ON THE 4-RANK OF CLASS GROUPS OF QUADRATIC NUMBER FIELDS

DJORDJO MILOVIC

ADVISOR: ÉTIENNE FOUVRY, UNIVERSITÉ PARIS-SUD 11

Abstract. We give an overview of the work of É. Fouvry and J. Klüners on the 4-rank of quadratic number fields. We give a slight generalization of their work by proving that the Cohen-Lenstra conjectures for the 4-rank of quadratic number fields hold when restricted to fields with discriminants in arithmetic progressions, and hence hold with specified splitting conditions at a finite number of primes. We also prove a Siegel-Walfisz-type theorem for the 4-rank in arithmetic progressions.

1. Introduction

Let d be a square-free rational integer and let $K = \mathbb{Q}(\sqrt{d})$ be a quadratic extension of $\mathbb{Q}$ with discriminant D, equal to d or $4d$ depending on whether $d \equiv 1 \pmod 4$ or $d \equiv 2$ or $3 \pmod 4$. Let $\mathcal{O}$ be the ring of integers of K. A fractional ideal of $\mathcal{O}$ is an $\mathcal{O}$-submodule $\mathfrak{a}$ of K such that there exists $\alpha \in K^\times$ satisfying $\alpha\mathfrak{a} \subset \mathcal{O}$. Every non-zero fractional ideal $\mathfrak{a}$ of $\mathcal{O}$ has a multiplicative inverse

$$\mathfrak{a}^{-1} = \{\alpha \in K : \ \alpha\mathfrak{a} \subset \mathcal{O}\},$$

which makes the set of non-zero fractional ideals of $\mathcal{O}$ a group under multiplication. We denote this group by $\mathrm{Div}(K)$, and we call its elements divisors.

Let $\mathcal{P}$ denote the subgroup of principal divisors, i.e. divisors of the from

$$(\beta) = \{\beta\alpha : \ \alpha \in \mathcal{O}\},$$

where $\beta \in K^\times$. Geometry of numbers can be used to show that $\mathrm{Div}(K)/\mathcal{P}$ is a finite group, called the class group. We will use the notation $\mathrm{Cl}_D = \mathrm{Div}(K)/\mathcal{P}$ to emphasize its dependence on D.

A closely related group is the narrow class group. Let $\mathcal{P}^+$ be the subgroup of $\mathcal{P}$ consisting of principal divisors that can be generated by a totally positive element, i.e. principal divisors of the form (β) with $\sigma(\beta) > 0$ for every real embedding $\sigma : K \hookrightarrow \mathbb{R}$. Note that in the case of quadratic number fields, a divisor can be generated by a totally positive element if and only if it can be generated by an element of positive norm. We define the narrow class group to be $\mathrm{C}_D = \mathrm{Div}(K)/\mathcal{P}^+$.

Class groups of quadratic number fields have been studied for a long time, and there are many very interesting (and very difficult) problems concerning their behavior. A particular quantity of interest is, for any prime p, the p-rank of a class group Cl_D, defined as $rk_p(\mathrm{Cl}_D) = \dim_{\mathbb{F}_p}(\mathrm{Cl}_D/\mathrm{Cl}_D^p)$ (note the multiplicative notation). We can similarly define, for any positive integer k, the p^k-rank of an abelian group G by setting $rk_{p^k}(G) = rk_p(G^{p^{k-1}})$.

In this paper, we will first present what is commonly called Gauss's genus theory, which describes the 2-rank of the narrow class group. Next, we will briefly discuss the heuristics of H. Cohen and H. W. Lenstra as well as F. Gerth's extension thereof, which are a collection of conjectures about the asymptotic behavior as $D \to \pm\infty$ of various quantities associated to class groups. One such quantity is the q-rank, where $q = p$ if p is an odd prime or $q = 4$.

In [FK07], É. Fouvry and J. Klüners prove the conjectures about the 4-rank. The proof involves first expressing the 4-rank of the narrow class group in a form friendly to the methods of analytic number theory and then handling the main and error terms with clever analytic and combinatorial arguments. Fouvry and Klüners have adapted their methods to obtain many interesting results. Perhaps most notably, in [FK10] they make first significant progress toward a conjecture of Stevenhagen [Ste93]. This work suggests that their methods are very robust. In this paper, we will give another example of the robustness of their methods by studying the 4-rank for quadratic number fields with discriminants in arithemetic progressions. We first modify their proof slightly to show that the Cohen-Lenstra conjecures still hold when restricted to arithmetic progressions. More precisely, we prove

Theorem 1. *Let $k \geq 1$, let $\varepsilon > 0$. Define $R(X, \varepsilon, k) = X(\log X)^{-2^{-k}+\varepsilon}$. Let $q \geq 1$ be odd and squarefree and let $(a, q) = 1$. Finally, let $\mathcal{N}(k, 2)$ denote the number vector subspaces of $\mathbb{F}_2^k$. Then*

$$\sum_{\substack{-X<D<0 \\ D\equiv 1 \ (mod\ 4) \\ D\equiv a \ (mod\ q)}} 2^{k rk_4(Cl_D)} = \mathcal{N}(k,2)\left(\sum_{\substack{-X<D<0 \\ D\equiv 1 \ (mod\ 4) \\ D\equiv a \ (mod\ q)}} 1\right) + O_{\varepsilon,k,q}(R(X,\varepsilon,k)). \tag{1}$$

The case $q = 1$ recovers the main result in [FK07].

Remark. Specifying how a finite number of non-ramified primes $p_1, \ldots, p_m$ split in a quadratic extension $\mathbb{Q}(\sqrt{d})$ is equivalent to specifying the quadratic character of D modulo each of the primes, which in turn is equivalent to specifying the congruence class of D modulo the product $p_1 \cdots p_m$. Hence, Theorem 1 shows that the Cohen-Lenstra heuristics are stable under restricting to quadratic fields with specified splitting conditions at a finite number of unramified primes, thus positively answering a question posed by Melanie Wood.

Remark. Note that for $D < 0$, the narrow class group coincides with the ordinary class group, so it suffices to prove Theorem 1 with C_D in place of Cl_D. In case $D > 0$, it is possible that the 4-ranks of Cl_D and C_D differ by 1, but this happens for only $\ll X/\sqrt{\log X}$ discriminants, and so in this case it would also suffice to prove the statement for C_D instead of Cl_D. For details, see [FK07, Lemma 10 and Corollary 1].

Finally, we extend the study of the asymptotics of the 4-rank in arithmetic progressions by proving a Siegel-Walfisz -type result for the first moment of the 4-rank.

Theorem 2. *Let $q \geq 1$ be odd and squarefree and let $(a, q) = 1$. Then for any $C > 0$, we have*

$$\sum_{\substack{-X<D<0 \\ D\equiv 1 \ (mod\ 4) \\ D\equiv a \ (mod\ q)}} 2^{rk_4(Cl_D)} = \frac{1}{\varphi(q)} \sum_{\substack{-X<D<0 \\ D\equiv 1 \ (mod\ 4) \\ (D,q)=1}} 2^{rk_4(Cl_D)} + O_{C,q}(X(\log X)^{-C}). \tag{2}$$

Remark. In the statements above, we restrict the sum to the set of discriminants that are negative and congruent to 1 modulo 4. Analogous statements hold true when we restrict to discriminants that are positive and/or congruent to 0 or 4 modulo 8. However, the formulas for the 4-rank in these cases become slightly more complicated and the subsequent arguments require casework that does not shed any additional light on the methods involved. Hence, for the sake of simplicity, we do not treat these other cases. A careful treatment of these cases can be found in [FK07].

2. Gauss's genus theory

Gauss's genus theory gives an explicit answer for the 2-rank of a class group of a quadratic number field. Here we summarize the argument from [Has80], both because it is rather simple and because it uses Dirichlet's theorem on primes in arithmetic progressions in a key way. The result somewhat naturally follows from attempting to determine when a divisor is a principal divisor. If a divisor $\mathfrak{b} = (\beta) = \prod_i \mathfrak{p}_i^{e_i}$ is principal, then the absolute norm $N(\mathfrak{b}) = \prod_i N(\mathfrak{p}_i)^e_i = \prod_i (\#(\mathcal{O}/\mathfrak{p}_i)^{e_i})$ is equal to $\pm \mathrm{N}_{K/\mathbb{Q}}(\beta)$. The converse is not true; that is, if $N(\mathfrak{b})$ is equal to $\pm \mathrm{N}_{K/\mathbb{Q}}(\beta)$ for some β in K, it need not be the case that $\mathfrak{b} = (\beta)$. Nonetheless, we can study the set (in fact the group) of divisors whose absolute norms are norms of elements of K, up to sign. This group certainly contains the principal divisors and hence might tell us something about the class group.

In light of our new goal, we introduce the Hilbert symbols (a, b) and $(a, b)_v$ for non-zero rational numbers a and b and a place v for $\mathbb{Q}$ (i.e., $v = p$ for some rational prime p or $v = \infty$). We set $(a, b) = 1$ if b is the norm of an element of $\mathbb{Q}(\sqrt{a})$ and -1 otherwise; similarly, set $(a, b)_v = 1$ if b is the norm of an element of $\mathbb{Q}_v(\sqrt{a})$ and -1 otherwise. In other words, $(a, b) \in \{\pm 1\}$ (or $(a, b)_v \in \{\pm 1\}$) and is equal to 1 if and only if the quadratic form $x^2 - ay^2 - bz^2$ represents 0 in $\mathbb{Q}$ (or $\mathbb{Q}_v$), that is, if and only if there exists a non-zero $(x, y, z) \in \mathbb{Q}^3$ (or $\mathbb{Q}_v^3$) such that $x^2 - ay^2 - bz^2 = 0$. This second way of viewing the Hilbert symbol is convenient because Hasse-Minkowski's theorem (see [Ser73]) now implies that b is a norm in $\mathbb{Q}(\sqrt{a})$ if and only if b is a norm in $\mathbb{Q}_v(\sqrt{a})$ for each place v. We now state some properties of the Hilbert symbol that we will need in our analysis. Their proofs can be found in [Has80].

Theorem 3. *Let $a, b \in \mathbb{Q}^\times$. Then*
(i) $\prod_v (a, b)_v = 1$.
(ii) If p is an odd prime, $p \nmid a$, and $p \nmid b$, then $(a, b)_p = 1$.
(iii) If p is an odd prime, $p \nmid a$, and $p | b$, then $(a, b)_p = \left(\frac{a}{p}\right)$.

(iv) $(a,b)_\infty = 1$ *unless both* a *and* b *are negative.*
(v) If $2|a$ *and* p *is a prime conguent to* 1 *modulo* 8*, then* $(a,p)_2 = 1$.
(vi) If d *is an even fundamental discriminant and* d_2 *is the exact power of* 2 *dividing* d*, then for any odd integer* b*,* $(d,b)_2$ *depends only on the congruence class of* b *modulo* d_2*; moreover, there exist congruence classes* b_- *and* b_+ *modulo* d_2 *such that* $(d,b_-)_2 = -1$ *and* $(d,b_+)_2 = 1$.
(vii) The local Hilbert symbol is multiplicative in each of the two arguments.

Remark. Part (i) above is called the product formula and is essentially a consequence of the quadratic reciprocity law. An interesting consequence of the product formula is that in order to prove that b is a norm in $\mathbb{Q}(\sqrt{a})$, it suffices to show that b is a norm in $\mathbb{Q}_v(\sqrt{a})$ for all but one place v.

Since $\mathcal{O}$ is a Dedekind domain, for any divisor $\mathfrak{b}$, there is a unique decompostion into prime ideals $\mathfrak{b} = \prod_p \mathfrak{p}^{b_p}\mathfrak{p}'^{b'_p} \prod_q \mathfrak{q}^{b_q} \prod_r \mathfrak{r}^{b_r}$; here the b_i are integers, the product over p is over the primes that decompose, the product over q is over the inert primes, the product over r is over the ramified primes, and $\mathfrak{p}$ and $\mathfrak{p}'$ are conjugate. Hence $N(\mathfrak{b}) = \prod_p p^{b_p+b'_p} \prod_q q^{2b_q} \prod_r r^{b_r}$. By part (iii) of Theorem 3, we have, for each p and q,

$$(d, N(\mathfrak{b}))_p = \left(\frac{d}{p}\right)^{b_p+b'_p} = 1$$

and

$$(d, N(\mathfrak{b}))_q = \left(\frac{d}{q}\right)^{2b_q} = 1.$$

Moreover, since $N(\mathfrak{b}) > 0$, part (iv) of Theorem 3 implies that $(d, N(\mathfrak{b}))_\infty = 1$. We now see that there exists $\beta \in \mathbb{Q}(\sqrt{d})$ such that $N(\mathfrak{b}) = \mathrm{N}_{K/\mathbb{Q}}(\beta)$ if and only if $(d, N(\mathfrak{b}))_r = 1$ for every ramified prime r. The set of $\mathfrak{b}$ whose norms are element norms is called the narrow principal genus and denoted by G^+:

$$G^+ = \left\{\mathfrak{b} \in \mathrm{Div}(K) | \ \exists\ \beta \in K \text{ such that } N(\mathfrak{b}) = \mathrm{N}_{K/\mathbb{Q}}(\beta)\right\}.$$

Clearly G^+ is a group under multiplication with $\mathcal{P}^+$ as a subgroup. However, in case that the class group differs from the narrow class group, $\mathcal{P}$ is not contained in G^+; for this reason, it is simpler to first study the 2-rank of the narrow class group and then make modifications to obtain the results for the class group.

Let $t = \omega(D)$, the number of distinct prime divisors of D, and let $\mathfrak{r}_1, \dots, \mathfrak{r}_t$ denote the prime ideals of $\mathcal{O}$ lying above the ramified primes $r_1, \dots, r_t$. For each $1 \le i \le t$, we may define a character

$$\chi_i : \mathrm{Div}(K)/G^+ \to \{\pm 1\}$$

by $\chi_i(\mathfrak{b}) = (D, N(\mathfrak{b}))_{r_i}$. This is indeed a character by part (vii) of Theorem 3. By the analysis above, we conclude that the group homomorphism

$$\chi = (\chi_1, \dots, \chi_t) : \mathrm{Div}(K)/G^+ \to \{\pm 1\}^t$$

is injective. Let T denote the set of t-tuples $(e_1, \dots, e_t) \in \{\pm 1\}^t$ such that $\prod_i e_i = 1$. By Hilbert's product formula, the image of χ injects into T. The following is called the main theorem on genera:

Theorem 4. *The group homomorphism* $\chi : Div(K)/G^+ \to T$ *is an isomorphism.*

Proof. It remains to show that χ is surjective. Suppose that $(e_1, \ldots, e_t) \in \{\pm 1\}^t$ with $\prod_i e_i = 1$. We wish to show that there is a divisor $\mathfrak{b}$ such that $(D, N(\mathfrak{b}))_{r_i} = e_i$ for each $1 \leq i \leq t$. By parts (iii) and (vi) of Theorem 3, the symbol $(D, b)_{r_i}$ depends only on the residue class of b modulo D_{r_i}, where $D_{r_i} = r_i$ if r_i is odd, $D_2 = 4$ (resp. 8) if D is congruent to 4 (resp. 0) modulo 8. Parts (iii) and (vi) of Theorem 3 ensure that there exist (invertible) congruence classes b_{r_i} modulo D_{r_i} such that $(D, N(\mathfrak{b}))_{r_i} = e_i$. There congruence classes determine (an invertible) congruence class b modulo D, and now Dirichlet's theorem on primes in arithmetic progressions ensures that there is a prime p congruent to b modulo D. If D is odd, we may choose p to also be congruent to 1 modulo 8.

By construction, we have that $(D, p)_{r_i} = e_i$ for each $1 \leq i \leq t$. By parts (ii) and (v) of Theorem 3, we deduce that $(D, p)_{p'} = 1$ for all primes p' not equal to p or one of the r_i. Finally, by part (i) of Theorem 3 and the assumption that $\prod_i e_i = 1$, we deduce that also $(D, p)_p = 1$. Part (iii) of Theorem 3 now implies that $\left(\frac{D}{p}\right) = 1$, so that p decomposes in $\mathbb{Q}(\sqrt{d})$. This means that there exists a prime ideal $\mathfrak{p}$ of $\mathcal{O}$ such that $N(\mathfrak{p}) = p$. Hence $(D, N(\mathfrak{p}))_{r_i} = e_i$ for all i. □

Theorem 4 shows that 2^{t-1} divides the order of the narrow class group C_D. To finish the proof that the 2-rank of C_D is $t-1$, it suffices to show that $G^+/\mathcal{P}^+ = \mathrm{C}_D^2$. We write $\mathfrak{b}^s$ for the conjugate of a divisor $\mathfrak{b}$. First note that if $c = [\mathfrak{b}]$ is a narrow class, then $c^{1+s} = c \cdot c^s = [(N(\mathfrak{b}))] = 1$. Hence $c^{1-s} = c^{1-s} \cdot c^{1+s} = c^2$, which means that $\mathrm{C}_D^2 = \mathrm{C}_D^{1-s}$. We will now prove that $G^+/\mathcal{P}^+ = \mathrm{C}_D^{1-s}$.

First, for any divisor $\mathfrak{b}$, $N(\mathfrak{b}^{1-s}) = N(\mathfrak{b})/N(\mathfrak{b}^s) = 1$, so that $\mathfrak{b}^{1-s}$ belongs to G^+. It remains to show that other inclusion. So suppose that $\mathfrak{b}$ belongs to G^+. Then $N(\mathfrak{b}) = \mathrm{N}_{K/\mathbb{Q}}(\beta)$ for some non-zero $\beta \in \mathbb{Q}(\sqrt{d})$. As $N(\mathfrak{b}) > 0$, so also $\mathrm{N}_{K/\mathbb{Q}}(\beta) > 0$, which means that $\mathfrak{b}$ and $\mathfrak{b}/\beta$ define the same narrow class. But $N(\mathfrak{b}/\beta) = 1$, so we must have $\mathfrak{b}/\beta = \prod_p \mathfrak{p}^{b_p} \mathfrak{p}'^{-b_p}$ where the product is over the primes p which decompose in $\mathbb{Q}(\sqrt{d})$. Since $\mathfrak{p}' = \mathfrak{p}^s$, we deduce that

$$[\mathfrak{b}] = [\mathfrak{b}/\beta] = \left[\prod_p \mathfrak{p}^{b_p}\right]^{1-s}$$

in the narrow sense. This concludes the proof that $rk_2(\mathrm{C}_D) = \omega(D) - 1$.

For class groups, the answer is a little more complicated. If $D < 0$ or $D > 0$ and D has no prime divisors $p \equiv 3 \pmod 4$, then $rk_2(\mathrm{Cl}_D) = rk_2(\mathrm{C}_D)$. Else, $rk_2(\mathrm{Cl}_D) = rk_2(\mathrm{C}_D) - 1$. These results can be deduced by following the same argument as above, except with $(D, \pm N(\mathfrak{b}))_v$ in place of $(D, N(\mathfrak{b}))_v$.

3. Cohen-Lenstra heuristics

If Gauss's genus theory gives a formula for the 2-rank of a class group Cl_D, then what is true about the p-rank for odd primes p? Although no such explicit formulas exist, Cohen and Lenstra [CL84] developed a heuristic model that is compatible with the numerical data gathered on class groups of quadratic number fields. Their model is based on the assumption that the "weight" of an isomorphism class G of finite abelian groups should be inversely proportional to the number of automorphisms of G. One way to explain this assumption is as follows. If $\#G = n$ and E is a set of n elements, then the number of abelian group structures on E that are isomorphic to G is $n!/\#Aut(G)$. Therefore, at least among groups of the same size, those with smaller automorphism groups should occur more often. This gives a probability measure on the set of all finite abelian groups of the same size, and Cohen and Lenstra's assumption is that, at least in the case of imaginary quadratic fields, the odd part of the class group is "random" in this sense.

More precisely, let G' denote the prime-to-2 part of G, and let f be a reasonable function on isomorphism classes of abelian groups. (Here "reasonable" is not really well-defined, but this is a heuristic model after all.) The conjectures are as follows. If $D < 0$, then Cl'_D behaves like the prime-to-2 part of a random group:

$$\lim_{X\to\infty} \frac{\sum_{-X<D<0} f(\mathrm{Cl}'_D)}{\sum_{-X<D<0} 1} = \lim_{X\to\infty} \frac{\sum_{G,\#G\leq X} f(G')/\#Aut(G)}{\sum_{G,\#G\leq X} 1/\#Aut(G)}.$$

If $D > 0$, then Cl'_D behaves like the prime-to-2 part of a random group quotiented by the cyclic subgroup generated by a random element:

$$\lim_{X\to\infty} \frac{\sum_{0<D<X} f(\mathrm{Cl}'_D)}{\sum_{0<D<X} 1} = \lim_{X\to\infty} \frac{\sum_{G,\#G\leq X} (1/\#G) \sum_{g\in G} f((G/\langle g\rangle)')/\#Aut(G)}{\sum_{G,\#G\leq X} 1/\#Aut(G)}.$$

In their paper, Cohen and Lenstra compute the predicted average values of $f(\mathrm{Cl}'_D)$ for certain choices of f.

If we fix an integer $\alpha \geq 0$ and an odd prime p, and we set

$$f_\alpha(G) = \prod_{0\leq i<\alpha} \left(p^{rk_p(G)} - p^i\right),$$

then the Cohen-Lenstra heuristics predict that the average value of f_α over negative discriminants is 1 and over positive discriminants is $p^{-\alpha}$.

Recall that we will use $\mathcal{N}(\alpha, p)$ denote the number vector subspaces of $\mathbb{F}_p^\alpha$. Fouvry and Klüners [FK07, Proposition 1] have shown that the above conjectures are true for all $0 \leq \alpha \leq \alpha_0$ if and only if the average value of

$$f'_\alpha(G) = p^{\alpha rk_p(G)}$$

over negative discriminants (resp. positive discriminants) is $\mathcal{N}(\alpha, p)$ (resp. $p^{-\alpha}(\mathcal{N}(\alpha+1, p) - \mathcal{N}(\alpha, p)))$ for all $0 \leq \alpha \leq \alpha_0$. Although these conjectures about the functions f_α and f'_α are in some sense equivalent, the interpetation of Fouvry and Klüners proves to be very useful in their work on the 4-rank.

Gerth [Ger87] extended the Cohen-Lenstra conjectures to the 2-part of class groups. The 2-rank is obviously not random in the sense of the Cohen-Lenstra heuristics, as it is dictated by Gauss's genus theory, but the 4-rank does seem to be random. More succinctly, both the 2-part and the prime-to-2 part of Cl_D^2 behave according to Cohen-Lenstra heuristics. Thus, one expects that, for negative discriminants,

$$\lim_{X\to\infty} \frac{\sum_{-X<D<0} p^{\alpha rk_p(\mathrm{Cl}_D^2)}}{\sum_{-X<D<0} 1} = \mathcal{N}(\alpha, p)$$

for all primes p.

So what results are actually known about these conjectures? Gerth [Ger84] used the theory of Rédei matrices, whose ranks are related to the 4-rank, to show that the expected 4-rank of Cl_D over discriminants D that are a product of a fixed number of primes t approaches the value predicted by Cohen-Lenstra heuristics as $t \to \infty$. Note, however, that for any fixed t, the discriminants D satisfying $\omega(D) = t$ form a set of density 0 in the entire set of discriminants of quadratic number fields.

H. Davenport and H. Heilbronn [DH71] showed that the average value of $3^{rk_3(\mathrm{Cl}_D)}$ is in accordance with Cohen-Lenstra heuristics. In fact, this result precedes Cohen-Lenstra heuristics, and Cohen and Lenstra actually cited this result in support of their heuristics).

Cohen-Lenstra heuristics have been extended to the imaginary quadratic extensions of $\mathbb{F}_q(t)$ for any prime power q, and J. Ellenberg, A. Venkatesh, and C. Westerland [EVW09] have shown that the class groups of these extensions match the predictions of Cohen-Lenstra heuristics as $q \to \infty$.

Finally, Theorem 1 with $q = 1$ is the result of Fouvry and Klüners [FK07] which proves the f'_α conjectures about the 4-rank. Moreover, Let $\eta_m(t) = \prod_{j=1}^m (1 - t^{-j})$ for m a non-negative integer or ∞. Fouvry and Klüners [FK06] have shown that their theorem implies that the density of negative (resp. positive) fundamental discriminants D such that $rk_4(\mathrm{Cl}_D) = r$ is equal to $2^{-r^2}\eta_\infty(2)\eta_r(2)^{-2}$ (resp. $2^{-r(r+1)}\eta_\infty(2)\eta_r(2)^{-1}\eta_{r+1}(2)^{-1}$). Morally, "the moments determine the density."

This result is stronger than that of Gerth because here there is no condition on the number of prime factors of D. Davenport and Heilbronn's theorem only gives the first moment for the 3-rank, which is not sufficient to deduce the density. Higher moments seem beyond reach at this time. Ellenberg, Venkatesh, and Westerland are missing finitely many moments for any particular q, and hence cannot deduce the densities predicted by Cohen-Lenstra heuristics except in the limit as $q \to \infty$. Thus, in some sense, the result of Fouvry and Klüners is the most complete case of Cohen-Lenstra heuristics currently known.

4. Proof of Theorem 1

The proof can be roughly divided into four main steps:

1. Algebraic number theory: derive a formula for the 4-rank which is "friendly" to the tools of analytic number theory

2. Indexing trick: derive a formula for the kth moment which is "friendly" to analytic number theory

3. Analytic number theory: isolate the main term from the error terms

4. Combinatorics: use linear algebra over $\mathbb{F}_2$ to compute the coefficient of the

main term

We will summarize the main arguments in this order.

4.1. **Step 1: algebraic number theory.** Suppose $D = p_1 \cdots p_t$ and let $\mathfrak{p}_i$ denote the prime ideals of $\mathcal{O}_K$ lying above p_i. Let

$$\mathcal{B} = \{\mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_t^{e_t} : e_i \in \{0,1\}\}.$$

It can be shown that $\mathcal{B} \to \mathrm{C}_D[2]$ is a $2:1$ map. Hence

$$2^{rk_4(\mathrm{C}_D)} = \#(\mathrm{C}_D^2/\mathrm{C}_D^4) = \#(\mathrm{C}_D^2[2]) = \frac{1}{2}\#\{\mathfrak{b} \in \mathcal{B} : \mathfrak{b} \in \mathrm{C}_D^2\}.$$

Now, Fouvry and Klüners [FK07, Theorem 5] show that the map $\mathfrak{b} \to N(\mathfrak{b})$ gives a 1-to-1 correspondence between the sets $\{\mathfrak{b} \in \mathcal{B} : \mathfrak{b} \in \mathrm{C}_D^2\}$ and $\{b > 0 : \ \mu^2(b) = 1,\ b|D,\ (D,b) = 1\}$. It can be shown that $(D,b) = (b, -d/b)$, so that the final formula becomes

$$2^{rk_4(\mathrm{C}_D)} = \frac{1}{2}\#\{b > 0 : \ \mu^2(b) = 1,\ b|D,\ (b,-d/b) = 1\}. \tag{3}$$

Legendre's theorem for ternary quadratic forms [FK07, Lemma 6] implies that if a and b are coprime squarefree integers and $b > 0$, then $(a,b) = 1$ if and only if a is a square modulo b and b is a square modulo $|a|$.

For $D < 0$, $D \equiv 1 \pmod 4$, $D = d$ is squarefree, and so equation (3) becomes

$$\begin{aligned} 2^{rk_4(\mathrm{C}_D)} = \frac{1}{2}\#\{(a,b) :\ & a,b \geq 1, -D = ab, \\ & a \text{ is a square modulo } b, \\ & b \text{ is a square modulo } a\}, \end{aligned}$$

The reason that this formula is "friendly" to the methods of analytic number theory is that we can use Legendre symbols to detect the conditions above. In fact, for a and b odd, coprime and squarefree, the sum

$$\frac{1}{2^{\omega(b)}} \sum_{c|b} \left(\frac{a}{c}\right) = \frac{1}{2^{\omega(b)}} \prod_{p|b} \left(1 + \left(\frac{a}{p}\right)\right)$$

is 1 if a is a square modulo b and 0 otherwise. Hence we deduce that

$$2^{rk_4(\mathrm{C}_D)} = \frac{1}{2 \cdot 2^{\omega(-D)}} \sum_{-D=ab} \left(\sum_{c|b} \left(\frac{a}{c}\right)\right) \left(\sum_{d|a} \left(\frac{b}{d}\right)\right).$$

4.2. **Step 2: indexing trick.** Set $a = D_{10}D_{11}$, $b = D_{00}D_{01}$, $c = D_{00}$, and $d = D_{11}$ to obtain

$$2^{rk_4(\mathrm{C}_D)} = \frac{1}{2 \cdot 2^{\omega(-D)}} \sum_{-D = D_{00}D_{01}D_{10}D_{11}} \prod_{(u,v) \in \mathbb{F}_2^4} \left(\frac{D_u}{D_v}\right)^{\phi_1(u,v)},$$

where $\phi_1(u,v) = (u_1 + v_1)(u_1 + v_2)$. Here $u = (u_1, u_2)$ and $v = (v_1, v_2)$, so that ϕ_1 is a function $\phi_1 : \mathbb{F}_2^4 \to \mathbb{F}_2$.

Suppose

$$-D = \prod_{u^1 \in \mathbb{F}_2^2} D^1_{u^1} = \cdots = \prod_{u^k \in \mathbb{F}_2^2} D^k_{u^k}.$$

Set $D_{u^1,\ldots,u^k} = gcd(D^1_{u^1},\ldots,D^k_{u^k})$, and note that

$$D^l_{u^l} = \prod_{(u_1,\ldots,\hat{u^l},\ldots,u^k)\in\mathbb{F}_2^{2k-2}} D_{u^1,\ldots,u^l,\ldots,u^k}.$$

This is convenient because, instead of specifying k different ways of writing $-D$ as a product of 4 integers, we now simply need to specify one way of writing $-D$ as a product of 4^k integers.

We thus have

$$2^{krk_4(\mathrm{C}_D)} = \frac{1}{2^k \cdot 2^{k\omega(-D)}} \sum_{D_{u^1,\ldots,u^k}} \prod_{j=1}^{k} \prod_{(u^j,v^j)\in\mathbb{F}_2^4} \left(\frac{D^j_{u^j}}{D^j_{v^j}}\right)^{\phi_1(u^j,v^j)}.$$

We can now rearrange the product using multiplicative properties of the Jacobi symbol and sum over $-X < D < 0$ such that $D \equiv a \pmod q$. We thus obtain what we will refer to as the "main formula:"

$$\sum_{\substack{-X<D<0\\ D\equiv 1 \pmod 4\\ D\equiv a \pmod q}} 2^{krk_4(\mathrm{C}_D)} = 2^{-k} \sum_{(D_u)} \left(\prod_{u\in\mathbb{F}_2^{2k}} 2^{-k\omega(D_u)}\right) \prod_{u,v\in\mathbb{F}_2^{2k}} \left(\frac{D_u}{D_v}\right)^{\phi_k(u,v)}, \tag{4}$$

where the sum is over 4^k-tuples of squarefree, positive, coprime integers (D_u) with $u = (u^1,\ldots,u^k) \in \mathbb{F}_2^{2k}$ satisfying

$$\prod_u D_u \leq X, \prod_u D_u \equiv -1 \pmod 4, \prod_u D_u \equiv -a \pmod q,$$

and where $\phi_k(u,v) = \phi_1(u^1,v^1) + \cdots + \phi_1(u^k,v^k)$.

As we remarked before, the case when D is negative and congruent to 1 modulo 4 is the simplest and most appropriate for communicating the ideas of the proof. To demonstrate this, we now briefly explain what changes if, for instance, D is positive and congruent to 1 modulo 4. In this case, equation (3) becomes

$$\begin{aligned} 2^{rk_4(\mathrm{C}_D)} = \frac{1}{2}\#\{(a,b):\ & a,b\geq 1, -D = ab,\\ & -a \text{ is a square modulo } b,\\ & b \text{ is a square modulo } a\}, \end{aligned}$$

and the analogue of the main formula (equation (4)) is

$$\sum_{\substack{-X<D<0\\ D\equiv 1 \pmod 4\\ D\equiv a \pmod q}} 2^{k rk_4(\mathrm{Cl}_D)} = 2^{-k}\sum_{(D_u)}\left(\prod_{u\in\mathbb{F}_2^{2k}} 2^{-k\omega(D_u)}\right) \tag{5}$$

$$\cdot\left(\prod_{u\in\mathbb{F}_2^{2k}}\left(\frac{-1}{D_u}\right)^{\lambda_k(u)}\right)\prod_{u,v\in\mathbb{F}_2^{2k}}\left(\frac{D_u}{D_v}\right)^{\phi_k(u,v)}, \tag{6}$$

where $\lambda_k(u)=\lambda_1(u^1)+\cdots+\lambda_1(u^k)$, $\lambda_1(u^j)=u_1^j u_2^j$, and the other conditions are as above. The factor with $\left(\frac{-1}{D_u}\right)$ creates extra difficulties later on in the computation of the coefficient of the main term. Other cases of discriminants create even more problems, most of which require modifications of the combinatorial arguments.

4.3. **Step 3: analytic number theory.** To exploit the indexing tricks, we define the notion of linked indices.

Let P be the quadratic form over $\mathbb{F}_2^{2k}$ defined by

$$P(w)=\sum_{j=1}^{k} w_{2j-1}(w_{2j-1}+w_{2j}).$$

Then for any two indices $u,v\in\mathbb{F}_2^{2k}$,

$$P(u+v)=\phi_k(u,v)+\phi_k(v,u).$$

We say that u and v are "linked" if $P(u+v)=1$, and "unlinked" otherwise.

In other words, u and v are linked if and only if exactly one of $\left(\frac{D_u}{D_v}\right)$ and $\left(\frac{D_v}{D_u}\right)$ appears with an odd exponent in the main formula.

Example. In the case that $k=1$, the main formula simplifies to

$$\sum_{\substack{-X<D<0\\ D\equiv 1 \pmod 4\\ D\equiv a \pmod q}} 2^{rk_4(\mathrm{C}_D)} = \frac{1}{2}\sum_{(D_u)} 2^{-\omega(D_{00}D_{01}D_{10}D_{11})}\left(\frac{D_{00}}{D_{11}}\right)\left(\frac{D_{11}}{D_{00}}\right)\left(\frac{D_{01}}{D_{11}}\right)\left(\frac{D_{10}}{D_{00}}\right),$$

where the sum is over 4-tuples of squarefree, positive, coprime integers $(D_u)=(D_{00},D_{01},D_{10},D_{11})$ satisfying

$$D_{00}D_{01}D_{10}D_{11}\le X,\ \equiv -1 \pmod 4),\ \equiv -a \pmod q).$$

Hence, in this case, there are two pairs of linked indices, namely $\{00,\ 10\}$ and $\{01,\ 11\}$, while the other four pairs are unlinked.

We also define the parameter $\Omega = e4^k(\log\log X + B_0)$, where B_0 is an absolute constant that will be defined shortly. Then we can show that the contribution to the right-hand-side of the main formula of those (D_u) such that there exists a D_u with $\omega(D_u)>\Omega$ is negligible. The point is that the 4^k-tuples with factors that have many prime divisors are relatively rare. Let Σ_1 denote this contribution. We now use a trivial estimate. More precisely, using the triangle inequality, bounding the Jacobi symbols trivially, completing the sum to be over all integers less than

X (not just those that are discriminants in an arithmetic progression), and noting that a squarefree integer n can be written as a product of 4^k integers in $4^{k\omega(n)}$ ways, we have

$$\begin{aligned}\Sigma_1 &\ll \sum_{n\le X} \mu^2(n)4^{k\omega(n)}2^{-k\omega(n)}\\ &\ll \left(\sum_{\substack{n\le X\\ \omega(n)>\Omega}} \mu^2(n)\right)^{1/2}\left(\sum_{n\le X} 4^{k\omega(n)}\right)^{1/2}\end{aligned}$$

The constant B_0 is defined in the following lemma [FK07, Lemma 11], which is used to give an upper bound for the first factor above:

Lemma 1. *There exists an absolute constant B_0 such that for every $X \ge 3$ and every $l \ge 0$, we have*

$$\#\{n \le X : \omega(n) = l, \mu^2(n) = 1\} \le B_0\frac{X}{\log X}\frac{(\log\log X + B_0)^l}{l!}.$$

Using this lemma as well as Stirling's formula, we can bound the first factor by $\ll \left(\frac{X}{\log X}4^{-k\omega(n)}\right)^{1/2}$.

To bound the second factor above, we use the following lemma [FK07, Lemma 12]:

Lemma 2. *Let $\gamma \in \mathbb{R}_{>0}$. Then*

$$\sum_{X-Y\le n\le X} \gamma^{\omega(n)} \ll Y(\log X)^{\gamma-1}$$

uniformly for $2 \le X\exp(-\sqrt{\log X}) \le Y < X$.

Hence the second factor above is $\ll \left(X(\log X)^{4^k-1}\right)^{1/2}$. Combining the above and using the definition of Ω again gives that $\Sigma_1 \ll X(\log X)^{-1}$, which is a lot smaller than the error term in Theorem 1.

The next big idea is to dissect the sum in the right-hand side of the main formula into pieces that can be handled separately. For this, define $\Delta = 1+\log^{-2^k} X$. For each $u \in \mathbb{F}_2^{2k}$, let $A_u \in \{1, \Delta, \Delta^2, \ldots\}$, and for each $A = (A_u)_{u\in\mathbb{F}_2^{2k}}$, define the sum

$$S(X, k, A) = 2^{-k}\sum_{(D_u)}\left(\prod_u 2^{-k\omega(D_u)}\right)\prod_{u,v}\left(\frac{D_u}{D_v}\right)^{\phi_k(u,v)},$$

where now, in addition to the conditions above, the (D_u) must satisfy

$$A_u \le D_u < \Delta A_u,\ \omega(D_u) \le \Omega.$$

Hence we have

$$\sum_{\substack{-X<D<0\\ D\equiv 1 \pmod 4\\ D\equiv a \pmod q}} 2^{k\mathrm{rk}_4(\mathrm{C}_D)} = \sum_A S(X, k, A) + O(X(\log X)^{-1}),$$

where the sum is over A such that $\prod_u A_u \leq X$.

We say that A satisfies condition (C1) if $\prod_u A_u \geq \Delta^{-4^k} X$. Then, using a trivial estimate similar to the one above along with Lemma 2,

$$\begin{aligned}\sum_{A \text{ sat. } (C1)} |S(X,k,A)| &\leq 2^{-k} \sum_{\Delta^{-4^k} X \leq n \leq X} \mu^2(n) 4^{k\omega(n) 2^{-k\omega(n)}} \\ &\ll (1-\Delta^{-4^k}) X (\log X)^{2^k - 1}\end{aligned}$$

Now, $\Delta^{-4^k} = 1 - 4^k \log^{-2^k} X + O(\log^{-2^{k+1}} X)$, so that

$$\sum_{A \text{ sat. } (C1)} |S(X,k,A)| \ll X (\log X)^{-1}.$$

Define $X^{\ddagger}$ to be the least $\Delta^l \geq \exp(\log^{2^{-k}\varepsilon} X)$ (we take $\varepsilon > 0$; this is the same ε as in the statement of the theorem). We now say that A satisfies (C2) if at most $2^k - 1$ of the A_u are larger than $X^{\ddagger}$.

Let r parametrize the number of A_u that are larger than $X^{\ddagger}$, let n denote the product of those D_u and m the product of the remaining ones. Then, again using a trivial estimate

$$\begin{aligned}\sum_{A \text{ sat. } (C2)} \leq \sum_{r=0}^{2^k - 1} &\sum_{m \leq (X^{\ddagger})^{4^k - r}} \mu^2(m) (4^k - r)^{\omega(m)} 2^{-k\omega(m)} \\ &\cdot \sum_{n \leq X/m} \mu^2(n) r^{\omega(n)} 2^{-k\omega(n)}\end{aligned}$$

By Lemma 2, the above is

$$\ll X \left(\sum_{r=0}^{2^k - 1} (\log X)^{r 2^{-k} - 1} \right) \left(\sum_{m \leq (X^{\ddagger})^{4^k}} \frac{\mu^2(m) 2^{k\omega(m)}}{m} \right).$$

A Mertens-style estimate can be used to show that the second factor above is $\ll (\log X^{\ddagger})^{2^k}$, which then gives the bound

$$\sum_{A \text{ sat. } (C2)} |S(X,k,A)| \ll X (\log X)^{\varepsilon - 2^{-k}}.$$

Note that the contribution from the A satisfying (C2) is one of the contributions to the largest error term.

Now we start exploiting the work we did in step 2 to give some non-trivial estimates. Let $X^{\dagger} = (\log X)^{3(1+4^k(1+2^k))}$. We now say that A satisfies (C3) if $\prod_u A_u < \Delta^{-4^k} X$ and there exist two linked indices u and v with $A_u, A_v \geq X^{\dagger}$.

We will make use of a result on double oscillation of characters [FK07, Lemma 15]:

Lemma 3. *Let a_m and b_n be complex numbers of modulus < 1. Then for every $M, N \geq 1$ and every $\varepsilon > 0$, we have*

$$\sum_{m\leq M}\sum_{n\leq N} a_m b_n \mu^2(2m)\mu^2(2n)\left(\frac{m}{n}\right) \ll_\varepsilon MN(M^{-1/2+\varepsilon}+N^{-1/2+\varepsilon}).$$

Fix $(D_w)_w$ and suppose without loss of generality that $\phi_k(u,v)=1$ and $\phi_k(v,u)=0$. Then apply the lemma above with $M=\Delta A_u$, $N=\Delta A_v$, $\varepsilon=1/6$ to give the bound

$$\left|\sum_{D_u}\sum_{D_v} a(u)b(v)\left(\frac{D_u}{D_v}\right)\right| \ll A_u A_v(A_u^{-1/3}+A_v^{-1/3}),$$

where

$$a(u)=2^{-k\omega(D_u)}\prod_{w\neq u,v}\left(\frac{D_u}{D_w}\right)^{\phi_k(u,w)}\prod_{w\neq u,v}\left(\frac{D_w}{D_u}\right)^{\phi_k(w,u)}$$

and

$$b(v)=2^{-k\omega(D_v)}\prod_{w\neq u,v}\left(\frac{D_v}{D_w}\right)^{\phi_k(v,w)}\prod_{w\neq u,v}\left(\frac{D_w}{D_v}\right)^{\phi_k(w,v)}.$$

Now if A satisfies (C3), then $|S(X,k,A)|$ is

$$\begin{aligned}&\leq 2^{-k}\sum_{(D_w)_{w\neq u,v}}\prod_{w\neq u,v}2^{-k\omega(D_w)}\left|\sum_{D_u}\sum_{D_v}a(u)b(v)\left(\frac{D_u}{D_v}\right)\right|\\ &\ll X(X^{\dagger})^{-1/3}\end{aligned}$$

There are at most $O((\log X)^{4^k(1+2^k)})$ choices for A, which then allows us to make the estimate

$$\sum_{A\text{ sat. }(C3)}|S(X,k,A)| \ll X(\log X)^{-1}.$$

We now say that A satisfies (C4) if $\prod_u A_u < \Delta^{-4^k}X$ and there exist two linked indices u and v with $A_u \geq X^{\ddagger}$ and $\Delta \leq A_v \leq X^{\dagger}$. For such A, $S(X,k,A)$ is

$$\ll \max_{\substack{(c,4)=1\\(d,q)=1}}\sum_{(D_w)_{w\neq u,v}}\sum_{D_v}\sum_{l=0}^{\Omega}2^{-kl}\left|\sum_{\substack{\omega(D_u)=l\\D_u\equiv c \pmod 4\\D_u\equiv d \pmod q}}\mu^2(2q\prod_w D_w)\left(\frac{D_u}{D_v}\right)\right|.$$

To give a bound for the inner sum above, we appeal to a version of Siegel-Walfisz theorem [FK07, Lemma 13]:

Theorem 5. *For every $r\geq 2$, for every primitive character χ modulo r, and for every $C>0$, we have*

$$\sum_{y\leq p\leq x}\chi(p) \ll_C \sqrt{r}x\log^{-C}x$$

uniformly for $x\geq y\geq 2$.

Writing $D_u = p_1 \cdots p_l$ with p_i in increasing order and using the theorem with $r = 4qD_v \leq 4q\Delta X^\dagger$ and $x = \frac{\Delta A_u}{p_1 \cdots p_{l-1}}$, we see that the inner sum is

$$\ll \max_{\substack{(c',4)=1 \\ (d',q)=1}} \sum_{p_1 \cdots p_{l-1} \leq (\Delta A_u)^{1-1/l}} \left| \sum_{\substack{p_l \equiv c' \pmod 4 \\ p_l \equiv d' \pmod q}} \mu^2(2qp_1 \cdots p_l \prod_{w \neq u} D_w) \left(\frac{p_l}{D_v} \right) \right|$$
$$\ll_C q^{3/2} A_v^{1/2} A_u (\log X)^{-C2^{-k-1}\varepsilon}$$

Choosing C large in terms of k, we can once again show that

$$\sum_{A \text{ sat. } (C4)} |S(X,k,A)| \ll X(\log X)^{-1}.$$

Note that in this estimate, it was useful for each D_u to have at most Ω prime factors.

The previous estimates now imply that

$$\sum_{\substack{-X<D<0 \\ D\equiv 1 \pmod 4 \\ D \equiv a \pmod q}} 2^{krk_4(\mathrm{Cl}_D)} = \sum_A S(X,k,A) + O\left(X(\log X)^{\varepsilon - 2^{-k}} \right),$$

where the sum is over A satisfying:

– $\prod_u A_u < \Delta^{-4^k} X$

– at least 2^k indices u satisfy $A_u \geq X^\ddagger$

– two indices u, v with $A_u, A_v > X^\dagger$ are always unlinked

– if A_u and A_v with $A_v \leq A_u$ are linked, then either $A_v = 1$

or $\Delta \leq A_v < X$ and $A_v \leq A_u < X^\ddagger$

At this point, we make use of the following lemma, which is proved via the theory of symmmetric bilinear forms [FK07, Lemma 18]:

Lemma 4. *Let $k \geq 1$ and let $\mathcal{U} \subset \mathbb{F}_2^{2k}$ be a set of unlinked indices. Then $\#\mathcal{U} \leq 2^k$ and for any $c \in \mathbb{F}_2^{2k}$, $c + \mathcal{U}$ is also a set of unlinked indices. If $\#\mathcal{U} = 2^k$, then either $\mathcal{U}$ is a vector subspace of $\mathbb{F}_2^{2k}$ of dimension k or a coset of such a subspace of dimension k.*

This lemma allows us to simplify the conditions above to:

– $\prod_u A_u < \Delta^{-4^k} X$

– $\mathcal{U} = \{u : A_u \geq X^\ddagger\}$ is a maximal subset of unlinked indices

– if $A_u \notin \mathcal{U}$, then $A_u = 1$

If $\mathcal{U}$ is any maximal subset of unlinked indices, we say that A is "admissable" for $\mathcal{U}$ if A satisfies the three conditions above.

We finally make use of the structure of unlinked indices by applying the law of quadratic reciprocity. For two unlinked indices u and v, $\phi_k(u,v) = \phi_k(v,u)$ so that

$$\left(\frac{D_u}{D_v}\right)^{\phi_k(u,v)} \left(\frac{D_v}{D_u}\right)^{\phi_k(v,u)} = (-1)^{\phi_k(u,v)\cdot \frac{D_u-1}{2}\frac{D_v-1}{2}}.$$

Not that the right-hand-side is determined completely by the congruence class of D_u and D_v modulo 4.

Let $\mathcal{U}$ be a maximal subset of unlinked indices and suppose A is admissable for $\mathcal{U}$. Then using the Chinese remainder theorem and the law of quadratic reciprocity, we deduce that

$$\begin{aligned} &S(X,k,A) \\ &= 2^{-k}\sum_{(g_u)}\sum_{(h_u)}\left(\sum_{(D_u)}\mu^2(\prod_{u\in\mathcal{U}} D_u)\prod_{u\in\mathcal{U}} 2^{-k\omega(D_u)}\right)\prod_{\{u,v\}\subset\mathcal{U}}(-1)^{\phi_k(u,v)\cdot\frac{h_u-1}{2}\frac{h_v-1}{2}}, \end{aligned}$$

where now the sum over (g_u) is a sum over 2^k-tuples (g_u) such that each g_u is a congruence class modulo q coprime to q and such that $\prod_{u\in\mathcal{U}} g_u \equiv -a \pmod{q}$, the sum over (h_u) is a sum over 2^k-tuples (h_u) such that each h_u is a congruence class modulo 4 coprime to 4 and such that $\prod_{u\in\mathcal{U}} h_u \equiv -1 \pmod 4$, and the sum over (d_u) is a sum over 2^k-tuples (d_u) such that $A_u \le D_u < \Delta A_u$, $\omega(D_u) \le \Omega$, $D_u \equiv g_u \pmod q$, and $D_u \equiv h_u \pmod 4$.

To eliminate the congruence conditions on D_u in the statement above, we use the following lemma on squarefree numbers in arithmetic progressions:

Lemma 5. *Let $q \ge 1$ be an odd squarefree integer. Let $l \ge 0$, let g be a congruence class modulo q coprime to q, let h be a congruence class modulo 4 coprime to 4, let n_0 be a squarefree integer coprime to $2q$. Then for all $C > 0$ and $Y \ge y \ge 1$, we have*

$$\begin{aligned} \sum_{\substack{y\le n\le Y \\ n\equiv h\ (mod\ 4) \\ n\equiv g\ (mod\ q) \\ \omega(n)=l}} \mu^2(n_0 n) =& \frac{1}{2\varphi(q)}\sum_{\substack{y\le n\le Y \\ \omega(n)=l}}\mu^2(2qn_0n) \\ &+ O_C\left((l+1)^{C+1}Y(\log Y)^{-C} + \omega(n_0)Y^{1-1/l}\right). \end{aligned}$$

This lemma is analogous to [FK07, Lemma 19] and its proof is exactly the same. To deal with the sum over (D_u) above, enumerate the indices in some way to obtain $\mathcal{U} = \{u_1, \ldots, u_{2^k}\}$, and then use the lemma (with C large in terms of k and ε) to get

$$\begin{aligned} \sum_{D_{u_i}}\mu^2(\prod_j D_{u_j})2^{-k\omega(D_{u_i})} =& \frac{1}{2\varphi(q)}\sum_{\substack{A_{u_i}\le D_{u_i}\le \Delta A_{u_i} \\ \omega(D_{u_i})\le\Omega}}\mu^2(2q\prod_j D_{u_j}) \\ &+ O\left(A_{u_i}(\log X)^{-1-4^k(1+2^k)}\right) \end{aligned}$$

In conclusion, repeating this process for each u_i and summing over all A admissable for a particular $\mathcal{U}$, we get that

$$\sum_{A \text{ adm. for } \mathcal{U}} S(X,k,A) = 2^{-k-2^k}\varphi(q)^{-1}\left(\sum_{(h_u)}\prod_{\{u,v\}\subset\mathcal{U}}(-1)^{\phi_k(u,v)\frac{h_u-1}{2}\frac{h_v-1}{2}}\right)$$
$$\cdot\left(\sum_{(D_u)}\mu^2(2q\prod_{u\in\mathcal{U}}D_u)\prod_{u\in\mathcal{U}}2^{-k\omega(D_u)}\right)+O(X(\log X)^{-1}),$$

where now the sum is over (D_u) such that there is an A admissable for $\mathcal{U}$ with $A_u \le D_u \le \Delta A_u$ and such that $\omega(D_u)\le\Omega$. Note that the congruence conditions have disappeared. Also, note that it was crucial in this argument that the number of prime factors of each D_u is bounded.

It can be shown using methods similar to those used to handle the contribution of A satisfying (C2) that

$$\sum_{(D_u)}\mu^2(2q\prod_{u\in\mathcal{U}}D_u)\prod_{u\in\mathcal{U}}2^{-k\omega(D_u)} = \sum_{n\le X}\mu^2(2qn)+O(X(\log X)^{\varepsilon-2^{-k}}).$$

Now note that for all $C>0$

$$\frac{1}{2\varphi(q)}\sum_{n\le X}\mu^2(2qn) = \sum_{\substack{n\le X\\ n\equiv -1 \pmod 4\\ n\equiv -a \pmod q}}\mu^2(n)+O_C(X(\log X)^{-C}).$$

Summing over all maximal subsets $\mathcal{U}$ of unlinked indices, we obtain

$$\sum_{\substack{-X<D<0\\ D\equiv 1 \pmod 4\\ D\equiv a \pmod q}} 2^{krk_4(\mathrm{C}_D)} = 2^{1-k-2^k}\left(\sum_{\mathcal{U}}\gamma(\mathcal{U})\right)\sum_{\substack{-X<D<0\\ D\equiv 1 \pmod 4\\ D\equiv a \pmod q}} 1$$
$$+O(X(\log X)^{\varepsilon-2^{-k}}),$$

where $\gamma(\mathcal{U}) = \sum_{(h_u)}\prod_{\{u,v\}\subset\mathcal{U}}(-1)^{\phi_k(u,v)\frac{h_u-1}{2}\frac{h_v-1}{2}}$.

It now remains to compute $\sum_{\mathcal{U}}\gamma(\mathcal{U})$.

4.4. **Combinatorics.** We now briefly outline the method used to compute $\sum_{\mathcal{U}}\gamma(\mathcal{U})$. This is where the proof changes the most for other families of discriminants (i.e. those that are not negative and congruent to 1 modulo 4). Although the forms over $\mathbb{F}_2$ used for these other families are different, the methods in these other cases are similar.

For our case, first define

$$\gamma(\mathcal{U},\nu) = \sum_{\substack{S\subset\mathcal{U}\\ \#S=s\equiv\nu \pmod 2}}(-1)^{e(S)},$$

where

$$e(S) = \sum_{\{u,v\}\subset S} \phi_k(u,v).$$

Notice that $\gamma(\mathcal{U}) = \gamma(\mathcal{U}, 1)$.

We can decompose ϕ_k into

$$\phi_k(u,v) = L(u,v) + P(v) + \Lambda(u) + \Lambda(v),$$

where L is a bilinear form, P is a quadratic form (same P as before), Λ is a linear form, and $L(u,u) = P(u)$.

By viewing the power set of $\mathcal{U}$ as an abelian group with the symmetric difference operation and by constructing a multiplicative character on this group, one can prove the following key lemma [FK07, Lemma 24]:

Lemma 6. *Write $\mathcal{U} = c + \mathcal{U}_0$, where $\mathcal{U}_0$ is a vector subspace. If either $\gamma(\mathcal{U}, 0) \neq 0$ or $\gamma(\mathcal{U}, 1) \neq 0$, then L is identically 0 on $\mathcal{U}_0 \times \mathcal{U}_0$.*

We say that a subspace V of $\mathbb{F}_2^{2k}$ is good if the dimension of V is k and if L is identically zero on $V \times V$. Note that for any good V and any $c \in \mathbb{F}_2^{2k}$, $c + V$ is a maximal set of unlinked indices.

Another key lemma is the following [FK07, Lemma 25]:

Lemma 7. *Write $\mathcal{U} = c + \mathcal{U}_0$ as before. If $\mathcal{U}_0$ is good and $S \subset \mathcal{U}$ with $s = \#S$ odd, then $e(S) = 0$.*

Putting the two lemmas together, we get that

$$\begin{aligned}\sum_{\mathcal{U}} \gamma(\mathcal{U}) &= \sum_{\mathcal{U}_0 \text{ good}} \sum_{\mathcal{U}=c+\mathcal{U}_0} \gamma(\mathcal{U}) + \sum_{\mathcal{U}_0 \text{ not good}} \sum_{\mathcal{U}=c+\mathcal{U}_0} \gamma(\mathcal{U}) \\ &= \sum_{\mathcal{U}_0 \text{ good}} \sum_{\mathcal{U}=c+\mathcal{U}_0} \#\{S \subset \mathcal{U} : s \text{ odd}\} = 2^k 2^{2^k-1} \sum_{\mathcal{U}_0 \text{ good}} 1\end{aligned}$$

The proof is now concluded by establishing a bijection between good subspaces $\mathcal{U}_0$ and vector subspaces of $\mathbb{F}_2^k$. (We write $\mathbb{F}_2^{2k} = X \oplus Y$ for a particular choice of k-dimensional $\mathbb{F}_2$ vector spaces X and Y such that $L(\pi_X(u), \pi_Y(v)) = L(u,v)$ and such that for any subspace F of X, there is exactly one subspace $\mathcal{U}_0$ of $\mathbb{F}_2^{2k}$ such that L is identically zero on $\mathcal{U}_0 \times \mathcal{U}_0$ and $\pi_X(\mathcal{U}_0) = F$.) □

5. Proof of Theorem 2

In Theorem 2, we are only concerned with the first moment this time (i.e. $k = 1$), and so the main formula simplifies considerably (see example on page 10). Step 1 is the same as before, and this time, there is no need for the indexing trick from step 2. Likewise, we entirely avoid step 4 because the combinatorics are easy for the case $k = 1$.

Let $C > 0$. We will re-define the parameters from the proof of Theorem 1 as follows. We set $\Omega = 4eC(\log\log X + B_0)$, $\Delta = 1 + \log^{-C} X$, $X^\dagger = (\log X)^{3(4+5C)}$, and $X^{\ddagger} = \exp(\log^{1/2} X)$.

With this choice of Ω, and for $C \geq 1$, the same arguments as on page 11 can be used to show that, on both the left-hand side and the right-hand side of equation (2) in Theorem 2, the contributions from 4-tuples with a factor having many prime divisors are $\ll X(\log X)^{1-4eC}$.

Using the same argument as on page 12, we see that the contribution from the sums over A satisfying conditions (C3) on both the left-hand side and the right-hand side of equation (2) is $\ll X(\log X)^{1-C}$.

With the new choice of Δ, there are at most $O((\log X)^{4(1+C)})$ choices for A. Thus, with the new choice of $X^{\dagger}$, following the the same argument as on page 13 gives that the contribution from the sums over A satisfying condition (C3) on both the left-hand side and the right-hand side of equation (2) is $\ll X(\log X)^{-C}$.

With the new choice of $X^{\ddagger}$, the same argument as on pages 13-14 with the constant in Theorem 5 taken to be $56C$ shows that the contribution from the sums over A satisfying condition (C4) on both the left-hand side and the right-hand side of equation (2) is $\ll X(X^{\dagger})^{5/6}(\log X)^{-14C} \ll X(\log X)^{10-C}$.

As we have just seen, all of the above contributions belong to the error term in Theorem 2 separately for the left-hand side and the right-hand side of equation (2). However, a little more work needs to be done for the sums over A satisfying condition (C2) and for those that contribute to the main term.

The main terms for the left-hand side and the right-hand side of the equation in Theorem 2 are given by

$$\sum_{A} S(X,1,A) \quad \text{and} \quad \frac{1}{\varphi(q)} \sum_{A} S(X,1,A)^{*}$$

where both sums are over A satisfying the conditions on the bottom of page 14; where

$$S(X,1,A) = \frac{1}{2} \sum_{(D_u)} 2^{-\omega(D_{00}D_{01}D_{10}D_{11})} \left(\frac{D_{00}}{D_{11}}\right) \left(\frac{D_{11}}{D_{00}}\right) \left(\frac{D_{01}}{D_{11}}\right) \left(\frac{D_{10}}{D_{00}}\right)$$

is as in the proof of Theorem 1; and where $S(X,1,A)^{*}$ is the same as $S(X,1,A)$ except that the condition $\prod_u D_u \equiv -a \pmod{q}$ is replaced by the condition $(\prod_u D_u, q) = 1$. An application of the Chinese remainder theorem and Lemma 5 as on pages 15-16 shows that the difference between the main terms is bounded by

$$\ll (\Omega+1)^{C+2} X(\log X)^{-C/2},$$

which suffices for Theorem 2.

It remains is to handle the contribution from A satisfying condition (C2). For this, first note that, for both the left-hand side and the right-hand side of equation (2), the contributions from the A where none of the A_u is larger than $X^{\ddagger}$ are $\ll (X^{\ddagger})^4 \log X \ll X^{\delta}$ for any $\delta > 0$. Hence we must compare the contributions where exactly one of the A_u is larger than $X^{\ddagger}$.

First suppose that this A_u is A_{10}. Since D_{10} is linked with D_{00}, we may suppose that $A_{00} = 1$. (Otherwise, these would satisfy either condition (C3) or (C4).) Both A_{01} and A_{11} must be less than $X^{\ddagger}$, and, as they are linked, we may suppose that one of them, say A_{01} is less than $X^{\dagger}$ (otherwise, these would satisfy condition (C3)). In summary, we have the following conditions on A: $A_{00} = 1$, $1 \leq A_{01} \leq X^{\dagger}$, $A_{10} > X^{\ddagger}$, and $1 \leq A_{11} \leq X^{\ddagger}$.

Thus, the difference $S(X, 1, A) - S(X, 1, A)^*/\varphi(q)$ for A satisfying these conditions is

$$\begin{aligned}&\sum_{\substack{D_{01}\\(D_{01},q)=1}}\ \sum_{\substack{D_{10}\\(D_{01},q)=1}}\ \sum_{\substack{D_{11}\\D_{11}\equiv -a/(D_{01}D_{10})\ (q)}} 2^{-\omega(D_{01}D_{10}D_{11})}\left(\frac{D_{01}}{D_{11}}\right)\\ &-\frac{1}{\varphi(q)}\sum_{\substack{D_{01}\\(D_{01},q)=1}}\ \sum_{\substack{D_{10}\\(D_{01},q)=1}}\ \sum_{\substack{D_{11}\\(D_{11},q)=1}} 2^{-\omega(D_{01}D_{10}D_{11})}\left(\frac{D_{01}}{D_{11}}\right).\end{aligned}$$

where the sums are over odd, squarefree, coprime D_u satisfying the usual conditions $A_u \leq D_u < \Delta A_u$, $\omega(D_u) < \Omega$, $\prod_u A_u < X$, and $\prod_u D_u \equiv -1 \pmod 4$. We handle the difference above by rearranging the order of summation so that the inner sums are over D_{10}. Since the number of prime divisors of D_{10} is bounded, Lemma 5 can be used as before to give the bound

$$\sum_{\substack{D_{10}\\D_{10}\equiv -a/(D_{01}D_{11})\ (q)}} 2^{-\omega(D_{10})} - \frac{1}{\varphi(q)}\sum_{\substack{D_{10}\\(D_{10},q)=1}} 2^{-\omega(D_{10})} \ll A_{10}(\log X)^{-4-5C}.$$

Now trivially estimating the Jacobi symbols $\left(\frac{D_{01}}{D_{11}}\right)$ and the sums over D_{01} and D_{11}, we get that the difference above is $\ll (\prod_u A_u)(\log X)^{-4-5C} \ll X(\log X)^{-4-5C}$. Finally, summing over all A satisfying the conditions above, we get that the contribution of this difference is $\ll X(\log X)^{-C}$.

The case where A_{01} is the large variable is symmetric to the situation above. The other two remaining cases are when $A_{00} > X^{\ddagger}$ and $A_{11} > X^{\ddagger}$. These two last cases are symmetric, so we may assume without loss of generality that $A_{00} > X^{\ddagger}$. This forces $A_{10} = 1$ (otherwise such an A would satisfy condition (C3) or (C4)). The extra complication now is that the factor $\left(\frac{D_{00}}{D_{11}}\right)\left(\frac{D_{11}}{D_{00}}\right)$ does not disappear. Hence, to proceed as before, we must fix the congruence classes of D_{00} and D_{11} modulo 4 and apply the law of quadratic reciprocity to fix this factor. Lemma 5 still applies when congruences modulo 4 are imposed on the D_u, and so the previous argument handles this case as well. There are only finitely many ways to fix the congruence classes of D_{00} and D_{11} modulo 4 and so the difference from all the contributions in the case A_{00} is the large variable is still $\ll X(\log X)^{-C}$. This finishes the estimate of the difference of contributions to the left-hand and the right-hand sides of equation (2) coming from those A satisfying (C2), and thus also finishes the proof of Theorem 2.

6. Acknowledgements

I would like to thank Professor Étienne Fouvry for his guidance and help throughout my work on this master's thesis. I would also like to thank Professor Melanie Wood for providing inspiration for the generalization to arithmetic progressions.

7. References

[CL84] Cohen, H. and Lenstra, H.W., *Heuristics on class groups of number fields*, *Number Theory, Noordwijkerhout 1983*, volume 1068 of *Lecture Notes in Math.*, pages 33-62. Springer, Berlin, 1984.

[DH71] Davenport, H. and Heilbronn, H., *On the density of discriminants of cubic fields II*, *Proc. Roy. Soc. London*, Ser. A 322(1551):405-420, 1971.

[EVW09] Ellenberg, J.S., Venkatesh, A., and Westerland, C., *Homological stability for Hurwitz spaces and the Cohen-Lenstra conjecture over function fields*, preprint: arXiv:0912.0325, 2009.

[FK06] Fouvry, É. and Klüners, J., *Cohen-Lenstra heuristics of quadratic number fields*, *Algorithmic Number Theory*, volume 4076 of *Lecture Notes in Computer Science*, pages 40-55. Springer, Berlin, 2006.

[FK07] Fouvry, É. and Klüners, J., *On the 4-rank of class groups of quadratic number fields*, *Invent. Math.*, 167(3):455-513, 2007.

[FK10] Fouvry, É. and Klüners, J., *On the negative Pell equation*, *Annals of Mathematics*, 172(3):2035-104, 2010.

[Ger84] Gerth III, F., *The 4-class ranks of quadratic fields*, *Invent. Math.*, 77(3):489-515, 1984.

[Ger87] Gerth III, F., *Extension of conjectures of Cohen and Lenstra*, *Exposition. Math.*, 5(2):181-184, 1987.

[Has80] Hasse, H., *Number Theory*, 3rd ed., Springer-Verlag, 1980.

[Ser73] Serre, J.P., *A Course in Arithmetic*, Graduate Texts in Math. 7, Springer, New York, 1973.

[Ste93] Stevenhagen, P., *The number of real quadratic fields having units of negative norm*, *Experiment. Math.*, 2:121-136, 1993.