THE NEGATIVE PELL EQUATION

ERICK KNIGHT AND STANLEY YAO XIAO

Abstract. By applying methods developed by A. Smith in [6], we show that the density of square-free integers d in $[1, N]$ for which the negative Pell equation $x^2 - dy^2 = -1$ has a solution is as predicted by the model of Stevenhagen.

Contents

1. Introduction 1
2. The narrow Reidi data and the negative Pell equation 5
3. Governing expansions and sets of raw cocycles 6
4. Raw cocycles for narrow class groups of real quadratic fields 19
5. Additive restrictive systems 23
6. Ramsey theory 29
7. Analytic tools 34
8. Equidistribution of Legendre symbols: the seed distribution 45
9. The Markov process: completion of the proof of Theorem 3 52
References 59

1. Introduction

It is a well known fact that for any square-free positive integer $D > 1$ the equation

$$x^2 - Dy^2 = 1, \tag{1.1}$$

misattributed by Leonhard Euler to John Pell, has infinitely many integer solutions. Moreover, as is now easy to see from algebraic number theory, that the solutions are naturally generated by a single fundamental solution.

The analogous equation, the *negative Pell equation*

$$x^2 - Dy^2 = -1, \tag{1.2}$$

is much more mysterious. If $p|D$, then reducing modulo p reveals that the congruence $x^2 \equiv -1 \pmod{p}$ must be soluble, whence $p \equiv 1 \pmod{4}$ or $p = 2$. Therefore one must conclude that D is a *sum of two squares*, by a theorem of Fermat. However not all such D has the property that (1.2) is soluble: the smallest counterexample is $D = 34$.

A well-known criterion for the solubility of (1.2) has to do with continued fractions. Indeed, (1.2) is soluble if and only if the period of the continued fraction expansion of $\sqrt{D}$ is odd. However, the period of the continued fraction of $\sqrt{D}$ as

Date: July 28, 2019.

a function of D appears to be extremely mysterious, and it is not clear how to use this criterion to understand (1.2). In [9], C. Tsang and the second author proved an alternative criterion for the solubility (1.2) in terms of the reduction theory of indefinite, irreducible integral binary quadratic forms, but like the continued fractions criterion, this does not appear to shed further light.

For a positive number N put

$$\mathfrak{S}(N) = \#\{n \leq N : n \text{ square-free}, \exists x, y \in \mathbb{Z} \text{ s.t. } x^2 + y^2 = n\}$$

for the counting function of square-free integers which are sums-of-two-squares up to N. It is well-known that there exists a positive number c_0 such that

$$\mathfrak{S}(B) \sim c_0 N (\log N)^{-1/2},$$

a fact already known to Landau.

In view of the fact that D being a sum of two squares is a necessary condition for the solubility of (1.2), it is natural to compare the number of $D \leq N$ for which (1.2) is soluble to $\mathfrak{S}(N)$. Put

$$\mathcal{S}(N) = \#\{D \leq N : D \text{ square-free}, (1.2) \text{ is soluble}\}. \tag{1.3}$$

A natural question is to ask whether or not $\mathcal{S}(N)$ is the same order of magnitude as $\mathfrak{S}(N)$, namely whether there exist positive numbers c_1, c_2, c_3 such that

$$c_1 \mathfrak{S}(N) < \mathcal{S}(N) < c_2 \mathfrak{S}(N) \text{ whenever } N > c_3. \tag{1.4}$$

This was answered by E. Fouvry and J. Klüners in [3]. They proved that, for any $\varepsilon > 0$, that (1.4) holds with $c_1 = \alpha - \varepsilon$ and $c_2 = 2/3 + \varepsilon$, where

$$\alpha = \prod_{j=1}^{\infty} (1 + 2^{-j})^{-1}. \tag{1.5}$$

In this paper, we prove the following theorem:

Theorem 1. The function $\mathcal{S}(X)$ satisfies the asymptotic relation

$$\mathcal{S}(N) \sim (1 - \alpha)\mathfrak{S}(N). \tag{1.6}$$

In the paper [7], Stevenhagen made the following conjecture, which is evidently equivalent to Theorem 1:

Conjecture 2. One has that

$$\lim_{X \to \infty} \frac{|\mathcal{S}(X)|}{\mathfrak{S}(X)} = 1 - \alpha.$$

The number $1 - \alpha$ is approximately equal to 58.1%, and is also the probability that a large random symmetric matrix over $\mathbb{F}_2$ is singular. The previous best known result towards this conjecture is in [3] where they show the lim inf is at least α and at most $2/3$.

The principal new ingredient is the breakthrough made by A. Smith [6], on the 2^∞-rank of class groups of imaginary quadratic fields and 2^∞-Selmer rank of elliptic curves. In particular, Smith proved in [6] that for a given elliptic curve $E/\mathbb{Q}$ with full rational 2-torsion, for 100% of quadratic twists E_d of E, the Mordell-Weil rank of E_d is at most one. This, in addition to the so-called parity conjecture for elliptic

curves, confirms a well-known conjecture of Goldfeld for such curves. Recently it has been announced by Smith that such a conclusion also applies, for example, to elliptic curves $E/\mathbb{Q}$ where the splitting field of 2-torsion has Galois group isomorphic to the symmetric group S_3, thereby confirming Goldfeld's conjecture for 100% of elliptic curves.

The remainder of this introduction will be an outline of the argument for the proof of Theorem 1.

1.1. **An outline of Smith's work in** [6]**.** In [6], Smith introduced radical new ideas extending the classical idea of controlling the behaviour of class groups using so-called *governing fields*. Roughly speaking, genus theory gives a canonical basis of the 2-part of the narrow class group of quadratic fields (since we are focusing on narrow class groups, it does not matter if the quadratic field is real or imaginary), and in some cases, namely those primes corresponding to the 4-part or the 8-part of the narrow class group, these primes can be identified by their splitting behaviour in an auxiliary field, called the governing field.

Unfortunately, in this rigid sense governing fields for the 2^k-part of the narrow class group, $k > 3$, conjecturally do not exist. Smith's key idea is that one can still make use of the idea of governing fields, by showing that they exist "on average". Since splitting behaviour of primes in fields can be shown to behave randomly thanks to theorems in analytic number theory such as Chebotarev's density theorem and the large sieve inequality, one can then use this idea to demonstrate that $\mathbb{F}_2$-matrices of Legendre symbols behaves randomly.

Indeed, the strategy of Smith, extending earlier observations due to Reidi, Stevenhagen, and Swinnerton-Dyer [8], is to show that the 2^k-rank of a narrow class group behaves like a Markov chain: that is, the distribution of the 2^k-rank of narrow class groups given the 2^ℓ-ranks for all $\ell < k$ is the same as the distribution given only the 2^{k-1}-rank.

In order to realize this strategy, Smith introduces three key constructs: governing expansions, sets of raw cocycles, and additive-restrictive systems. Roughly speaking, governing expansions consist of collections of functions from $G_\mathbb{Q} = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ to $\mathbb{F}_2$, whose fields of definition give analogues of governing fields. Sets of raw cocycles correspond to elements of $C^1(G_\mathbb{Q}, N[2^k])$ where N is some $G_\mathbb{Q}$-module, which models the behaviour of class groups. Finally, additive-restrictive systems are auxiliary systems that can be attached to governing expansions or sets of raw cocycles, that, when well-defined, simplify the arithmetic complexity of these objects by replacing them with essentially purely combinatorial constructs. After doing this, then our arithmetic statements essentially become probabilistic.

1.2. **Understanding the difference between the class group and narrow class group of real quadratic fields.** In this subsection, we explain how to translate Smith's machinery to deal with the negative Pell equation (1.2). The main obstacle is that Smith's machinery heavily relies on the fact that there is a basis for $\mathrm{Cl}(K)[2]$, with K an imaginary quadratic field, which is given by a set of primes. This is only valid for imaginary quadratic fields, but is no longer true for

real quadratic fields.

Indeed, the difference between the class group $\mathrm{Cl}(K)$ and the *narrow class group* $\mathfrak{cl}(K)$ of a real quadratic field lies at the heart of the matter for us. Indeed, one has that there is a solution to the negative Pell equation for n_0 if and only if the Frobenius at infinity in $\mathfrak{cl}(\mathbb{Q}(\sqrt{n_0}))$ is trivial[1]. This motivates us to look at the structure of the 2-power torsion inside $\mathfrak{cl}(\mathbb{Q}(\sqrt{n_0}))$.

Letting $n_1, \dots, n_m$ be a sequence of nonnegative non-increasing integers. Put

$$S^{n_1,\dots,n_m}(N) = \{D \in \mathcal{S}(N) | \dim 2^{k-1}\mathfrak{cl}(\mathbb{Q}(\sqrt{D}))/2^k\mathfrak{cl}(\mathbb{Q}(\sqrt{D})) = n_k, k = 1, \dots, m\}.$$

Additionally, define

$$S^{n_1,\dots,n_m,-}(N) = \{D \in S^{n_1,\dots,n_m}(N) | \operatorname{Frob}_\infty \in 2^m\mathfrak{cl}(\mathbb{Q}(\sqrt{D}))\}$$

and

$$S^{n_1,\dots,n_m,+}(N) = \{D \in S^{a_2,\dots,a_n}(N) | 0 \neq \operatorname{Frob}_\infty \in 2^{n-1}\mathfrak{cl}(\mathbb{Q}(\sqrt{D}))\}.$$

Further, define

$$S^{n_1,\dots,n_m,\pm}(N) = S^{n_1,\dots,n_m,-}(N) \cup S^{n_1,\dots,n_m,+}(N) = S^{n_1,\dots,n_{m-1},-}(N) \cap S^{n_1,\dots,n_m}(N).$$

The choice of $+$ and $-$ may seem strange but it is meant to align with whether or not there may be a solution to the negative Pell equation.

For positive integers k_1, k_2 and a ring R, put $\mathrm{Mat}_{k_1\times k_2}(R)$ for the set of $k_1 \times k_2$ matrices over R. Then put

(1.7)

$$P(a; k_1, k_2) = \text{Probability that an element } M \in \mathrm{Mat}_{k_1\times k_2}(\mathbb{F}_2) \text{ has } \dim\ker M = a,$$

where each entry of M is sampled with respect to a uniform distribution. Similarly, we denote by $P^{\mathrm{Sym}}(k_1; k_2)$ to denote the probability that a random $k_2 \times k_2$ symmetric matrix, with each entry is an i.i.d Bernoulli random variable, has kernel having rank k_1. Then the main result here is the following:

Theorem 3. Let $n_1, \dots, n_m$ be a nonnegative, non-increasing sequence of integers. Then

$$\lim_{N\to\infty} \frac{|S^{n_1,\dots,n_m,\pm}(N)|}{|S^{n_1,\dots,n_{m-1},-}(N)|} = P(n_m, n_{m-1}, n_{m-1}+1),$$

and

$$\lim_{N\to\infty} \frac{|S^{n_1,\dots,n_m,-}(N)|}{|S^{n_1,\dots,n_m,\pm}(N)|} = 2^{-n_m}.$$

Section 2 explains how theorem 3 implies theorem 1. This argument is essentially due to Stevenhagen in [7]. The remainder of the paper will be dedicated to proving theorem 3. This argument will be very similar to the arguement by Smith in [6] where he proves a similar result for the class groups of imaginary quadratic fields.

[1]The Frobenius at infinty is trivial if and only if the narrow class group is equal to the class group. That happens if and only if the fundamental unit in $\mathbb{Q}(\sqrt{D})$ has norm -1. If D is even, then that is exactly what the negative Pell equation is looking for. If D is odd, then one has that $(\mathbb{Z}[\frac{1+\sqrt{D}}{2}]^\times)/(\mathbb{Z}[\sqrt{D}]^\times)$ injects into $(\mathbb{Z}[\frac{1+\sqrt{D}}{2}]/2)^\times$ which is a group of size 1 or 3, which shows the equivalence of looking in $\mathbb{Z}[\frac{1+\sqrt{D}}{2}]$ or $\mathbb{Z}[\sqrt{D}]$.

Finally, we note that the standard way to index this problem is not by the value of D but rather by the discriminant of the field $\mathbb{Q}(\sqrt{D})$; that is one looks at the set of all quadratic fields of the form $\mathbb{Q}(\sqrt{D})$ whose discriminant is less than X and such that no prime congruent to 3 (mod 4) divides D. However, the strategy of proof will actually prove that we get the correct ratio when we look at the set of all D that are of the form $\{ND'\}$ where N is a fixed integer not divisible by any prime that is 3 (mod 4) and D' is a varying integer that is the product of primes that are all 1 (mod 4). Applying this result to $N = 1$ and $N = 2$, one gets theorem 1 for both that ordering as well as the more standard ordering.

Acknowledgements: We would like to thank Arul Shankar for many insightful conversations and helping us get started on this project, and Alex Smith for answering some questions by the first author as well as developing the ideas in [6] which made this paper possible.

2. The narrow Reidi data and the negative Pell equation

Let $K = \mathbb{Q}(\sqrt{p_1 \cdots p_n})$. Define V to be the free vector space over $\mathbb{F}_2$ with basis vectors $e_1, \ldots, e_n$. There is a natural map from $V \to \mathfrak{cl}(K)$ given by sending $e_i \mapsto [(p_i, \sqrt{p_1 \cdots p_n})]$. This map is two-to-one, and we will call the generator of the kernel of this map r. The map is surjective onto $\mathfrak{cl}(K)[2]$ by genus theory. The element $e_1 + \cdots + e_n$ maps to the Frobenius at infinity. Consequently, one has that the negative Pell equation has a solution if and only if $\mathrm{Cl}(K) = \mathfrak{cl}(K)$ if and only if $r = e_1 + \cdots + e_n$.

Now, define $V_1 = V$, and consider the map $r_1 : V_1 \to \mathfrak{cl}(K)/2\mathfrak{cl}(K)$. Define $V_2 = \ker(r_1)$, and observe that the image of the map from $V_2 \to \mathfrak{cl}(K)/4\mathfrak{cl}(K)$ lands inside $2\mathfrak{cl}(K)/4\mathfrak{cl}(K)$. This map will be denoted r_2, and then this process iterates to give $V_k = \ker(r_{k-1})$ and $r_k : V_k \to 2^{k-1}\mathfrak{cl}(K)/2^k\mathfrak{cl}(K)$ for all n. The data of V_k and r_k for all n will be called the narrow Reidi data. The finiteness of the class group shows that $\cap V_n = \langle r \rangle$, and so one gets the following two criteria surrounding for the negative Pell equation:

Theorem 2.0.1. *One has the following two criteria:*

1. *If $V_k = \langle e_1 + \cdots + e_n \rangle$ for some k, then the negative Pell equation has a solution.*
2. *If $e_1 + \cdots + e_n \notin V_k$ for some k, then the negative Pell equation has no solution.*

Moreover, there exists a value of k for which one of these two must happen.

We also will present an alternative way to view Reidi maps. Write V' for the free vector space over $\mathbb{F}_2$ with basis $e'_1, \ldots, e'_n$. Then $V' \to \mathfrak{cl}^\vee(K)[2]$ by sending e'_i to the function that sends $\mathfrak{p}$ to $\left(\frac{N(\mathfrak{p})}{p_i}\right)$ for $\mathfrak{p}$ unramified in K. This map is two-to-one again, but this time the kernel is just $\langle e'_1 + \cdots + e'_n \rangle$. Now, the map from $\mathfrak{cl}(K)[2] \to \mathfrak{cl}(K)/2\mathfrak{cl}(K)$ is the same data as a pairing $\mathfrak{cl}(K)[2] \times \mathfrak{cl}^\vee(K)[2] \to \mathbb{F}_2$. Defining $V'_1 = V'$ as before, we will write $R_1 : V_1 \times V'_1 \to \mathbb{F}_2$ to be the pullback of the pairing above.

To proceed inductively, write V'_k to be the preimage of $2^{k-1}(\mathfrak{cl}^\vee(K)[2^k]) \subset \mathfrak{cl}^\vee(K)[2]$ in V. One sees that the map $r_k : V_k \to 2^{k-1}\mathfrak{cl}(K)/2^k\mathfrak{cl}(K)$ is again

the same data as a pairing $V_k \times 2^{k-1}(\mathfrak{cl}^\vee(K)[2^k]) \to \mathbb{F}_2$, and as before, we will pull this back to V_k' to get a pairing $R_k : V_k \times V_k' \to \mathbb{F}_2$. The definitions of these spaces imply that the left kernel of R_k is V_{k+1} and the right kernel is V_{k+1}. This set of data (the V_k's, the V_k's, and the R_k's) are the same as the narrow Reidi data, and this data is in some ways easier to control.

Since $\mathfrak{cl}(K)/2\mathfrak{cl}(K)$ corresponds to the extension $K(\sqrt{p_1}, \ldots, \sqrt{p_{n-1}})$, one has that $e_1 + \cdots + e_n \in V_2$. The conjecture of Stevenhagen for the negative Pell equation comes from two predictions that R_1 is a random symmetric pairing (identifying the basis e_i with e_i'; symmetric comes from quadratic reciprocity) over $\mathbb{F}_2$ (which gives a distribution on the dimensions of the various V_2s as K varies), and that r is a random non-zero element of V_2 as K varies. The paper [3] shows that the moments of the size of V_2 are consistent with the first fact. Theorem 3's goal is to show that the matricies R_i for $i > 1$ are random matricies such that $R_i(\cdot, e_1' + \cdots e_n') = 0$ (no longer symmetric as we no longer have anything like quadratic reciprocity) over $\mathbb{F}_2$.

One can then imagine a Markov process where at stage i one is looking at V_i, and the releveant information is the pair $dim(V_i)$ and whether $Frob_\infty = 0$ in V_i; we will denote such a pair by $(d, \pm)$ with $+$ corresponding to $Frob_\infty \neq 0$ in V_i and $-$ corresponding to $Frob_\infty = 0$ in V_i. If $Frob_\infty$ is nonzero in V_i, it is never 0 in any future V_i (indeed it doesn't lie in any future V_i). Under the process where R_i is a random matrix as above, one gets that $Pr((a,-),(b,-)) = P(b, a-1, a) \times \frac{2^b-1}{2^a-1}$, $Pr((a,-),(b,+)) = P(b, a, a-1) \times \frac{2^a-2^b}{2^b-1}$, and the terminal states are $(a,+)$ where the negative Pell equation doesn't have a solution, and $(0,-)$ where it does. However, this Markov process is clearly ultimately choosing a random nonzero element in V_2, and it will end at $(0,-)$ with probability $\frac{1}{2^{dim(V_2)}-1}$, which is exactly Stevenhagen's prediction. Finally, his paper contains the calculations to show that this implies theorem 1.

3. Governing expansions and sets of raw cocycles

A classical governing field for an integer d is a finite extension $L/\mathbb{Q}$ with the property that the splitting behaviour of a prime p in $\mathcal{O}_L$ determines the rank of some piece of the class group of the quadratic field $\mathbb{Q}(\sqrt{dp})$.

In this section, we wish to generalize the concept of a governing field. In fact we will give two different constructions of (a set of) these fields, which do not necessarily coincide. The existence and uniqueness of classical governing fields should then be interpreted as a small-dimensional phenomenon, where there is simply not enough degrees of freedom for differing objects to exist. While in general this is not true, we can give conditions that guarantee these constructions coincide.

We develop the following formalism, in line with [6]. Suppose we have a pairwise disjoint sets of odd primes $X_1, \cdots, X_d$. We write X for their Cartesian product.

To formalize the behaviour where we wish to fix some set of primes while allowing others to vary, we consider, for each subset $S \subset \{1, \cdots, d\}$, the product

$$\overline{X}_S = \left(\prod_{i \in S} X_i \times X_i \right) \times \prod_{j \notin S} X_j.$$

Put ρ_1, ρ_2 for the projection maps from a Cartesian product $T \times T$ to the first and second coordinates, respectively. Then for each such $\overline{x}$, we put

$$K(\overline{x}) = \prod_{i \in S} \mathbb{Q}\left(\sqrt{\rho_1(\pi_i(\overline{x})) \cdot \rho_2(\pi_i(\overline{x}))}\right). \tag{3.1}$$

This field has the right shape as the one that we wish to model.

We shall see eventually that in order to construct governing fields consistently the various sets of primes indexed by $\overline{x}$ need to satisfy some intricate compatibility conditions, which necessitates references to a larger index set.

To construct our governing fields, we actually construct certain functions $\phi : G_{\mathbb{Q}} \to \mathbb{F}_2$, where $G_{\mathbb{Q}} = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, and our governing fields will simply be the fields of definition of these functions.

To exercise the most amount of control and to enable us to actually perform calculations, we would require the set of functions to be relatively simple and well-behaved. We thus start with a rather nice collection of functions that are multiplicative products of characters. Fix $S \subset \{1, \cdots, d\}$ and $\overline{x} \in \overline{X}_S$. Define, for $T \subset S$,

(3.2)

$$\chi_{T,\overline{x}}(\sigma) = \begin{cases} 1 & \text{if } \sigma\left(\sqrt{\rho_1(\pi_i(\overline{x})) \cdot \rho_2(\pi_i(\overline{x}))}\right) = -\sqrt{\rho_1(\pi_i(\overline{x})) \cdot \rho_2(\pi_i(\overline{x}))} \text{ for } i \in T \\ 0 & \text{otherwise.} \end{cases}$$

The function $\chi_{T,\overline{x}}$ has the product decomposition

$$\chi_{T,\overline{x}}(\sigma) = \prod_{i \in T} \chi_{\{i\},\overline{x}}(\sigma), \tag{3.3}$$

which shows that indeed they are a product of characters.

We record a rather trivial observation, which will be used repeatedly later:

Lemma 3.0.1. *Let $\chi_{T,\overline{x}}$ be as in (3.2). Then $\chi_{T,\overline{x}}$ is supported on a finite abelian extension of $\mathbb{Q}$.*

Proof. It is clear from definition that $\chi_{T,\overline{x}}$ is supported on the field

$$L = \prod_{i \in T} \mathbb{Q}\left(\sqrt{\rho_1(\pi_i(\overline{x})) \cdot \rho_2(\pi_i(\overline{x}))}\right),$$

which is abelian over $\mathbb{Q}$ of finite degree since it is the compositum of finitely many quadratic fields. □

Moreover, we see that $\chi_{T,\overline{x}}$ satisfies the equation

$$\chi_{T,\overline{x}}(\sigma\tau) = \prod_{i \in T} \left(\chi_{\{i\},\overline{x}}(\sigma) + \chi_{\{i\},\overline{x}}(\tau)\right) = \sum_{U \subset T} \chi_{U,\overline{x}}(\sigma)\chi_{T-U,\overline{x}}(\tau). \tag{3.4}$$

Moreover, from group cohomology, we have

$$\begin{aligned} d\chi_{T,\overline{x}}(\sigma,\tau) &= \chi_{T,\overline{x}}(\sigma\tau) - \chi_{T,\overline{x}}(\sigma) - \chi_{T,\overline{x}}(\tau) \\ &= \sum_{\emptyset \neq U \subset T} \chi_{U,\overline{x}}(\sigma)\chi_{T-U,\overline{x}}(\tau). \end{aligned} \tag{3.5}$$

(3.4) and (3.5) are properties that we want our functions ϕ to satisfy in general. However, since (3.5) involves the condition that $T \neq \emptyset$, we must take care to address the case when $T = \emptyset$. Indeed, suppose we have a function $\phi_\emptyset : G_\mathbb{Q} \to \mathbb{F}_2$ which satisfies (3.5). The right hand side is then an empty sum, which is necessarily zero. Thus we must have

$$0 = \phi_{\emptyset,\overline{x}}(\sigma\tau) - \phi_{\emptyset,\overline{x}}(\sigma) - \phi_{\emptyset,\overline{x}}(\tau),$$

which implies that $\phi_\emptyset$ is a homomorphism from $G_\mathbb{Q}$ to $\mathbb{F}_2$. Therefore, it makes sense for our formalism to accept $\phi_\emptyset$ as a given homomorphism.

We now make the following definition which, roughly speaking, given a subset $S \subset \{1, \cdots, d\}$ and an element $\overline{x} \in \overline{X}_S$, a collection of functions $\phi_{S'} = \phi_{S',\overline{x},S}$ which satisfies properties (3.4) and (3.5).

Definition 3.0.2. Let $\phi_\emptyset : G_\mathbb{Q} \to \mathbb{F}_2$ be a homomorphism. Let $S \subset \{1, \cdots, d\}$ and $\overline{x} \in \overline{X}_S$ be given. For each subset $S' \subset S$, let $\phi_{S'} = \phi_{S',\overline{x},S}$ be a function from $G_\mathbb{Q}$ to $\mathbb{F}_2$. Suppose that $\phi_{S'}$ satisfies

$$d\phi_{S'}(\sigma,\tau) = \sum_{\emptyset \neq T \subset S'} \chi_{T,\overline{x}}(\sigma) \cdot \phi_{S'-T,\overline{x}}(\tau). \tag{3.6}$$

We then say that $\phi_{S'}$ is a $(S', \overline{x}, S)$-expansion of $\phi_\emptyset$.

We remark that we can drop the dependence on S in the definition above, since $\overline{x}$ necessarily determines S.

The next proposition is very important, in that it demonstrates how one uses class field theory to build expansions of larger subsets of S from smaller ones.

Proposition 3.0.3. *Let $X_1, \cdots, X_d$ be pairwise disjoint sets of odd primes, and let S be a subset of $\{1, \cdots, d\}$. Let $\overline{x} \in \overline{X}_S$ and $\phi_\emptyset$ be a homomorphism from $G_\mathbb{Q}$ to $\mathbb{F}_2$. Suppose that we have $(S - \{i\}, \overline{x})$-expansions of $\phi_\emptyset$ for all $i \in S$. Put M_i for the field of definition of $\phi_{S-\{i\}}$ and*

$$M = K(\overline{x}) \prod_{i \in S} M_i.$$

Suppose that for all $i \in S$ we have

1. *$\rho_1(\pi_i(\overline{x})), \rho_2(\pi_i(\overline{x}))$ split completely in $M_i/\mathbb{Q}$; and*
2. *$\rho_1(\pi_i(\overline{x}))\rho_2(\pi_i(\overline{x}))$ is a square at 2 and at all primes where $M_i/\mathbb{Q}$ is ramified.*

Then $\phi_\emptyset$ has an $(S, \overline{x})$-expansion ϕ_S whose field of definition is unramified above M at all finite places.

Proof. Put

$$\psi = \sum_{\emptyset \neq T \subset S} \chi_{T,\overline{x}}(\sigma) \cdot \phi_{S-T}(\tau).$$

We wish to confirm that the image of ψ in $H^2(G_\mathbb{Q}, \mathbb{F}_2)$ is zero. We identify $\mathbb{F}_2$ with ± 1, and obtain the exact sequence

$$1 \to \{\pm 1\} \to \overline{\mathbb{Q}}^\times \xrightarrow{2} \overline{\mathbb{Q}}^\times \to 1,$$

where we derive the exact sequence

$$0 = H^1\left(G_{\mathbb{Q}}, \overline{\mathbb{Q}}^{\times}\right) \to H^2(G_{\mathbb{Q}}, \mathbb{F}_2) \to H^2\left(G_{\mathbb{Q}}, \overline{\mathbb{Q}}^{\times}\right).$$

The left equality follows from Hilbert's Theorem 90. We know that the map

$$H^2\left(G_{\mathbb{Q}}, \overline{\mathbb{Q}}^{\times}\right) \to \prod_v H^2\left(G_{\mathbb{Q}_v}, \overline{\mathbb{Q}}_v^{\times}\right)$$

is injective, where the product is over all places of $\mathbb{Q}$. Furthermore, the hypotheses of the proposition imply that the invariant $\mathrm{inv}_v(\psi) = 0$ for all places v. Therefore, ψ is the image of a 1-cochain. This cochain corresponds to a $\mathbb{F}_2$ central extension of M.

Write this extension as $M(\sqrt{\alpha})/M$. This extension is Galois over $\mathbb{Q}$, so if $M(\sqrt{\alpha})/\mathbb{Q}$ is ramified at some place p other than 2 or ∞ where $M/\mathbb{Q}$ is unramified, we can clear the ramification by simply multiplying α by p. Now suppose that $M/\mathbb{Q}$ is also ramified at p. We see that the local conditions force ψ to be trivial on $G_{\mathbb{Q}_p}$, so $M_p(\sqrt{\alpha})/\mathbb{Q}_p = (M(\sqrt{\alpha}) \otimes_{\mathbb{Q}} \mathbb{Q}_p)/\mathbb{Q}_p$ has Galois group

$$(\mathbb{Z}/2\mathbb{Z}) \times \mathrm{Gal}(M_p/\mathbb{Q}_p)$$

if $M_p(\sqrt{\alpha}) \neq M_p$. But the inertia group cannot contain $(\mathbb{Z}/2\mathbb{Z})^2$ for $p \neq 2$, so $M_p(\sqrt{\alpha})/M_p$ is unramified. At $p = 2$, we can avoid ramification by multiplying α by ± 2 or -1. □

3.1. **Governing expansions.** In this section, we define *governing expansions*, which give rise to the generalized governing fields alluded to earlier. Here we diverge from the exposition of [6] slightly.

Definition 3.1.1. Let $X_1, \cdots, X_d$ be a pairwise disjoint collection of odd primes, and put X for their product. For $i_a \leq d$ and $\overline{Y}_\emptyset \subset X$, an $\left(i_a, \overline{Y}_\emptyset\right)$*-data set* is a collection of objects:

(1) A subset $\overline{Y}_S \subseteq \overline{X}_S$ for each $S \subseteq \{1, \cdots, d\}$ containing i_a; and
(2) A continuous function $\phi_{\overline{x}} : G_{\mathbb{Q}} \to \mathbb{F}_2$ for each set S containing i_a and $\overline{x} \in \overline{Y}_S$.

For each function $\phi_{\overline{x}}$ as in Definition 3.1.1 we let $M(\overline{x})$ denote the field of definition of $\phi_{\overline{x}}$. These will eventually be our governing fields.

Given an element $\overline{x} \in \overline{X}_S$ and a subset T of S, we write $U = S - T$ and we define

(3.7)
$$\hat{x}(T) = \{\overline{y} \in \overline{X}_T : \pi_i(\overline{y}) \in \pi_i(\overline{x}) \text{ for } i \in U \text{ and } \pi_{\{1,\cdots,d\}-U}(\overline{y}) = \pi_{\{1,\cdots,d\}-U}(\overline{x})\}.$$

We then have:

Definition 3.1.2. Suppose we are given a $(i_a, \overline{Y}_\emptyset)$-data set. Let $\mathfrak{G}(i_a, \overline{Y}_\emptyset)$ be the functions $\phi_{\overline{x}}$ contained therein. We say that $\mathfrak{G}$ is a *pre-governing expansion* if it satisfies the following:

(1) $\phi_{\overline{x}} = \chi_{\{i_a\},\overline{x}}$ whenever $\overline{x} \in \overline{Y}_{\{i_a\}}$.
(2) If S contains i_a and $\overline{x} \in \overline{Y}_S$, then

$$\hat{x}(T) \subset \overline{Y}_T \text{ for } i_a \in T \subset S \text{ or for } T = \emptyset.$$

(3) For any $\overline{x} \in \overline{Y}_S$, we have that

$$d\phi_{\overline{x}}(\sigma, \tau) = \sum_{i_a \notin T \subset S} \chi_{T,\overline{x}}(\sigma) \cdot \phi_{\overline{x}_{S-T}}(\tau)$$

holds for all $\overline{x}_{S-T} \in \hat{x}(S-T)$.

(4) If $\overline{x}_1, \overline{x}_2 \in \overline{X}_S$ satisfy

$$\{\rho_1(\pi_i(\overline{x}_1)), \rho_2(\pi_i(\overline{x}_1))\} = \{\rho_1(\pi_i(\overline{x}_2)), \rho_2(\pi_i(\overline{x}_2))\}$$

and

$$\hat{x}_1(\emptyset) \cup \hat{x}_2(\emptyset) \subseteq \overline{Y}_\emptyset,$$

then $\overline{x}_1 \in \overline{Y}_S$ if and only if $\overline{x}_2 \in \overline{Y}_S$. Moreover, when both are in $\overline{Y}_S$, then $\phi_{\overline{x}_1} = \phi_{\overline{x}_2}$.

One should interpret Definition 3.1.2 as demanding that the functions $\phi_{\overline{x}}$ satisfy certain necessary compatibility conditions in order for the corresponding fields $M(\overline{x})$ to interact appropriately, in a well-defined manner.

While it seems that our choice of functions $\phi_{\overline{x}}$ can be quite arbitrary, it so happens that the conditions we have imposed are quite restrictive. This is revealed by taking iterated commutators. For each S containing i_a and $\overline{x}$ we define an operator $\beta_{|S|,\overline{x}}$ on the space of $|S|$-tuples $(\sigma_1, \cdots, \sigma_k)$, $\sigma_i \in G_{\mathbb{Q}}$ for $i = 1 \cdots, k$, with $|S| = k$, by

$$\beta_{|S|}\phi_{\overline{x}}(\sigma_1, \cdots, \sigma_k) = \phi_{\overline{x}}([\sigma_1, [\sigma_2, [\cdots, [\sigma_{k-1}, \sigma_k] \cdots]]]). \tag{3.8}$$

The operators β_k can be evaluated in a nice way, provided that the functions $\phi_{\overline{x}}$ satisfy (3.6).

Lemma 3.1.3. *Let β_k be given as in (3.8). Suppose that the functions $\phi_{\overline{x}}$ satisfy (3.6). Then*

$$\beta_k\phi_{\overline{x}}(\sigma_1, \cdots, \sigma_k) = \sum_g \prod_{i \leq k} \chi_{g(i),\overline{x}}(\sigma_i),$$

where the sum runs over bijections $g : \{1, \cdots, k\} \to S$ satisfying $g(k) = i_a$ or $g(k-1) = i_a$.

Proof. By the definition of group cohomology, we have for any function $\phi : G_{\mathbb{Q}} \to \mathbb{F}_2$

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

Since the target group is $\mathbb{F}_2$, the sign is immaterial. Therefore

$$\begin{aligned} d\phi_{\overline{x}}([\sigma, \tau], \tau\sigma) &= \phi_{\overline{x}}([\sigma, \tau]\tau\sigma) - \phi_{\overline{x}}([\sigma, \tau]) - \phi_{\overline{x}}(\tau\sigma) \\ &= \phi_{\overline{x}}(\sigma\tau) - \phi_{\overline{x}}([\sigma, \tau]) - \phi_{\overline{x}}(\tau\sigma). \end{aligned}$$

By (3.6), we have by Lemma 3.0.1 that it vanishes on every commutator, whence $d\phi_{\overline{x}}([\sigma, \tau], \tau\sigma) = 0$. It thus follows that

$$\phi_{\overline{x}}([\sigma, \tau]) = \phi_{\overline{x}}(\sigma\tau) + \phi_{\overline{x}}(\tau\sigma).$$

From the fact that our target group is $\mathbb{F}_2$, we see that

$$\begin{aligned} d\phi_{\overline{x}}(\sigma, \tau) + d\phi_{\overline{x}}(\tau, \sigma) &= \phi_{\overline{x}}(\sigma\tau) - \phi_{\overline{x}}(\sigma) - \phi_{\overline{x}}(\tau) + \phi_{\overline{x}}(\tau\sigma) - \phi_{\overline{x}}(\tau) - \phi_{\overline{x}}(\sigma) \\ &= \phi_{\overline{x}}(\sigma\tau) + \phi_{\overline{x}}(\tau\sigma). \end{aligned} \tag{3.9}$$

We now use again the property that the functions $\phi_{\overline{x}} = \phi_{S,\overline{x}}$ satisfy the coboundary condition (3.6). We further recall that $\phi_\emptyset$ is a given homomorphism.

We calculate $\beta_3\phi_{\overline{x}}(\sigma_1,\sigma_2,\sigma_3)$, from which the general phenomenon should be clear. We have

$$\begin{aligned}\beta_3\phi_{\overline{x}}(\sigma_1,\sigma_2,\sigma_3) &= \phi_{\overline{x}}([\sigma_1,[\sigma_2,\sigma_3]])\\ &= \phi_{\overline{x}}(\sigma_1[\sigma_2,\sigma_3]) + \phi_{\overline{x}}([\sigma_2,\sigma_3]\sigma_1)\\ &= d\phi_{\overline{x}}(\sigma_1,[\sigma_2,\sigma_3]) + d\phi_{\overline{x}}([\sigma_2,\sigma_3],\sigma_1).\end{aligned}$$

Observe that the very last term vanishes, since $\chi_{T,\overline{x}}$ is supported on an abelian extension. We then have

$$\begin{aligned}d\phi_{\overline{x}}(\sigma_1,[\sigma_2,\sigma_3]) &= \sum_{\emptyset\neq T\subset S}\chi_{T,\overline{x}}(\sigma_1)\cdot\phi_{S-T,\overline{x}}([\sigma_2,\sigma_3])\\ &= \sum_{\emptyset\neq T\subset S}\chi_{T,\overline{x}}(\sigma_1)\left(d\phi_{S-T,\overline{x}}(\sigma_2,\sigma_3) + d\phi_{S-T,\overline{x}}(\sigma_3,\sigma_2)\right).\end{aligned}$$

Since $|S|=3$, we write $S=\{i_1,i_2,i_3\}$. The decomposition above will then depend on whether $|T|=1$ or $|T|=2$. We treat the first case, say $T=\{i_1\}$. First we note that

$$\begin{aligned}d\phi_{\{i_2,i_3\},\overline{x}}(\sigma_2,\sigma_3) &= \chi_{\{i_2\},\overline{x}}(\sigma_2)\cdot\phi_{\{i_3\},\overline{x}}(\sigma_3) + \chi_{\{i_3\},\overline{x}}(\sigma_2)\cdot\phi_{\{i_2\},\overline{x}}(\sigma_3)\\ &= \chi_{\{i_2\},\overline{x}}(\sigma_2)\chi_{\{i_3\},\overline{x}}(\sigma_3) + \chi_{\{i_3\},\overline{x}}(\sigma_2)\chi_{\{i_2\},\overline{x}}(\sigma_3).\end{aligned}$$

It thus follows that

$$\chi_{\{i_1\},\overline{x}}(\sigma_1)\left(d\phi_{\{i_2,i_3\},\overline{x}}(\sigma_2,\sigma_3) + d\phi_{\{i_2,i_3\},\overline{x}}(\sigma_3,\sigma_2)\right) = \sum_g\prod_{j\leq k}\chi_{\{g(j)\},\overline{x}}(\sigma_j),$$

where the sum runs over all bijections from $\{1,2,3\}$ to $\{i_1,i_2,i_3\}$ which sends either i_2 or i_3 to i_1. The calculations above readily generalize to give

$$\beta_k\phi_{\overline{x}}\left(\sigma_1,\cdots,\sigma_k\right) = \sum_g\prod_{i\leq k}\chi_{g(i),\overline{x}}\left(\sigma_i\right),$$

where g runs over all bijections from $\{1,\cdots,|S|\}$ to S with the property that $g(|S|-1)$ or $g(|S|)$ is equal to i_a, as desired. □

The following definition will be convenient:

Definition 3.1.4. We say $\overline{x}_1,\overline{x}_2,\overline{x}_3\in\overline{Y}_S$ are *3-cyclic* if for all $i\in S$, we have

$$\pi_{S-\{i\}}(\overline{x}_1) = \pi_{S-\{i\}}(\overline{x}_2) = \pi_{S-\{i\}}(\overline{x}_3)$$

and

$$\pi_i(\overline{x}_1) = (p_1,p_2), \pi_i(\overline{x}_2) = (p_2,p_3), \pi_i(\overline{x}_3) = (p_1,p_3).$$

This has the following consequence:

Lemma 3.1.5. *Suppose $\overline{x}_1,\overline{x}_2,\overline{x}_3\in\overline{Y}_S$ are 3-cyclic. Then*

$$\beta_k\phi_{\overline{x}_1} + \beta_k\phi_{\overline{x}_2} = \beta_k\phi_{\overline{x}_3}.$$

To see this, it suffices to understand what happens when $|S|=1$. In this case we need to verify that

$$\chi_{\{i\},\overline{x}_1}(\sigma) + \chi_{\{i\},\overline{x}_2}(\sigma) = \chi_{\{i\},\overline{x}}(\sigma)$$

for all $\sigma \in G_{\mathbb{Q}}$. This follows from considering the cases. For example, suppose that $\sigma(\sqrt{p_1p_2}) = -\sqrt{p_1p_2}$ and $\sigma(\sqrt{p_2p_3}) = -\sqrt{p_2p_3}$. Then the LHS above vanishes, and

$$\sigma(\sqrt{p_1p_3}) = \sigma(p_2\sqrt{p_1p_3}) = \sigma(\sqrt{p_1p_2} \cdot \sqrt{p_2p_3}) = \sqrt{p_1p_3},$$

so the RHS also vanishes.

We make one further definition:

Definition 3.1.6. We say that a pre-governing expansion $\mathfrak{G}$ with initial data i_a and $\overline{Y}_\emptyset$ is *quadratically consistent* if for S containing i_a and for all $\overline{x} \in \overline{X}_S$, we have $\overline{x} \in \overline{Y}_S$ whenever the following are satisfied:

- $\hat{x}(\emptyset) \subseteq \overline{Y}_\emptyset$,
- For all $i \in S - \{i_a\}$ we have $\hat{x}(S - \{i\}) \subseteq \overline{Y}_{S-\{i\}}$,
- For each $i \in S$ and $\overline{x}_i \in \hat{x}(S - \{i\})$, we have $\rho_1(\pi_i(\overline{x}))$ and $\rho_2(\pi_i(\overline{x}))$ split completely in $M(\overline{x})$ and

  $$\rho_1(\pi_i(\overline{x}))\rho_2(\pi_i(\overline{x}))$$

  is a quadratic residue at 2 and at all primes ramifying in $K(\overline{x}_i)/\mathbb{Q}$.

Proposition 3.1.7. *Let $X_1, \cdots, X_d$ be a pairwise disjoint collection of odd primes. For any $i_a \in \{1, \cdots, d\}$ and $\overline{Y}_\emptyset \subseteq X$, there is a set $\mathfrak{G}$ of pre-governing expansions with initial data $i_a, \overline{Y}_\emptyset$ which satisfy:*

1. *For any S containing i_a and $\overline{x}_1, \overline{x}_2, \overline{x}_3 \in \overline{Y}_S$ 3-cyclic, we have $\phi_{\overline{x}_1} + \phi_{\overline{x}_2} = \phi_{\overline{x}_3}$.*
2. *For all $\overline{x} \in \overline{Y}_S$, the extension $M(\overline{x})K(\overline{x})/K(\overline{x})$ is unramified at all finite places.*
3. *$\mathfrak{G}$ is quadratically consistent.*

Proof. For each $S \supseteq \{i_a\}$, take W_S to be the $\mathbb{F}_2$-vector space generated by $\phi_{\overline{x}}, \overline{x} \in \overline{Y}_S$. We note that to confirm the first property it suffices to show that the $\phi_{\overline{x}}$ can be chosen so that $\beta_{|S|}$ is injective on W_S, since by Lemma 3.1.5 we have $\beta_{|S|}\phi_{\overline{x}_1} + \beta_{|S|}\phi_{\overline{x}_2} = \beta_{|S|}\phi_{\overline{x}_3}$. Therefore, if $\beta_{|S|}$ is injective, then this would imply the desired relation.

This is clear for $S = \{i_a\}$ because β_1 is the identity. Now suppose by induction that we have found $\phi_{\overline{y}}$ satisfying this property for all $\overline{y} \in \overline{Y}_T$ and proper subsets T of S which contain i_a, and we wish to prove the result for S. Proposition 3.0.3 shows that we can certainly find expansions $\phi_{\overline{x}}$ for each $\overline{x} \in \overline{Y}_S$. The only question is whether we can make the map from W_S injective.

Take M to be the Hilbert class field of $K(X) = \prod_{\overline{x} \in \overline{X}_{\{1,\cdots,d\}}} K(\overline{x})$. For each prime p that ramifies in $K(X)/\mathbb{Q}$, choose $\mathfrak{P}$ to be a prime of M lying above p, and take $\sigma_{\mathfrak{p}}$ to be the nontrivial inertia element corresponding to $\mathfrak{P}$. By twisting $\phi_{\overline{x}}$ by the quadratic characters $\chi_{\pm p}$ if necessary, we can assume $\phi_{\overline{x}}(\sigma_p) = 0$. Note that by construction 2 does not ramify in $K(X)/\mathbb{Q}$, so p is odd; whence one of p or $-p$ is congruent to 1 (mod 4), whence we can choose the sign so that $\phi_{\overline{x}}$ remains unramified at 2.

We may now assume that $\phi_{\overline{x}}$ vanishes at σ_p for each rational prime p which ramifies in $K(X)/\mathbb{Q}$. We shall see, together with our coboundary conditions, that

this is enough to guarantee that $\beta_{|S|}$ is injective on W_S.

Let $k = |S|$. Suppose that $\beta_k \phi = 0$, where

$$\phi = \sum_j c_j \phi_{\overline{x}_j}$$

for some constants $c_j \in \mathbb{F}_2$. Let τ be the iterated commutator of $\sigma_2, \cdots, \sigma_k$. Then we have

$$0 = \beta_k \phi(\sigma_1, \cdots, \sigma_k) = d\phi(\sigma_1, \tau) + d\phi(\tau, \sigma_1), \tag{3.10}$$

linearity and repeated application of (3.9).

We now suppose that $k > 2$. We then see that $d\phi(\tau, \sigma_1) = 0$, by Lemma 3.0.1. Therefore

$$\beta_k \phi = d\phi(\sigma_1, \tau) = \sum_j \sum_{i \in S - \{i_a\}} c_j \chi_{\{i\}\overline{x}_j}(\sigma_1) \cdot \phi_{\overline{x}_j, S-\{i\}}(\tau).$$

Using the independence of the sets of characters corresponding to each i, we get that

$$\sum_j c_j \chi_{i,\overline{x}_j}(\sigma_1) \cdot \phi_{\overline{x}_j, S-\{i\}}\left([\sigma_2, [\cdots, [\sigma_{k-1}, \sigma_k]\cdots]]\right) = 0$$

for all $i \in S - \{i_a\}$. This can be rewritten as

$$\sum_j c_j \chi_{\{i\},\overline{x}_j}(\sigma_1) \cdot \beta_{k-1} \phi_{\overline{x}_j, S-\{i\}}(\sigma_2, \cdots, \sigma_k) = 0.$$

By the induction hypothesis, we have

$$\sum_j c_j \chi_{i,\overline{x}_j}(\sigma) \cdot \phi_{\overline{x}_j, S-\{i\}}(\tau) = 0$$

for all $\sigma, \tau \in G_{\mathbb{Q}}$. Taking coboundaries then gives

$$\sum_j \sum_{i \in T \subseteq S - \{i_a\}} \chi_{T,\overline{x}_j}(\sigma) \cdot \phi_{\overline{x}_j, S-T}(\tau) = 0.$$

Again by the independence of characters, we see that

$$\sum_j c_j \chi_{T,\overline{x}_j}(\sigma) \cdot \phi_{\overline{x}_j, S-T}(\tau) = 0$$

for any $T \subseteq S - \{i_a\}$. Adding these together then gives that $d\phi = 0$.

When $k = 2$ we have $\tau = \sigma_2$, and thus (3.10) turns int

$$0 = d\phi(\sigma_1, \sigma_2) + d\phi(\sigma_2, \sigma_1).$$

This asserts that $d\phi(\sigma_1, \sigma_2) = d\phi(\sigma_2, \sigma_1)$, since both are valued in $\mathbb{F}_2$. By linearity, it suffices to perform the calculation with a single $\phi_{\overline{x}}$. By our coboundary condition, we have

$$\begin{aligned} d\phi_{\overline{x}}(\sigma_1, \sigma_2) &= \chi_{\{i_1\}}(\sigma_1)\chi_{\{i_2\}}(\sigma_2) + \chi_{\{i_2\}}(\sigma_1)\chi_{\{i_1\}}(\sigma_2) \\ &= 0. \end{aligned}$$

We thus conclude again that $d\phi = 0$.

Therefore, we see that ϕ is a Galois cocycle and hence corresponds to a quadratic extension of $\mathbb{Q}$. But, from the hypothesis that $\phi(\sigma_p) = 0$ for all primes which ramify in $K(X)$, we see that this quadratic extension is necessarily unramified at all finite primes, which means it is the trivial extension and $\phi = 0$. Therefore, $\beta_{|S|}$ is injective on W_S, as claimed. □

Definition 3.1.8. A set $\mathfrak{G}(i_a, \overline{Y}_\emptyset)$ of pre-governing expansions which satisfies the properties of Proposition 3.1.7 is called a set of *governing expansions*.

One final result needed for sets of governing expansions is to show that the fields $M(\overline{x})$, corresponding to the functions $\phi_{\overline{x}}$ in the expansion, are sufficiently distinct.

Proposition 3.1.9. *Let $X_1, \cdots, X_d$ be a pairwise disjoint collection of odd primes, and let X be their Cartesian product. Let $i_a \in S \subseteq \{1, \cdots, d\}$ be a fixed set. Put $M_S(X)$ for the Hilbert class field of $K(X)$, given in Proposition 3.1.7. For $i \in S$, let $T_i = (X_i, Z_i)$ be an ordered tree, where $Z_i \subseteq X_i \times X_i$ is the edge set.*

Suppose we have a set of governing expansions on X such that

$$\pi_S\left(\overline{Y}_S\right) \supseteq Z = \prod_{i \in S} Z_i.$$

For $z \in Z$, put $\overline{x}(z)$ for a preimage of z under π_S. Then for any $z_0 \in Z$ we have that

$$M_S(X) \prod_{z \neq z_0} M(\overline{x}(z))$$

does not contain the field $M(\overline{x}(z_0))$.

Proof. To ease notation, we put $\overline{z} = \overline{x}(z)$ for $z \in Z_i$. It suffices to check that $d\phi_{\overline{z}_1}$ cannot be written as a $\mathbb{F}_2$-linear combination of the other $d\phi_{\overline{z}}$ inside $H^2\left(\mathrm{Gal}(M_S(X)/\mathbb{Q}), \mathbb{F}_2\right)$. Since $\mathrm{Gal}(M_S(X)/\mathbb{Q})$ has nilpotence degree $|S| - 1$, we see that the map

$$\beta(\psi)(\sigma_1, \cdots, \sigma_k) = \psi(\sigma_1, \tau) + \psi(\tau, \sigma_1),$$

where τ is the iterated commutator of $\sigma_2, \cdots, \sigma_k, k = |S|$, is trivial on any 2-coboundary.

It suffices to check that $\beta(d\phi_{\overline{z}_1}) = \beta_k \phi_{\overline{z}_1}$ is not in the span of the other $\beta_k \phi_{\overline{z}}$. Put

$$K_i(X) = \prod_{\overline{x} \in \overline{X}_{\{1, \cdots, d\} - \{i\}}} K(\overline{x}),$$

we define V_i to be the associated $\mathbb{F}_2$-vector space $\mathrm{Gal}(K(X)/K_i(X))$. With this notation, we consider $\beta_k \phi_{\overline{z}}$ restricted to the product

$$V_{i_1} \times \cdots \times V_{i_{k-1}} \times V_{i_a},$$

where $S = \{i_1, \cdots, i_{k-1}, i_a\}$. We then find that we can express $\beta_k \phi_{\overline{z}}$ as a tensor product, namely

$$\beta_k \phi_{\overline{z}} = \chi_{i_1, \overline{z}} \otimes \cdots \otimes \chi_{i_{k-1}, \overline{z}} \otimes \chi_{\{i_a\}, \overline{z}},$$

valid for any $\overline{z} \in Z$. Since we assumed that Z is the edge set of a tree, it follows that there are no 3-cycles, whence the set

$$\{\chi_{i, \overline{z}} : \overline{z} \in Z\}$$

is a linearly independent set; that is, once all duplicate entries are removed, the remaining characters are linearly independent. Since each $\overline{z}$ corresponds to a distinct tuple of characters, the tensor product structure implies that $\phi_{\overline{z}_1}$ is independent of the other $\phi_{\overline{z}}$. This completes the proof. □

3.2. **Raw cocycles.** In this subsection, we describe Smith's notion of *sets of raw cocycles*. Unlike the governing expansions described in the previous subsection, raw cocycles are objects that model the behaviour of *narrow* class groups very precisely. Indeed, by showing that governing expansions can be related to sets of raw cocycles, we obtain control over narrow class groups with sets of number fields, fulfilling the governing field philosophy.

In this subsection we consider a $G_{\mathbb{Q}}$-module N which is isomorphic to some power of $\mathbb{Q}_2/\mathbb{Z}_2$, if the $G_{\mathbb{Q}}$-structure is forgotten. Take $X_1, \cdots, X_d$ to be pairwise disjoint sets of odd primes such that N is not ramified, and let X be their Cartesian product. For $x \in X$ we put $N(x)$ for the quadratic twist of N by the quadratic character of the field

$$\mathbb{Q}\left(\sqrt{\pi_1(x)\cdots\pi_d(x)}\right)/\mathbb{Q}.$$

Note that 2-torsion is preserved: that is, for all $x \in X$ we have $N(x)[2] = N[2]$.

For $x, x' \in X$ put $\Upsilon(x, x')$ for the isomorphism $N(x) \to N(x')$ defined over the smallest possible number field. That is, it is the isomorphism which preserves the Galois structure above

$$K(x, x') = \mathbb{Q}\left(\sqrt{\pi_1(x)\pi_1(x')\cdots\pi_d(x)\pi_d(x')}\right).$$

We denote the associated multiplicative quadratic character by $\chi(x, x')$.

It is well-known from genus theory that 2-torsion of narrow class groups of quadratic fields (real or imaginary) is never trivial; in fact it can easily be seen from the same theory that 2-torsion can be arbitrarily large, namely roughly equal to the number of distinct odd prime divisors of the discriminant. Therefore our set-up should take this into consideration. We therefore assume that our Galois module N has non-trivial 2-torsion.

Definition 3.2.1. Given a $G_{\mathbb{Q}}$-module N, $X = X_1 \times \cdots \times X_d$, and a function

$$\mathrm{rk} : X \to \mathbb{N} \cup \{0\} \cup \{\infty\},$$

we define for each $x \in X$ and $k \leq \mathrm{rk}(x)$ an element

$$\psi_k(x) \in C^1\left(G_{\mathbb{Q}}, N(x)[2^k]\right),$$

where C^1 denotes the set of 1-cocycles of $G_{\mathbb{Q}}$ in $N(x)[2^k]$. The collection of objects $(\mathrm{rk}, \psi_k(x))$ will be called a set of *raw cocycles* on X if whenever $x \in X$ and $k < \mathrm{rk}(x)$, we have

$$2\psi_{k+1}(x) = \psi_k(x).$$

One should interpret the objects in Definition 3.2.1 as basic models for elements of the 2^k-part of a narrow class group. We introduce some further refinements to improve the model.

Definition 3.2.2. Let $\mathfrak{R} = (\mathrm{rk}, \psi_k(x))$ be a set of raw cocycles. For a subset $S \subset \{1, \cdots, d\}$, we say that $\mathfrak{R}$ is consistent over S if

$$\psi_1(x) = \psi_1(x') \text{ whenever } x, x' \in X \text{ satisfy } \pi_{\{1,\cdots,d\}-S}(x) = \pi_{\{1,\cdots,d\}-S}(x').$$

For $i_a \in S$, we say $\mathfrak{R}$ is i_a-consistent (over S) if there is some injection of Galois modules $\iota : \mathbb{F}_2 \to N[2]$ such that

$$\psi_1(x) - \psi_1(x') = \iota \circ \chi_{\pi_{\{i_a\}}(x)\pi_{\{i_a\}}(x')}.$$

Further, we write $\mathrm{rk}(\mathfrak{R})$ and $\psi_k(\mathfrak{R}, x)$ to denote the function and 1-cocycles attached to $\mathfrak{R}$. Similarly, $i_a(\mathfrak{R})$ and $\iota(\mathfrak{R})$ denote the given index and the injection for an i_a-consistent $\mathfrak{R}$.

As the notation suggests, we are interested in the situation when

$$i_a(\mathfrak{R}) = i_a(\mathfrak{G}), \tag{3.11}$$

where $\mathfrak{R}$ is an i_a-consistent set of raw cocycles and $\mathfrak{G}$ a governing expansion with data i_a.

The next definition is intended to capture the phenomenon where the 2^k-part of narrow class groups tend to be trivial for most such groups.

Definition 3.2.3. Let $\mathfrak{R}$ be a set of raw cocycles on X and let S be a non-empty subset of $\{1, \cdots, d\}$ such that for all $\overline{x} \in \overline{X}_S$ we have $\mathrm{rk}(\mathfrak{R})(x) \geq |S|$ for all $x \in \hat{x}(\emptyset)$. Fix $x_0 \in \hat{x}(\emptyset)$. Put

$$\psi(\mathfrak{R}, \overline{x}) = \sum_{x \in \hat{x}(\emptyset)} \Upsilon(x, x_0) \circ \psi_{|S|}(\mathfrak{R}, x).$$

If in addition $\mathfrak{R}$ is consistent over S, we say that $\mathfrak{R}$ is *minimal* at $\overline{x}$ if $\psi(\mathfrak{R}, \overline{x}) = 0$.

Throughout, we take x_0 to be a fixed element of $\hat{x}(\emptyset)$.

Proposition 3.2.4. *Take $\mathfrak{R}$ to be a set of raw cocycles on X. Let $S \subset \{1, \cdots, d\}$ be a non-empty subset and such that $\mathfrak{R}$ is consistent over S and there exists $\overline{x} \in \overline{X}_S$ such that $\psi(\overline{x}) = \psi(\mathfrak{R}, \overline{x})$ is defined.*

Suppose that for all proper subsets $T \subset S$ and $\overline{x}_1 \in \hat{x}(T)$, $\mathfrak{R}$ is minimal at $\overline{x}_1$. Then $\psi(\overline{x})$ maps into $N[2]$ and its coboundary is zero.

In particular, $\psi(\overline{x})$ corresponds to an element of $C^1(G_{\mathbb{Q}}, N[2])$.

Proof. By our minimality assumption for $\overline{x}_1$, we see that for all sets $S_j = S - \{j\}$ and $\overline{x}_1 \in S_j$ we have

$$\psi(\mathfrak{R}, \overline{x}_1) = \sum_{x \in \hat{x_1}(\emptyset)} \Upsilon(x, x_0) \circ \psi_{|S|-1}(\mathfrak{R}, x) = 0.$$

By the definition of $\overline{x}_1, \hat{x_1}(\emptyset)$, and $\psi(\mathfrak{R}, x)$ we see that

$$\begin{aligned} 2\psi(\mathfrak{R}, x) &= \sum_{x \in \hat{x}(\emptyset)} \Upsilon(x, x_0) \circ \psi_{|S|-1}(\mathfrak{R}, x) \\ &= \sum_{1 \leq j \leq d} \sum_{\overline{x}_1 \in \overline{X}_{S_j}} \sum_{x \in \hat{x_1}(\emptyset)} \Upsilon(x, x_0) \circ \psi_{|S|-1}(\mathfrak{R}, x) \\ &= 0, \end{aligned}$$

since the inner sum is always zero by assumption. To show that it is also a coboundary, we calculate

$$d\left(\Upsilon(x,x_0)\circ\psi_{|S|}(x)\right)(\sigma,\tau) = (\sigma\Upsilon(x,x_0) - \Upsilon(x,x_0)\sigma)\circ\psi_{|S|}(x)(\tau)$$
$$= \begin{cases} \sigma\Upsilon(x,x_0)\circ\psi_{|S|-1}(x)(\tau) & \text{if } \chi(x,x_0)=-1 \\ 0 & \text{otherwise.} \end{cases}$$

By linearity, we see that

$$d\psi(\overline{x})(\sigma,\tau) = \sum_{\chi(x,x_0)(\sigma)=-1} \sigma\Upsilon(x,x_0)\circ\psi_{|S|-1}(x)(\tau).$$

For $i\in S$, let H_i be the subset of $x\in\hat{x}(\emptyset)$ such that $\pi_i(x_0)\neq\pi_i(x)$. For a subset $T\subseteq S$, put

$$H_T = \bigcap_{i\in T} H_i.$$

Any element $\sigma\in\mathrm{Gal}(K(\overline{x})/\mathbb{Q})$ can be expressed in the form

$$\sigma = \sum_{i\in T_\sigma}\sigma_i, \tag{3.12}$$

where σ_i is the unique non-trivial element of $\mathrm{Gal}(K(\overline{x})/\mathbb{Q})$ that fixes $\sqrt{\rho_1(\pi_j(\overline{x})\rho_2(\pi_j(\overline{x})}$ for all $j\neq i$ in S, and where T_σ is a subset of S. We claim that, for $x\in\hat{x}(\emptyset)$, the identity

$$\sum_{\substack{\emptyset\neq T\subseteq T_\sigma \\ H_T\ni x}} (-2)^{|T|-1} = \begin{cases} 1 & \text{if } \chi(x,x_0)(\sigma)=-1 \\ 0 & \text{otherwise.} \end{cases} \tag{3.13}$$

For a given $x\in\hat{x}(\emptyset)$, put T_x for the largest subset T of S for which $x\in H_T$. Then the right hand side of (3.13) is equal to one if $|T_x\cap T_\sigma| = m$ is odd. The left hand side of (3.13) is then

$$\sum_{\emptyset\neq T\subseteq T_\sigma\cap T_x} (-2)^{|T|-1} = \sum_{k=1}^{m}\binom{m}{k}(-2)^{k-1} = \left(1-(-1)^m\right)/2$$

by the binomial theorem, and this is readily equal to the right hand side of (3.13). It thus follows that

$$d\psi(\overline{x})(\sigma,\tau) = \sum_{\substack{\emptyset\neq T\subseteq T_\sigma \\ H_T\ni x}} (-2)^{|T|-1}\sigma\Upsilon(x,x_0)\circ\psi_{|S|-1}(x)(\tau) \tag{3.14}$$
$$= \sum_{\emptyset\neq T\subseteq T_\sigma} (-2)^{|T|-1}\sigma\left(\sum_{x\in H_T}\Upsilon(x,x_0)\circ\psi_{|S|-|T|}(x)(\tau)\right),$$

but the inner sum vanishes due to our minimality hypotheses, whence the coboundary vanishes as desired. □

An immediate consequence of Proposition 3.2.4 is that if ψ is minimal at $\overline{x}\in\overline{X}_S$, it is minimal at $\overline{y}$ for all $\overline{y}\in\overline{X}(T)$ and $T\subseteq S$.

Next, we need to establish the language to compare sets of raw cocycles and governing expansions.

Definition 3.2.5. Let $\mathfrak{R}$ be a set of raw cocycles on X and $\mathfrak{G}$ a set of governing expansions on X. Let S be a given subset of $\{1, \cdots, d\}$ and $\overline{x} \in \overline{X}_S$ a fixed element.

If $i_a \in S$ and $\mathfrak{R}$ is $i_a(\mathfrak{G})$-consistent over S, we say that $\mathfrak{R}$ *agrees* with $\mathfrak{G}$ at $\overline{x}$ if $\psi(\mathfrak{R}, \overline{x})$ and $\phi_{\overline{x}}(\mathfrak{G})$ exist

$$\psi(\mathfrak{R}, \overline{x}) - \iota \circ \phi_{\overline{x}}(\mathfrak{G}) = 0.$$

If S does not contain $i_a(\mathfrak{G})$ and if $\mathfrak{R}$ is consistent over S, we say that $\mathfrak{R}$ *agrees* with $\mathfrak{G}$ at $\overline{x}$ if it is minimal at $\overline{x}$.

The next proposition gives conditions for when a set of raw cocycles $\mathfrak{R}$ and a set of governing expansions $\mathfrak{G}$ to agree at $\overline{x}$, provided that they agree on elements of $\hat{x}(T)$ for proper subsets $T \subset S$.

Proposition 3.2.6. *Let $\mathfrak{R}$ be a set of raw cocycles on X and $\mathfrak{G}$ a set of governing expansions on X. Suppose that for some $S \subseteq \{1, \cdots, d\}$, we have $\mathfrak{R}$ is $i_a(\mathfrak{G})$-consistent over S and that there exists $\overline{x} \in \overline{X}_S$ such that $\psi(\mathfrak{R}, \overline{x})$ and $\phi_{\overline{x}}(\mathfrak{G})$ both exist. Then*

$$\psi(\overline{x}) - \iota \circ \phi_{\overline{x}} \in C^1\left(G_{\mathbb{Q}}, N[2]\right)$$

whenever $\mathfrak{R}$ agrees with $\mathfrak{G}$ at $\overline{x}_1$, for all $\overline{x}_1 \in \hat{x}(T)$ and proper subsets $T \subset S$.

Proof. We have shown that the minimality hypotheses imply that $2\psi(\overline{x}) = 0$. Therefore, we only need to check the cocycle condition. We can unpack (3.14) as

$$d\psi(\overline{x})(\sigma, \tau) = \sum_{\emptyset \neq T \subseteq S} \chi_{T,\overline{x}}(\sigma) \cdot \left(\sum_{x \in H_T} \Upsilon(x, x_0) \circ \psi_{|S|-|T|}(x)(\tau) \right).$$

By hypothesis, we have that for any proper subset $T \subset S$ containing i_a and $\overline{x}_1 \in \hat{x}(T)$ that

$$\psi(\mathfrak{R}, \overline{x}_1) - \iota \circ \phi_{\overline{x}_1}(\mathfrak{G}) = 0.$$

Now choose $\overline{x}_1 \in \hat{x}(S - T)$ such that $\pi_i(\overline{x}_1) \neq \pi_i(x_0)$ for $i \in T$. It is then clear from definition that

$$\sum_{x \in H_T} \Upsilon(x, x_0) \circ \psi_{|S|-|T|}(x) = \sum_{x \in \hat{x_1}(\emptyset)} \Upsilon(x, x_0) \circ \psi_{|S|-|T|}(x),$$

and the right hand side is manifestly equal to $\psi(\mathfrak{R}, \overline{x}_1)$. Since for $i_a \notin S - T$ we have that $\psi(\mathfrak{R}, \overline{x}_1)$ is minimal, it thus follows that

$$\sum_{\emptyset \neq T \subseteq S} \chi_{T,\overline{x}}(\sigma) \cdot \left(\sum_{x \in H_T} \Upsilon(x, x_0) \circ \psi_{|S|-|T|}(x)(\tau) \right)$$

$$= \iota \circ \sum_{i_a \notin T \subseteq S} \chi_{T,\overline{x}}(\sigma) \cdot \phi_{\overline{x}_{S-T}}(\tau) = \iota \circ d\phi_{\overline{x}}(\sigma, \tau).$$

By (3.14), it follows that $d(\psi(\overline{x}) - \iota \circ \phi_{\overline{x}}) = 0$, as desired. □

4. Raw cocycles for narrow class groups of real quadratic fields

In this section, we emulate Smith's construction of *raw cocycles for class groups* ([6], Section 2.3), with the necessary change that we discuss *narrow class groups*. Indeed, our key insight is the observation that Smith's arguments essentially translate verbatim to the narrow class group setting for real quadratic fields.

Let $K/\mathbb{Q}$ be a real quadratic field, say $K = \mathbb{Q}(\sqrt{n_0})$ with $n_0 \in \mathbb{N}$ a non-square. Suppose, as before, that $X_1, \cdots, X_d$ are pairwise disjoint sets of odd primes that are unramified in K. We then define

$$K(x) = \mathbb{Q}\left(\sqrt{n_0 \prod_{i \leq d} \pi_i(x)}\right)$$

for $x \in X$. We shall assume, as we must for our application to the negative Pell equation, that for each $i \leq d$ and $p \in X_i$ we have that $p \equiv 1 \pmod 4$. We will take $N(x)$ to be the module $\mathbb{Q}_2/\mathbb{Z}_2$ twisted by the quadratic character corresponding to the extension $K(x)/\mathbb{Q}$.

Let $T_a \subseteq \{1, \cdots, d\}$ and let Δ_a be a square-free integer dividing $2n_0$. Define the character $\psi_1(x) : G_{\mathbb{Q}} \to N[2]$ by

$$\psi_1(x) = \chi_{\Delta_a} + \sum_{i \in T_a} \chi_{\pi_i(x)}. \tag{4.1}$$

We assume that the field of definition of $\psi_1(x)$ is unramified above $K(x)$ for all $x \in X$. We then see that $\psi_1(x)$ corresponds to an element of the dual narrow class group $\mathfrak{cl}^\vee K(x)[2]$.

Proposition 4.0.1. *Take $\psi_1(x)$ as in (4.1), and let $K(x)^{\mathrm{ur}}$ be the maximal extension of $K(x)$ which is unramified at all finite primes . Then for $k > 0$ we have*

$$\psi_1(x)|_{\mathrm{Gal}(\overline{\mathbb{Q}}/K(x))} \in 2^{k-1}\mathfrak{cl}^\vee K(x)[2^k]$$

if and only if, for some

$$\psi_k(x) \in C^1\left(\mathrm{Gal}\left(K(x)^{\mathrm{ur}}/\mathbb{Q}\right), N(x)[2^k]\right),$$

we have

$$\psi_1(x) = 2^{k-1}\psi_k(x).$$

Proof. We see that $\psi_k(x)$, when restricted to $G_{K(x)}$, is in $\mathfrak{cl}^\vee K(x)[2^k]$, whence the sufficiency of finding such a $\psi_k(x)$ is clear. Conversely, given a map

$$\psi_k(x)' \in \mathfrak{cl}^\vee K(x)[2^k],$$

we know that the field of definition L of $\psi_k(x)'$ is dihedral over $\mathbb{Q}$, with its unique order 2^k-cyclic subgroup corresponding to the intermediate field $K(x)$. To continue, we need to extend the character $\psi_k(x)'$ from $\mathrm{Gal}(L/K(x))$ to a cocycle $\psi_k(x)$ on $\mathrm{Gal}(L/\mathbb{Q})$. Choosing some $\tau \in \mathrm{Gal}(L/\mathbb{Q})$ so that we have a coset decomposition

$$\mathrm{Gal}(L/\mathbb{Q}) = \mathrm{Gal}(L/K(x)) + \tau \cdot \mathrm{Gal}(L/K(x)),$$

and choosing some $\alpha \in N(x)$ with $2^{k-1}\alpha = \psi_1(\tau)$, we can find such a $\psi_k(x)$ by setting

$$\psi_k(x)(\sigma) = \psi_k(x)'(\sigma) \text{ and } \psi_k(x)(\tau \cdot \sigma) = \alpha - \psi_k(x)'$$

for all $\sigma \in \mathrm{Gal}(L/K(x))$. Note that for $\sigma_1, \sigma_2 \in G_{K(x)}$ we have

$$\begin{aligned} d\psi_k(x)(\sigma_1, \tau\sigma_2) &= \psi_k(x)(\sigma_1\sigma_2) - \psi_k(x)(\sigma_1) - \psi_k(x)(\tau\sigma_2) \\ &= \psi_k(x)(\tau(\tau^{-1}\sigma_1\tau)\sigma_2) - \psi_k(x)(\sigma_1) - \alpha + \psi_k(x)(\sigma_2) \\ &= \alpha - \psi_k(x)(\tau^{-1}\sigma_1\tau) - \psi_k(x)(\sigma_2) - \psi_k(x)(\sigma_1) - \alpha + \psi_k(x)(\sigma_2) \\ &= -\psi_k(x)(\tau^{-1}\sigma_1\tau) - \psi_k(x)(\sigma_1), \end{aligned}$$

and the last line vanishes because the dihedral group law shows that $\psi_k(x)(\tau^{-1}\sigma_1\tau) = -\psi_k(x)(\sigma_1)$. Next we check

$$\begin{aligned} d\psi_k(x)(\tau\sigma_1, \sigma_2) &= \psi_k(x)(\tau\sigma_1\sigma_2) - \psi_k(x)(\tau\sigma_1) - (\tau\sigma_1) \circ \psi_k(x)(\sigma_2) \\ &= \alpha - \psi_k(x)(\sigma_1\sigma_2) - \alpha + \psi_k(x)(\sigma_1) + \psi_k(x)(\sigma_2) \\ &= -\psi_k(x)(\sigma_1) - \psi_k(x)(\sigma_2) + \psi_k(x)(\sigma_1) + \psi_k(x)(\sigma_2) \\ &= 0. \end{aligned}$$

Finally, we see that

$$\begin{aligned} d\psi_k(x)((\tau\sigma_1), (\tau\sigma_2)) &= \psi_k((\tau\sigma_1)(\tau\sigma_2) - \psi_k(x)(\tau\sigma_1) - \psi_k(x)(\tau\sigma_2) \\ &= \psi_k(x)(\tau^2(\tau^{-1}\sigma_1\tau)\sigma_2) - \alpha + \psi_k(x)(\sigma_1) + \alpha - \psi_k(x)(\sigma_2) \\ &= -\psi_k(x)(\sigma_1) + \psi_k(x)(\sigma_2) + \psi_k(x)(\sigma_1) - \psi_k(x)(\sigma_2) \\ &= 0. \end{aligned}$$

We thus checked that $\psi_k(x)$ obeys the cocycle condition, completing the proof. □

We may therefore define

$$\overline{\mathfrak{cl}}^{\vee} K(x)[2^k] = C^1\left(\mathrm{Gal}(K(x))^{\mathrm{ur}}/\mathbb{Q}), N(x)[2^k]\right). \tag{4.2}$$

Note that we always have

$$\overline{\mathfrak{cl}}^{\vee} K(x)[2^k] \cong \mathfrak{cl}^{\vee} K(x)[2^k] \oplus (\mathbb{Z}/2^k\mathbb{Z}).$$

We denote by $w_a = (T_a, \Delta_a)$ corresponding to the element of $\overline{\mathfrak{cl}}^{\vee} K(x)[2]$, given by the function $\psi_1(x)$ in (4.1). We then make the following definition:

Definition 4.0.2. For $w_a = (T_a, \Delta_a)$ corresponding to the element of $\overline{\mathfrak{cl}}^{\vee} K(x)[2]$ given by the function $\psi_1(x)$. Denote by $\mathfrak{R}(w_a)$ for a set of raw cocycles on X so that, for all $x \in X$, we have

$$\psi_1(\mathfrak{R}, x) = \psi_1(x).$$

Let $\mathrm{rk}(\mathfrak{R})(x)$ be the largest integer k such that $\psi_1(x)$ corresponds to an element of $2^{k-1}\overline{\mathfrak{cl}}^{\vee} K(x)[2^k]$, with

$$\psi_k(\mathfrak{R}|x) \in \overline{\mathfrak{cl}}^{\vee} K(x)[2^k].$$

For a given $w_b = (T_b, \Delta_b)$, where T_b is any subset of $[d]$ and Δ_b is a positive square-free divisor of $2n_0$. For any $x \in X$, define the ideal

$$w_b(x) = \prod_{p|\Delta_b} \mathfrak{P}(p) \cdot \prod_{i \in T_b} \mathfrak{P}(\pi_i(x)), \tag{4.3}$$

where $\mathfrak{P}(p)$ is the unique prime above p in $K(x)$. Taking $\overline{\mathfrak{cl}}K(x)[2]$ to be the set of ideals with square-free norm dividing the discriminant of $K(x)/\mathbb{Q}$, we see that there is a natural surjective map

$$\overline{\mathfrak{cl}}K(x)[2] \to \mathfrak{cl}K(x)[2]$$

which has kernel isomorphic to $\mathbb{Z}/2\mathbb{Z}$. We write $2^{k-1}\overline{\mathfrak{cl}}K(x)[2^k]$ for the preimage of $2^{k-1}\mathfrak{cl}K(x)[2^k]$ under this map.

For a cocycle ψ_k, let $L(\psi_k)$ be the definition of ψ_k over $K(x)$. If $\psi_k(x)$ exists, then the Artin symbol

$$\left[\frac{L(\psi_k)/K(x)}{\mathfrak{P}}\right]$$

lies in the order 2 subgroup of $\mathrm{Gal}(L(\psi_k)/K(x))$ at any $\mathfrak{P}$ dividing the discriminant of $K(x)/\mathbb{Q}$. We identify this subgroup with $\mathbb{Z}/2\mathbb{Z}$.

We note that the Artin symbol induces a *pairing* on the set of w_b's, by simply defining

$$\langle w_a, w_b\rangle = \left[\frac{L(\psi_k^{w_a})/K(x)}{w_b(x)}\right]. \tag{4.4}$$

Theorem 4.0.3. *Let $X = X_1 \times \cdots \times X_d$, with $\{X_i\}$ a pairwise disjoint collection of odd primes. Let n_0 be a fixed positive integer. For $x \in X$, suppose that $w_a = (T_a, \Delta_a)$ correspond to an element of $\overline{\mathfrak{cl}}^{\vee}K(x)[2]$. Let $\mathfrak{G}$ be a set of governing expansions on X and $\mathfrak{R}(w_a)$ a set of raw cocycles. Let $S \subset \{1, \cdots, d\}$ be a subset of cardinality at least three containing $i_a(\mathfrak{G})$.*

Suppose that $w_b = (T_b, \Delta_b)$ is such that $w_b(x) \in 2^{|S|-2}\overline{\mathfrak{cl}}K(x)[2^{|S|-1}]$ for all $x \in \hat{x}(\emptyset)$ and that there exists $i_b \in S, i_b \neq i_a$ such that

$$S \cap T_b \subseteq \{i_b\} \text{ and } S \cap T_a \subseteq \{i_a\}. \tag{4.5}$$

(1) If $S \cap T_a = \emptyset$ or $S \cap T_b = \emptyset$, and for each $i \in S, i \neq i_a, i_b$, $\overline{z}_i \in \hat{x}(S - \{i\})$, and $\overline{y} \in \hat{z}_i(S - \{i, i_a, i_b\})$, $\mathfrak{R}$ is minimal at $\overline{y}$. Then $\psi_{|S|-1}(\mathfrak{R}, x)$ exists for all $x \in \hat{x}(\emptyset)$ and

$$\sum_{x \in \hat{x}(\emptyset)} \left[\frac{L(\psi_{|S|-1}(\mathfrak{R}, x))/K(x)}{w_b(x)}\right] = 0.$$

(2) Suppose that $S \cap T_a = \{i_a\}$ and $S \cap T_b = \{i_b\}$. Suppose that there exists $\overline{z} \in \hat{x}(S - \{i_b\})$ such that $\phi_{\overline{z}}(\mathfrak{G})$ exists. Suppose further that for every $i \in S$ different than i_b, each $\overline{z}_i \in \hat{x}(S - \{i\})$, and $\overline{y} \in \hat{z}_i(S - \{i, i_b\})$, we have $\mathfrak{R}$ agrees with $\mathfrak{G}$ at $\overline{y}$. Then $\psi_{|S|-1}(\mathfrak{R}, x)$ exists for all $x \in \hat{x}(\emptyset)$ and

$$\sum_{x \in \hat{x}(\emptyset)} \left[\frac{L(\psi_{|S|-1}(\mathfrak{R}, x))/K(x)}{w_b(x)}\right] = \phi_{\overline{z}}(\mathfrak{G})\left(\mathrm{Frob}(\rho_1(\pi_{i_b}(\overline{x}))) \cdot \mathrm{Frob}(\rho_2(\pi_{i_b}(\overline{x})))\right).$$

For a function $\phi : G_{\mathbb{Q}} \to \mathbb{F}_2$ supported on a finite Galois extension $L/\mathbb{Q}$ and a rational prime p which is not inert in L, put

$$\mathrm{inv}_p(\phi) = \mathrm{inv}_p(L),$$

where

$$\mathrm{inv}_p(L) = \begin{cases} 0 & \text{if } p \text{ splits over } L \\ 1/2 & \text{if } p \text{ ramifies over } L. \end{cases}$$

Proof. Let $\overline{z} \in \hat{x}(S - \{i_a, i_b\})$. For $x \in \hat{z}(\emptyset)$, our choice of w_b implies that the Artin symbol

$$\left[\frac{L(\psi_{|S|-1}(\mathfrak{R}, x))/K(x)}{w_b(x)}\right]$$

depends only on w_a, w_b and x, but not on the choice of raw cocycles $\mathfrak{R}$.

Write $b(x)$ for the norm of the ideal $w_b(x)$ for $x \in \hat{z}(\emptyset)$. Our restrictions on T_b implies that $b(x)$ does not depend on x; we shall simply denote it by b. Let p be a prime divisor of b, and let $\mathfrak{P}$ be the prime above p in $K(x)$. Put $\Delta(x)$ for the discriminant of $K(x)/\mathbb{Q}$. If $\psi_{|S|-1}(\mathfrak{R}, x)$ exists, then we can write it locally at p as χ or $\chi + \chi_{\Delta(x)}$, where χ is an unramified character. Therefore

$$\left[\frac{L(\psi_{|S|-1}(\mathfrak{R}, x))/K(x)}{\mathfrak{P}}\right] = \mathrm{inv}_p(\chi \cup \chi_b).$$

Further we see that $\mathrm{inv}_p(\chi_{\Delta(x)} \cup \chi_b) = 0$ from our assumptions on w_b, whence

$$\left[\frac{L(\psi_{|S|-1}(\mathfrak{R}, x))/K(x)}{\mathfrak{P}}\right] = \mathrm{inv}_p\left(\psi_{|S|-1}(x) \cup \chi_b\right). \tag{4.6}$$

Take x_0 to be the element of $\hat{z}(\emptyset)$ outside of all $\hat{z}_i(\emptyset)$, and write $\overline{y}_i$ for the element in $\hat{z}(S - \{i, i_a, i_b\}) \cap \hat{z}_i(S - \{i, i_a, i_b\})$.

To prove the first assertion, consider

$$\psi = -\sum_{x \in \hat{z}(\emptyset) - \{x_0\}} \Upsilon(x, x_0) \circ \psi_{|S|-1}(\mathfrak{R}, x).$$

By Proposition 3.2.4, we see that ψ is a cocycle mapping to $N(x_0)$, and we then see that

$$2^{|S|-2}\psi = \psi_1(x_0).$$

From our minimality assumption we find that

$$2\psi = -\sum_{x \in \hat{z}(\emptyset) - \hat{y}_i(\emptyset) - \{x_0\}} \Upsilon(x, x_0) \circ \psi_{|S|-2}(\mathfrak{R}, x)$$

for each $i \in S - \{i_a, i_b\}$. We see that the field of definition of 2ψ is unramified at each $\pi_i(\overline{z}_i)$ for $i \in S - \{i_a, i_b\}$, whence 2ψ must have field of definition unramified above $K(x_0)$, so some quadratic twist of ψ is unramified above $K(x_0)$. This shows that $\psi_{|S|-1}(\mathfrak{R}, x_0)$ exists, and via (4.6), we have

$$\sum_{x \in \hat{z}(\emptyset)} \left[\frac{L(\psi_{|S|-1}(\mathfrak{R}, x)/K(x)}{w_b(x)}\right] = \sum_{p|b} \mathrm{inv}_p\left(\psi(\overline{z}) \cup \chi_b\right).$$

The assumption on w_b implies that the choice of $\psi_{|S|-1}(\mathfrak{R}, x_0)$ does not affect the value of this sum, so we can take $\psi(\overline{z})$ to be a quadratic character. By Hilbert reciprocity, this equals

$$\sum_{p|b} \mathrm{inv}_p\left(\psi(\overline{z}) \cup \chi_b\right).$$

However, χ_b is locally trivial at all primes ramifying in any $K(x)$ that do not divide b, so this sum is zero. This completes the proof of the first part of the theorem.

For the second part, we instead consider

$$\psi = \iota \circ \phi_{\overline{z}} - \sum_{x \in \hat{z}(\emptyset) - \{x_0\}} \Upsilon(x, x_0) \circ \psi_{|S|-1}(\mathfrak{R}, x).$$

By Proposition 3.2.6 we see that this is a cocycle mapping to $N(x_0)$, and we again find that $2^{|S|-2}\psi = \psi_1(x_0)$. Furthermore, we have

$$2\psi = \iota \circ \phi_{\overline{y}_i} - \sum_{x \in \hat{z}(\emptyset) - \hat{y}_i(\emptyset) - \{x_0\}} \Upsilon(x, x_0) \circ \psi_{|S|-2}(\mathfrak{R}, x)$$

for each $i \in S - \{i_a, i_b\}$, where we are taking $\phi_{\overline{y}_{i_a}} = 0$. Then then field of definition of 2ψ must be unramified above $K(x_0)$. Thus $\psi_{|S|-1}(\mathfrak{R}, x_0)$ exists and can be taken to be a quadratic twist of ψ. Following the same argument as in the proof of the first part, we can ignore this quadratic twist, and we find

$$\sum_{x \in \hat{z}(\emptyset)} \left[\frac{L(\psi_{|S|-1}(\mathfrak{R}, x))/K(x)}{w_b(x)} \right] = \sum_{p|b} \mathrm{inv}_p \left(\phi_{\overline{z}} \cup \chi_b \right).$$

Applying the same argument to all $\overline{z} \in \hat{x}(S - \{i_b\})$, we find

$$\sum_{x \in \hat{x}(\emptyset)} \left[\frac{L(\psi_{|S|-1}(\mathfrak{R}, x))/K(x)}{w_b(x)} \right] = \mathrm{inv}_{p_{1b}} \left(\phi_{\overline{z}} + \cup \chi_{p_{1b}} \right) + \mathrm{inv}_{p_{2b}} \left(\phi_{\overline{z}} \cup \chi_{p_{2b}} \right),$$

where $p_{ib} = \rho_i(\pi_{i_b}(\overline{x}))$ for $i = 1, 2$. This is equivalent to the desired conclusion. □

5. Additive restrictive systems

We now introduce the notion of an additive-restrictive system, a construction that allows us to gain additional control over sets of governing expansions and raw cocycles.

Definition 5.0.1. Let $X_1, \cdots, X_d$ be pairwise disjoint sets of odd primes. An *additive restrictive system* on this collection is a sequence of tuples

$$\left((\overline{Y}_S, F_S, A_S) : S \subseteq \{1, \cdots, d\} \right)$$

where for each S, A_S is an abelian group and $F_S : \overline{Y}_S \to A_S$ a function, such that:

- If $S \neq \emptyset$, then

  $$\overline{Y}_S = \{\overline{x} \in \overline{X}_S : \hat{x}(T) \subset \ker F_T \text{ for all } T \text{ a proper subset of } S\}.$$

- Whenever $\overline{x}_1, \overline{x}_2, \overline{x}_3 \in \overline{Y}_S$ is a 3-cycle, we have

  $$F_S(\overline{x}_1) + F_S(\overline{x}_2) = F_S(\overline{x}_3).$$

We will denote by $\overline{Y}_S^\circ = \ker F_S$, and the letter $\mathfrak{A}$ to denote an additive-restrictive system. We shall write $\overline{Y}_S(\mathfrak{A}), F_S(\mathfrak{A})$, etc. to denote the associated data of $\mathfrak{A}$.

The key property of additive restrictive systems is that we can bound how quickly the sets $\overline{Y}_S^\circ$ shrink as S increases through additive combinatorics.

Proposition 5.0.2. *Suppose $X = X_1 \times \cdots \times X_d$ is a product of finite sets of odd primes, and suppose*

$$\left((\overline{Y}_S, F_S, A_S) : S \subseteq \{1, \cdots, d\} \right)$$

is an additive restrictive system on X. Write δ for the density $|\overline{Y}_\emptyset^\circ|/|X|$ and write $\mathcal{A}$ for the maximum size of a group A_S in the sequence. Then for any $S \subseteq \{1, \cdots, d\}$, the density of $\overline{Y}_S^\circ$ in $\overline{X}_S$ is at least

$$\delta^{2^{|S|}} \mathcal{A}^{-3^{|S|}}.$$

Proof. Write δ_S for the density of $\overline{Y}_S^\circ$ in $\overline{X}_S$. Our goal is to compare the density of $\overline{Y}_S^\circ$ in $\overline{X}_S$ with that of $\overline{Y}_{S-\{s\}}^\circ$ in $\overline{X}_{S-\{s\}}$, say. To that end, we put, for $s \in S$ and $\overline{x}_0 \in \overline{X}_S$,

$$m(\overline{x}_0) = \pi^{-1}_{\{1,\cdots,d\}-\{s\}} \left(\pi_{\{1,\cdots,d\}-\{s\}}(\overline{x}_0)\right),$$
$$V_{\overline{x}_0} = \overline{Y}^\circ_{S-\{s\}} \cap m(\overline{x}_0),$$

and

$$W_{\overline{x}_0} = \overline{Y}^\circ_S \cap m(\overline{x}_0).$$

We see that W naturally injects into $V \times V$. Note that two elements in $\overline{Y}_{S-\{s\}} \times \overline{Y}_{S-\{s\}}$ which differ only at the s-th coordinate can be glued together to give an element of $\overline{X}_S$. By definition, elements of V only differ at the s-th coordinate. It thus follows that there is a map from $V \times V$ into $\overline{X}_S \cap m(\overline{x}_0)$, say $\nu_{\overline{x}_0}$.

We now define the relation $\overline{x}_1 \sim \overline{x}_2$ if and only if $\nu_{\overline{x}_0}(\overline{x}_1, \overline{x}_2) \in W_{\overline{x}_0}$. We claim that $\sim$ is an equivalence relation. First, by using additivity, we see that for an element $\overline{x} \in \overline{Y}_S$ with the property that $\pi_{\{s\}}(\overline{x}) = (p, p)$, then $(\overline{x}, \overline{x}, \overline{x})$ forms a 3-cycle, whence

$$F_S(\overline{x}) + F_S(\overline{x}) = F_S(\overline{x}),$$

which implies that $F_S(\overline{x}) = 0$. This shows that $\sim$ is reflexive. Symmetry follows similarly, since if $\pi_{\{s\}}(\overline{x}_1) = (p_1, p_2)$ and $\pi_{\{s\}}(\overline{x}_2) = (p_2, p_1)$. Now put $\overline{x} \in \overline{Y}_S \cap m(\overline{x}_0)$ be such that $\pi_{\{s\}}(\overline{x}) = (p_2, p_2)$. Then $(\overline{x}_1, \overline{x}, \overline{x}_2)$ is a 3-cycle, whence

$$F_S(\overline{x}_1) + F_S(\overline{x}) = F_S(\overline{x}_2)$$
$$F_S(\overline{x}_1) + 0 = F_S(\overline{x}_2),$$

which shows that $\sim$ is symmetric. Finally, for any $\overline{x}_1, \overline{x}_2, \overline{x}_3 \in V$ we have

$$\overline{y}_1 = \nu_{\overline{x}_0}(\overline{x}_1, \overline{x}_2), \overline{y}_2 = \nu_{\overline{x}_0}(\overline{x}_2, \overline{x}_3), \overline{y}_3 = \nu_{\overline{x}_0}(\overline{x}_3, \overline{x}_1)$$

forms a 3-cycle. If $\overline{x}_1 \sim \overline{x}_2$ and $\overline{x}_2 \sim \overline{x}_3$, then additivity implies that

$$F_S(\overline{y}_1) + F_S(\overline{y}_2) = F_S(\overline{y}_3).$$

Our hypothesis implies that the left hand side is zero, hence $F_S(\overline{y}_3)$, whence $\overline{y}_3 \in W_{\overline{x}_0}$ and thus $\overline{x}_1 \sim \overline{x}_3$. This confirms transitivity and thus shows that $\sim$ is an equivalence relation on $V \times V$.

Recall that all coordinates of an element $\overline{x}_1 \in V$ are fixed (and equal to that of $\overline{x}_0$) other than the s-th coordinate. Therefore, an equivalence class is determined by assigning coordinates corresponding to $T \subset S$, with T containing s, to $0 \in A_S$ and then assigning the remaining coordinates to arbitrary elements in A_S. It follows that the number of such equivalence classes satisfy

$$\prod_{s \in T \subseteq S} |A_S|^{2^{|S|-|T|}} \leq \prod_{i=0}^{|S|-1} \mathcal{A}^{\binom{|S|-1}{i} 2^i} = \mathcal{A}^{3^{|S|-1}}. \tag{5.1}$$

Write $\delta_{\overline{x}_0}$ for the density of V in $\overline{X}_{S-\{s\}} \cap m(\overline{x}_0)$. Then the density of $V \times V$ in $\overline{X}_S \cap m(\overline{x}_0)$ is $\delta_{\overline{x}_0}^2$, and by (5.1), the density of W inside $\overline{X}_S \cap m(\overline{x}_0)$ satisfies

$$\begin{aligned}\frac{|W|}{|\overline{X}_S \cap m(\overline{x}_0)|} &= \frac{|W|}{|V \times V|} \cdot \frac{|V \times V|}{\overline{X}_S \cap m(\overline{x}_0)} \\ &\geq \mathcal{A}^{-3^{|S|-1}} \delta_{\overline{x}_0}^2.\end{aligned}$$

The average of the $\delta_{\overline{x}_0}$ is $\delta_{S-\{s\}}$, and $\overline{Y}_S^{\circ}$ is given by the union of $W_{\overline{x}_0}$ over all $\overline{x}_0$, whence

$$\delta_S \geq \mathcal{A}^{-3^{|S|-1}} \cdot \delta_{S-\{s\}}^2.$$

Applying this repeatedly we see that

$$\delta_S \geq \mathcal{A}^{-3^{|S|-1}(1+\frac{2}{3}+\frac{4}{9}+\cdots)} \cdot \delta_{\emptyset}^{2^{|S|}} = \delta^{2^{|S|}} \mathcal{A}^{-3^{|S|}}.$$

□

5.1. **Additive restrictive systems for governing expansions.** We now construct additive restrictive systems attached to governing expansions.

Proposition 5.1.1. *Let $\mathfrak{G}$ be a set of governing expansions on $X = X_1 \times \cdots \times X_d$, and let $S_{\max}$ be a subset of $\{1, \cdots, d\}$ which contains $i_a = i_a(\mathfrak{G})$. Then there is an additive-restrictive system $\mathfrak{A}$ on X satisfying the property that for all $i_a \in S \subseteq S_{\max}$, we have*

$$\overline{Y}_S(\mathfrak{A}) = \overline{Y}_S(\mathfrak{G}).$$

Furthermore, for all $S \subseteq \{1, \cdots, d\}$, the abelian group A_S attached to S in $\mathfrak{A}$ satisfies

$$|A_S(\mathfrak{A})| \leq 2^{|S_{\max}|+1}.$$

Proof. Given $\mathfrak{G}$, we will construct abelian groups $A_S(\mathfrak{A})$ and functions $F_S(\mathfrak{A}) : \overline{Y}_S(\mathfrak{G}) \to A_S(\mathfrak{A})$ which satisfies the conditions of the Proposition for all $S \subseteq S_{\max}$. We start with the case when $S = \{j\}$ is a singleton. We then take $F_S(\overline{x}) = 0$ if and only if

$$\rho_1(\pi_j(\overline{x}))\rho_2(\pi_j(\overline{x}))$$

is a quadratic residue at 2 and at all primes in $\pi_{S_{\max}-\{j\}}(\overline{x})$. First observe that $(\mathbb{Z}/8\mathbb{Z})^*/\{\square\} \cong (\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z})$ and for any odd prime p we have $(\mathbb{Z}/p\mathbb{Z})^*/\{\square\} \cong \mathbb{Z}/2\mathbb{Z}$. Therefore, the target group for F_S can be chosen to be $A_S = (\mathbb{Z}/2\mathbb{Z})^{|S_{\max}|+1}$.

Suppose now that $|S| > 1$ and $i_a \in S$. Then we want to choose F_S so that $F_S(\overline{x}) = 0$ if and only if $\phi_{\overline{x}}(\mathfrak{G})$ is trivial at $\pi_i(\overline{x})$ for all $i \in S_{\max} - S$. By definition, $\phi_{\overline{x}}$ is an unramified quadratic character at each such prime, so for each $i \in S_{\max} - S$, $\phi_{\overline{x}}$ takes on one of two possible values at $\pi_i(\overline{x})$. Therefore we may choose the target group to be

$$A_S = (\mathbb{Z}/2\mathbb{Z})^{|S_{\max}-S|}.$$

If S is not a subset of $S_{\max}$, or if S does not contain i_a and $|S| > 1$, then we let A_S be the trivial group.

We now claim that this defines an additive-restrictive system $\mathfrak{A}$ which satisfies $\overline{Y}_S(\mathfrak{A}) = \overline{Y}_S(\mathfrak{G})$ for all $S \subseteq S_{\max}$. We first check that

$$\mathfrak{A} = \{(\overline{Y}_S(\mathfrak{G}), F_S, A_S) : S \subseteq \{1, \cdots, d\}\}$$

does in fact define an additive-restrictive system.

First consider the case when $S = \{j\}$. We check that both of the conditions for an additive-restrictive system are satisfied in this case. Note that $F_\emptyset$ maps into the trivial group, and hence is identically zero, whence $\overline{Y}_\emptyset^\circ = Y_\emptyset$. Therefore, condition (2) in Definition 3.1.2 implies the first condition of additive-restrictive systems. Next, we must check additivity. Suppose $\overline{x}_1, \overline{x}_2, \overline{x}_3 \in \overline{Y}_S$ is a 3-cycle. The second condition follows from the fact that $(p_1p_2)(p_2p_3)$ and p_1p_3 have the same residue modulo squares for every prime in $\pi_{S_{\max}-\{j\}}(\overline{x})$.

Now suppose that $i_a \in S$ and $|S| > 1$. We must check that

$$\overline{Y}_S(\mathfrak{G}) = \{\overline{x} \in \overline{X}_S : \hat{x}(T) \subseteq \overline{Y}_T^\circ \text{ for all } T \subsetneq S\}.$$

Observe that $\overline{Y}_T^\circ = \overline{Y}_T$ whenever $i_a \notin T$ and $|T| > 1$, so this condition is vacuous by the definition of a governing expansion in these cases. For the singletons $T = \{j\}$, note that $\overline{Y}_T(\mathfrak{G})$ is not defined, so again the condition is vacuous. Therefore, it remains only to check the condition for proper subsets $T \subset S$ which contain i_a.

We demonstrate this in the case when $|S| = 2$. Suppose that $\overline{x} \in \overline{Y}_S$. By the definition of a governing expansion, it follows that $\hat{x}(\{i_a\}) \subseteq \overline{Y}_{\{i_a\}}$. Then the first condition follows from the fact that $\mathfrak{G}$ is square residue compatible (Definition 3.1.6). Additivity follows similarly. □

5.2. **Additive restrictive systems for raw cocycles.** It takes considerably more effort to attach an additive-restrictive system to a set of raw cocycles. Put

$$\mathcal{D}_{(2)}^\vee = \{w_a = (T_a, \Delta_a) : \exists x_0 \in X \text{ s.t. } \psi_1(x_0) \in 2\overline{\mathfrak{cl}}^\vee K(x_0)[4]\} \tag{5.2}$$

and

$$\mathcal{D}_{(2)} = \{w_b = (T_b, \Delta_b) : \exists x_0 \in X \text{ s.t. } w_b(x_0) \in 2\overline{\mathfrak{cl}} K(x_0)[4]\}. \tag{5.3}$$

For each $w \in \mathcal{D}_{(2)}^\vee$, attach a set of raw cocycles $\mathfrak{R}(w)$ as in Definition 4.0.2.

We then make the definitions:

Definition 5.2.1. Let $X_1, \cdots, X_d$ be pairwise disjoint sets of odd primes, and put X for their Cartesian product. Let $i_a \in \{1, \cdots, d\}$ and $S \subseteq \{1, \cdots, d\}$, and $\mathfrak{G}(i_a)$ a set of governing expansions. For $w \in \mathcal{D}_{(2)}^\vee$, let $\mathfrak{R}(w)$ be a set of raw cocycles. Suppose that $\overline{x} \in \overline{X}_S$ satisfies

$$\mathrm{rk}(\mathfrak{R}(w))(x) \geq |S| + 1 \text{ for all } x \in \hat{x}(\emptyset).$$

If $\mathfrak{R}(w)$ is consistent over S, we say $\mathfrak{R}(w)$ is *acceptably ramified at* $(\overline{x}, i)$, where $i \in \{1, \cdots, d\} - S$, if

$$\sum_{x \in \hat{x}(\emptyset)} \Upsilon(x, x_0) \circ \psi_{|S|+1}\left(\mathfrak{R}(w), x\right)(\sigma_{\pi_i(\overline{x})}) = 0.$$

If $i_a \in S$ and $\mathfrak{R}(w)$ is i_a-consistent over S, then we say $\mathfrak{R}(w)$ is *acceptably ramified at* $(\overline{x}, i)$ for $i \in \{1, \cdots, d\} - S$ if there exists $\overline{z} \in \overline{X}_{S \cup \{i\}}$ satisfying $\overline{x} \in \overline{z}(S)$, $\phi_{\overline{z}}(\mathfrak{G}(i_a))$ is defined and satisfies $\rho_1(\pi_{i_a}(\overline{z})) \neq \rho_2(\pi_{i_a}(\overline{z}))$, and

$$\sum_{x \in \hat{x}(\emptyset)} \Upsilon(x, x_0) \circ \psi_{|S|+1}(\mathfrak{R}(w), x)(\sigma_{\pi_i(\overline{x})}) = \phi_{\overline{z}}(\sigma_{\pi_i(\overline{x})}).$$

For $j_1, j_2 \leq n_0$, we define an additive-restrictive system $\mathfrak{A}(j_1, j_2)$, attached to both a set of governing expansions and a set of raw cocycles, as follows. First choose a positive integer $m \geq 2$, and choose a set $S(j_1, j_2) \subseteq \{1, \cdots, d\}$ having cardinality $m+1$. For $S \subseteq \{1, \cdots, d\}$, let A_S be the trivial group if S is not a subset of $S(j_1, j_2)$ or if $|S| > |S(j_1, j_2)| - 2$. In particular, we have $\ker F_S = \overline{Y}_S$ in such cases.

We attach a set of governing expansions to $\mathfrak{A}(j_1, j_2)$ as follows. First choose an index $i_a(j_1, j_2)$ such that $i_a \in S(j_1, j_2)$. Let $\mathfrak{G}(i_a)$ be a set of governing expansions such that for every S with $|S| = |S(j_1, j_2)| - 2$ and $i_a \in S$, and for every $\overline{x} \in \overline{X}_S$, the expansion

$$\phi_{\overline{x}}(\mathfrak{G}(i_a(j_1, j_2))$$

exists. Further, we assume that it is trivial at 2 and ∞, as well as at all primes in $\pi_{\{1,\cdots,d\}-S}(\overline{x})$, and at all primes dividing n_0.

Choose filtrations of $\mathbb{F}_2$-vector spaces

$$\mathcal{D}^{\vee}_{(2)} \supseteq \mathcal{D}^{\vee}_{(3)} \supseteq \cdots \supseteq \mathcal{D}^{\vee}_{(m)}$$

and

$$\mathcal{D}_{(2)} \supseteq \mathcal{D}_{(3)} \supseteq \cdots \supseteq \mathcal{D}_{(m)},$$

where $\mathcal{D}^{\vee}_{(m)}$ contains the non-trivial element of the kernel of the map

$$\mathcal{D}^{\vee}_{(2)} \to \mathfrak{cl}^{\vee}(K(x_0))[4] \tag{5.4}$$

respectively. Choosing a pairing $\mathrm{Art}_{(k)}$ on $\mathcal{D}^{\vee}_{(k)} \times \mathcal{D}_{(k)}$ with the property that for each k, the left kernel of $\mathrm{Art}_{(k)}$ is $\mathcal{D}^{\vee}_{(k+1)}$ and the right kernel is $\mathcal{D}_{(k+1)}$.

For each $2 \leq k \leq m$, put $n_k = \dim \mathcal{D}^{\vee}_{(k)}$. Further, we choose bases

$$w_{a,1}, \cdots, w_{a,n_2} \in \mathcal{D}^{\vee}_{(2)} \tag{5.5}$$

and

$$w_{b,1}, \cdots, w_{b,n_2} \in \mathcal{D}_{(2)} \tag{5.6}$$

so that in each case, the first n_k vectors are a basis for $\mathcal{D}^{\vee}_{(2)}, \mathcal{D}_{(2)}$ respectively for $2 \leq k \leq m$. Moreover, we shall assume that there is an index $i_b \in S(j_1, j_2)$ such that $T(w_{a,i}), T(w_{b,j})$ do not contain i_b for $i, j = 1, \cdots, n_2$.

Now for each $j \leq n_2$, we attach a set of raw cocycles $\mathfrak{R}(w_{a,j})$ in accordance to Definition 4.0.2. We have the following lemma:

Lemma 5.2.2. *Let $S(j_1, j_2) \subseteq \{1, \cdots, d\}$ be such that $|S(j_1, j_2)| = m+1$. Then for each $S \subseteq \{1, \cdots, d\}$, one can choose sets $\overline{Y}_S \subseteq \overline{X}_S$, abelian groups A_S, and functions $F_S : \overline{Y}_S \to A_S$ which satisfy:*

(1) $A_S = \{1\}$ if $S \not\subset S(j_1, j_2)$ or $|S| > |S(j_1, j_2)| - 2$;
(2) $A_S = (\mathbb{Z}/2\mathbb{Z})^{n_2(n_2+m=2)}$ if $S \subset S(j_1, j_2)$ and $|S| \leq |S(j_1, j_2) - 2$, and for each $\overline{x} \in \ker F_S$, we have $\mathfrak{R}(w_{a,j})$ is minimal at $\overline{x}$ if $j \neq j_1$, $\mathfrak{R}(w_{a,j})$ agrees with $\mathfrak{G}(i_a(j_1, j_2))$ at $\overline{x}$, and $\mathfrak{R}(w_{a,j})$ is acceptably ramified at $(\overline{x}, i)$ for all $i \in S(j_1, j_2) - S$.

Proof. Suppose $\overline{x} \in \overline{Y}_S(j_1, j_2)$ for some subset $S \subset S(j_1, j_2)$ having cardinality at most $|S(j_1, j_2)| - 2$. Then Propositions 3.2.4 and 3.2.6 imply that $\psi(\mathfrak{R}(w), \overline{x})$ or $\psi(\mathfrak{R}(w), \overline{x}) + \phi_{\overline{x}}(\mathfrak{G}(i_a))$ is a cocycle for each $w \in \mathcal{D}_{(2)}^{\vee}$. We denote this cocycle by ψ.

Note that ψ is a quadratic character. The acceptable ramification conditions prevent ψ from being ramified at any prime in $\psi_S(\overline{x})$, so it is an unramified character over any $K(x)$ with $x \in \hat{x}(\emptyset)$. Since $\mathrm{rk}(\mathfrak{R}(w)) > |S|$ for each $x \in \hat{x}(\emptyset)$, and from the local triviality assumptions we made for our governing expansions $\mathfrak{G}(i_a(j_1, j_2))$, we find that ψ is trivial over any $K(x)$ at all primes where $K(x)/\mathbb{Q}$ ramified, aside those in $\pi_S(x)$. If ψ is trivial over $K(x)$ at all primes in $\pi_S(x)$, we then see that ψ corresponds to an element of $\mathcal{D}_{(2)}^{\vee}$. Observe that there are $2^{|S|}$ possibilities for the behaviour at primes in $\pi_S(x)$ and 2^{n_2+1} elements in $\mathcal{D}_{(2)}^{\vee}$. The acceptable ramification are given by one of $2^{|S(j_1,j_2)-S|}$ possibilities. These conditions are additive, and we have one set of such conditions for each of the n_2 vectors $w_{a,j}$, whence we may choose

$$A_S = (\mathbb{Z}/2\mathbb{Z})^{n_2(n_2+1+m+1-|S|+|S|)} = (\mathbb{Z}/2\mathbb{Z})^{n_2(n_2+m+1)}$$

as desired. Finally, fixing a particular behaviour in each of these cases is tantamount to choosing a function F_S mapping to the identity of A_S. This completes the proof. □

Finally, we may state and prove our main result on additive-restrictive systems:

Proposition 5.2.3. *Let $K = \mathbb{Q}(\sqrt{n_0})$ be a real quadratic field, and let $\mathfrak{A}(j_1, j_2)$ be an additive-restrictive system arising from Lemma 5.2.2. Put $S = S(j_1, j_2)$ and let $\overline{x} \in \overline{X}_S$. Suppose that for each $i \in S$ there exists $\overline{z}_i \in \hat{x}(S - \{i\})$ so that*

$$\overline{z}_i \in \overline{Y}^{\circ}_{S-\{i\}}\left(\mathfrak{A}(j_1, j_2)\right).$$

Then we have that $\hat{x}(\emptyset) \subseteq \overline{Y}^{\circ}_{\emptyset}$.

Furthermore, writing $(p_{1,b}, p_{2,b}) = \pi_{i_b}(\overline{x})$ and $i_a = i_a(j_1, j_2)$, we have

$$\sum_{x \in \hat{x}(\emptyset)} \left[\frac{L(\psi_m(\mathfrak{R}(w_{a,j_3}), x))/K(x)}{w_{b,j_4}(x)} \right]$$

$$= \begin{cases} \phi_{\overline{z}_{i_b}}(\mathfrak{G}(i_a))\left(\mathrm{Frob}(p_{1,b}) \cdot \mathrm{Frob}(p_{2,b})\right) & \text{if } (j_3, j_4) = (j_1, j_2), \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Take an arbitrary $x_0 \in \hat{x}(\emptyset)$ such that $x_0 \notin \overline{z}_i$ for any i. We need to check that the Artin pairings corresponding to x_0 is given by $\mathrm{Art}_{(k)}$ for $k < m$. We do this by considering the value of the pairing on (w_{a,j_3}, w_{b,j_4}) or $(w_{a,j_3}, t_b + w_{b,j_2})$ when $j_2 = j_4$. Recall that t_b is the generator of the kernel of the map

$$\mathcal{D}_{(2)} \to \mathfrak{cl}K(x_0)[4].$$

The value of the pairing at these tuples determines the pairing everywhere by bilinearity. However, due to the minimality restrictions on $\mathfrak{R}(w)$, Theorem 4.0.3 implies that the Arting pairings for $k < m$ at x_0 is equal to the sum of the Artin pairings at all other vertices in $\hat{x}(\emptyset)$. This shows that $x_0 \in Y^{\circ}_{\emptyset}$.

The pairings at $k = m$ follow similarly except at $(w_{a,j_1}, t_b + w_{b,j_2})$. At this tuple, the second part of Theorem 4.0.3 applies, and the claimed result follows. □

For future use, we define *variable indices* associated to $\mathfrak{A}(j_1, j_2)$ as:

Definition 5.2.4. For $i_b \leq d$ and $j_1, j_2 \leq n_0$, a set of *variable indices* consists of a tuple $(i_b, i_a(j_1, j_2), S(j_1, j_2))$ where $S(j_1, j_2)$ is as in Lemma 5.2.2 and in addition contains both i_b and $i_a(j_1, j_2)$. Further, we insist that for all $j \leq n$ $T(w_{a,j}), T(w_{b,j})$ do not contain i_b, and $S(j_1, j_2)$ is disjoint from $T(w_{a,j})$ and $T(w_{b,j})$ for all $j \leq n$ except for j_1, j_2. Moreover, we shall insist that

$$T(w_{b,j_1}) \cap S(j_1, j_2) = T(w_{a,j_2}) \cap S(j_1, j_2) = \emptyset,$$

$$T(w_{a,j_1} \cap S(j_1, j_2) = \{i_a(j_1, j_2)\},$$

and

$$S(j_1, j_2) \subseteq T(w_{b,j_2}) \cup \{i_b\}.$$

6. Ramsey theory

In Proposition 5.2.3 we found a condition on $\overline{x} \in \overline{X}_S$ so that the sum

$$\sum_{x \in \hat{x}(\emptyset)} \left[\frac{L(\psi_m(\mathfrak{R}(w_{a,j_3}, x))/K(x)}{w_{b,j_4}(x)} \right] \in \mathbb{F}_2$$

was determined by an Artin symbol in the field of definition of some governing expansion $\phi_{\overline{z}_{i_b}}$. This is not sufficient to determine the value of the pairing at any particular $x \in \hat{x}(\emptyset)$. However, if we have enough choices of $\overline{x}$ where we can evaluate this sum, we can still prove that the value of the pairing is forced to be 1 on about half of the vertices of $\overline{Y}^{\circ}_{\emptyset}$.

The first task is to check that there exists $\overline{x}$ whose vertices lie in $\overline{Y}^{\circ}_{\emptyset}$. This is a question in *Ramsey theory*: we can prove that such $\overline{x}$ exists provided that $\overline{Y}^{\circ}_{\emptyset}$ is large enough.

Proposition 6.0.1. *Let $d \geq 2$ be an integer, and let δ be a positive number satisfying $0 < \delta < 2^{-d-1}$. Let $X_1, \cdots, X_d$ be finite sets with cardinality at least $n > 0$. Suppose that Y is a subset of $X = X_1 \times \cdots \times X_d$ of cardinality at least $\delta|X|$. Then, for any positive integer r satisfying*

$$r \leq \left(\frac{\log n}{5 \log \delta^{-1}} \right)^{1/(d-1)},$$

there exists a choice of subsets $Z_i \subseteq X_i, i = 1, \cdots, d$ each of cardinality r such that

$$Z_1 \times \cdots \times Z_d \subseteq Y.$$

Proof. We can find subsets $X_i' \subset X_i$, so that $|X_i'| = n$ and Y has density at least δ in $X_i' \times \cdots \times X_d'$. We may therefore assume that $|X_i| = n$ for $i = 1, \cdots, d$.

Write $N(r, Y)$ for the number of ways of choosing subsets $Z_i \subset X_i$ for all $i \leq d$, each of cardinality r, so that $Z_1 \times \cdots \times Z_d \subset Y$. Write $N(n, r, \delta)$ for the minimum of $N(r, Y)$ over all Y of cardinality at least $\delta|X|$. To prove the proposition, we will

prove the stronger claim that whenever n, d, δ, r are positive numbers satisfying $n \geq r \geq 2$ and

$$\left(2^{-d-1}\delta\right)^{2r^{d-1}} \cdot nr^{-1} \geq 1, \tag{6.1}$$

we have

$$N_d(n, r, \delta) \geq \left(2^{-d-1}\delta\right)^{\frac{r^{d+1}-r}{r-1}} \frac{n^{rd}}{(r!)^d}. \tag{6.2}$$

We prove this by induction. Suppose that $d = 1$. Then

$$N_1(n, r, \delta) \geq \frac{(\delta n - r)^r}{r!}.$$

For $r \leq \delta n/2$, this gives

$$N_1(n, r, \delta) \geq \left(\frac{\delta}{2}\right)^r \frac{n^r}{r!}.$$

This is the base case for (6.2).

Now consider the case $d > 1$, and choose Y with $N(r, Y)$ minimal. Take X_{thick} to be the subset of $x \in X_1$ so that

$$Y_x = Y \cap (\{x\} \times X_2 \times \cdots \times X_d)$$

has density at least $\delta/2$ in $\{x\} \times X_2 \times \cdots \times X_d$. X_{thick} has denisty at least $\delta/2$ in X_1.

Take $\mathcal{L}$ to be the set of choices of subsets $Z_2, \cdots, Z_d, Z_i \subseteq X_i$, such that each Z_i has cardinality r. We have

$$|\mathcal{L}| \leq \frac{n^{r(d-1)}}{(r!)^{d-1}}.$$

For

$$\mathbf{z} = (Z_2, \cdots, Z_d) \in \mathcal{L},$$

put $n_{\mathbf{z}}$ for the number of $x \in X_{\text{thick}}$ such that Y contains

$$\{x\} \times Z_2 \times \cdots \times Z_d.$$

Then

$$N_d(n, r, \delta) = N(r, Y) \geq \sum_{\substack{\mathbf{z} \in \mathcal{L} \\ n_{\mathbf{z}} \geq r}} \frac{(n_{\mathbf{z}} - r)^r}{r!} \geq \sum_{\mathbf{z} \in \mathcal{L}} \frac{n_{\mathbf{z}}^r}{2^r r!} - \frac{r^r}{r!}.$$

We thus have

$$\sum_{\mathbf{z} \in \mathcal{L}} n_{\mathbf{z}} \geq |X_{\text{thick}}| \cdot N_{d-1}(n, r, \delta/2) \geq \frac{\delta n}{2} \left(2^{-d-1}\delta\right)^{\frac{r^d - r}{r-1}} \frac{n^{r(d-1)}}{(r!)^{d-1}},$$

whence

$$\frac{\sum_{\mathbf{z} \in \mathcal{L}} n_{\mathbf{z}}}{\sum_{\mathbf{z} \in \mathcal{L}} 1} \geq \frac{\delta n}{2} \left(2^{-d-1}\delta\right)^{\frac{r^d - r}{r-1}} \geq 4n \left(2^{-d-1}\delta\right)^{\frac{r^d - 1}{r-1}}.$$

Applying Cauchy-Schwarz, we then get

$$N_d(n, r, \delta) \geq \frac{n^{r(d-1)}}{(r!)^{d-1}} \left(-\frac{r^r}{r!} + \frac{(4n)^r (2^{-d-1}\delta)^{\frac{r^{d+1}-r}{r-1}}}{2^r r!}\right).$$

But, for $r \geq 2$, we have

$$\frac{r^{d+1}-r}{r-1} \leq 2r^d,$$

so (6.1) implies

$$N_d(n,r,\delta) \geq (2^{-d-1}\delta)^{\frac{r^{d+1}-r}{r-1}} \frac{n^{rd}}{(r!)^d},$$

as claimed. □

We now define a *generic differential* on $\overline{X}_S$ for $S \subseteq \{1,\cdots,d\}$.

Definition 6.0.2. Let $X_1,\cdots,X_d$ be pairwise disjoint finite, non-empty set and let X be their Cartesian product. Choose a subset $S \subset \{1,\cdots,d\}$ of cardinality at least two, and some subset $Z \subseteq X$ such that $\pi_{\{1,\cdots,d\}-S}(Z)$ is a point. For a function $F : Z \to \mathbb{F}_2$, define

$$dF : \{\overline{x} \in \overline{X}_S : \hat{x}(\emptyset) \subseteq Z\} \to \mathbb{F}_2$$

by

$$dF(\overline{x}) = \begin{cases} \sum_{x \in \hat{x}(\emptyset)} F(x) & \text{if } |\hat{x}(\emptyset)| = 2^{|S|} \\ 0 & \text{otherwise.} \end{cases}$$

Write $\mathfrak{C}_S(Z)$ for the image of this map $d(\cdot)$. In addition, for $\varepsilon > 0$, put $\mathfrak{C}_S(\varepsilon, Z)$ for the set of $g \in \mathfrak{C}_S(Z)$ expressible in the form $g = dF$ for some F that equals 1 on more than $(1/2+\varepsilon)|Z|$ or fewer than $(1/2-\varepsilon)|Z|$ points in Z.

Proposition 6.0.3. *For X and Z as in Definition 6.0.2, choose $\delta > 0$ so that*

$$|Z| \geq \delta \cdot |\pi_S(X)|.$$

Suppose that $|X_i| \geq n$ for each $i \in S$. Then, for all $\varepsilon > 0$, we have

$$\frac{|\mathfrak{C}_S(\varepsilon, Z)|}{|\mathfrak{C}_S(Z)|} \leq \exp\left(|\pi_S(X)| \cdot \left(-\delta\varepsilon^2 + 2^{|S|+2} \cdot n^{-1/2^{|S|}}\right)\right).$$

Proof. Take Z' to be a maximal subset of Z so that there is no $\overline{z} \in \overline{X}_S$ satisfying $|\hat{z}(\emptyset)| = 2^{|S|}$ and $\hat{z}(\emptyset) \subseteq Z'$. We see that the kernel of the map $d : \mathbb{F}_2^Z \to \mathfrak{C}_S(Z)$ then has size at most $2^{|Z'|}$. By applying (6.1) with $r = 2$, we then have

$$|Z'| \leq |\pi_S(X)| \cdot 2^{|S|+2} \cdot N^{-1/2^{|S|}}.$$

We then have

$$|\mathfrak{C}_S(Z)| \geq 2^{|Z|} \cdot \exp\left(-|\pi_S(X)| \cdot 2^{|S|+2} \cdot N^{-1/2^{|S|}}\right).$$

On the other hand, from Hoeffding's inequality, the number of F equalling 1 on more than $(1/2+\varepsilon)|Z|$ or fewer than $(1/2-\varepsilon)|Z|$ points in Z is bounded by

$$2^{|Z|+1} \exp\left(-2\varepsilon^2|Z|\right).$$

Then $\mathfrak{C}_S(\varepsilon, Z)$ is bounded by

$$|\mathfrak{C}_S(\varepsilon, Z)| \leq 2^{|Z|} \exp\left(-2\varepsilon^2|Z|\right).$$

Taking ratios of these estimates then gives the result. □

There are two issues to be addressed when attempting to apply Proposition 6.0.3. First is that we do not necessarily have control over $\overline{Y}^{\circ}_{\emptyset}$. Second, it is not enough that $\hat{x}(\emptyset)$ lie in $\overline{Y}^{\circ}_{\emptyset}$ to conclude that $dF(\overline{x}) = g(\overline{x})$ for some relevant F, g; we must insist that $\hat{x}(T)$ meet $\overline{Y}^{\circ}_{T}$ for each proper subset T of S.

However, the robustness of the structure of additive-restrictive systems both of these issues can be circumvented. First, Proposition 5.2.3 gives a criterion for x_0 to be in $\overline{Y}^{\circ}_{\emptyset}$ provided there is a nice cube $\overline{x} \in \overline{X}_S$ with all over vertices in $\overline{Y}^{\circ}_{\emptyset}$. Therefore, we do not need to consider all possible $Z = \overline{Y}^{\circ}_{\emptyset}$ but only those that satisfies this regularity condition. Second, by Proposition 5.0.2 we can bound the density of $\overline{Y}^{\circ}_{S}$ in $\overline{Y}_S$ from below in terms of the density of $\overline{Y}^{\circ}_{\emptyset}$ in $\overline{Y}_{\emptyset}$. This is enough to get around the second issue.

We make the following definition:

Definition 6.0.4. Let $X_1, \cdots, X_d$ and S as in Definition 6.0.2. For $a \geq 2$ and $\varepsilon > 0$, we say that a triple $(Z, F, \mathfrak{A})$ where $Z \subseteq \overline{X}_S$, $F : Z \to \mathbb{F}_2$, and $\mathfrak{A}$ an additive-restrictive system on X is (ε, a)-*acceptable* if the following holds:

(1) The image of Z under $\pi_{\{1,\cdots,d\}-S}$ is a point;
(2) For each $T \subseteq S$, we have $|A_T(\mathfrak{A})| \leq a$;
(3) The equalities

$$\overline{Z}_S = \bigcap_{T \subsetneq S} \{\overline{x} \in \overline{X}_S : \hat{x}(T) \cap \overline{Y}^{\circ}_{T}(\mathfrak{A}) \neq \emptyset\}$$

and

$$Z = \overline{Y}^{\circ}_{\emptyset}(\mathfrak{A})$$

hold; and
(4) The function F is equal to 1 on more than $|Z|/2 + \varepsilon|\pi_S(X)|$ or fewer than $|Z|/2 - \varepsilon|\pi_S(X)|$ of the points in Z.

Put $\mathfrak{C}_S(\varepsilon, a, X)$ for the subset of $\mathfrak{C}_S(\pi_S(\overline{X}_S))$ consisting of those g for which there is some (ε, a)-acceptable $(Z, F, \mathfrak{A})$ such that $\hat{x}(\emptyset) \subseteq Z$ and

$$dF(\overline{x}) = g(\overline{x})$$

whenever $\overline{x} \in \overline{Z}_S$.

This is summarized in the following proposition:

Proposition 6.0.5. *Let $X_1, \cdots, X_d$ and S as in Definition 6.0.2 and put $n = \min_{i \in S} |X_i|$. For $a \geq 2$ and $\varepsilon > 0$, let $\mathfrak{C}_S(\varepsilon, a, X)$ be as in Definition 6.0.4. Then there exists an absolute constant A such that whenever $\varepsilon < a^{-1}$ and*

$$A \cdot 6^{|S|} \log \varepsilon^{-1} \geq \log n, \tag{6.3}$$

we have

$$\frac{\mathfrak{C}_S(\varepsilon, a, X)|}{|\mathfrak{C}_S(\pi_S(X))|} \leq \exp\left(-|\pi_S(X)| \cdot n^{-1/2}\right).$$

Proof. Let $g \in \mathfrak{C}_S(\varepsilon, a, X)$. For $x_0 \in Z$, define $Z(x_0)$ to be the set of $x \in Z$ for which there exists some $\overline{x} \in \overline{X}_S$ with $x, x_0 \in \hat{x}(\emptyset)$ such that, whenever T is a proper subset of S and $\overline{y} \in \hat{x}(T)$ that contains the vertex x_0, then $\overline{y}$ is in $\overline{Y}^{\circ}_{T}$. From

Proposition 5.0.2 we see that there is some sequence $x_1, \cdots, x_r$ of points in Z so that

$$Z'(x_j) = Z(x_j) - Z(x_{j-1}) - \cdots - Z(x_1) \tag{6.4}$$

has density at least $(a^{-1}\varepsilon/2)^{3^{|S|}} \geq \varepsilon^{3^{|S|+1}}$ for $j \geq 1$. Each $Z(x_j)$ is determined by the sequence of structures

$$Z(x_j) \cap \pi^{-1}_{S-\{i\}}(x_j)$$

as i varies through S. The number of unspecified positions is at most

$$|\pi_S(X)| \cdot \sum_{i \in S} |X_i|^{-1} \leq |\pi_S(X)| \cdot |S| \cdot n^{-1}$$

by the definition of n. There are at most $\varepsilon^{-3^{|S|+1}}$ elements x_j, so with $x_1, \cdots, x_r$ given the $Z(x_j)$'s can be specified with at most

$$\varepsilon^{-3^{|S|+1}} |\pi_S(X)| \cdot |S| \cdot n^{-1}$$

bits. We find that there must be a j so that F equals 1 on at least

$$|Z'(x_j)|(1+\varepsilon)/2$$

vertices in $Z'(x_j)$.

The conditions on $Z'(x_j)$ imply that, whenever $x \in Z'(x_j)$, there is a cube $\overline{x} \in \overline{X}_S$ with $x, x_j \in \hat{x}(\emptyset)$ such that $dF(\overline{x}) = g(\overline{x})$. Using the additivity of dF and g we find that, whenever $\overline{x} \in \overline{X}_S$ has $\hat{x}(\emptyset)$ contained in $Z'(x_j)$, then $dF(\overline{x}) = g(\overline{x})$. Proposition 6.0.3 then implies that the number of $g \in \mathfrak{C}_S(\varepsilon, a, X)$ corresponding to this choice of $Z'(x_j)$ is bounded by

$$|\mathfrak{C}_S(\pi_S(X))| \cdot \exp\left(|\pi_S(X))| \cdot \left(-\varepsilon^{4+3^{|S|+1}} + 2^{|S|+2} \cdot n^{-1/2^{|S|}}\right)\right).$$

For sufficiently large A, (6.3) implies that

$$|\mathfrak{C}_S(\pi_S(X))| \cdot \exp\left(-|\pi_S(X))| \cdot \varepsilon^{5+3^{|S|+1}}\right).$$

Summing over all possible choices of $(x_1, \cdots, x_r)$, associated choices of $Z(x_i)$, and over all choices of j, we find that the ratio being estimated in the proposition is bounded by

$$r|\pi_S(X)|^r \exp\left(|\pi_S(X)| \cdot \left(-\varepsilon^{5+3^{|S|+1}} + \varepsilon^{-3^{|S|+1}} n^{-1}|S|\right)\right).$$

Observe that for any positive number $y \geq 1$ and any $\varepsilon' > 0$, we have

$$ry^r \ll_{\varepsilon'} \exp(\varepsilon' y).$$

Thus, by choosing A sufficiently large and taking $\varepsilon' = 1$ say, we find the bound

$$\exp\left(-|\pi_S(X)| \cdot \varepsilon^{7+3^{|S|+1}}\right),$$

which is enough to prove the proposition. □

7. Analytic tools

7.1. **Distribution of integers with fixed number of prime factors.** In this section, we wish to describe the properties of square-free integers n having exactly r prime factors. For given positive numbers N, D, put

$$S_r(D;N) = \{n = p_1 \cdots p_r \leq N, p_i > D \text{ for } i = 1, \cdots, r, \}. \tag{7.1}$$

$S_r(D;N)$ is a familiar object in sieve theory, where we have eliminated those numbers with "small" prime factors. We wish to consider "typical" sets $S_r(D;N)$, and we make the following definition.

Definition 7.1.1. Let $N > 30, D > 3$ be real numbers and r a positive integer. We say that $S_r(D;N)$ is *archetypal* if $D < \exp\left((\log N)^{1/4}\right)$ and r satisfies

$$|r - \log\log N + \log\log D| \leq \frac{2}{3}\left(\log\log N - \log\log D\right). \tag{7.2}$$

In our application, we will in fact be taking $D = \log\log\log N$, so the hypothesis that $D < \exp((\log N)^{1/4})$ will be immaterial. One should also recognize that (7.2) is the assertion that $S_r(D;N)$ contains numbers which have close to the average number of prime factors.

Our next definition is to capture the phenomenon that, except for a negligible number of circumstances, an archetypal $S_r(D;N)$ will satisfy one of a few nice properties, listed below:

Definition 7.1.2. Let $S_r(D;N)$ be an archetypal set of numbers. Let $D_1 > D$. For $n \in S_r(D;N)$, let $p_1 < \cdots < p_r$ be the prime factors of n.

We say that $n \in S_r(D;N)$ is *comfortably spaced* above D_1 if, for all $i < r$ with $p_i > D_1$, we have

$$4D_1 < 2p_i < p_{i+1}.$$

For $C_0 > 1$, we say n is C_0-*regular* if for all $i \leq r/3$, we have

$$|\log\log p_i - \log\log D - i| < C_0^{1/5} \cdot \max\{i, C_0\}^{4/5}.$$

Finally, we say n is *extravagantly spaced* if there exists $m \in (r^{1/2}/2, r/2)$ such that

$$\log p_m \geq \log\left(\frac{\log p_m}{\log D}\right) \cdot (\log\log\log N)^{1/2} \cdot \sum_{i=1}^{m-1} \log p_i. \tag{7.3}$$

The following result demonstrates that most integers $n \in S_r(N;D)$ are comfortably spaced above D_1, C_0-regular, or extravagantly spaced, provided that D_1, C_0 are chosen appropriately as functions of N.

One necessary modification to $S_r(D;N)$ for our application to the negative Pell equation is that we must demand that the numbers $n \in S_r(D;N)$ satisfy the property that $p|n \Rightarrow p \equiv 1 \pmod 4$ (note that since the prime factors of n exceed $D > 3$, n is necessarily odd). We thus put

$$S_r^*(D;N) = \{n \in S_r(D;N) : p|n \Rightarrow p \equiv 1 \pmod 4\}.$$

Proposition 7.1.3. *Let $S_r^*(D;N)$ be an archetypal set of numbers. Consider the uniform distribution on $S_r^*(D;N)$. Then there is an absolute constant $c > 0$ such that:*

(1) *The probability that n is not comfortably spaced above D_1 is*

$$O\left((\log D_1)^{-1} + (\log N)^{-1/2}\right),$$

where the implied constant is absolute;

(2) *For all $C_0 > 0$ the probability that n is not C_0-regular is*

$$O\left(\exp(-c \cdot C_0) + \exp(-c(\log\log N)^{1/2})\right),$$

where the implied constant is absolute; and

(3) *The probability that n is not extravagantly spaced is*

$$O\left(\exp(-c(\log\log\log N)^{1/2})\right),$$

with absolute implied constant.

We note that in our application we will take $D = \log\log\log N$, $D_1 = D^{(\log\log N)^{c'}}$ and $C_0 = c''\log\log\log N$, so that all of the probabilities in Proposition 7.1.3 will tend to zero.

For $x > 0$, put

$$F(x) = \sum_{\substack{p \le x \\ p \equiv 1 \pmod 4}} \frac{1}{p}.$$

Using Dirichlet's theorem applied to the non-trivial character mod 4, we see that there are absolute constants $A, c > 0$ such that whenever $x \gg 1$, we have

$$\left|F(x) - \frac{\log\log x}{2} - B_1\right| \le A\exp\left(-c(\log x)^{1/2}\right), \tag{7.4}$$

where B_1 is an absolute constant.

The proof of Proposition 7.1.3 will depend on the following three lemmas, where the first lemma is exactly claim (1) in Proposition 7.1.3.

Lemma 7.1.4. *Let $S_r^*(D;N)$ be an archetypal set of numbers. Sampling with respect to the uniform distribution, the probability that n is not comfortably spaced above D_1 is $O\left((\log D_1)^{-1} + (\log N)^{-1/2}\right)$.*

Proof. We note that the number of uncomfortably spaced elements in $S_r^*(D;N)$ is bounded by

$$\sum_{\substack{D_1 < p < N \\ p \equiv 1 \pmod 4}} \sum_{p < q < 2p} \left|S_{r-2}^*(D; N/pq)\right|.$$

We divide the range for p int $p < (\log N)^{1/4}$ and $p \ge (\log N)^{1/4}$. In the first case we have the bound

$$O\left(|S_{r-2}^*(D;N)| \sum_{\substack{D_1 < p < N \\ p \equiv 1 \pmod 4}} \sum_{p < q < 2p} (pq)^{-1}\right) = O\left(|S_r^*(D;N)|(\log D_1)^{-1}\right).$$

For the second case, note that the sum over $p \geq (\log N)^{1/4}$ can be bounded by $O(N(\log N)^{-1})$, so we are done since

$$\frac{(\log\log N - F(D) + B_1)^{r-1}}{(r-1)!} \geq (\log N)^{1/2}$$

for all sufficiently large N, by our restriction on r. □

Lemma 7.1.5. *Let $S_r^*(D;N)$ be an archetypal set of numbers. Put*

$$u = \frac{\log N}{\exp(F(D) - B_1)}.$$

Then for all N sufficiently large, we have

$$\frac{c_1 N}{\log N}\frac{(\log u)^{r-1}}{2^r(r-1)!} < |S_r^*(D;N)| < \frac{c_2 N}{\log N}\frac{(\log u)^{r-1}}{2^r(r-1)!} \tag{7.5}$$

provided that r satisfies (7.2).

To prove Lemma 7.1.5 and to state our final lemma, we require some terminology. Suppose that T is a collection of r-tuples of odd primes congruent to 1 modulo 4. Define $\mathcal{G}(T) \subset \mathbb{R}^r$ to be the union

$$\mathcal{G}(T) = \bigcup_{(p_1,\cdots,p_r)\in T} \prod_{i\leq r} \left[F(p_i) - \frac{1}{p_i} - B_1, F(p_i) - B_1\right]. \tag{7.6}$$

Suppose $V \subseteq \mathbb{R}^r$

$$\mathcal{L}(T) = \{((\log\log p_1)/2, \cdots, (\log\log p_r)/2) : (p_1, \cdots, p_r) \in T\}.$$

For $\mathbf{x} = (x_1, \cdots, x_r) \in \mathbb{R}^r$, define

$$\tau(\mathbf{x}) = \prod_{i\leq r}\left[x_i - A\exp\left(-c\cdot e^{x_i/2}\right), x_i + A\exp\left(-c\cdot e^{x_i/2}\right)\right].$$

Now define

$$V^\dagger = \bigcup_{\mathbf{x}\in V} \tau(\mathbf{x}), V^{\dagger\dagger} = \{x \in \mathbb{R}^r : x_i \geq -B_1 \text{ for } i = 1, \cdots, r, \tau(\mathbf{x}) \subseteq V\}.$$

For appropriate choices of A, c, we see that whenever T is the maximal set of prime tuples of primes congruent to 1 mod 4 such that $\mathcal{L}(T) \subseteq T$, we have

$$V^{\dagger\dagger} \subseteq \mathcal{G}(T) \subseteq V^\dagger. \tag{7.7}$$

Observe that

$$\mathrm{Vol}(\mathcal{G}(T)) = \sum_{(p_1,\cdots,p_r)\in T} \frac{1}{p_1\cdots p_r}.$$

Put

$$V_r(u) = \left\{(x_1, \cdots, x_r) \in \mathbb{R}^r : e^{2x_1} + \cdots + e^{2x_r} \leq u\right\}.$$

We then see that

$$\exp\left(x + A\exp(-c\cdot e^{x/2})\right) - \exp(x) \leq \kappa$$

for some κ depending on A, c but not x. Then $V_r(u)^\dagger$ is contained in $V_r(u + r\kappa)$, while $V_r(u)^{\dagger\dagger}$ contains

$$V_r(u - r\kappa) \cap (-B_1, \infty)^r.$$

At the same time, we see that for $B \in \mathbb{R}$,

$$\mathrm{Vol}\left(V_r(u) \cap (B, \infty)^r\right) = 2^{-r} I_r(e^{-B}u),$$

where

$$I_r(u) = \frac{1}{2\pi i}\int_{b-i\infty}^{b+i\infty} \frac{e^s}{s}\left(\int_1^\infty \frac{e^{-ts/u}}{t}dt\right)^k ds, b>0. \tag{7.8}$$

It is known that $I_r(u)$ satisfies the estimate

$$I_r(u) = \frac{e^{-\gamma\alpha}}{\Gamma(1+\alpha)}(\log u)^k + O\left((\alpha+1)(\log u)^{k-1}(\log\log u)^3\right), \tag{7.9}$$

where $\alpha = k/\log u$, γ is the Euler-Mascheroni constant, and the implied constant is absolute; see Lemma 5.1 in [6].

Proof of Lemma 7.1.5. Let

$$F_r(D;N) = \sum_{\substack{D,p_1<\cdots,p_r\\ p_1\cdots p_r\le N\\ p_i\equiv 1 \pmod 4, 1\le i\le r}} \frac{1}{p_1\cdots p_r}.$$

Then for $D, N > 0$ and $r\in\mathbb{N}$ we have

$$I_r\left(\frac{\log N - r\kappa}{\exp\left(F(D)-B_1\right)}\right) \le 2^r F_r(D;N) \le I_r\left(\frac{\log N + r\kappa}{\exp\left(F(D)-B_1\right)}\right).$$

If $r^2 < A\log N$ and $\log\log N - F(D) + B_1 > 1$, we have

$$\begin{aligned} 2^r F_r(D;N) &= I_r\left(\frac{\log N}{\exp(F(D)-B_1)}\right)\\ &\quad + O\left(\frac{r^2}{\log N}\cdot(\log\log N - F(D) + B_1)^{r-1}\right). \end{aligned}$$

We restrict to the case when $\log\log N > 2\log\log D$, so that the ratio

$$u = \frac{\log N}{\exp(F(D)-B_1)}$$

is at least 3, and that

$$\frac{r}{3} < \log u < \frac{5r}{3}$$

by (7.2). Put

$$G_r(D;N) = \sum_{\substack{D<p_1,\cdots,p_r\\ p_1\cdots p_r\le N\\ p_i\equiv 1 \pmod 4, 1\le i\le r}} \log(p_1\cdots p_r) \tag{7.10}$$

and

$$H_r(D;N) = \sum_{\substack{D<p_1,\cdots,p_r\\ p_1\cdots p_r\le N\\ p_i\equiv 1 \pmod 4, 1\le i\le r}} 1. \tag{7.11}$$

We claim that

$$G_r(D;N) = r2^{-r}I_{r-1}(u) + O_\varepsilon\left(\frac{N}{\log N}(\log u)^{r+3}\right) \tag{7.12}$$

and

$$H_r(D;N) = \frac{r2^{-r}N}{\log N}I_{r-1}(u) + O_\varepsilon\left(\frac{N}{\log^2 N}(\log u)^{r+3}\right). \tag{7.13}$$

To see (7.12), we first note that, upon rearranging the sum and observing that $\log(p_1 \cdots p_r) = \log(p_1) + \cdots + \log(p_r)$, we have

$$\begin{aligned} G_r(D;N) &= \sum_{\substack{D<p_1,\cdots,p_r \\ p_1\cdots p_r \leq N \\ p_i \equiv 1 \pmod 4, 1\leq i\leq r}} (\log p_1 + \cdots \log p_r) \\ &= r \sum_{\substack{D<p_1,\cdots,p_r \\ p_1\cdots p_r \leq N \\ p_i \equiv 1 \pmod 4, 1\leq i\leq r}} \log p_1. \end{aligned}$$

To ameliorate notation, we shall denote by $\sum^\sharp$ for the condition that $p_i \equiv 1 \pmod 4$. We then observe that

$$\sum_{\substack{D<p_1,\cdots,p_r \\ p_1\cdots p_r \leq N}}^\sharp \log p_1 = \sum_{\substack{D<p_1,\cdots,p_{r-1} \\ p_1\cdots p_{r-1} \leq N/P}}^\sharp \sum_{\substack{p>D \\ p\equiv 1 \pmod 4}}^{N/P} \log p,$$

where $P = p_1 \cdots p_{r-1}$. By Dirichlet's theorem for primes congruent to 1 mod 4 with an explicit zero-free region estimate, we have

$$\sum_{\substack{p>D \\ p\equiv 1 \pmod 4}}^{N/P} \log p = \frac{N}{2P}\left(1 + O\left(\exp\left(-c\sqrt{NP^{-1}}\right)\right)\right) - \sum_{\substack{p<D \\ p\equiv 1 \pmod 4}} \log p.$$

It thus follows that

$$G_r(D;N) = \frac{rN}{2} F_{r-1}(D;N/D) - \frac{r}{2} H_r(D;N/D) \sum_{\substack{p<D \\ p\equiv 1 \pmod 4}} \log p +$$

$$O\left(rNe^{-c\sqrt{\log D}}\left(F_{r-1}(D;N) - F_r(D;N_0)\right) + rNe^{-c\sqrt{\log N/N_0}} F_r(D;N_0)\right)$$

for any $N > N_0 > D$. We choose $N_0 = N\exp\left(\left(c^{-1}\log\log N\right)^2\right)$, we see that this is an acceptable error term for the purpose of (7.12). To see (7.13), we do partial summation applied to (7.12).

Note that $H_r(D;N)$ and $S_r^*(D;N)$ are virtually the same set: indeed, $S_r^*(D;N)$ have the extra stipulation that ordering of the prime factors matters and that each element is square-free. One easily checks that the density of square-free elements among numbers of the form $\{n = p_1 \cdots p_r \leq N : D < p_1, \cdots, p_r, p_i \equiv 1 \pmod 4\}$ is at least $1/2$, whence

$$\frac{|H_r(D;N)|}{2\cdot r!} \leq |S_r^*(D;N)| \leq \frac{|H_r(D;N)|}{r!}. \tag{7.14}$$

From this and (7.13) we conclude that there exist positive numbers c_1, c_2 such that (7.5) holds for N sufficiently large. □

To proceed, we put

$$S_{r,k}(D;N) = \{p_1 \cdots p_r \in S_r^*(D;N) : p_i < N', 1 \leq i \leq k, p_i > N', i > k\}, \tag{7.15}$$

where

$$N' = \exp\left(\sqrt{\log N \exp\left(F(D) - B_1\right)}\right).$$

We now state the next lemma we need:

Lemma 7.1.6. *Let D, N, r be as in Definition 7.1.1, and $S_r^*(D;N)$ be an archetypal set of numbers. Let $S_{r,k}(D;N)$ be as in (7.15). Then the density of the set*

$$\bigcup_{|r-k/2|>r^{2/3}} S_{r,k}(D;N)$$

in $S_r^*(D;N)$ is bounded by

$$O\left(\exp(-c(\log\log N)^{1/2})\right).$$

If $|r-k/2| \le r^{2/3}$, then for any two sets T_1, T_2 of k-tuples of primes with size bounded by N_1 in increasing order, we have

$$\frac{|S_{r,k}(D;N,T_1)|}{|S_{r,k}(D;N,T_2)|} = O\left(\frac{\mathrm{Vol}(\mathcal{G}(T_1))}{\mathrm{Vol}(\mathcal{G}(T_2))}\right),$$

where $S_{r,k}(D;N,T_i)$ is the subset of $S_{r,k}(D;N)$ consisting of those numbers whose k smallest prime factors form a vector $(p_1,\cdots,p_k) \in T_i$, $i=1,2$.

Proof. By definition, we have

$$|S_{r,k}(D;N)| = \sum_{\substack{D<p_1<\cdots<p_k<N_1 \\ p_i\equiv 1 \pmod 4, 1\le i\le k}} \left|S_{r-k}\left(N_1; \frac{N}{p_1\cdots p_k}\right)\right|.$$

Note that by definition $N_1 = O_\varepsilon(N^\varepsilon)$ for any $\varepsilon > 0$. Therefore, we can certainly assume that N is large enough so that $N_1^k < N^{1/2}$. This implies that

$$\frac{(\log\log N - F(N_1) + B_1)^{r-1}}{(\log\log(N/P) - F(N_1) + B_1)^{r-1}} \le C$$

for some absolute constant C for any choice of $P = p_1\cdots p_k$. The same is true of the ratio $(\log N)/(\log N - \log P)$. Thus we may find positive numbers c_1, c_2 such that

$$\frac{c_1}{P}|S_{r-k}(N_1;N)| < |S_{r-k}(N_1,N/P)| < \frac{c_2}{P}|S_{r-k}(N_1,N)| \tag{7.16}$$

for sufficiently large N. We then see that

$$|S_{r,k}(D;N)| = O\left(\frac{N}{\log N}\frac{(\log\log N - F(N_1)+B_1)^{r-1}}{2^{2r}(r-k+1)!k!}\right)$$

for $r > k$. Hoeffding's inequality gives the first part of the proposition if we remove the case $r = k$ from the union. The case $r = k$ is insignificant as its contribution is bounded by N_1^r, which we know is $O_\varepsilon(N^\varepsilon)$ for any $\varepsilon > 0$. The first part of the lemma then follows. The second part follows from (7.16) in terms of the corresponding statements for $S_{r,k}(D;N)$, and we are done. □

7.2. **Proof of Proposition 7.1.3.** We first consider $S_{r,k}(D;N)$ with $|k-r/2| \le r^{2/3}$. The remaining k will contribute only to the error term.

Let T_2 be the set of k-tuples of distinct primes from the interval (D, N_1) congruent to 1 mod 4. We then find that

$$\mathrm{Vol}(\mathcal{G}(T_2)) \ge c(\log\log N_1 - \log\log D)^k.$$

Take T_1 to be the set of non-C_0-regular prime tuples in T_2. The majorization of the grid of T_1 consists of elements that are not $C_0 - \kappa$ regular for some $\kappa > 0$, independent of C_0. We can estimate the volume of this majorization to be bounded by

$$O\left(\exp(-c \cdot C_0) \cdot (\log\log N_1 - \log\log D)^k\right).$$

The result then follows from Lemma 7.1.6.

For (3), take T_1 to be the set of prime tuples which, for $m > k^{1/2}$, we have

$$\log p_m \leq \log\left(\frac{\log p_m}{\log D}\right) \cdot (\log\log\log N)^{1/2} \cdot \left(\sum_{i=1}^{m-1} \log p_i\right).$$

The majorization of this grid consists of tuples $(x_1, \cdots, x_k)$ so, for $m > k^{1/2}$, we have

$$e^{x_m} \leq A x_m \cdot (\log\log\log N)^{1/2} \cdot \left(\sum_{i=1}^{m-1} e^{x_i}\right)$$

for some $A > 0$. The volume of this majorization is

$$O\left(\exp(-c \cdot (\log\log\log N)^{1/2}) \cdot (\log\log N_1 - \log\log D)^k\right).$$

We are again done by Lemma 7.1.6.

7.3. **Chebotarev's density theorem and the large sieve.** There are two main tools to predict the distribution of the Legendre symbol $\left(\frac{d}{p}\right)$ over a given set of primes p. If d is small relative to the primes [, then we can use Chebotarev's density theorem (a refinement of the Siegel-Walfisz theorem) to predict the distribution. On the other hand, if d is similar in size to the primes p, we can use the large sieve results due to Jutila to predict the distribution of these symbols on average, over a large range of d [5]. This is a standard set-up in analytic number theory.

We start with the form of the Chebotarev density theorem that we will be using.

Proposition 7.3.1. *Suppose $M/\mathbb{Q}$ is a Galois extension and $\mathrm{Gal}(M/\mathbb{Q})$ is a 2-group. Suppose $M = KE$, where $E/\mathbb{Q}$ is Galois of degree d and $K/\mathbb{Q}$ is an elementary abelian extension. Suppose that the discriminants $\Delta_E = \Delta(E/\mathbb{Q})$ and $\Delta_K = \Delta(K/\mathbb{Q})$ are co-prime. Let K_0 be a subfield of K so that Δ_{K_0} is maximal among all subfields of K.*

Put $G = \mathrm{Gal}(M/\mathbb{Q})$ and let $F : G \to [-1, 1]$ be a class function of G with average zero. Then there is an absolute constant $c > 0$ such that

$$\sum_{p \leq x} F\left(\left[\frac{M/\mathbb{Q}}{p}\right]\right) \log p =$$

$$O\left(x^{\beta}|G| + x|G| \exp\left(\frac{-cd^{-4}\log x}{\sqrt{\log x} + 3d\log|\Delta_{K_0}\Delta_E|}\right)(d^2 \log|x\Delta_{K_0}\Delta_E|)^4\right)$$

for $x \geq 3$, where β is the maximal real zero of an Artin L-function defined for G, where we ignore the term if no such zero exists.

Proof. Since G is a 2-group, it is nilpotent and hence supersolvable, whence the Artin conjecture is true for any non-trivial irreducible representation ρ of G. In particular, the Artin L-function $L(\rho, s)$ is entire. The representations $\rho \otimes \rho$ and $\rho \otimes \overline{\rho}$ also satisfy the Artin conjecture, hence

$$L(\rho \otimes \overline{\rho}, s)$$

is entire except for a simple pole at $s = 1$ and

$$L(\rho \otimes \rho, s)$$

is entire unless ρ is isomorphic to $\overline{\rho}$.

The bounds are then consequences of well-known zero-free regions of L-functions. For example, Theorem 5.10 of [4] shows that $L(\rho, s)$ has no zeroes in the region

$$\Re(s) \geq 1 - \frac{c}{d^4 \log(\mathfrak{q}(\rho)(|\Im(s)| + 3))},$$

where $\mathfrak{q}(\rho)$ is the so-called *analytic conductor* of L (see (5.7) in [4]). In the case of Artin L-functions, we have the inequality

$$\mathfrak{q}(\rho) \leq q(\rho) 4^d,$$

where $q(\rho)$ is the (usual) conductor of L. Theorem 5.13 of [4] then gives

$$\sum_{p \leq x} \chi_\rho \left(\left[\frac{M/\mathbb{Q}}{p} \right] \right) \log p = \tag{7.17}$$

$$O\left(x^\beta + x \exp\left(\frac{-cd^{-4} \log x}{\sqrt{\log x} + 3 \log \mathfrak{q}(\rho)} \right) (d \log x \mathfrak{q}(\rho))^4 \right).$$

Now observe that ρ is defined on $\mathrm{Gal}(K_0 E/\mathbb{Q})$ for some quadratic extension $K_0/\mathbb{Q}$ inside K, whence its degree is bounded by $2d$ and the conductor of L is bounded by the discriminant of $K_0 L/\mathbb{Q}$, which is then bounded by

$$\Delta_{K_0}^d \Delta_E^2.$$

Then the upper bound in (7.17) may be replaced by

$$O\left(x^\beta + x \exp\left(\frac{-cd^{-4} \log x}{\sqrt{\log x} + 3d \log |\Delta_{K_0} \Delta_E|} \right) \left(d^2 \log x |\Delta_{K_0} \Delta_E| \right)^4 \right).$$

Now we may write $F = \sum_\rho a_\rho \chi_\rho$, the sum running over irreducible representations of G. Then

$$\begin{aligned} \sum_\rho |a_\rho| &= \sum_\rho \left| \frac{1}{|G|} \sum_{g \in G} F(g) \cdot \overline{\chi_\rho}(g) \right| \\ &\leq \sum_\rho \left(\frac{1}{|G|} \sum_{g \in G} F(g) \cdot \overline{F}(g) \right)^{1/2} \cdot \left(\frac{1}{|G|} \sum_{g \in G} \chi_\rho(g) \overline{\chi_\rho}(g) \right)^{1/2} \\ &\leq \sum_\rho 1 \leq |G|. \end{aligned}$$

This completes the proof. □

Next we give the form of the large sieve inequality that we will use:

Proposition 7.3.2. *Let X_1, X_2 be disjoint sets of odd primes such that for all $p \in X_i$, we have $p \leq t_i$, for $i = 1, 2$. Then, for any $\varepsilon > 0$, we have*

$$\sum_{x_1 \in X_1} \left| \sum_{x_2 \in X_2} \left(\frac{x_1}{x_2} \right) \right| = O_\varepsilon \left(t_1 t_2^{3/4+\varepsilon} + t_2 t_1^{3/4+\varepsilon} \right).$$

Proof. See, for example, the proof of Lemma 15 in [2]. □

7.4. **Boxes of integers.** In this section, we give the definitions and results that enable us to move from a set of positive integers less than a certain bound to a product space of primes. As before, $S_r^*(D; N)$ denotes the square-free integers less than N with exactly r prime factors, all of which are congruent to 1 mod 4 and greater than D.

Let $3 \leq D \leq D_1 \leq N$ be real numbers, and r an integer satisfying (7.2). Let W be a subset of elements in $S_r^*(D; N)$ that is comfortably spaced above D_1.

Let $k \leq r$ be a non-negative integer, and choose an increasing sequence of primes congruent to 1 mod 4 satisfying

$$D < p_1 < \cdots < p_k < D_1.$$

Take

$$D_1 < t_{k+1} < t_{k+2} < \cdots < t_r$$

to be an increasing sequence of real numbers. For $i > k$, put

$$t_i' = \left(1 + \frac{1}{e^{i-k} \log D_1} \right) t_i.$$

Take $X_i = \{p_i\}$ for $i \leq k$ and X_i to be the set of primes in the interval (t_i, t_i') congruent to 1 mod 4 for $i > k$.

If $t_i' < t_{i+1}$ for all $r > i > k$, then there is a bijection from X to a subset of $S_r^*(D; N)$ when N is sufficiently large, by simply taking a prime p_i from each X_i and then multiplying. By abuse of notation, we denote this subset of $S_r^*(D; N)$ by X as well. For a subset $W \subset S_r^*(D; N)$, we say that X *meets* W if $X \cap W$ is non-empty.

The restriction to comfortably spaced W means that, if $X \cap W$ is non-empty, then the X_i's are automatically disjoint sets and none of them contain any prime below D_1.

Proposition 7.4.1. *Let $3 \leq D \leq D_1 \leq N$ be positive number satisfying $2 \log\log D_1 \leq \log\log N$ and r a positive integer satisfying (7.2). Take W to be a subset of $S_r^*(D; N)$ that is comfortably spaced above D_1. Suppose V is any other subset of $S_r^*(D; N)$, and that there are numbers $\delta, \varepsilon > 0$ such that*

$$|W| > (1 - \varepsilon) \cdot |S_r^*(D; N)|$$

and for any box X meeting W, we have

$$(\delta - \varepsilon)|X| < |V \cap X| < (\delta + \varepsilon)|X|.$$

Then

$$|V| = \delta |S_r^*(D; N)| + O\left((\varepsilon + (\log D_1)^{-1}) \cdot |S_r^*(D; N)| \right).$$

Proof. Let $\mathfrak{D}_k$ to be the space of tuples

$$\mathbf{t} = (p_1, \cdots, p_k, t_{k+1}, \cdots, t_r)$$

corresponding to boxes meeting W; we write the corresponding box as $X(\mathbf{t})$. Consider

$$\int_{\mathfrak{D}_k} |V \cap X(\mathbf{t})| \cdot \frac{dp_1 \cdots dp_k dt_{k+1} \cdots dt_r}{t_{k+1} \cdots t_r},$$

where the measure corresponding to $dp_1, \cdots, dp_k$ is the indicator function on primes congruent to 1 mod 4, and the measure corresponding to dt_j is Lebesgue measure. If $n \in W$ has exactly r prime factors less than N_1 and corresponds to the triple $(q_1, \cdots, q_r)$, then n is in $X(\mathbf{t})$ if

$$(q_1, \cdots, q_k) = (p_1, \cdots, p_k)$$

and for $i > k$ we have

$$t_i \le p_i \le t_i \left(1 + \frac{1}{e^{i-k} \log D_1}\right).$$

Then, whenever

$$\prod_{i=k+1}^{r} \left(1 + \frac{1}{e^{i-k} \log D_1}\right) n < N, \tag{7.18}$$

we have that the measure of the subset of $\mathfrak{D}_k$ corresponding to boxes containing n is

$$\prod_{i=k+1}^{r} \log \left(1 + \frac{1}{e^{i-k} \log D_1}\right).$$

If n is outside W but in $S_r^*(D; N)$ with exactly k prime factors below N_1, or if n is in W but does not satisfy (7.18), then the measure of boxes containing n is still bounded by this product. Any n not satisfying (7.18) is in the range

$$N\left(1 - A(\log D_1)^{-1}\right) \le n \le N$$

where A is some positive number.

Taking $H_r(D; N)$ as in (7.11) and using (7.13), together with Lemma 7.1.5, we find that for $c \in (0, 1)$

$$\left(c + O\left(\frac{(\log\log N)^4}{\log N}\right)\right) |S_r^*(D; N)| \ll \frac{H_r(D; N) - H_r(D; (1-c)N)}{2^r r!}$$
$$\ll \left(c + O\left(\frac{(\log\log N)^4}{\log N}\right)\right) |S_r^*(D; N)|,$$

where the implied constants are absolute. From here we see that the number of n not satisfying (7.18) is $O\left(|S_r^*(D; N)| / \log D_1\right)$.

Next we see that

$$\sum_{k \ge 0} \prod_{i=k+1}^{r} \log \left(1 + \frac{1}{e^{i-k} \log D_1}\right)^{-1} \int_{\mathfrak{D}_k} |V \cap B(\mathbf{t})| \cdot \frac{dp_1 \cdots dp_k dt_{k+1} \cdots dt_r}{t_{k+1} \cdots t_r}$$

is at least as large as

$$|V \cap W| - O\left((\log D_1)^{-1} \cdot |S_r^*(D; N)|\right)$$

and is no larger than $|V|$. The estimates on $|V \cap X(\mathbf{t})|$ relative to $|X(\mathbf{t})|$ then give the desired conclusion. □

In applications, the W in Proposition 7.4.1 should be interpreted as being some "nice" set, which includes the notions of comfortable spacing, regularity, and extravagant spacing. We would need to include one more condition on controlling Siegel zeroes in order to make our results fully unconditional.

Proposition 7.4.2. *Take $d_1, d_2, \cdots$ be a potentially infinite sequence of distinct square-free integer satisfying*

$$d_i^2 < |d_{i+1}|$$

for all $i \geq 1$. Let $3 \leq D \leq D_1 \leq N$ be real numbers satisfying $2 \log\log D_1 \leq \log\log N$, $D_1 = D^{(\log\log N)^{1/10}}$, and r satisfying (7.2).

For each $i \in \mathbb{N}$, put d_i' for the product of primes dividing d_i which exceed D, and $\mathbf{d}'$ to be the subset of the d_i' for which $|d_i|$ is greater than D_1. Put

$$V = \bigcup_{X \cap \mathbf{d}' \cdot \mathbb{Z} \neq \emptyset} X \tag{7.19}$$

where the union is over all boxes X in $S_r^*(D;N)$ that contain some element n divisible by an element of $\mathbf{d}'$. We further assume that $2D \log D < \log D_1$. Then

$$|V| = O\left(\frac{|S_r^*(D;N)|}{\log D_1}\right).$$

Proof. For $d_i' \in S_r^*(D;N)$, write it as $d_i' = p_1 \cdots p_r$. Suppose that some element in X is divisible by d_i'. Taking $n \in X$, we see that there are prime factors $q_1, \cdots, q_r$ of n such that

$$q_i = p_i \text{ if } p_i < D_1$$

and

$$\frac{q_i}{2} < p_i < 2q_i \text{ otherwise.}$$

If $d_i' < N^{2/3}$, there is then an absolute constant A so that the number of n sharing a box with a multiple of d_i' is bounded by

$$A^r \cdot \prod_{p_i \leq D_1} p_i^{-1} \prod_{p_i > D_1} (\log p_i)^{-1} |S_r^*(D;N)| = O\left(\frac{|S_r^*(D;N)|}{\log d_i'}\right).$$

We can also bound the contribution from $d_i' \geq N^{2/3}$ by $O\left(N(\log N)^{-1}\right)$.

Removing finitely many terms and renumbering, we may assume that $|d_1| < D_1$. We then get $|d_i| > D_i^{2^{i-1}}$ for $i \geq 1$, so $|d_i'| > D_1^{2^{i-1}}$; since

$$\prod_{p \leq D} p < e^{2D} \ll_\varepsilon D_1^\varepsilon$$

for all $\varepsilon > 0$. Therefore the contribution from $d_i' < N^{2/3}$ is

$$O\left(|S_r^*(D;N)| \cdot \sum) i \geq 1 \frac{1}{2^i \log D_1}\right) = O\left(\frac{|S_r^*(D;N)|}{\log D_1}\right),$$

as desired. □

We say that a box X is *Siegel-free above* D_1 if it is not contained in the set V in (7.19) with respect to the sequence $\{d_k\}$ in Definition 8.0.2 and satisfying (8.1).

8. Equidistribution of Legendre symbols: the seed distribution

Definition 8.0.1. Let $\mathcal{P}$ be an arbitrary set of prime numbers congruent to. For $r > 0$, let $\mathcal{M}$ be a subset of

$$\{\{i, j\} : i, j \in \{1, \cdots, r\}\}.$$

and let $\mathcal{M}_{\mathcal{P}}$ be some subset of $\{1, \cdots, r\} \times V_{\mathcal{P}}$, where $V_{\mathcal{P}}$ is the set of square-free numbers whose prime divisors are in $\mathcal{P}$. Let a be an arbitrary function from $\mathcal{M} \cup \mathcal{M}_{\mathcal{P}}$ to $\{\pm 1\}$.

Let $X_1, \cdots, X_r$ be pairwise disjoint sets of odd primes which are also disjoint with $\mathcal{P}$, and put X for their Cartesian product. We put $X(a)$ for the set of $(x_1, \cdots, x_r) \in X$ satisfying

$$\left(\frac{x_i}{x_j}\right) = a(\{i, j\}) \text{ for all } i < j \text{ with } \{i, j\} \in \mathcal{M}$$

and

$$\left(\frac{d}{x_j}\right) = a((i, d)) \text{ for all } (i, d) \in \mathcal{M}_{\mathcal{P}}.$$

Our goal is to find situations where $|X(a)|$ is well approximated by $2^{-|\mathcal{M}_P \cup \mathcal{M}|}|X|$. To do this unconditionally, we need to account for the possibility of Siegel zeroes in the L-functions of the associated quadratic characters. We use the following definition of Siegel zeroes:

Definition 8.0.2. Let c be a positive real number. Put $\mathrm{Sieg}(c)$ for the set of square-free integers d so that the quadratic character χ_d associated with $\mathbb{Q}(\sqrt{d})/\mathbb{Q}$ has Dirichlet L-function having a real root s satisfying

$$1 \geq s \geq 1 - c(\log 2d)^{-1}.$$

We can order $\mathrm{Sieg}(c)$ by increasing magnitude, getting a sequence $d_1, d_2, \cdots$ (of course, since we expect that there are no L-functions with Siegel zeroes, this sequence is likely empty). By Landau's theorem (Theorem 5.28 in [4]), we can choose c sufficiently small so that

$$d_i^2 < |d_{i+1}| \tag{8.1}$$

for all $i \geq 1$. We say an integer d is *Siegel-less* if $d \neq d_i$ for all $i \geq 1$.

Let $c_1, \cdots, c_8$ be positive numbers which satisfy

$$\frac{1}{c_1} + \frac{c_2 c_4}{4} + \frac{c_7 \log 2}{2} + c_8 < \frac{1}{8}, c_3 > 1, c_5 > 3. \tag{8.2}$$

Let A be a positive number which depends on the numbers c_i in (8.2), and $t, t_1, t_1', \cdots, t_r, t_r'$ positive real numbers satisfying

$$A < t < t_1 < t_1' < \cdots < t_r < t_r'. \tag{8.3}$$

Let $X_1, \cdots, X_r$, a, and P be as in Definition 8.0.1, and suppose that the X_i's satisfy the property that

$$X_i \subset (t_i, t_i') \text{ for } i \leq r. \tag{8.4}$$

Suppose that there is an index $1 \leq k \leq r$ such that whenever $\mathcal{M}_P$ contains an element of the form (i,d), then $i > k$. When $i > k$, we assume that

$$X_i = \left\{ t_i < p < t_i' : \left(\frac{d}{p}\right) = a((i,d)) \text{ for all } (i,d) \in \mathcal{M}_P \right\}. \tag{8.5}$$

We shall also make the following assumption:

(8.6)

Suppose that D_1 is a product of primes in P and D_2 a square-free number such that each prime divisor of D_2 is in X_i for some i, and at most one prime factor of D_2 is in X_i for each i. We assume that $D_1 D_2$ is Siegel-less whenever $|D_1 D_2| > t$.

We then have the following:

Proposition 8.0.3. *Let $c_1, \cdots, c_8$ be as in (8.2), $A, t, t_1, t_1', \cdots, t_r, t_r'$ be as in (8.3), and $X_1, \cdots, X_r$, a, and $\mathcal{P}$ as in (8.4) and (8.5). Suppose further that the following holds:*

1. *$p < t_1'$ for all $p \in \mathcal{P}$;*
2. *$t_1' > r^{c_1}$ and $t_k' < \exp\left((t_1')^{c_2}\right)$;*
3. *For $1 \leq i \leq r$, we have*
$$|X_i| \geq \frac{2^{ic_3} \cdot t_i'}{(\log t_i')^4} \text{ and } |P| \leq \log t_i' - i;$$
4. *If $k \neq r$, then*
$$t_{k+1}' > \exp\left((\log t_1')^{c_5}\right), \exp(t^{c_6});$$
5. *$k < c_7 \log t_1'$ and that, for any $i \leq r$ and j satisfying $r \geq j \geq i - 2 + c_7 \log t_i'$, we have*
$$\exp\left((\log t_i')^{c_5}\right) < t_j'.$$

Then the inequality

$$\left| |X(a)| - 2^{-|\mathcal{M}|}|X| \right| \leq (t_1')^{-c_8} \cdot 2^{-|\mathcal{M}|}|X|.$$

Proof. We will show that, subject to the hypotheses of the proposition, that

$$\left| |X(a) - 2^{-|\mathcal{M}|}|X| \right| \leq r \cdot (t_1')^{-c_8 - c_1^{-1}} \cdot 2^{-|\mathcal{M}|} A|X|.$$

The bound on t_1' shows that this implies the desired conclusion.

We proceed by induction on r. The statement is obvious for $r = 1$, where $\mathcal{M}$ is empty. Now suppose we now know the result for every product of length $r-1$, and we wish to show the result for $X = X_1 \times \cdots \times X_r$. To this end, for $x_1 \in X_1$, put $X_i(a, x_1)$ for the subset of $x_i \in X_i$ satisfying

$$\left(\frac{x_1}{x_i}\right) = a(\{1,i\})$$

whenever $\{1,i\} \in \mathcal{M}$.

If $\{1,i\} \in \mathcal{M}$ for $i \leq k$, we apply Proposition 7.3.2 to obtain

$$\sum_{x_1 \in X_1} \left| \sum_{x_i \in X_i} \left(\frac{x_1}{x_i}\right) \right| = O_\varepsilon\left(t_i' \cdot (t_1')^{3/4+\varepsilon}\right).$$

Then, for any $\varepsilon > 0$, the assumptions on the size of X_i's then gives

$$\sum_{x_1 \in X_1} \left| \sum_{x_i \in X_i} \left(\frac{x_1}{x_i} \right) \right| < (t_1')^{-1/4 + c_2 c_4 + \varepsilon} \cdot |X_1| \cdot |X_i|$$

for sufficiently large A. Choosing $c_a, c_b > 0$ with

$$c_a + c_b < \frac{1}{4} - c_2 c_4,$$

we expect that, for all $x_1 \in X_1$, we have

$$||X_i(a, x_1)| - |X_i|/2| < (t_i')^{-c_b} \cdot |X_i| \text{ for all } \{1, i\} \in \mathcal{M}$$

with at most $k \cdot (t_1')^{-c_a} \cdot |X_1|$ exceptions. Write X_1^{bad} for the set of exceptional x_1. We choose

$$c_a > c_7 \log 2 + c_8 + c_1^{-1}$$

and

$$c_b > c_8 + c_1^{-1}.$$

One checks that (8.2) means that it is possible to choose such c_a, c_b.

Suppose $\{1, i\} \in \mathcal{M}$ is such that $i > k$. We apply Proposition 7.3.1 to the field M generated by $\sqrt{x_1}$ and by $\sqrt{p}$ with $p \in P$. Take $F : \text{Gal}(M/\mathbb{Q}) \to [-1, 1]$ to equal $1 - 2^{-|P|-1}$ for σ corresponding to the Frobenius class of the elements of $X_i(a, x_1)$, and otherwise equal to $-2^{-|P|-1}$. We are interested in bounding

$$\sum_{p \leq t_i'} F\left(\left[\frac{M/\mathbb{Q}}{p} \right] \right) \log p.$$

For any $c > 0$, Theorem 5.28 of [4] shows that we can choose A large enough so that

$$\beta < 1 - t^{-c}$$

whenever β is a Siegel zero of the L-function corresponding to some χ_D with $|D| < t$. Applying Proposition 7.3.1, we find that

$$x^\beta < t_i' \exp\left(-(\log t_i')^{1 - c c_6^{-1}} \right).$$

Since $\log |\Delta_{K_0}| = O\left((\log t_1')^2\right)$, we as obtain the bound

$$\exp\left(\frac{-cd^{-4} \log x}{\sqrt{\log x} + 3d \log |\Delta_{K_0} \Delta_E|} \right) \leq \exp\left(-(\log t_i')^{1/3 + \varepsilon} \right)$$

for some $\varepsilon > 0$. From $|G| \leq t_1'$ we then find the upper bound

$$\sum_{p \leq t_i'} F\left(\left[\frac{M/\mathbb{Q}}{p} \right] \right) \log p \leq t_i' \exp\left(-(\log t_i')^{1/3 + \varepsilon} \right)$$

for sufficiently large A. For $i \geq k$, we can always find

$$|X_i(a, x_1)| - |X_i|/2| < |X_i|/t_1'.$$

Write $X^{\text{bad}}(a)$ for the subset of $X(a)$ with $x_1 \in X_1^{\text{bad}}$. Choose $x_1 \in X^{\text{bad}}(a)$, and add it to $\mathcal{P}$, shifting its conditions from $\mathcal{M}$ to $\mathcal{M}_{\mathcal{P}}$. Then consider the product

$$X_2 \times \cdots \times X_k \times X_{k+1}(a, x_1) \times \cdots \times X_r(a, x_1).$$

This product now has length $r-1$, so the inductive hypothesis holds. Once we shift up k, it obeys the hypotheses of the proposition, hence the inductive step tells us that the subset of $X(a)$ starting with x_1' has size at most

$$2^{-|\mathcal{M}|+k+1}\frac{|X|}{|X_1|}.$$

Then $X^{\text{bad}}(a)$ has size bounded by

$$2^{k+1}\cdot(t_1')^{-c_a}\cdot 2^{-|\mathcal{M}|}|X|.$$

On the other hand, we look at the product

$$X_2(a,x_1)\times\cdots\times X_r(a,x_1)$$

and find that the subset of $X(a)$ starting with a good x_1 has size at most

$$2^{-|\mathcal{M}|}\frac{|X|}{|X_1|}\cdot\left(1+(r-1)(t_1')^{-c_8-c_1^{-1}}\right)\cdot\left(1+(t_1')^{-c_b}\right)^k\cdot\left(1+(t_1')^{-1}\right)^r$$

and at least

$$2^{-|\mathcal{M}|}\frac{|X|}{|X_1|}\cdot\left(1-(r-1)(t_1')^{-c_8-c_1^{-1}}\right)\cdot\left(1-(t_1')^{-c_b}\right)^k\cdot\left(1-(t_1')^{-1}\right)^r.$$

We have $k<c_7\log t_1'$, whence $\left(1+(t_1')^{-c_b}\right)^k$ is an error term from our lower bound on c_b. Similarly, the term $\left(1+(t_1')^{-1}\right)^r$ gives an error term whenever $1>c_8+2c_1^{-1}$, which is always satisfied by our hypothesis.

Finally, from our assumptions on c_a and (8.2), we find that the contribution from $X^{\text{bad}}(a)$ can be absorbed into the error term. This completes the proof. □

The lower bound assumed of t_1 in Proposition 8.0.3 are essential; we do not have sufficiently strong control over Jacobi symbols involving only small primes to give the equidistribution result we need. However, there is a combinatorial trick to get around this bad behaviour. We make the following definition:

Definition 8.0.4. Given $X=X_1\times\cdots\times X_r$ and a as in Definition 8.0.1, and given a permutation $\sigma:\{1,\cdots,r\}\to\{1,\cdots,r\}$, we put

$$X(\sigma,a)=(X_{\sigma_1}\times\cdots\times X_{\sigma(r)})(a).$$

Given $k_2\leq r$, define $\mathfrak{S}(k_2)$ to be the set of permutations of $\{1,\cdots,r\}$ that fix the points $i, i>k_2$.

The reason for this definition is twofold: first, the ordering of the primes do not affect the rank: hence, the class structure of a point in $X(\sigma,a)$ is independent of σ. Secondly, this has the effect of mixing the "bad" corner of a Legendre symbol matrix in with the rest.

We may now state and prove the following refinement of Proposition 8.0.3:

Theorem 8.0.5. *Let $c_1,\cdots,c_8$ be positive numbers satisfying (8.2), $c_9,\cdots,c_{12}$ positive numbers satisfying*

$$c_{10}\log 2+2c_{11}+c_{12}<1 \text{ and } c_{11}+c_{12}<c_9,$$

and $A,t,t_1,t_1',\cdots,t_r,t_r'$ be numbers satisfying (8.3). Let X_i be the set of prime numbers congruent to 1 mod 4 in the interval (t_i,t_i') and let X be their Cartesian

product. Let a be as in Definition 8.0.1. Choose non-negative integers k_0, k_1, k_2 satisfying

$$0 \le k_0 \le k_1 < k_2 \le r \text{ and } t'_{k_0+1} > t.$$

Suppose that Assumption (8.6) holds and $k_2 > A$. Put $\mathfrak{t} = t'_{k_0+1}$ and that the following hold:

(1) $p < \mathfrak{t}$ for all $p \in \mathcal{P}$;
(2) $\mathfrak{t} > r^{c_1}$ and $t'_{k_1} < \exp(\mathfrak{t}^{c_2})$;
(3) For $i > k_0$,

$$|X_i| \ge \frac{2^{|\mathcal{P}|+ic_3} \cdot k_2^{c_9} \cdot t'_i}{(\log t'_i)^{c_4}} \text{ and } |\mathcal{P}| \le \log t'_i i.$$

(4) If $k_1 \ne r$, then

$$t'_{k_1+1} > \exp\left((\log t'_1)^{c_5}\right), \exp\left(t^{c_6}\right);$$

(5) $k_1 - k_0 < c_7 \log \mathfrak{t}$ and that, for any $k_0 < i \le r$ and any j satisfying $i - 2 + c_7 \log t'_i \le j \le r$, we have

$$\exp\left((\log t'_i)^{c_5}\right) < t'_j;$$

(6) $c_{10} \log k_2 > |\mathcal{P}| + k_0$ and $c_{11} \log k_2 > \log k_1$.

Then, for any choice of $\mathcal{M}, \mathcal{M}_{\mathcal{P}}$ we have

$$\sum_{a \in \mathbb{F}_2^{\mathcal{M} \cup \mathcal{M}_{\mathcal{P}}}} \left| 2^{-|\mathcal{M} \cup \mathcal{M}_P|} \cdot k_2! \cdot |X| - \sum_{\sigma \in \mathfrak{S}(k_2)} |X(\sigma, a)| \right|$$
$$\le \left(\left(k_2^{-c_{12}} + \mathfrak{t}^{-c_8} \right) \cdot k_2! \cdot |X| \right).$$

Let $X, \mathcal{P}, \mathcal{M}, \mathcal{M}_{\mathcal{P}}$ be as in Definition 8.0.1. Assume that $\mathcal{M}, \mathcal{M}_{\mathcal{P}}$ are maximal given r and $\mathcal{P}$. Choose integers $0 \le k_0 \le k_1 \le k_2 \le r$ so that

$$2^{|\mathcal{P}|+k_0+1} \cdot k_1^2 < k_2.$$

For σ a permutation of $\{1, \cdots, r\}$ and a as in Definition 8.0.1, put $X_C(\sigma, a)$ for the set of $x = (x_1, \cdots, x_r) \in X$ so that

$$\left(\frac{d}{x_j}\right) = a\left(\left(\sigma^{-1}(j), d\right)\right) \text{ for all } (j, d) \in [k_1] \times \mathcal{P}$$

and

$$\left(\frac{x_i}{x_j}\right) = a\left(\left\{\sigma^{-1}(i), \sigma^{-1}(j)\right\}\right)$$

whenever $i, j \le k_1$, $\sigma^{-1}(i) \le \sigma^{-1}(j)$, and either $i \le k_0$ or $j \le k_0$.

Put m_C for the number of Legendre symbol conditions specified; that is,

$$m_C = k_1|\mathcal{P}| + \frac{1}{2}(k_0^2 - k_0) + k_0(k_1 - k_0).$$

The proof of Theorem 8.0.5 will follow from Proposition 8.0.3 and the following:

Proposition 8.0.6. *In the set-up above, for any $x \in X$ we have*

$$\sum_{a \in \mathbb{F}_2^{\mathcal{M} \cup \mathcal{M}_{\mathcal{P}}}} \left(2^{-m_C} \cdot k_2! - \{\sigma \in \mathfrak{S}(k_2) : x \in X_C(\sigma, a)\}\right)^2$$

$$\leq \frac{2^{|\mathcal{P}|+k_0+1} \cdot k_1^2}{k_2} \cdot 2^{-2m_C+|\mathcal{M}\cup\mathcal{M}_{\mathcal{P}}|} \cdot (k_2!)^2.$$

Proof. Put

$$W(a) = \{\sigma \in \mathfrak{S}(k_2) : x \in X_C(\sigma, a)\}.$$

We see that the average size of $W(a)$ over all a is $2^{-m_C} \cdot k_2!$, as the condition that $x \in X_C(\sigma, a)$ for a given x and σ is given by m_C binary conditions on a.

We now consider the average of $|W(a)|^2$. We see that $|W(a)|^2$ is the number of pairs of permutations (σ_1, σ_2) so that $x \in X_C(\sigma_1, a) \cap X_C(\sigma_2, a)$. Write $W(\sigma_1, \sigma_2)$ for the set of a so that x is in this intersection.

The maximal number of conditions on a in $W(\sigma_1, \sigma_2)$ is $2m_C$; a lower bound on the number of conditions depends on σ_1, σ_2. Let d_1 be the number of $i \in \{1, \cdots, r\}$ so that $\sigma_1^{-1}(i)$ and $\sigma_2^{-1}(i)$ are both at most k_1. Then we see that $W(\sigma_1, \sigma_2)$ is determined by at least

$$2m_C - d_1(|\mathcal{P}| + k_0)$$

conditions. Therefore

$$|W(\sigma_1, \sigma_2)| \leq 2^{-2m_C+d_1(|\mathcal{P}|+k_0)+|\mathcal{M}\cup\mathcal{M}_{\mathcal{P}}|}.$$

At the same time, the number of ways to choose a permutation τ of $\{1, \cdots, k_2\}$ so that

$$|\tau(\{1, \cdots, k_1\}) \cap \{1, \cdots, k_1\}| \geq d$$

for a given positive integer d is bounded by the number of ways to choose two cardinality d subsets from $\{1, \cdots, k_1\}$, a bijection between these sets, and a bijection between their complements in $\{1, \cdots, k_2\}$. This is bounded by

$$d! \cdot \binom{k_1}{d}^2 \cdot (k_2 - d)! \leq \left(\frac{k_1^2}{k_2}\right)^d \cdot k_2!.$$

The mean value of $|W(a)|^2$ is therefore bounded by

$$\sum_{d \geq 0} 2^{-2m_C+d(|\mathcal{P}|+k_0)} \left(\frac{k_1^2}{k_2}\right)^d \cdot (k_2!)^2.$$

Combining this with the average of $|W(a)|$ gives the desired conclusion. □

8.1. **Proof of Theorem 8.0.5.** Without loss of generality, we may assume that $\mathcal{M}, \mathcal{M}_{\mathcal{P}}$ are both maximal as in Proposition 8.0.6. We use the same notation for $m_C, X_C(\sigma, a)$. Further, we assume that $X_1, \cdots, X_{k_0}$ are singletons $x_1, \cdots, x_{k_0}$.

We prove the theorem by bounding

$$\sum_a \left| k_2! \cdot |X| - 2^{m_C} \cdot \sum_{\sigma \in \mathfrak{S}(k_2)} |X_C(\sigma, a)| \right|$$
$$+ \sum_{\sigma \in \mathfrak{S}(k_2)} \sum_a \left| 2^{m_C} \cdot |X_C(\sigma, a)| - 2^{|\mathcal{M}\cup\mathcal{M}_{\mathcal{P}}|} |X(\sigma, a)| \right|.$$

The former sum can be bounded via Proposition 8.0.6 by

$$k_2^{-c_{12}} \cdot |X| \cdot 2^{|\mathcal{M}\cup\mathcal{M}_{\mathcal{P}}|} \cdot k_2!.$$

For the second sum, fix a σ and a choice of a outside of the values referenced in the definition of $X_C(\sigma, a)$. There are then 2^{m_C} of a, and these a partition X into sets

$$X_C(\sigma, a) = \{x_1\} \times \cdots \times \{x_{k_0}\} \times X_{k_0+1}(a) \times \cdots \times X_{k_1}(a) \times X_{k_1+1} \times \cdots \times X_r,$$

with $X_i(a)$ the subset of X_i consistent with the choice of $\mathcal{P}$ and $x_1, \cdots, x_{k_0}$.

Given $k_0 < i \leq k_1$, the union of all $X_C(\sigma, a)$ for which

$$|X_i(a)| \leq \frac{|X_i|}{2^{|\mathcal{P}|+k_0} \cdot k_2^{c_9}}$$

has order at most $k_2^{-c_9}|X|$. Because of this, we can restrict the sum to be over only (σ, a) that do not satisfy this inequality at all such i, introducing an error with magnitude bounded by

$$k_1 k_2^{-c_9} \cdot |X| \cdot 2^{|\mathcal{M} \cup \mathcal{M}_{\mathcal{P}}|} \cdot k_2!.$$

Once restricted, each summand can be bounded by Proposition 8.0.3 and can be seen to be less than

$$\mathfrak{t}^{-c_8} \cdot 2^{m_C}|X|,$$

which gives the desired conclusion.

8.2. **4-narrow class groups of real quadratic fields behave like random matrices.** For the remainder of the paper, we put

$$D = \log\log\log N, D_1 = D^{(\log\log N)^{1/10}}, C_0 = C \cdot \log\log\log N, \tag{8.7}$$

where $C > 0$ is a fixed constant.

Definition 8.2.1. Let D, D_1, N, C_0 be as in (8.7) and let r be given by (7.2). We say that X is *acceptable* if it is Siegel-free and comfortably spaced above D_1, C_0-regular.

We will now prove the following theorem, which is critical as it provides the "seed" distribution necessary to run our Markov chain argument. Recall that for positive integers k_0, k_1, k_2 that $P(k_0; k_1, k_2)$ denotes the probability that a random $k_1 \times k_2$, where each of the entries is an independent Bernoulli random variable, has kernel of rank k_0.

Theorem 8.2.2. *For a positive number $N > 3$, put $\mathcal{S}(N)$ for the set of positive square-free integers up to N which can be written as a sum of two squares, so that $\mathfrak{S}(N) = |\mathcal{S}(N)|$. Then there exists $c > 0$ such that for all $n_1 \geq 0$ we have*

$$\left| \frac{|\mathcal{S}(N) \cap S^{n_1,\pm}(N)|}{\mathfrak{S}(N)} - \lim_{n\to\infty} P^{\mathrm{Sym}}(n_1; n+n_1, n+n_1) \right| = O\left((\log\log N)^{-c}\right).$$

Proof. By Propositions 7.1.3, 7.4.1, and 7.4.2 we find that it suffices to prove this on acceptable boxes in $S_r^*(D; N)$, with D larger than the largest bad prime of the number field associated to $S^{n_1,\pm}$.

We apply Proposition 8.0.5 to our acceptable box X, where k_0 is the smallest positive integer so that $\mathfrak{t} = t'_{k_0+1}$ is larger than D_1. Taking $t = D_1$, we have that the Siegel-less condition holds. Now choose k_1 minimal so that t'_{k_1+1} is larger than $\exp(D_1^{c_6})$, and take $k_2 = r$. Finally, let $\mathcal{P}$ be the set of all primes congruent to 1 mod 4 less than D, and take $\mathcal{M}, \mathcal{M}_{\mathcal{P}}$ maximal. We need to check that, at least for

sufficiently large N and some appropriate choice of $c_1, \cdots, c_{12}, C$, that the conditions of Proposition 8.0.5 hold. This is a simple numerical exercise.

Thus, considered up to permutation, the Legendre symbol matrices in our acceptable box X are equidistributed with error within the acceptable bound. Now, we apply an analogue of Swinnerton-Dyer's work in [8] to symmetric matricies. As in the proof of Corollary 6.11 in [6], a little work on the modified Markov process gives the necessary error estimates, showing the theorem. □

9. The Markov process: completion of the proof of Theorem 3

To prove Theorem 3, it suffices to prove

(9.1)
$$\left|\left|\mathcal{T}(N) \cap S^{n_2,\cdots,n_{m+1},\pm}(N)\right| - P(n_{m+1}, n_m, n_m - 1)\left|\mathcal{T}(N) \cap S^{n_2,\cdots,n_m,\pm}(N)\right|\right|$$
$$= O\left(N(\log N)^{-1/2}(\log\log\log\log N)^{-c_m}\right),$$

where $c_m > 0$ depends only on m.

We shall see that (9.1) will follow by applying Proposition 9.0.1, (9.1) is a consequence of applying Propositions 7.1.3 and 7.4.2 to Proposition 7.4.1, in conjunction with the following proposition; it moves the question from $\mathcal{T}(N)$ to boxes $S_r^*(D; N)$.

Proposition 9.0.1. *There exists $c > 0$ such that, for any quadratic field $\mathbb{Q}(\sqrt{n_0})$, there is some $A > 0$ so that the following holds: for D, D_1, N as in (8.7), r as in (7.2) and D larger than the largest prime factor of n_0, we have that for any X of $S_r^*(D; N)$ which is extravagantly spaced and Siegel-free above D_1, and $(\log\log\log N)^{1/2}$-regular. Then, for any $m \geq 1$ and any sequence $n_1 \geq \cdots \geq n_{m+1}$ of non-negative integers of the same parity, we have*

$$\left|\left|X \cap S^{n_1,\cdots,n_{m+1},\pm}(N)\right| - P(n_{m+1}, n_m, n_m + 1) \cdot \left|X \cap S^{n_1,\cdots,n_m,\pm}(N)\right|\right|$$
$$\leq A|X|(\log\log\log\log N)^{-c_m}.$$

The next step is to address the following issue: as $x \in X$ varies, the pairs $(w_a(x), w_b(x))$ corresponding to elements of $\mathfrak{cl}^\vee(K(x_0))[2] \times \mathfrak{cl}(K(x_0))[2]$ change. We aim to restrict our attention to sets $X(a)$, where this issue is no longer present. This reduction is cumbersome, since there are some choices of a where we are unable to find suitable sets of raw cocycles, governing expansions, or additive-restrictive systems.

We now consider X, N, n_0 as in Proposition 9.0.1. Suppose that the extravagant spacing that occurs in the hypothesis of the Proposition occurs between indices k_{gap} and $k_{\text{gap}} + 1$. Let $\mathcal{P}$ be the set of prime numbers less than D congruent to 1 mod 4. In the context of Definition 8.0.1, take $\mathcal{M}, \mathcal{M}_{\mathcal{P}}$ to be maximal, and let a be any function in $\mathbb{F}_2^{\mathcal{M} \cup \mathcal{M}_{\mathcal{P}}}$.

Under these assumptions, any $\overline{x} \in \overline{X}_{\{1,\cdots,r\}}$ entirely contained in $X(a)$ is then quadratically consistent (Definition 3.1.6). Take $\text{Art}_{(1)}, \cdots, \text{Art}_{(m-1)}$ to be a choice of lower pairings as in (5.4), and choose bases $w_{a,1}, \cdots, w_{a,n_1}$ and $w_{b,1}, \cdots, w_{b,n_1}$ as

in (5.6). We assume that $i_b > k_{\mathrm{gap}}$, and write S_{gap} for the union of the $S(j_1, j_2) - \{i_b\}$ that appear in Lemma 5.2.2. We assume that

$$S_{\mathrm{gap}} \subseteq [k_{\mathrm{gap}}/2, k_{\mathrm{gap}}].$$

Let $\mathcal{P}_{\mathrm{gap}}$ be an element of $\prod_{i \in \{1, \cdots, k_{\mathrm{gap}}\} - S_{\mathrm{gap}}} X_i$. We assume that a is consistent with this choice of $\mathcal{P}_{\mathrm{gap}}$. Next we write

$$X_i(a, \mathcal{P}_{\mathrm{gap}})$$

for the subset of X_i consistent with a and the data associated with $\mathcal{P}_{\mathrm{gap}}$, and put $X(a, \mathcal{P}_{\mathrm{gap}})$ for the subset of $X(a)$ which equals $\mathcal{P}_{\mathrm{gap}}$ on $\{1, \cdots, k_{\mathrm{gap}}\} - S_{\mathrm{gap}}$. Finally, given our choice of pairs $\mathrm{Art}_{(1)}, \cdots, \mathrm{Art}_{(k)}$, put

$$X(a, \mathcal{P}_{\mathrm{gap}}, k)$$

for the subset of $i(a, \mathcal{P}_{\mathrm{gap}})$ whose first k Artin pairings agree with the given sequence.

We shall refer to the above objects as the *initial data*. We then have:

Proposition 9.0.2. *There exists $c > 0$ for which the following holds: for the initial data as chosen above, writing*

$$n_{\max} = \left\lfloor \sqrt{c_m \log\log\log\log\log N} \right\rfloor$$

and assume that $n_{\max} \geq n$. Next we suppose that

$$|X_i(a, P_{\mathrm{gap}})| > 4^{-k_{\mathrm{gap}}} \cdot |X_i| \tag{9.2}$$

for $i \in S_{\mathrm{gap}}$. Let $\mathrm{Art}_{(m)}$ to be any $n_m \times n_m$ matrix with coefficients in $\mathbb{F}_2$. Then there is some number $A > 0$ depending on n_0 so that

$$\Big| |X(a, \mathcal{P}_{\mathrm{gap}}, m)| - 2^{-n_m(n_m+1)} \cdot |X(a, \mathcal{P}_{\mathrm{gap}}, m-1)| \Big|$$
$$\leq A \cdot |X(a, \mathcal{P}_{\mathrm{gap}})| \cdot (\log\log\log\log N)^{-c_m}.$$

We now have most of the ingredients to apply the structures which appear in Sections 3 and 4. The next thing to add is a set of governing expansions. To do so, we introduce some additional objects below.

Choose the initial data as above, which obeys the conditions of Proposition 9.0.2. For each $i \in S_{\mathrm{gap}}$, choose a subset $Z_i \subset X_i$. For each set $S(j_1, j_2)$ and data as in Definition 5.2.4, choose a set of governing expansions $\mathfrak{G}(i_a(j_1, j_2))$ on the product Z_{gap} of the Z_i's. For any set S of the form $S(j_1, j_2) - \{i_b\}$ and any $\overline{x} \in \overline{(Z_{\mathrm{gap}})}_S$, we assume that $\phi_{\overline{x}}(\mathfrak{G}(i_a(j_1, j_2)))$ exists.

For $x \in Z_{\mathrm{gap}}$, put $L(x)$ for the composition of all quadratic fields ramified only at ∞, the places of $\mathcal{P}$, and the places of $\mathcal{P}_{\mathrm{gap}}$. Put $M(j_1, j_2)$ for the composition of the fields of definition for the set of $\phi_{\overline{x}}$ with $\overline{x} \in \overline{(Z_{\mathrm{gap}})}_{S(j_1,j_2)-\{i_b\}}$ and $M_\circ(j_1, j_2)$ for the composition of the fields of definition for the set of $\phi_{\overline{x}}$ with $\overline{x} \in \overline{(Z_{\mathrm{gap}})}_S$ for some proper subset S of $S(j_1, j_2) - \{i_b\}$.

We assume that, for each $S(j_1, j_2)$, the field $M_\circ(j_1, j_2)/\mathbb{Q}$ splits completely at all primes in $\mathcal{P}, \mathcal{P}_{\mathrm{gap}}$, and in any Z_i with i outside of $S(j_1, j_2) - \{i_b\}$.

Next, take $\mathfrak{M}$ to be the composition of any $L(x)$ with the set of $M(j_1, j_2)$, and take $M_\circ$ to be the composition of any $L(x)$ with the set of $M_\circ(j_1, j_2)$. We write $X_i(M_\circ)$ to be the subset of primes in X_i so p is consistent with the choice of a and $\mathcal{P}_{\text{gap}}$, and the prime p splits completely in each $M_\circ(j_1, j_2)$. Note that $X_i(M_\circ)$ is described alternatively as the subset of X_i mapping under the Frobenius map to one specific central element of $\text{Gal}(M_\circ/\mathbb{Q})$. Finally, put

$$Z = \{\mathcal{P}_{\text{gap}}\} \times Z_{\text{gap}} \times \prod_{i > k_{\text{gap}}} X_i(M_\circ). \tag{9.3}$$

We shall denote the content above as the *starting data*.

Proposition 9.0.3. *Let the starting data be as above, and put*

$$M = \left\lfloor (\log\log\log\log N)^{1/5(m+1)} \right\rfloor,$$

with $M > 0$. We shall assume that $|Z_i| = M$ for all i.

Then there exists an absolute constant $c > 0$ and a number A depending only on n_0 so that

$$\Big| |Z \cap X(a, \mathcal{P}_{\text{gap}}, m)| - 2^{-n_m(n_m+1)} \cdot |Z \cap X(a, \mathcal{P}_{\text{gap}}, m-1)| \Big|$$
$$\leq A \cdot |Z \cap X(a, \mathcal{P}_{\text{gap}})| \cdot (\log\log\log\log N)^{-c_m}.$$

We shall now prove Proposition 9.0.3.

Proof of Proposition 9.0.3. Take F to be a non-zero multiplicative character of the vector space of $n_m \times n_m$ matrices with coefficients in $\mathbb{F}_2$. For $x \in Z \cap X(a, \mathcal{P}_{\text{gap}}, m-1)$, write $\text{Art}(x)$ for the Artin pairing on $\mathcal{D}_{(m)}^\vee \times \mathcal{D}_{(m)}$. To prove the proposition, it is enough to prove that

$$\sum_{x \in Z \cap X(a, \mathcal{P}_{\text{gap}}, m-1)} F(\text{Art}(x))$$
$$= O\left(|Z \cap X(a, \mathcal{P}_{\text{gap}})| \cdot (\log\log\log\log N)^{-c_m} \right)$$

for each F. Take $j_1 < j_2 \leq n_0$ so that F depends on the value of $\text{Art}(x)_{j_1 j_2}$, and take $S = S(j_1, j_2)$. From Proposition 3.1.9 we find that there is a natural bijection

$$\text{Gal}(M(j_1, j_2)M_\circ/M_\circ) \cong \mathfrak{C}_{S - \{i_b\}}(\pi_{S-\{i_b\}}(Z))$$

of $\mathbb{F}_2$-vector spaces, with our notation in Definition 6.0.2. For σ in this Galois group, we take $X_{i_b}(\sigma)$ to be the subset of $X_{i_b}(M_\circ)$ mapping under Frobenius to σ. From the Chebotarev density theorem, we find

$$|X_{i_b}(\sigma)| = 2^{-(M-1)^{m+1}} \cdot |X_{i_b}(M_\circ)| \cdot \left(1 + O\left(e^{-k_{\text{gap}}}\right)\right).$$

Choose $x_i \in X_i(M_\circ)$ for $i > k_{\text{gap}}$ and $i \neq i_b$ such that the set of x_i is consistent with a, writing this tuple as $\mathcal{P}_{\text{gap},+}$. From Propositions 7.3.2 and 8.0.3, we see that, outside of a negligible set of choices of $\mathcal{P}_{\text{gap},+}$, if we write $X_{i_b}(\mathcal{P}_{\text{gap},+})$ for the subset of X_{i_b} consistent with a, we have

$$|X_{i_b}(\sigma) \cap X_{i_b}(\mathcal{P}_{\text{gap},+})| \tag{9.4}$$
$$= 2^{-(M-1)^{m+1}} \cdot |X_{i_b}(M_\circ) \cap X_{i_b}(\mathcal{P}_{\text{gap},+})| \cdot (1 + O(e^{-k_{\text{gap}}}))$$

for each σ.

On the product

$$Z_{AR} = Z_{\text{gap}} \times (X_{i_b}(M_\circ) \cap X_{i_b}(\mathcal{P}_{\text{gap},+})),$$

we can find full additive-restrictive inputs as in Definition 5.2.4. The corresponding additive-restrictive system has abelian groups with orders bounded by

$$2^{n_{\max}(n_{\max}+2m+6)}.$$

We then apply Proposition 6.0.5 to the additive restrictive system $\mathfrak{A}(j_1, j_2)$. By Propositions 5.2.3 and 6.0.5, whenever

$$\varepsilon < 2^{-n_{\max}(n_{\max}+2m+6)} \tag{9.5}$$

and

$$\log M \geq A \cdot 6^{m+2} \log \varepsilon^{-1}, \tag{9.6}$$

there is a choice of $\sigma_1, \cdots, \sigma_M$ in $\text{Gal}(M(j_1, j_2)M_\circ/M_\circ)$ so that, for any σ in this Galois group and any choice of $Z'_{AR} = Z_{\text{gap}} \times \{x_1, \cdots, x_M\}$ with

$$x_i \in X_{i_b}(\sigma + \sigma_i) \cap X_{i_b}(\mathcal{P}_{\text{gap}})$$

for all $i \leq M$, we have

$$\sum_{x \in Z'_{AR}} F(\text{Art}(x)) \leq \varepsilon \cdot |Z'_{AR}|.$$

From (9.4), we see that Z_{AR} can be split into subproducts Z'_{AR} with the remainders that can be absorbed into the error term. Therefore, equidistribution applies to Z_{AR} as well. Choosing

$$\varepsilon = (\log\log\log\log\log N)^{\frac{-c}{(m+1)6^m}}$$

we find that ε satisfies (9.5) and (9.6). This completes the proof of this proposition, with

$$c_m = \frac{c}{(m+1)6^m}, \tag{9.7}$$

where we recall that c is an absolute constant. □

It remains to show how one uses Proposition 9.0.3 implies Proposition 9.0.1. indeed, we will see that

$$\text{Proposition 9.0.3} \Rightarrow \text{ Proposition 9.0.2} \Rightarrow \text{ Proposition 9.0.1.}$$

9.1. **Proposition 9.0.2 implies Proposition 9.0.1.** The argument is reduced to controlling bad types of pairs $(a, \mathcal{P}_{\text{gap}})$. First, we need to avoid the case where n is not less than $n_{\max}$. Second, we need to avoid those a where, for some choice of pairings, we cannot find suitable initial data for Proposition 9.0.2. Finally, we need to avoid those $(a, \mathcal{P}_{\text{gap}})$ for which (9.2) fails to hold for some $i \in S_{\text{gap}}$. We claim that the union of $X(a, \mathcal{P}_{\text{gap}})$ over all three kinds of bad pairs $(a, \mathcal{P}_{\text{gap}})$ can be absorbed into the error term of Proposition 9.0.1.

The assertion that the union of $X(a)$ for which $n_m \geq n_{\max}$ can be absorbed into the error term is a consequence of Theorem 8.2.2.

Next, consider those a such that, for some choice of pairings $\text{Art}_{(k)}$ and a basis, there is no suitable choice of $S(j_1, j_2)$ as in Lemma 5.2.2. We claim that the union

of $X(a)$ for such a can also be absorbed into the error term.

The proportion of a for which there are elements $w_1, w_2 \in \mathfrak{cl}(K(x_0))$ so that either w_1 or w_2 is non-zero, and

$$|(T_1(w_1) + T_2(w_2)) \cap [k_{\text{gap}}/2, k_{\text{gap}}]| > \left(\frac{1}{4} + 2^{-10n_{\max}}\right) \cdot k_{\text{gap}} \tag{9.8}$$

has density at most

$$O\left(\left(\frac{15}{16}\right)^r + \exp\left(2^{-20n_{\max}} \cdot k_{\text{gap}}\right)\right)$$

in the space $\mathbb{F}_2^{\mathcal{M} \cup \mathcal{M}_{\mathcal{P}}}$. Here, $T_1 + T_2$ denotes the symmetric difference.

Call a generic if there is no non-zero tuple w of $X(a)$ for which $T_1(w) + T_2(w) = \{1, \cdots, r\}$, and if there are no pairs of non-zero tuples (w_1, w_2) with $w_1 + w_2$ also non-zero, but $T_1(w_1) + T_2(w_1) + T_1(w_2) + T_2(w_2) = \{1, \cdots, r\}$. We then see, after some effort, that the conditions of Lemmas 4-6 in [8] are satisfied, so that are non-generic a due to the condition on w can be bounded by their consequences. It thus follows we obtain the bound

$$O\left(2^{2|\mathcal{P}|} \cdot (3/4)^r\right).$$

For the condition on (w_1, w_2), we can use Lemma 7 in [8] instead, after noting that the condition $u_1' = u_2''$ can be weakened to $u_1'/u_2'' \in X_{\mathcal{S}}$, in the notation of [8], without any change. Therefore we conclude that the number of non-generic a is bounded by

$$A^{|\mathcal{P}|} \cdot \left(\frac{15}{16}\right)^r$$

for some absolute constant $A > 1$.

We may now suppose that w is a generic tuple. We can then conclude, from genericity, that the local conditions at the r primes coming from X are independent, and so the proportion of a where w is an admissible tuple for $X(a)$ is bounded by $O\left(4^{-r}\right)$. Similarly, if (w_1, w_2) is generic as above, the probability that w_1 and w_2 are both admissible for $X(a)$ is bounded by $O\left(4^{-2r}\right)$ by independence.

Hoeffding's inequality is sufficient to complete the estimate of the density of $a \in \mathbb{F}_2^{\mathcal{M} \cup \mathcal{M}_{\mathcal{P}}}$ not satisfying (9.8) for some w_1, w_2.

For a other than those in this set, it is easy to find sets admissible indices $S(j_1, j_2)$ if $n_{\max}$ is sufficiently large in terms of n_0. Choose $i_b > k_{\text{gap}}$, so that i_b is not in $T_i(w_j)$ for any $i = 1, 2$ and j. Then each $S(j_1, j_2) - \{i_b, i_a(j_1, j_2)\}$ can be taken to be an arbitrary subset of size m inside of

$$T_2(w_{j_2}) \cap (\{1, \cdots, r\} - T_1(w_{j_2})) \cap \bigcap_{j \neq j_2} (\{1, \cdots, r\} - (T_1(w_j) \cup T_2(w_j))).$$

The assumptions on a give that this intersection has density about 4^{-n_1} on the integers in the interval $[k_{\text{gap}}/2, k_{\text{gap}}]$, which will be larger than m for sufficiently large $n_{\max}$.

If $k_2 < k_{\text{gap}}/2$, we see that permuting the first k_2 indices do not change whether (9.8) holds for a given a. Then, by Proposition 8.0.5, we find that the union of $X(a)$ over all a for which it may be impossible to find a set of acceptable $S(j_1, j_2)$ can be absorbed into the error of Proposition 9.0.1.

Finally, we claim that the union of $X(a, \mathcal{P}_{\text{gap}})$ over all $(a, \mathcal{P}_{\text{gap}})$ for which (9.2) does not hold for some i can be absorbed into the error term of Proposition 9.0.1. We will work in the context of Proposition 8.0.3. To do this, add the primes $p_1, \cdots, p_k$ of the box to the set P; taking $X_i(a, \mathcal{P})$ to be the subset of X_i consistent with $\mathcal{P}$ an the choice of a. We attempt to apply the argument of Proposition 8.0.3 to

$$X_1(a, \mathcal{P}) \times \cdots \times X_r(a, \mathcal{P}).$$

This will work only if no $X_i(a, \mathcal{P})$ is smaller than $\frac{|X_i|}{(\log t_1')^{c'}}$ for some choice of c'. For such a choice, outside a set of choices of a over which the union of the $X(a)$'s can be absorbed into the error term of Proposition 9.0.1, we always have

$$|X_i(a, \mathcal{P})| \geq \frac{|X_i|}{(\log t_1')^{c'}}.$$

Suppose we have such an a. Then a choice of $\mathcal{P}_{\text{gap}}$ for which (9.2) does not hold would be exceptional in the sense of the proof of Proposition 8.0.3. In that proof, the union of all such exceptional can be seen to be absorbed into the error term of Proposition 9.0.1.

We now note that there are at most $2^{mn_{\max}^2}$ sequences of pairings $\text{Art}_{(k)}$. The conclusion of Proposition 9.0.2 then implies

$$\left|\left|X(a, \mathcal{P}_{\text{gap}}) \cap S^{n_1, \cdots, n_{m+1}, \pm}(N)\right| - P(n_{m+1}; n_m, n_m + 1) \cdot \left|X(a, \mathcal{P}_{\text{gap}}) \cap S^{n_1, \cdots, n_m, \pm}(N)\right|\right|$$
$$\leq A \cdot 2^{mn_{\max}^2} \cdot |X(a, \mathcal{P}_{\text{gap}})| \cdot (\log\log\log\log N)^{-c_m}.$$

A routine calculation shows that the sum of this error over all a and $\mathcal{P}_{\text{gap}}$ is then acceptable for the error term of Proposition 9.0.1.

9.2. **Proposition 9.0.3 implies Proposition 9.0.2.** Choose *initial data* as right before the statement of Proposition 9.0.2. Write V_{gap} for the subset of $\prod_{i \in S_{\text{gap}}} X_i$ consistent with $\mathcal{P}_{\text{gap}}$ and the conditions of a. Take

$$R = \lfloor \exp\exp(k_{\text{gap}}/5) \rfloor,$$

and assume that $R > 0$. We choose $t \geq 0$, and for each $i \in S_{\text{gap}}$, we choose sequences of subsets

$$Z_i^1, \cdots, Z_i^t \subseteq X_i(a, \mathcal{P}_{\text{gap}})$$

each with cardinality M. We take

$$Z_{\text{gap}}^{\ell} = \prod_{i \in S_{\text{gap}}} Z_i^{\ell}.$$

We suppose that these subsets satisfy the following:

- For $\ell \neq \ell'$, we have $|Z_{\text{gap}}^{\ell} \cap Z_{\text{gap}}^{\ell'}| \leq 1$;
- Each Z_{gap}^{ℓ} is a subset of V_{gap} and any point in V_{gap} is in at most R of the Z_{gap}^{ℓ};
- The set Z_{gap}^{ℓ} can be used as *starting data* for Proposition 9.0.3.

Furthermore, we assume that the sequence of Z^ℓ_{gap} cannot be extended under these requirements to a sequence of $t+1$ subgrids.

Write

$$X_{\mathrm{gap}} = \prod_{i \in S_{\mathrm{gap}}} X_i(a, \mathcal{P}_{\mathrm{gap}}).$$

Take $V^{\mathrm{bad}}_{\mathrm{gap}}$ to be the set of points in V_{gap} that are consistent with the choice of a and $\mathcal{P}_{\mathrm{gap}}$ and that are in fewer than R of the Z^ℓ_{gap}. Write δ for the density of $V^{\mathrm{bad}}_{\mathrm{gap}}$ in X_{gap}. By a greedy algorithm, one can choose a subset W of $V^{\mathrm{bad}}_{\mathrm{gap}}$ of density at least δ/RM^{m+1} such that no point in W is in more than two of the Z^ℓ_{gap}.

By adjoining splitting behaviour at the primes in $\mathcal{P}_{\mathrm{gap}}$ to the system constructed in Proposition 5.1.1, we can then define an additive-restrictive system on X_{gap} with $\overline{Y}^\circ_\emptyset = W$ and where, if $\overline{x} \in \overline{Y}^\circ_{S_{\mathrm{gap}}}$, then the governing expansions defined at $\overline{x}$ are as required for Proposition 9.0.3. The maximal size of the abelian groups in this additive-restrictive system is bounded by $2^{k_{\mathrm{gap}}+|\mathcal{P}|}$. Then, by Proposition 5.0.2, the density of $\overline{Y}^\circ_{S_{\mathrm{gap}}}$ in $X_{\mathrm{gap}} \times X_{\mathrm{gap}}$ is at least

$$\left(\frac{\delta}{2^{k_{\mathrm{gap}}|\mathcal{P}|} \cdot RM^{m+1}}\right)^{3^{|S_{\mathrm{gap}}|}}.$$

We note $|S_{\mathrm{gap}}| \leq (m+1)n_0^2$. In addition, for sufficiently large N, we always have

$$|X_i(a, \mathcal{P}_{\mathrm{gap}})| > \exp\exp(0.3 \cdot k_{\mathrm{gap}})$$

for $i \in S_{\mathrm{gap}}$. Applying Proposition 6.0.1 and the assumptions on t, we then have

$$M^{2m} > \frac{\exp(0.3 \cdot k_{\mathrm{gap}})}{(m+1)^{3(m+1)n_m^2} \cdot (\exp(k_{\mathrm{gap}}/4) + \log \delta^{-1})}$$

for sufficiently large N. We can then bound δ by $\exp(-\exp(k_{\mathrm{gap}}/4))$ for sufficiently large N. Then, following the proof of Proposition 8.0.3, we see that the subset of $x \in X(a, \mathcal{P}_{\mathrm{gap}})$ for which $\pi_{S_{\mathrm{gap}}}(x)$ is in $V^{\mathrm{bad}}_{\mathrm{gap}}$ can be absorbed into the error term in Proposition 9.0.2.

We associate grids Z^ℓ_{gap} with fields M^ℓ and $M^\ell_\circ$ and a supergrid Z^ℓ as in (9.3). For $x \in X(a, \mathcal{P}_{\mathrm{gap}})$ with $\pi_{S_{\mathrm{gap}}}(x)$ outside of $V^{\mathrm{bad}}_{\mathrm{gap}}$, write $\Theta(x)$ for the number of $\ell \leq t$ for which x is in Z^ℓ. Write d_{ML} for the degree of M^ℓ over some $L(x)$ with $x \in Z^\ell_{\mathrm{gap}}$; from Proposition 3.1.9, we find this degree is independent of x and ℓ. For $i > k_{\mathrm{gap}}$, put $X_i(L(x))$ for the subset of $X_i(a, \mathcal{P}_{\mathrm{gap}})$ consistent with the choice of x. By Proposition 7.3.1 and the definition of extravagant spacing, we then have

$$|X_i(M^\ell_\circ)| = d_{ML}^{-1} \cdot |X_i(L(x))|(1 + O(\exp(-2k_{\mathrm{gap}})))$$

for $i > k_{\mathrm{gap}}$. Proposition 8.0.3 then gives that the subset of

$$\prod_{i > k_{\mathrm{gap}}} X_i(M^\ell_\circ)$$

consistent with a has order

$$d_{ML}^{-(r-k_{\mathrm{gap}})} \cdot |X(a, \mathcal{P}_{\mathrm{gap}} \cap \pi^{-1}_{S_{\mathrm{gap}}}(x)| \cdot (1 + O(\exp(-k_{\mathrm{gap}}))).$$

From this, we see that $\Theta(x)$ has average order

$$d_{ML}^{-(r-k_{\text{gap}})} R \cdot (1 + O(\exp(-k_{\text{gap}}))).$$

Similarly, from the requirements on $Z^\ell \cap Z^{\ell'}$ and Proposition 3.1.9, we see that $M_\circ^\ell M_\circ^{\ell'}$ has degree d_{ML}^2 over $L(x)$ for $Z^\ell, Z^{\ell'}$ distinct grids containing x. Then the average order of $\Theta(x)$ is seen to be

$$d_{ML}^{-2(r-k_{\text{gap}})} \cdot R^2 (1 + O(\exp(-k_{\text{gap}}))).$$

Thus, outside of a set of density $O(\exp(-k_{\text{gap}}/2))$ in the support of Θ, we find that $\Theta(x)$ over the mean value of Θ is within $\exp(-k_{\text{gap}}/4)$ of 1. The effect of the set of low density can be absorbed into the error term of Proposition 9.0.2, and the variance between $\Theta(x)$ can also be absorbed into the error term. Proposition 9.0.2 then follows from Proposition 9.0.3.

References

[1] M. Bhargava, I. Varma, *On the mean number of 2-torsion elements in the class groups, narrow class groups, and ideal groups of cubic orders and fields*, Duke. Math. J., **164** (2015), 1911-1933.

[2] E. Fouvry, J. Klüners, *On the 4-rank of class groups of quadratic number fields*, Invent. Math. **167** (2007), 455-513.

[3] E. Fouvry, J. Klüners, *On the negative Pell equation*, Ann. of. Math. (3) **172** (2010), 2035-2104.

[4] H. Iwaniec, E. Kowalski, *Analytic Number Theory*, American Mathematical Society Colloquium Publications, **53**, Providence, RI, 2004.

[5] M. Jutila, *On mean values of Dirichlet polynomials with real characters*, Acta. Arith. **27** (1975), 1253-1279.

[6] A. Smith, 2^∞*-Selmer groups,* 2^∞*-class groups, and Goldfelds' conjecture*, arXiv:1702.02325 [math.NT].

[7] P. Stevenhagen, *The Number of Real Quadratic Fields having Units of Negative Norm*, Experiment. Math. **2** (1993) 121-136.

[8] P. Swinnerton-Dyer, *The effect of twisting on the 2-Selmer group*, Math. Proc. Cambridge. Philo. Soc. **145** (2008), 513-526.

[9] C. Tsang, S. Y. Xiao, *Binary quartic forms with bounded invariants and small Galois groups*, to appear in Pac. J. Math.

Department of Mathematics, University of Toronto, Toronto, BA6256
E-mail address: eknight@math.toronto.edu

Department of Mathematics, University of Toronto, Toronto, HU 1015
E-mail address: syxiao@math.toronto.edu