0)}[2] = 2^{-1}\mathbb{Z}/\mathbb{Z}$ (a character), so L/L_0 is at most quadratic.

With the twist, $\psi + \chi_c$ is still a cocycle with $2(\psi + \chi_c) = 2\psi$. The previous paragraph, applied to $\psi + \chi_c$ instead of ψ, yields a Galois tower $L^c/L_0/\mathbb{Q}$ with L^c/L_0 quadratic. □

Proof of existence of twist. Suppose $L/K(\mathbf{x}_0)$ is ramified along a tower of primes $\mathfrak{Q}/\mathfrak{q}/\mathfrak{p}/p$, with $\mathfrak{q}/\mathfrak{p}$ unramified but $\mathfrak{Q}/\mathfrak{q}$ ramified. Assumption (1) on ψ says $L \le \prod K(\mathbf{x})^{\mathrm{ur}}$, so

- Lemma B.3 implies $p \in \bigcup_{i\in[d]}(X_i \setminus \pi_i(\mathbf{x}_0))$, because $L/K(\mathbf{x}_0)$ is ramified over p; while
- if we choose $\mathbf{x}^p \in \mathbf{X}$ with p ramified in $K(\mathbf{x}^p)$, say $\pi_i(\mathbf{x}^p) = p$, and $\pi_j(\mathbf{x}^p) = \pi_j(\mathbf{x}_0)$ for $j \in [d] - i$, then Lemma B.3 implies that $LK(\mathbf{x}^p)/K(\mathbf{x}^p)$ is unramified over p.

By the first point, $\mathfrak{p}/p$, hence $\mathfrak{q}/p$, is unramified. Let $p_0 = \pi_i(\mathbf{x}_0)$ and $c = p^*p_0^*$, so $L(\sqrt{c}) = LK(\mathbf{x}^p)$ by simplicity of $\mathbf{X}$. Clearly no new primes of $\mathbb{Z}$ can ramify in L^c. If we show that L^c/L_0 is unramified over p, then the desired twist will exist by induction.

Now, $\psi|_{G_{L_0}}$ and $\chi_c|_{G_{L_0}}$ are both quadratic characters, with kernels G_L and $G_{L_0(\sqrt{c})}$, respectively. If $L = L_0(\sqrt{\alpha})$, then the kernel of $\psi|_{G_{L_0}} + \chi_c|_{G_{L_0}}$ is $G_{L_0(\sqrt{\alpha c})}$, so $L^c = L_0(\sqrt{\alpha c})$, which is **Galois** over $\mathbb{Q}$ by the first half of the lemma. By the second point above, $LK(\mathbf{x}^p)/K(\mathbf{x}^p)$ is unramified over p, so $e_p(LK(\mathbf{x}^p)/\mathbb{Q}) = 2$. Yet $e_p(L/L_0) = 2$ in the Galois tower $L(\sqrt{c})/L/L_0/\mathbb{Q}$, so $L(\sqrt{c})/L = LK(\mathbf{x}^p)/L$ must be unramified over p. Now restrict attention to the biquadratic extension $L(\sqrt{c})/L_0$, all of which is Galois over $\mathbb{Q}$. Since $\mathfrak{q}/p$ is unramified, $\mathfrak{q}$ must ramify in $L_0(\sqrt{c})$. In the local Galois picture, $L(\sqrt{c})/L_0$ satisfies Lemma B.5, so $L_0(\sqrt{\alpha c})/L_0 = L^c/L_0$ is unramified over p. □

Remark C.4. I got stuck trying to prove this lemma while reading [8]; thanks to Alex for explaining the details to me, especially for the field of definition. Below is what Alex suggested for the twist proof; it differs a little from the proof above.

Once we have $\mathfrak{q}/p$ unramified, the *normal* subgroup $\mathrm{Gal}(L/L_0) \cong \mathbb{Z}/2$ of $\mathrm{Gal}(L/\mathbb{Q})$ must be the inertia group of *every* prime $\mathfrak{Q}/p$ of L. Since $e_p(LK(\mathbf{x}^p)/\mathbb{Q}) = 2$, the inertia group of every prime of $L(\sqrt{c})$ over $\mathbb{Q}$ is also of size two. It also always lies in the pullback $\mathrm{Gal}(L(\sqrt{c})/L_0)$ of $I_p(L/\mathbb{Q}) = \mathrm{Gal}(L/L_0)$ under $\mathrm{Gal}(L(\sqrt{c})/\mathbb{Q}) \twoheadrightarrow \mathrm{Gal}(L/\mathbb{Q})$. But $\mathrm{Gal}(L(\sqrt{c})/\mathbb{Q}) \twoheadrightarrow \mathrm{Gal}(L/\mathbb{Q})$ induces a surjection, hence isomorphism, of the equally-sized inertia groups, with target $I_p(L/\mathbb{Q})$. Consequently, if σ is a nontrivial inertia element in $\mathrm{Gal}(L(\sqrt{c})/\mathbb{Q})$, then $\psi(\sigma) = \psi(\sigma|_L)$ is nonzero, or else L would be equal to L_0 by definition of L/L_0. Similarly, $\chi_c(\sigma) = \chi_c(\sigma|_{K(\mathbf{x}^p)}) \ne 0$. Thus $(\psi + \chi_c)(\sigma) = 2^{-1} + 2^{-1} = 0$, and $\psi + \chi_c$ kills the twisted inertia group $I_p(L^c/\mathbb{Q})$. So $L^c/K(\mathbf{x}_0)$ must be unramified over p.

Appendix D. Results on minimality

First, minimality is stable under restriction. There might be a conceptual reason (dihedral intuition?). For now, see the formal argument below (downwards induction on d).

View the target of $\psi_d(\mathbf{X})$ as a trivial $G_{\mathbb{Q}}$-module, so $d\psi_d(\mathbf{X})$ measures how far $\psi_d(\mathbf{X})$ is from being a group homomorphism. Compute the coboundary:

$$\begin{aligned} d\psi_d(\mathbf{X})(g,h) &= \sum_{\mathbf{x}\in\mathbf{X}} d\iota_{\mathbf{x}}\psi_d(\mathbf{x})(g,h) \\ &= \sum_{\mathbf{x}\in\mathbf{X}} g\iota_{\mathbf{x}}\psi_d(\mathbf{x})h - \iota_{\mathbf{x}}\psi_d(\mathbf{x})gh + \iota_{\mathbf{x}}\psi_d(\mathbf{x})g \\ &= \sum_{\mathbf{x}\in\mathbf{X}} g\iota_{\mathbf{x}}\psi_d(\mathbf{x})h - \iota_{\mathbf{x}}g\psi_d(\mathbf{x})h = \sum_{\mathbf{x}\in\mathbf{X}} \psi_d(\mathbf{x})h - g\psi_d(\mathbf{x})h = \sum_{\mathbf{x}:\delta_{K(\mathbf{x})}(g)=-1} \psi_{d-1}(\mathbf{x})h, \end{aligned}$$

where the last step uses the $G_{\mathbb{Q}}$-action $g \mapsto \delta_{K(\mathbf{x})}(g) = \pm 1$, and $2\psi_d = \psi_{d-1}$.

Remark D.1. If $d = 1$, we automatically get 0 coboundary with no hypotheses on $\psi_d(\mathbf{X})$, because the $\psi_1(\mathbf{x})$ cocycles are actually characters (homomorphisms).

What does $\{\mathbf{x} : \delta_{K(\mathbf{x})}(g) = -1\}$ look like? Let $E/\mathbb{Q}$ be the Galois extension defined by

$$G_E := \ker(g \mapsto (\delta_{K(\mathbf{x})}(g))_{\mathbf{x}\in\mathbf{X}}) = \bigcap_{\mathbf{x}\in\mathbf{X}} G_{K(\mathbf{x})}.$$

Observation D.2. *Fix $i \in [d]$ and $p_i \in X_i$, odd by definition. There exists $g = g_{i,p_i} \in G_{\mathbb{Q}}$, unique modulo G_E, such that $\delta_{K(\mathbf{x})}(g) = -1$ if and only if the ith component of $\mathbf{x}$ is p_i. As i, p_i vary, these elements generate $G_{\mathbb{Q}}/G_E = \mathrm{Gal}(E/\mathbb{Q})$.*

Proof. See Proposition B.2 and its proof, which places $E = \prod K(\mathbf{x})/\mathbb{Q}$ in E_{gen}. Adjoin $\sqrt{-4}$ for simplicity, and let $L/\mathbb{Q}$ be the resulting multiquadratic field of dimension t.

Choose $g \in \mathrm{Gal}(L/\mathbb{Q}) = \mathbb{F}_2^t$ acting nontrivially on $\mathbb{Q}(\sqrt{\Delta_{p_i}})/\mathbb{Q}$, but trivially on the remaining $t-1$ pieces $\mathbb{Q}(\sqrt{\Delta_p})/\mathbb{Q}$ of L, including $\mathbb{Q}(\sqrt{-4})/\mathbb{Q}$ for $p = 2$. Since $\Delta_{p_i} = \pm p_i$, the element g acts nontrivially on $\sqrt{p_i}$ but trivially on $\sqrt{q}$ if $q \mid \Delta_K$ or $q \in X_1 \cup \cdots \cup X_d \setminus p_i$. Thus $\delta_{K(\mathbf{x})}(g) = -1$ if and only if $\pi_i(\mathbf{x}) = p_i$. So g induces the desired $g_{i,p_i} \in G_{\mathbb{Q}}$.

Two different g_{i,p_i}'s in $G_{\mathbb{Q}}$ agree under the map $g \mapsto (\delta_{K(\mathbf{x})}(g))_{\mathbf{x}\in\mathbf{X}}$, so their ratio lies in G_E by definition. Thus $g_{i,p_i} \bmod G_E$ is unique.

Clearly $G_L \le G_E$, so $E \le L$. Take the explicit representatives $g_{i,p_i} = g \in \mathrm{Gal}(L/\mathbb{Q})$ defined earlier. To show generation, pick $\sigma \in \mathrm{Gal}(L/\mathbb{Q})$. Modulo the images of g_{i,p_i} in $\mathrm{Gal}(E/\mathbb{Q})$, we may assume σ acts trivially on $\mathbb{Q}(\sqrt{\Delta_{p_i}})$ for all primes $p_i \in X_1 \cup \cdots \cup X_d$. Then σ acts uniformly on the fields $K(\mathbf{x})$ as $\mathbf{x} \in \mathbf{X}$ varies. If the action is $+1$, then $\sigma \equiv 1 \bmod \mathrm{Gal}(L/E)$ as desired. Otherwise, if the action is -1, then

$$\sigma \equiv \prod_{p_1\in X_1} g_{1,p_1} \mod \mathrm{Gal}(L/E),$$

because $\prod_{p_1\in X_1} \delta_{K(\mathbf{x})}(g_{1,p_1}) = -1$ for all $\mathbf{x} \in \mathbf{X}$, by construction of the g_{1,p_1}'s. □

If $\psi_d(\mathbf{X}_{[d]}(K)) = 0$, then $d\psi_d(\mathbf{X}) = 0$, so $\psi_{d-1}(\mathbf{X}_S(K_{\mathbf{y}})) = 0$ for any index $i \in [d]$ with complementary variation set $S = [d] - i$, and any singleton $\mathbf{y}$ of X_i. In other words, $\mathfrak{R}(\mathbf{X})$ is minimal with respect to $\mathbf{y}, S$. By induction, minimality is stable under restriction. In fact, we only needed that $\psi_d(\mathbf{X})$ was a cocycle (or homomorphism) to conclude that $\mathfrak{R}(\mathbf{X})$ is minimal with respect to all proper subsets. What is the best converse statement?

Proposition D.3 ([8, Cf. Proposition 2.5])**.** *Let $\mathbf{X} = \mathbf{X}_{[d]}(K)$ represent a family.*

(1) *Assume* $\mathfrak{R}(\mathbf{X})$ *is a set of raw* cocycles *of level* $k \geq d$*, such that* $\mathfrak{R}(\mathbf{X})$ *is minimal with respect to* $p_i, [d]-i$ *for all* $i \in [d]$ *and* $p_i \in X_i$*. Then* $\psi_d(\mathbf{X})$ *is a* quadratic character.
(2) *Assume* $\mathfrak{R}(\mathbf{X})$ *is a set of raw* cochains *of level* $k \geq d$*, such that* $\psi_d(\mathbf{x})$ *is a* cocycle for all $\mathbf{x} \neq \mathbf{x}_0$*. Assume* $\mathfrak{R}(\mathbf{X})$ *is minimal with respect to* $p_i, [d]-i$ *for all* $i \in [d]$ *and* $p_i \in X_i \setminus \pi_i(\mathbf{x}_0)$*. If* $\psi_d(\mathbf{X})$ *is a* quadratic character*, then* $\psi_d(\mathbf{x}_0)$ *is in fact a cocycle.*

Remark D.4. Smith does not explicitly state the second version, but at least when $\psi_d(\mathbf{X}) = 0$ (trivial character), it is used in proving Theorem 4.15.

Smith gives a binomial theorem proof. Here is another perspective.

Proof. In the first case, $2\psi_d(\mathbf{X})$ breaks up (in any number of ways) into sums of ψ_{d-1} terms, where each sum vanishes by the minimality assumptions. So $2\psi_d(\mathbf{X}) = 0$ and $\psi_d(\mathbf{X})$ is a set map $G_\mathbb{Q} \to 2^{-1}\mathbb{Z}/\mathbb{Z}$. In the second case this is assumed.

Earlier, we computed the coboundary of $\psi_d(\mathbf{X})$ at $(g,h) \in G_\mathbb{Q}^2$ to be

$$d\psi_d(\mathbf{X})(g,h) = \sum_{\mathbf{x}\in\mathbf{X}} \psi_d(\mathbf{x})h - g\psi_d(\mathbf{x})h = \sum_{\mathbf{x}:\delta_{K(\mathbf{x})}(g)=-1} \psi_{d-1}(\mathbf{x})h$$

in the first case. In the second case, a similar coboundary calculation shows that

$$\begin{aligned} 0 = d\psi_d(\mathbf{X})(g,h) &= \sum_{\mathbf{x}\in\mathbf{X}} \psi_d(\mathbf{x})h - \psi_d(\mathbf{x})gh + \psi_d(\mathbf{x})g \\ &= d\psi_d(\mathbf{x}_0)(g,h) + \sum_{\mathbf{x}\in\mathbf{X}} \psi_d(\mathbf{x})h - g\psi_d(\mathbf{x})h \\ &= d\psi_d(\mathbf{x}_0)(g,h) + \sum_{\mathbf{x}:\delta_{K(\mathbf{x})}(g)=-1} \psi_{d-1}(\mathbf{x})h. \end{aligned}$$

In both cases, we wish to show that

$$\sum_{\mathbf{x}:\delta_{K(\mathbf{x})}(g)=-1} \psi_{d-1}(\mathbf{x}) = 0$$

for all $g \in G_\mathbb{Q}$, or equivalently that

$$\sum_{\mathbf{x}\in\mathbf{X}} \psi_d(\mathbf{x}) = \sum_{\mathbf{x}\in\mathbf{X}} g\psi_d(\mathbf{x}).$$

Certainly, the first equality holds for any g_{i,p_i} from the previous observation. In the first case, that's just minimality of $\mathfrak{R}(\mathbf{X})$ with respect to $p_i, [d]-i$. In the second case, if p_i is exceptional for i, then bundling up minimality with respect to $q_i, [d]-i$ for $q_i \neq p_i$, together with $2\psi_d(\mathbf{X}) = 0$, still recovers the desired first equality.

Now, let U be the subset of $G_\mathbb{Q}$ for which either equality holds. Clearly U contains G_E, and we have just shown $g_{i,p_i} \in U$. But if $g, g' \in U$, then

$$\begin{aligned} \sum_{\mathbf{x}\in\mathbf{X}} (gg'-1)\psi_d(\mathbf{x}) &= \sum_{\mathbf{x}\in\mathbf{X}} (gg'-g-g'+1)\psi_d(\mathbf{x}) \\ &= \sum_{\mathbf{x}\in\mathbf{X}} (1-g)(1-g')\psi_d(\mathbf{x}) = \sum_{\mathbf{x}:\delta_{K(\mathbf{x})}(g')=-1} (1-g)\psi_{d-1}(\mathbf{x}). \end{aligned}$$

If g' is one of the g_{i,p_i}, then the sum vanishes under the first hypothesis by minimality of $\mathfrak{R}(\mathbf{X})$ with respect to $p_i, [d]-i$;[10] and still under the second hypothesis by a minimality bundling argument, at least if the technical assumption

$$\sum_{\mathbf{x}\in\mathbf{X}} \psi_{d-1}(\mathbf{x}) = \sum_{\mathbf{x}\in\mathbf{X}} g\psi_{d-1}(\mathbf{x})$$

holds. Assuming this, $g, g' \in U$ implies $gg' \in U$ whenever $g' \in \{g_{i,p_i}\}$. So $U = G_{\mathbb{Q}}$, since $G_{\mathbb{Q}}/G_E$ is a **finite** group generated by the g_{i,p_i}.

Under the second hypothesis, it remains to verify the technical assumption. Set

$$L := \{\ell \geq 0 : \sum_{\mathbf{x}\in\mathbf{X}} 2^{\ell}\psi_d(\mathbf{x}) = \sum_{\mathbf{x}\in\mathbf{X}} g2^{\ell}\psi_d(\mathbf{x}) \quad \forall g \in G_{\mathbb{Q}}\}.$$

Above, we proved $0 \in L$ as long as $1 \in L$. The same argument with $\psi_1, \dots, \psi_d$ doubled shows that $1 \in L$ as long as $2 \in L$. Generally, $\ell \in L$ as long as $\ell + 1 \in L$. But certainly $d \in L$, so $d-1 \in L$, etc. and finally $0 \in L$. $\square$

Appendix E. Class group heuristics

For any set S of finite primes of K, we have $I_K^S = \mathbb{Z}^{|S|}$ canonically, while $\operatorname{rank}(U_K^S) = |S| + \operatorname{rank}(U_K)$ (easiest proof uses finiteness of class group: $\wp^h$ is principal for $\wp \in S$). If S is large enough (i.e. generates the class group), $\operatorname{Cl}(K) \xleftarrow{\sim} \operatorname{coker}(U_K^S \twoheadrightarrow P_K^S \hookrightarrow I_K^S)$.

As K varies in a natural family (of global fields), want to understand distribution of $\operatorname{Cl}(K)$; for simplicity, let's restrict attention to $\operatorname{Cl}_{p^\infty}(K) := \operatorname{Cl}(K)[p^\infty]$ separately for each prime p.

E.1. Random matrix formulation.

Fix $u := \operatorname{rank}(U_K)$ and let $n := |S| \to \infty$. For any K, S, consider the map $\iota = \iota_K^S \colon U_K^S \otimes \mathbb{Z}_p \to I_K^S \otimes \mathbb{Z}_p$: choosing bases on the left[11] and right, we get a matrix $A = A_K^S \colon \mathbb{Z}_p^{n+u} \to \mathbb{Z}_p^n$. By Smith normal form theory, the set of possible resulting matrices is precisely $\{A : \operatorname{coker} A \cong \operatorname{coker} \iota\}$. We want to know the resulting distribution on $\operatorname{coker} A \cong \operatorname{coker} \iota$, at least as $n \to \infty$.

The safest form of the Cohen–Lenstra heuristics roughly states:

Conjecture E.1. *Let $K/\mathbb{Q}$ vary among, say, degree d number fields with a given* unit rank $u := \operatorname{rank} U_K$, *containing no pth roots of unity. Then*

$$\mathbb{P}(\operatorname{Cl}_{p^\infty}(K) \cong P) = \lim_{n\to\infty} \mathbb{P}(\operatorname{coker} A_n \cong P) = |P|^{-u}|\operatorname{Aut}(P)|^{-1} \prod_{k\geq 1}(1 - p^{-k-u})$$

for any finite abelian p-group P, where $A_n \colon \mathbb{Z}_p^{n+u} \to \mathbb{Z}_p^n$ is a random matrix drawn with respect to Haar measure on $M_{n,n+u}(\mathbb{Z}_p)$.

Remark E.2. The μ_p assumption might be unnecessary sometimes, especially if $p = 2$?

Remark E.3. To understand the second equality, note that for large n and e with $p^e P = 0$, almost all maps $(\mathbb{Z}/p^e)^n \to P$ are surjective, so there are around $|\operatorname{Aut}(P)|^{-1}|P|^n$ subgroups of $(\mathbb{Z}/p^e)^n$, say—or better, open subgroups of $\mathbb{Z}_p^n$—with cokernel isomorphic to P. But there are Haar measure $\sim |P|^{-n-u}$ matrices $\mathbb{Z}_p^{n+u} \to \mathbb{Z}_p^n$ with a prescribed image of index $|P|$.[12] So there should be Haar measure $\sim |P|^{-u}|\operatorname{Aut}(P)|^{-1}$ matrices with cokernel isomorphic to P.

[10] Minimality implies coboundary zero, which implies $\sum \psi_{d-1}(\mathbf{x}_S) = \sum g\psi_{d-1}(\mathbf{x}_S)$.

[11] mod torsion (if K contains pth roots of unity)

[12] First show this for image *contained* in that prescribed subgroup, a la Ellenberg–Venkatesh–Westerland surjections perspective [2, 3], and then use inclusion-exclusion.

See Wood [10] for more details on the random matrix train of thought.

Remark E.4. If A has full rank, then $\operatorname{coker} A$ is unaffected up to isomorphism by small perturbations. For example, we can play around with Gaussian elimination on the perturbed Smith normal form of A. Alternatively, we can even show that $\operatorname{im} A = \operatorname{im} A'$ by noting that $\operatorname{im} A$ is finite-index, hence open (contains $p^r\mathbb{Z}_p^n$ for large r), in $\mathbb{Z}_p^n$, so $A'e_i \approx Ae_i$ lies in $\operatorname{im} A$ for all i means $\operatorname{im} A' \leq \operatorname{im} A$, and vice versa.

Remark E.5. Why expect Cohen–Lenstra? Can we choose $A = A_K^S$ equidistributed (or weaker, see [10])? Perhaps one can choose canonical bases on the left and right for which we have no known structure obstructing equidistribution. For example, the "canonical" choice $I_K^S \otimes \mathbb{Z}_p = \mathbb{Z}_p^n$ is justified precisely because the matrices A should automatically equidistribute in $\mathrm{GL}_n(\mathbb{Z}_p)A$ as K varies. Similarly, even though there might not be a nice identification $U_K^S \otimes \mathbb{Z}_p = \mathbb{Z}_p^{n+u}$, the ultimate distribution of A should be dense in $A\,\mathrm{GL}_{n+u}(\mathbb{Z}_p)$.[13]

Remark E.6. What if instead, we modeled the inclusions $\iota'\colon P_K^S \otimes \mathbb{Z}_p \hookrightarrow I_K^S \otimes \mathbb{Z}_p$ by random matrices $A'\colon \mathbb{Z}_p^n \to \mathbb{Z}_p^n$? (Requiring A' to be injective or not shouldn't matter: almost all matrices $\mathbb{Z}_p^n \to \mathbb{Z}_p^n$ are injective.) The resulting distribution $\operatorname{coker} A'$ (independent of u) would not match $\operatorname{coker} A$ (dependent on u) chosen above, except when $u = 0$. How do we rule out this alternative heuristic for $u > 0$?

References

[1] H. Cohen and H. W. Lenstra Jr., Heuristics on class groups of number fields.
[2] J. S. Ellenberg and A. Venkatesh, Statistics of Number Fields and Function Fields, `http://math.stanford.edu/~akshay/research/evicm.pdf` (ICM 2010).
[3] J. S. Ellenberg, A. Venkatesh, and C. Westerland, Homological stability for Hurwitz spaces and the Cohen-Lenstra conjecture over function fields, arXiv:0912.0325.
[4] F. Gerth, The 4-class ranks of quadratic fields (1984).
[5] J. Klys, The Distribution of p-Torsion in Degree p Cyclic Fields, arXiv:1610.00226.
[6] J. Neukirch, A. Schmidt, and K. Wingberg, Cohomology of Number Fields.
[7] A. Smith, Governing fields and statistics for 4-Selmer groups and 8-class groups, arXiv:1607.07860 (2016).
[8] ———, 2^∞-Selmer groups, 2^∞-class groups, and Goldfeld's conjecture, arXiv:1702.02325v2 (Jun 2017).
[9] W. C. Waterhouse, Pieces of Eight in Class Groups of Quadratic Fields.
[10] Wood, Random integral matrices and the Cohen Lenstra Heuristics, arXiv:1504.04391.

[13] In this case, one lazy way to see this is to vary S among sets of size $n \gg h_K$, and as long as U_K is not always identified in a stupidly consistent way in $\mathbb{Z}_p^{n+u}$, we should be OK.