Algebra & Number Theory

Volume 14
2020
No. 4

The distribution of p-torsion in degree p cyclic fields

Jack Klys

msp

ALGEBRA AND NUMBER THEORY 14:4 (2020)
https://doi.org/10.2140/ant.2020.14.815

The distribution of p-torsion in degree p cyclic fields

Jack Klys

We compute all the moments of the p-torsion in the first step of a filtration of the class group defined by Gerth (1987) for cyclic fields of degree p, unconditionally for $p = 3$ and under GRH in general. We show that it satisfies a distribution which Gerth conjectured as an extension of the Cohen–Lenstra–Martinet conjectures. In the $p = 3$ case this gives the distribution of the 3-torsion of the class group modulo the Galois invariant part. We follow the strategy used by Fouvry and Klüners (2007) in their proof of the distribution of the 4-torsion in quadratic fields.

1. Introduction

Let K be a number field of degree n. Let Cl_K denote the class group and $\mathrm{Cl}_{K,p}$ denote the p-part. Let S be the set of finite abelian p-groups. We are interested in the question: what is the probability of any $A \in S$ occurring as $\mathrm{Cl}_{K,p}$ for K of degree n? The Cohen–Lenstra heuristics [Cohen and Lenstra 1984] propose an answer to this question for quadratic fields.

We make the question more precise as follows. Let D_K denote the discriminant of K. Let $\mathcal{D}_X^{\pm}$ be the set of real (resp. complex) quadratic fields with $|D_K| < X$. For any X define

$$S_X^{\pm}(A) = \frac{|\{K \in \mathcal{D}_X^{\pm} \mid \mathrm{Cl}_{K,p} \cong A\}|}{|\mathcal{D}_X^{\pm}|}.$$

The probability of A occurring as $\mathrm{Cl}_{K,p}$ in the family of real (resp. complex) quadratic fields is $\lim_{X\to\infty} S_X^{\pm}(A)$. In general this is not known to exist. Cohen and Lenstra conjectured that it does and proposed a distribution on S which should equal this quantity.

For $s \in \mathbb{Z}_{\geq 1} \cup \{\infty\}$ let

$$\eta_s(p) = \prod_{i=1}^{s}\left(1 - \frac{1}{p^i}\right).$$

One can show [Cohen and Lenstra 1984; Hall 1938] that

$$\sum_{G \in S} \frac{1}{|\mathrm{Aut}\, G|} = \frac{1}{\eta_\infty(p)} < \infty.$$

MSC2010: primary 11R29; secondary 11R37, 11R45.
Keywords: Cohen–Lenstra heuristics, arithmetic statistics, class groups, cyclic fields.

Then for any $A \in S$ and $u \geq 0$ define

$$\mu_u(A) = \frac{\eta_\infty(p)}{|\mathrm{Aut}\, A||A|^u}.$$

This defines a probability measure on S, called the Cohen–Lenstra distribution. They originally considered the case $n = 2$, $p \neq 2$ and considered complex quadratic and real quadratic fields separately.

Conjecture 1.1 (Cohen–Lenstra). *For $A \in S$*

$$\mu_0(A) = \lim_{X\to\infty} S_X^-(A), \quad \mu_1(A) = \lim_{X\to\infty} S_X^+(A).$$

These conjectures were extended to higher degree number fields by Cohen and Martinet [1987] again for $p \nmid n$.

No cases of these conjectures are known in full strength, though there has been much recent work on the subject. In the setting of number fields there are results giving the average size of the class group or subgroup thereof. There is the classical result of Davenport and Heilbronn [1971] and Datskovsky and Wright [1988] for the average size of 3-torsion of quadratic fields. There are also partial results for 8 and 16 torsion of quadratic fields due to Milovic [2017; 2018].

Below we will discuss in more detail the work of Fouvry and Klüners [2007] on 4 torsion of quadratic fields. Recently Smith [2017] has proven the distribution of the whole 2^∞-torsion of complex quadratic fields, thus generalizing their work.

There are also nonabelian versions which have been studied by Alberts [2016] and Bhargava [2014]. The conjectures have also been studied in the setting of function fields which provides additional tools such as moduli schemes. Some results here are the work of Ellenberg, Venkatesh and Westerland [Ellenberg et al. 2016], Boston and Wood [2017] and Wood [2019].

The original conjectures ignored the case when p divides the degree of the number fields. Gerth proposed a way of extending them to p-torsion in degree p cyclic fields by considering a certain subgroup of $\mathrm{Cl}_K[p]$ (which he calls the narrow principal genus — see Section 4 of [Gerth 1987]). He proved theorems providing compelling evidence for these conjectures (Theorems 4.3 and 5.11 in [Gerth 1984] and Theorem 2 in [Gerth 1987]) For the case $p = 2$ and $n = 2$ Gerth's extension implies the conjectures should hold in their original form, but with Cl_K^2 instead of Cl_K. This was proved by Fouvry and Klüners [2007]. To state their result, let $\mathrm{rk}_4(\mathrm{Cl}_K) = \mathrm{rk}_2(\mathrm{Cl}_K^2)$ and for any $k \in \mathbb{Z}_{\geq 1}$ let

$$M_k^{\pm}(2) = \lim_{X\to\infty} \frac{\sum_{K, 0<\pm D_K<X} 2^{k\, \mathrm{rk}_4(\mathrm{Cl}_K)}}{\sum_{K, 0<\pm D_K<X} 1}.$$

Define $\mathcal{N}(k, p)$ to be the number of subspaces of $\mathbb{F}_p^k$.

Theorem 1.2 (Fouvry–Klüners). *For every $k \in \mathbb{Z}_{\geq 1}$*

$$M_k^-(2) = \mathcal{N}(k, 2), \quad M_k^+(2) = \mathcal{N}(k+1, 2) - \mathcal{N}(k, 2).$$

By a separate result Fouvry and Klüners [2006] deduce that these moments are enough to determine a distribution.

Theorem 1.3 (Fouvry–Klüners). *The density of complex quadratic fields K with* $\mathrm{rk}_4(\mathrm{Cl}_K) = s$ *is*

$$\frac{\eta_\infty(2)}{\eta_s^2(2)2^{s^2}}$$

and the density of real quadratic fields with $\mathrm{rk}_4(\mathrm{Cl}_K) = s$ *is*

$$\frac{\eta_\infty(2)}{\eta_s(2)\eta_{s+1}(2)2^{s(s+1)}}.$$

Gerth conjectured a distribution for a certain subgroup of $\mathrm{Cl}_K[p]$ of cyclic p fields for all p. To state it we first define some notation. Throughout the paper p will denote an odd prime. Let K be a cyclic field of degree p with Galois group $G = \langle\sigma_K\rangle$. Let $\varphi_K = 1-\sigma_K$ act on $\mathrm{Cl}_K[p]$. It can be shown (see Section 3) there is a filtration

$$\mathrm{Cl}_K[p]^G = \ker\varphi_K \subseteq \ker\varphi_K^2 \subseteq \cdots \subseteq \ker\varphi_K^{p-1} = \mathrm{Cl}_K[p]. \tag{1-1}$$

Then Gerth conjectured a distribution for the p-rank of $\varphi_K(\ker\varphi_K^2)$. Notice that for $p=3$ we have $\ker\varphi_K^2 = \mathrm{Cl}_K[3]$ and so the above filtration implies $\varphi_K(\ker\varphi_K^2) \cong \mathrm{Cl}_K[3]/\,\mathrm{Cl}_K[3]^G$. We prove the following theorem which verifies Gerth's conjecture for $p=3$:

Theorem 1.4. *The density of cyclic cubic fields with* $\mathrm{rk}_3(\mathrm{Cl}_K[3]/\,\mathrm{Cl}_K[3]^G) = s$ *is*

$$\frac{\eta_\infty(3)}{\eta_s(3)\eta_{s+1}(3)3^{s(s+1)}}.$$

We can extend this to all odd p under the assumption of GRH for Artin L-functions (we remark the L-functions we will consider are all known to be entire, and as such we do not need to assume Artin's holomorphy conjecture).

Theorem 1.5. *Assume GRH for Artin L-functions. Let p be odd. The density of degree p cyclic fields with* $\mathrm{rk}_p(\varphi_K(\ker\varphi_K^2)) = s$ *is*

$$\frac{\eta_\infty(p)}{\eta_s(p)\eta_{s+1}(p)p^{s(s+1)}}.$$

The above filtration (1-1) is analogous to the filtration

$$\mathrm{Cl}_{K,2}^G = \mathrm{Cl}_K[2] \subseteq \mathrm{Cl}_K[4] \subseteq \cdots \subseteq \mathrm{Cl}_{K,2}$$

when $p=2$ and the object $\varphi_K(\ker\varphi_K^2)$ is hence analogous to $\mathrm{Cl}_K[4]^2 \cong \mathrm{Cl}_K[4]/\,\mathrm{Cl}_K[2]$ from Theorem 1.2. Since the completion of this paper Koymans and Pagano [2018] have extended the methods of Smith [2017] to determine the distribution of $\mathrm{Cl}_K[p^\infty]/\,\mathrm{Cl}_K[p^\infty]^G$ for odd p (conditional on the generalized Riemann hypothesis). They in fact prove a refined result which implies the distribution of $\varphi_K(\ker\varphi_K^{j+1})/\varphi_K(\ker\varphi_K^j)$ for all j.

Before continuing we make some remarks about $\mathrm{Cl}_K[p]^G$. It is the part of $\mathrm{Cl}_K[p]$ corresponding by class field theory to the genus field of K, that is the maximal unramified extension of K which is abelian over $\mathbb{Q}$. It can be shown $|\mathrm{Cl}_K[p]^G| = p^{r-1}$ where r is the number of primes ramified in K and that the average of $\mathrm{rk}_p(\mathrm{Cl}_K[p]^G)$ is ∞.

In the case $p = 2$ this quantity is $\mathrm{Cl}_K[4]^G = \mathrm{Cl}_K[2]$ and hence

$$\mathrm{rk}_2 \mathrm{Cl}_K^2 = \mathrm{rk}_2(\mathrm{Cl}_K[4]/\mathrm{Cl}_K[2]),$$

that is removing this part corresponds to replacing the 2-rank of Cl_K by 4-rank as defined above.

We deduce Theorems 1.4 and 1.5 from the following theorem together with [Fouvry and Klüners 2006]. Define

$$M_k(p) = \lim_{X\to\infty} \frac{\sum_{K, D_K < X} p^{k\,\mathrm{rk}_p(\varphi_K(\ker \varphi_K^2))}}{\sum_{K, D_K < X} 1}.$$

Theorem 1.6. *Let $k \in \mathbb{Z}_{\geq 1}$. Then unconditionally for $p = 3$ and under the assumption of GRH for Artin L-functions for $p > 3$ we have*

$$M_k(p) = \mathcal{N}(k+1, p) - \mathcal{N}(k, p).$$

The proof of Theorem 1.6 follows the strategy of Fouvry and Klüners. For any degree p cyclic field K we express $|\mathrm{Cl}_K[p]|$ using a sum of idele class characters, and then sum over all degree p cyclic fields of discriminant up to X. We then study the asymptotics of this expression using techniques from analytic number theory.

In the $p = 3$ case we require several versions of a large sieve inequality for cubic characters to bound the error term. We prove one such version as well as applying several others from the literature, due to Heath-Brown [2000], Baier and Young [2010] and Iwaniec and Kowalski [2004]. The reason for assuming the generalized Riemann hypothesis in the general case is that certain versions of the large sieve are not yet available for order p characters. In particular we lack analogs of Propositions 6.3 and 6.4. This is the only obstacle to an unconditional proof for all p.

Finally we remark briefly about an equivalent formulation of the Cohen–Lenstra conjectures which is commonly used. The distribution μ_u is characterized by the fact (see for instance [Ellenberg et al. 2016, Lemma 8.2]) that for all $A \in S$

$$\mathbb{E}_{G\sim\mu_u}(|\mathrm{Sur}(G, A)|) = \sum_{G\in S} \mu_u(G) \cdot |\mathrm{Sur}(G, A)| = \frac{1}{|A|^u}. \tag{1-2}$$

This is often called the A-moment of μ_u and computing it only for certain A can still provide information about the distribution of elements in Cl_K.

It is clear that $|\mathrm{Hom}(G, (\mathbb{Z}/p\mathbb{Z})^k)| = p^{k\,\mathrm{rk}_p(G)}$. Furthermore

$$|\mathrm{Hom}(G, (\mathbb{Z}/p\mathbb{Z})^k)| = \sum_{i=0}^{k} n(k, i, p)|\mathrm{Sur}(G, (\mathbb{Z}/p\mathbb{Z})^i)|$$

where $n(k, i, p)$ is the number of i-dimensional subspaces of $\mathbb{F}_p^k$. Hence Theorem 1.6 can be rephrased as computing the A moments in the above sense for all the groups $A = (\mathbb{Z}/p\mathbb{Z})^k$.

2. Preliminaries

2A. *Class field theory.* For any number field K Galois over $\mathbb{Q}$ and rational prime l let $K_{\mathfrak{p}}$ be the completion of K at the prime $\mathfrak{p} \mid l$. Let $N_{\mathfrak{p}} : K_{\mathfrak{p}} \to \mathbb{Q}_l$ be the norm map. Denote $K_l = K \otimes_{\mathbb{Q}} \mathbb{Q}_l$ and $N_l = \prod_{\mathfrak{p} \mid l} N_{\mathfrak{p}}$. Note the isomorphism $K_l \cong \prod_{\mathfrak{p} \mid l} K_{\mathfrak{p}}$ which lets us view N_l as a function on K_l.

Let C_K denote the idele class group of K. Let $N_{C_K} : C_K \to C_{\mathbb{Q}}$ denote the norm map defined by $(N_{C_K}\alpha)_l = N_l\big(\prod_{\mathfrak{p} \mid l} \alpha_{\mathfrak{p}}\big) = \prod_{\mathfrak{p} \mid l} N_{\mathfrak{p}}\alpha_{\mathfrak{p}}$.

In several places we will use the following isomorphism of $C_{\mathbb{Q}}$ with $\mathbb{R}_+ \times \prod_l \mathbb{Z}_l^\times$. For $x \in C_{\mathbb{Q}}$ there exists a unique $a_x \in \mathbb{Q}^\times$ such that $a_x \cdot x \in \mathbb{R}_+ \times \prod_l \mathbb{Z}_l^\times$. It is not hard to see that $x \mapsto a_x \cdot x$ is well defined and bijective.

Define the morphism $\langle \cdot \rangle_{\mathfrak{p}} : K_{\mathfrak{p}} \to C_K$ by $\langle b \rangle_{\mathfrak{p}} = (\dots, 1, b, 1, \dots)$ the class of the element with b in the $\mathfrak{p}$-th coordinate and 1 elsewhere. We will need the following lemma in Section 4.

Lemma 2.1. *Let $b \in \mathbb{Q}^*$ and $\langle b \rangle_l \in C_{\mathbb{Q}}$. Then $b \in N_l K_l$ if and only if $\langle b \rangle_l \in N_{C_K} C_K$.*

Proof. If $b = N_l \alpha_l$ for some $\alpha_l \in K_l$ then clearly $\langle b \rangle_l = N_{C_K}(\dots, 1, \alpha_l, 1, \dots)$.

For the converse note that, under the natural embedding of $\mathbb{Q}_l^*$ into $C_{\mathbb{Q}}$, we have $N_{C_K} C_K \cap \mathbb{Q}_l^* = N_l K_l^*$ by Corollary 5.8 from [Neukirch 1999, Section VI.5, page 394]. Hence if $\langle b \rangle_l \in N_{C_K} C_K$ then it follows immediately from the definition of N_{C_K} that $b \in N_l K_l^*$. □

2B. *Cyclic degree p fields.* Let $K/\mathbb{Q}$ be a degree p cyclic extension (p an odd prime). Then the discriminant is of the form $D_K = (p_1 \cdots p_r)^{p-1}$ where each p_i is either a prime congruent to 1 mod p or equal to p^2 and they are distinct. Conversely every integer of this form is a discriminant of a degree p cyclic field [Mayer 1992].

By class field theory each such extension corresponds to a character $\hat{\chi}$ of $C_{\mathbb{Q}}$ with $\ker \hat{\chi} = N_{C_K}(C_K)$ an index p subgroup of $C_{\mathbb{Q}}$. Through the identification $C_{\mathbb{Q}} \cong \mathbb{R}_+ \times \prod_l \mathbb{Z}_l^\times$ $\hat{\chi}$ descends to a character

$$\chi : (1 + p\mathbb{Z}_p) \times \prod_{l \mid D_K, l \neq p} \mathbb{F}_l^\times \to \mu_p \tag{2-1}$$

where the $(1 + p\mathbb{Z}_p)$ factor appears if and only if $p \mid D_K$. The character χ is nontrivial on each factor.

By an order p character we will mean a character $\chi_l : \mathbb{F}_l^\times \to \mu_p$ for any prime $l \neq p$, or $\chi_p : (1 + p\mathbb{Z}_p) \to \mu_p$. For each prime l (including $l = p$) there are $p - 1$ such distinct nontrivial characters. Thus χ factors into a product of order p characters $\chi = \prod_{l \mid D_K} \chi_l$. Hence there are $(p-1)^{\omega(n)}$ distinct nontrivial characters defined on the domain in (2-1) where $\omega(n)$ denotes the number of distinct prime divisors of n. Furthermore two distinct characters χ, ψ have the same kernel if and only if $\chi = \psi^i$ for some $1 \leq i \leq p - 1$.

It follows from these facts that for a fixed integer n which is a discriminant of a degree p cyclic extensions of $\mathbb{Q}$, the number of such extensions $K/\mathbb{Q}$ with $D_K = n$ is $(p-1)^{\omega(n)-1}$.

The following asymptotic formula for the number of degree p cyclic fields with discriminant up to X^{p-1} is well-known [Wright 1989]

$$\sum_{K, D_K < X^{p-1}} 1 \sim cX.$$

2C. ***The field $\mathbb{Q}(\zeta_3)$ and cubic reciprocity.*** Let ζ_3 be a cube root of unity. The following facts about the field $\mathbb{Q}(\zeta_3)$ and the cubic residue symbol can be found for instance in [Baier and Young 2010, Section 2.1].

The extension $\mathbb{Q}(\zeta_3)/\mathbb{Q}$ is quadratic, and for $x \in \mathbb{Q}(\zeta_3)$ we denote its Galois conjugate by $\bar{x}$. The ring of integers of $\mathbb{Q}(\zeta_3)$ is $\mathcal{O} = \mathbb{Z}[\zeta_3]$. It is a principal ideal domain and every ideal $(n) \subset \mathcal{O}$ with $(n, 3) = 1$ has a unique generator n which satisfies $n \equiv 1 \bmod 3\mathcal{O}$. The only prime which ramifies is $(3) = ((1 - \zeta_3)^2)$. The primes of $\mathbb{Z}$ which split in $\mathcal{O}$ are exactly the ones congruent to $1 \bmod 3\mathbb{Z}$.

The set $\{1, \zeta_3\}$ is a basis for $\mathcal{O}$, so that every element of $\mathcal{O}$ can be written as $a + b\zeta_3$. Letting $N : \mathbb{Q}(\zeta_3) \to \mathbb{Q}$ denote the norm map we have the formula $N(a + b\zeta_3) = a^2 + b^2 - ab$. Using this it can be shown that $|\{a + b\zeta_3 \in \mathcal{O} \mid N(a + b\zeta_3) \leq X\}| = O(X)$ and for a fixed $b \in \mathbb{Z}$ that

$$|\{a \in \mathbb{Z} \mid N(a + b\zeta_3) \leq X\}| = O(X^{1/2}).$$

For $n, m \in \mathcal{O}$ coprime and $(m, 3) = 1$ denote by $\left(\frac{n}{m}\right)_3$ the cubic residue symbol. It satisfies

$$\left(\frac{\bar{n}}{\bar{m}}\right)_3 = \overline{\left(\frac{n}{m}\right)_3}, \quad \left(\frac{n}{m}\right)_3^2 = \overline{\left(\frac{n}{m}\right)_3} \tag{2-2}$$

and, if additionally $n, m \equiv 1 \bmod 3\mathcal{O}$ then there is the law of cubic reciprocity

$$\left(\frac{n}{m}\right)_3 = \left(\frac{m}{n}\right)_3. \tag{2-3}$$

3. Counting p-torsion in degree p cyclic fields

The first goal is to describe the subgroup of class group whose distribution we will be computing. See [Stevenhagen 1995] for a slightly different treatment of some of the material found in this section.

We recall that p will always denote an odd prime.

Let K be a degree p cyclic extension of $\mathbb{Q}$ with Galois group $G = \langle \sigma_K \rangle$. Let $\varphi_K = 1 - \sigma_K$. We view σ_K as a morphism acting on $\mathrm{Cl}_K[p^\infty]$. Let $NG = \sum_{i=0}^{p-1} \sigma_K^i$. Then since $NG : \mathrm{Cl}_K \to \mathrm{Cl}_\mathbb{Q}$ and the latter is trivial we have $\ker NG = \mathrm{Cl}_K$. Thus we can view $\mathrm{Cl}_K[p^\infty]$ as a module over the ring $\mathbb{Z}_p[\sigma_K]/\langle NG \rangle$.

It can be shown that in $\mathbb{Z}_p[\sigma_K]/\langle NG \rangle$ there is the relation $(\varphi_K^{p-1}) = (p)$. Thus there is a filtration

$$\mathrm{Cl}_K[p]^G = \ker \varphi_K \subseteq \ker \varphi_K^2 \subseteq \cdots \subseteq \ker \varphi_K^{p-1} = \mathrm{Cl}_K[p].$$

From this we can write down the exact sequence

$$1 \to \mathrm{Cl}_K[p]^G \to \ker \varphi_K^2 \to \varphi_K(\ker \varphi_K^2) \to 1$$

so that $|\ker \varphi_K^2| = |\mathrm{Cl}_K[p]|^G |\varphi_K(\ker \varphi_K^2)|$.

Note that as a special case of the ambiguous class number formula (see [Lemmermeyer 2013, Theorem 1]) we have $|\mathrm{Cl}_K| = p^{r-1}$. This implies $\mathrm{Cl}_K^G \subset \mathrm{Cl}_K[p^\infty]$.

Denote by N the norm map $N_{K/\mathbb{Q}}$ (both on ideals and elements of K). Let $\mathcal{J}$ be the group of fractional ideals of K. Furthermore let $P_1, \dots, P_r$ be the ramified primes of K, and let $\mathcal{B} = \{P_1^{e_k} \cdots P_r^{e_r} \mid e_i = 0, 1, \dots, p-1\}$. For any $I \in \mathcal{J}$ let $\bar{I}$ denote the natural projection to Cl_K.

Lemma 3.1. *Let $\overline{\mathcal{B}}$ be the projection of $\mathcal{B}$ to Cl_K. Then $|\overline{\mathcal{B}}| = p^{r-1}$.*

Proof. Clearly $\overline{\mathcal{B}} \subset \mathrm{Cl}_K[p]^G$. We will show $\overline{\mathcal{B}}$ generates Cl_K^G and the lemma will follow from $|\mathrm{Cl}_K^G| = p^{r-1}$.

Let $I \in \mathcal{J}$ such that $\bar{I} \in \mathrm{Cl}_K^G$, so that $I^{\sigma_K} = (\alpha)I$ for some $\alpha \in K$. Applying N to both sides gives $N(\alpha) = 1$, hence multiplying by -1 if necessary, we have $N\alpha = 1$ in K. By Hilbert's Theorem 90 there exists $\beta \in K$ such that $\alpha = \beta^{1-\sigma_K}$.

Thus $I^{\sigma_K} = (\beta)^{1-\sigma_K} I$ and rearranging $((\beta)I)^{\sigma_K} = (\beta)I$. So $(\beta)I$ is fixed by σ_K in $\mathcal{J}$. This implies gI is divisible only by ramified and rational primes in K. Thus $\overline{gI} \in \overline{\mathcal{B}}$. This completes the proof. □

Next we give another description of $\varphi_K(\ker \varphi_K^2)$.

Lemma 3.2. *Consider N acting on $\mathcal{J}$ the group of fractional ideals of K. Then*

$$\ker N = \varphi_K(\mathcal{J}).$$

Proof. It is clear that $\varphi_K(\mathcal{J}) \subset \ker N$. Suppose $NI = 1$ for some ideal $I \in \mathcal{J}$. Then I can only be divisible by split primes. Let $q \in \mathbb{Z}$ be a prime above which I is supported and let $Q_1, \dots, Q_p$ be all the prime ideals in K lying above q such that $Q_i^{\sigma_K} = Q_{i+1}$. Then $N(Q_1^{a_1} Q_2^{a_2} \cdots Q_p^{a_p}) = q^{\sum a_i}$ which implies that $\sum a_i = 0$. Then $Q_1^{a_1} Q_2^{a_2} \cdots Q_p^{a_p} = (Q_1^{a_1} Q_2^{a_1+a_2} \cdots Q_{p-1}^{a_1+\cdots+a_{p-1}})^{1-\sigma_K}$. Applying this to all primes q below I shows $I \in \varphi_K(\mathcal{J})$. □

Lemma 3.3. *For any $I \in \mathcal{J}$ such that $\bar{I} \in \mathrm{Cl}_K[p]^G$ we have*

$$\bar{I} \in \varphi_K(\ker \varphi_K^2) \iff NI = N(\alpha) \text{ in } \mathcal{J} \text{ for some } \alpha \in K$$

(*note this condition is independent of the ideal representing $\bar{I}$*).

Proof. Suppose first that $\bar{I} \in \varphi_K(\ker \varphi_K^2)$. So for some $\bar{J} \in \mathrm{Cl}_K$, we have $\bar{I} = \bar{J}^{1-\sigma_K}$. We have for some $\alpha \in K^\times$ that $(\alpha)I = J^{1-\sigma_K}$ in $\mathcal{J}$. Taking norm of this gives $NI = N(\alpha^{-1})$ which proves one direction.

Now suppose that $NI = N(\alpha)$ for some $\alpha \in K$. Hence $I = (\alpha)J$ for some ideal $J \in \ker N$. By Lemma 3.2 we have $J = J_0^{\varphi_K}$ for some ideal $J_0 \in \mathcal{J}$. So $\bar{I} = \bar{J}_0^{\varphi_K}$. Then $\bar{J}_0^{\varphi_K^2} = \bar{I}^{\varphi_K} = 1$ in Cl_K since $\bar{I} \in \mathrm{Cl}_K^G$. Thus $\bar{J}_0 \in \ker \varphi_K^2$. □

Let

$$\tilde{D}_K = \begin{cases} D_K & \text{if } p \nmid D_K, \\ D_K/p^{p-1} & \text{if } p \mid D_K. \end{cases} \tag{3-1}$$

Proposition 3.4. *With the above notation let $\Omega(\tilde{D}_K)$ be the set of positive integers dividing $\tilde{D}_K$. Then*

$$|\varphi_K(\ker \varphi_K^2)| = \tfrac{1}{p}|\{b \in \Omega(\tilde{D}_K) \mid b = N\alpha \text{ for some } \alpha \in K^\times\}|.$$

Proof. Firstly it is clear that the map $N : \mathcal{B} \to \Omega(\tilde{D}_K)$ is a bijection. Since $|\bar{\mathcal{B}}| = p^{r-1}$ by Lemma 3.1, let $\{(1), (l_2) \dots, (l_p)\} \subset \mathcal{B}$ be the principal ideals. To each $J \in \mathrm{Cl}_K[p]^G = \bar{\mathcal{B}}$ we associate the set $\omega_J = \{N(I_J), N(I_J l_2), \dots, N(I_J l_p)\} \subset \mathcal{J}$ where I_J is a choice of representative of J supported on the ramified primes. For every element of ω_J we may choose the unique generator which is a positive integer. With this identification we may assume $\omega_J \subset \Omega(\tilde{D}_K)$.

We claim that if $J_1 \neq J_2$ in $\bar{\mathcal{B}}$ then $\omega_{J_1} \cap \omega_{J_2} = \varnothing$. Suppose to the contrary that $x \in \omega_{J_1} \cap \omega_{J_2}$ so $N(I_{J_1} l_i) = N(I_{J_2} l_j)$ for some i, j, and hence $N(I_{J_1}) = N(I_{J_2} l_k)$ for some l_k since the principal ideals of $\mathcal{B}$ form a subgroup. Since N is injective on $\mathcal{B}$ we get $I_{J_1} = I_{J_2} l_k$, a contradiction. Thus we can write as a disjoint union

$$\Omega(\tilde{D}_K) = \bigcup_{J \in \bar{\mathcal{B}}} \omega_J .$$

Then also

$$\{b \in \Omega(\tilde{D}_K) \mid b = N\alpha \text{ for some } \alpha \in K^\times\} = \bigcup_J \omega_J$$

where the union is over all J such that some element of ω_J is of the form $N\alpha$ for some $\alpha \in K$. Note this condition is equivalent to every element of ω_J is of the form $N\alpha$ for some $\alpha \in K$ since if $N(I_J l_i) = N(\alpha)$ then $N(I_J l_j) = N(\alpha l_j l_i^{-1})$.

By Lemma 3.3 $|\varphi_K(\ker \varphi_K^2)|$ is the number of classes $J \in \mathrm{Cl}_K[p]^G$ such that $NI = N(\alpha)$ for some (any) representative I of J and some $\alpha \in K$. In the above notation this is the set of $J \in \bar{\mathcal{B}}$ such that $N\alpha \in \omega_J$ for some $\alpha \in K$. Thus

$$|\varphi_K(\ker \varphi_K^2)| = \tfrac{1}{p} \bigcup_J |\omega_J|$$

where the union is over all J such that every element of ω_J is of the form $N\alpha$ for some $\alpha \in K$. This completes the proof. □

4. The p-torsion as a character sum

The goal of this section will be to prove a formula for the size of $\varphi_K(\ker \varphi_K^2)$ defined in the previous section.

Let $K/\mathbb{Q}$ be a cyclic degree p extension with discriminant D_K and Galois group $\mathrm{Gal}(K/\mathbb{Q}) = \langle \sigma_K \rangle$. Let $\hat{\chi}$ be a character of C_K corresponding to K and χ its quotient (see Section 2B). Recall that χ factors into a product of order p characters $\chi = \prod_{l \mid D_K} \chi_l$.

For simplicity we will henceforth write $\mathrm{im}(\varphi_K)$ for $\varphi_K(\ker \varphi_K^2)$.

Recall the definition of $\tilde{D}_K$ in (3-1) and that $\omega(\tilde{D}_K)$ denotes the number of distinct prime divisors of $\tilde{D}_K$. For any prime l of $\mathbb{Q}$ we defined the morphism $\langle \cdot \rangle_l : K_l \to C_K$ by $\langle b \rangle_l = (\dots, 1, b, 1, \dots)$ the element with b in the l-th coordinate.

Proposition 4.1. *For each degree p cyclic field K let σ_K denote a generator of the Galois group and D_K the discriminant and $\hat{\chi}$ a corresponding character. Then*

$$|\mathrm{im}(\varphi_K)| = \tfrac{1}{p^{\omega(\tilde{D}_K)+1}} \sum_{(b_0,\dots,b_{p-1})} \prod_{i=0}^{p-1} \prod_{\substack{l \mid b_i \\ l \text{ prime}}} (1+\hat{\chi}+\cdots+\hat{\chi}^{p-1})(\langle B\rangle_l)$$

where the sum is over p-tuples of coprime positive integers (b_i) satisfying

$$\left(\prod_{i=0}^{p-1} b_i\right)^{p-1} = \tilde{D}_K \quad \text{and} \quad B = b_1 b_2^2 \cdots b_{p-1}^{p-1}.$$

Proof. Since $K/\mathbb{Q}$ is cyclic by the Hasse norm theorem $b \in \mathbb{Q}^\times$ is a global norm if and only if b is a local norm everywhere:

$$b = N\alpha \text{ for some } \alpha \in K \iff b = N_l \alpha_l \text{ for some } \alpha_l \in K_l, \text{ for all } l.$$

Recall $K_l = K \otimes_{\mathbb{Q}} \mathbb{Q}_l$. Hence by Proposition 3.4 we want to detect when $b \mid \tilde{D}_K$ satisfies $b \in N_l K_l$ for all l. If $l \nmid \tilde{D}_K$ this condition is trivial since b is then a local unit in $\mathbb{Q}_l$, and it is a standard fact from local class field theory that if l is unramified in $K/\mathbb{Q}$ then $N_l : K_l \to \mathbb{Q}_l$ surjects onto the local units. Hence we need to check the condition only for $l \mid \tilde{D}_K$.

By Lemma 2.1 $b \in N_l K_l$ if and only if the idele $\langle b\rangle_l \in C_{\mathbb{Q}}$ satisfies $\langle b\rangle_l \in N_{C_K} C_K = \ker \hat{\chi}$. Since $\hat{\chi}$ has order p this implies

$$(1+\hat{\chi}+\cdots+\hat{\chi}^{p-1})(\langle b\rangle_l) = \begin{cases} p & \text{if } \langle b\rangle_l \in \ker \hat{\chi}, \\ 0 & \text{else.} \end{cases}$$

Hence we arrive at the following expression which detects when b is a norm at l:

$$\left(\frac{1+\hat{\chi}+\cdots+\hat{\chi}^{p-1}}{p}\right)(\langle b\rangle_l) = \begin{cases} 1 & \text{if } b \in N_l K_l, \\ 0 & \text{else.} \end{cases} \tag{4-1}$$

Note that $\tilde{D}_K$ is a $p-1$ power. Write a divisor of $\tilde{D}_K$ as $b_1 b_2^2 \cdots b_{p-1}^{p-1}$ where the b_i are square-free and coprime. Let

$$G(b) = \begin{cases} 1 & \text{if } b \in NK^\times, \\ 0 & \text{else.} \end{cases}$$

Thus by Proposition 3.4 we have

$$p|\mathrm{im}(\varphi_K)| = \sum_{b_1 b_2^2 \cdots b_{p-1}^{p-1} \mid \tilde{D}_K} G(b_1 b_2^2 \cdots b_{p-1}^{p-1})$$

where the sum is over positive integers dividing $\tilde{D}_K$. In the following we will let $B = b_1 b_2^2 \cdots b_{p-1}^{p-1}$. By (4-1) we get

$$\begin{aligned} p|\mathrm{im}(\varphi_K)| &= \sum_{(b_0,\dots,b_{p-1})} \prod_{l \mid \tilde{D}_K} \left(\frac{1+\hat{\chi}+\cdots+\hat{\chi}^{p-1}}{p} \right)(\langle B \rangle_l) \\ &= \frac{1}{p^{\omega(\tilde{D}_K)}} \sum_{(b_0,\dots,b_{p-1})} \prod_i \prod_{l \mid b_i} (1+\hat{\chi}+\cdots+\hat{\chi}^{p-1})(\langle B \rangle_l) \end{aligned}$$

where the sum is over p-tuples of positive integers satisfying $(\prod_i b_i)^{p-1} = \tilde{D}_K$. This completes the proof. □

We define some notation to state the next proposition. Define the function $\Phi : \mathbb{F}_p^2 \to \mathbb{F}_p$ by

$$\Phi(u, v) = \Phi((u_1, u_2), (v_1, v_2)) = u_1(v_2 - u_2). \tag{4-2}$$

Under the identification $C_{\mathbb{Q}} \cong \mathbb{R}_+ \times \prod_l \mathbb{Z}_l^\times$ (see Section 2A) for any integer b the class of $\langle b \rangle_l$ maps to

$$\left(\dots, \frac{1}{l^i}, \frac{b}{l^i}, \frac{1}{l^i}, \dots \right)$$

where $i = \mathrm{ord}_l\, b$. Hence if χ decomposes as $\prod_{l \mid D_K} \chi_l$ acting on $(1 + p\mathbb{Z}_p) \times \prod_{l \mid D_K, l \neq p} \mathbb{F}_l^\times$ (see Section 2B) then

$$\hat{\chi}(\langle b \rangle_l) = \chi_l\left(\frac{b}{l^i}\right) \prod_{q \neq l} \chi_q\left(\frac{1}{l^i}\right). \tag{4-3}$$

Theorem 4.2. *For each degree p cyclic field K let σ_K denote a generator of the Galois group and D_K the discriminant and $\chi = \prod_{l \mid D_K} \chi_l$ a corresponding character. Then*

$$|\mathrm{im}(\varphi_K)| = \frac{1}{p^{\omega(\tilde{D}_K)+1}} \sum_{(D_u)} \prod_{v \in \mathbb{F}_p^2} \prod_{\substack{l \mid D_v \\ l \text{ prime}}} \chi_l\left(\prod_{u \in \mathbb{F}_p^2} D_u^{\Phi(u,v)} \right) \tag{4-4}$$

where the sum is over p^2-tuples of coprime positive integers (D_u) indexed by $u \in \mathbb{F}_p^2$ and satisfying $\left(\prod_{u \in \mathbb{F}_p^2} D_u\right)^{p-1} = \tilde{D}_K$.

Proof. For this proof we will denote $\tilde{D}_K = D^{p-1}$. By Proposition 4.1 we have

$$|\mathrm{im}(\varphi_K)| = \frac{1}{p^{\omega(D)+1}} \sum_{(b_0,\dots,b_{p-1})} \prod_{i=0}^{p-1} \prod_{l \mid b_i} (1+\hat{\chi}+\cdots+\hat{\chi}^{p-1})(\langle B \rangle_l) \tag{4-5}$$

where the sum is over p-tuples of positive integers satisfying $\prod_i b_i = D$ and $B = b_1 b_2^2 \cdots b_{p-1}^{p-1}$. We fix i and focus on the innermost product in (4-5). Expanding it we get

$$\prod_{l \mid b_i} (1+\hat{\chi}+\cdots+\hat{\chi}^{p-1})(\langle B \rangle_l) = \sum_{j_1,\dots,j_{\omega(b_i)}} \prod_{l \mid b_i} \hat{\chi}^{j_l}(\langle B \rangle_l) \tag{4-6}$$

where the sum is over tuples of integers with each $0 \leq j_k \leq p-1$. For any such fixed tuple $j_1, \ldots, j_{\omega(b_i)}$ and for each $0 \leq j \leq p-1$ define D_{ip+j} to be the product of all primes $l \mid b_i$ such that $j_l = j$. Then we can instead write (4-6) as

$$\sum_{j_1,\ldots,j_{\omega(b_i)}} \prod_{l \mid b_i} \hat{\chi}^{j_l}(\langle B\rangle_l) = \sum_{(D_{ip},D_{ip+1},\ldots,D_{ip+p-1})} \prod_{j=0}^{p-1} \prod_{l \mid D_{ip+j}} \hat{\chi}^{j}(\langle B\rangle_l)$$

where the sum on the right is over all p-tuples of positive integers satisfying $\prod_{j=0}^{p-1} D_{ip+j} = b_i$. Thus we get

$$|\mathrm{im}(\varphi_K)| = \frac{1}{p^{\omega(D)+1}} \sum_{(b_0,\ldots,b_{p-1})} \prod_{i=0}^{p-1} \Biggl(\sum_{(D_{ip},D_{ip+1},\ldots,D_{ip+p-1})} \prod_{j=0}^{p-1} \prod_{l \mid D_{ip+j}} \hat{\chi}^{j}(\langle B\rangle_l) \Biggr)$$

where the inner sum is over all p-tuples of positive integers satisfying $\prod_{j=0}^{p-1} D_{ip+j} = b_i$. If $l \mid b_i$ then $\mathrm{ord}_l\, B = i$ hence by (4-3) we have

$$\hat{\chi}^{j}(\langle B\rangle_l) = \hat{\chi}(\langle B^j\rangle_l) = \chi_l\biggl(\frac{B^j}{l^{ij}}\biggr) \prod_{q \mid D, q \neq l} \chi_q\biggl(\frac{1}{l^{ij}}\biggr).$$

Plugging this in and rearranging summations we get

$$|\mathrm{im}(\varphi_K)| = \frac{1}{p^{\omega(D)+1}} \sum_{(D_0,\ldots,D_{p^2-1})} \prod_{i,j=0}^{p-1} \prod_{l \mid D_{ip+j}} \biggl[\chi_l\biggl(\frac{B^j}{l^{ij}}\biggr) \prod_{q \mid D, q \neq l} \chi_q\biggl(\frac{1}{l^{ij}}\biggr) \biggr] \tag{4-7}$$

where the first sum is over all p^2-tuples of positive integers satisfying $\prod_{i=0}^{p^2-1} D_i = D$. Taking the last two products from (4-7) and rearranging them gives

$$\prod_{l \mid D_{ip+j}} \biggl[\chi_l\biggl(\frac{B^j}{l^{ij}}\biggr) \prod_{q \mid D, q \neq l} \chi_q\biggl(\frac{1}{l^{ij}}\biggr) \biggr]$$
$$= \biggl[\prod_{q \mid D/D_{ip+j}} \chi_q\biggl(\prod_{l \mid D_{ip+j}} \frac{1}{l^{ij}} \biggr) \biggr] i \biggl[\prod_{q \mid D_{ip+j}} \chi_q\biggl(\prod_{l \mid D_{ip+j}/q} \frac{1}{l^{ij}} \biggr) \biggr] \times \biggl[\prod_{l \mid D_{ip+j}} \chi_l\biggl(\frac{B^j}{l^{ij}}\biggr) \biggr]$$

and grouping products and renaming the variable l to q in the last term gives

$$= \biggl[\prod_{q \mid D/D_{ip+j}} \chi_q\biggl(\frac{1}{D_{ip+j}^{ij}}\biggr) \biggr] \biggl[\prod_{q \mid D_{ip+j}} \chi_q\biggl(\frac{q^{ij}}{D_{ip+j}^{ij}}\biggr) \biggr] \biggl[\prod_{q \mid D_{ip+j}} \chi_q\biggl(\frac{B^j}{q^{ij}}\biggr) \biggr]. \tag{4-8}$$

Define $A_{i,j}$, $B_{i,j}$, $C_{i,j}$ to be respectively the first, second, and third factors in (4-8).

Let $\tilde{D} = \prod_{i,j=0}^{p-1} D_{ip+j}^{ij}$, $\tilde{D}_j = \prod_{i=0}^{p-1} D_{ip+j}^{ij}$ and $\overline{D}_j = \prod_{i=0}^{p-1} D_{ip+j}$. Then the last two terms in (4-8) can be combined:

$$B_{i,j} \cdot C_{i,j} = \prod_{q \mid D_{ip+j}} \chi_q\biggl(\frac{q^{ij}}{D_{ip+j}^{ij}}\biggr) \prod_{q \mid D_{ip+j}} \chi_q\biggl(\frac{B^j}{q^{ij}}\biggr) = \prod_{q \mid D_{ip+j}} \chi_q\biggl(\frac{B^j}{D_{ip+j}^{ij}}\biggr). \tag{4-9}$$

Let $B_j = \prod_{i=0}^{p-1} B_{i,j} C_{i,j}$.

Taking the product over $i = 0, \dots, p-1$ of the first factor in (4-8) gives

$$\prod_{i=0}^{p-1} A_{i,j} = \Bigg[\prod_{q \mid D/\bar{D}_j} \chi_q\bigg(\frac{1}{\tilde{D}_j}\bigg)\Bigg]\Bigg[\prod_{i=0}^{p-1}\prod_{q \mid D_{ip+j}} \chi_q\bigg(\frac{D_{ip+j}^{ij}}{\tilde{D}_j}\bigg)\Bigg]. \tag{4-10}$$

Define $E_{1,j}$, $E_{2,j}$ to be respectively the first and second factor in (4-10).

So far we have shown

$$|\mathrm{im}(\varphi_K)| = \frac{1}{p^{\omega(D)+1}} \sum_{(D_0,\dots,D_{p^2-1})} \prod_{j=0}^{p-1} (E_{1,j} \cdot E_{2,j} \cdot B_j). \tag{4-11}$$

Now we can compute

$$\prod_{j=0}^{p-1} E_{1,j} = \prod_{j=0}^{p-1} \prod_{q \mid \bar{D}_j} \chi_q\bigg(\frac{\tilde{D}_j}{\tilde{D}}\bigg) \tag{4-12}$$

and

$$\prod_{j=0}^{p-1} (E_{2,j} B_j) = \prod_{i,j=0}^{p-1} \prod_{q \mid D_{ip+j}} \chi_q\bigg(\frac{D_{ip+j}^{ij}}{\tilde{D}_j}\bigg) \prod_{q \mid D_{ip+j}} \chi_q\bigg(\frac{B^j}{D_{ip+j}^{ij}}\bigg) = \prod_{j=0}^{p-1} \prod_{q \mid \bar{D}_j} \chi_q\bigg(\frac{B^j}{\tilde{D}_j}\bigg). \tag{4-13}$$

Plugging (4-12) and (4-13) into (4-11) we get

$$\begin{aligned}
|\mathrm{im}(\varphi_K)| &= \frac{1}{p^{\omega(D)+1}} \sum_{(D_0,\dots,D_{p^2-1})} \prod_{j=0}^{p-1} \prod_{q \mid \bar{D}_j} \chi_q\bigg(\frac{B^j}{\tilde{D}_j}\bigg) \prod_{q \mid \bar{D}_j} \chi_q\bigg(\frac{\tilde{D}_j}{\tilde{D}}\bigg) \\
&= \frac{1}{p^{\omega(D)+1}} \sum_{(D_0,\dots,D_{p^2-1})} \prod_{j=0}^{p-1} \prod_{l \mid \bar{D}_j} \chi_l\bigg(\frac{B^j}{\tilde{D}}\bigg) \\
&= \frac{1}{p^{\omega(D)+1}} \sum_{(D_0,\dots,D_{p^2-1})} \prod_{i,j=0}^{p-1} \prod_{l \mid D_{ip+j}} \chi_l\bigg(\frac{B^j}{\tilde{D}}\bigg).
\end{aligned}$$

From the definition of B we have

$$B = \prod_{i=0}^{p-1} (D_{ip} D_{ip+1} \cdots D_{ip+p-1})^i$$

hence for $0 \le u_1, u_2, v_1, v_2 \le p-1$ the exponent of $D_{u_1 p + u_2}$ in $B^{v_2}/\tilde{D}$ is

$$u_1 v_2 - u_1 u_2 = u_1(v_2 - u_2).$$

Let $u = (u_1, u_2) \in \mathbb{F}_p^2$ and $v = (v_1, v_2) \in \mathbb{F}_p^2$. Recall we defined $\Phi : \mathbb{F}_p^2 \to \mathbb{F}_p$ by

$$\Phi(u, v) = u_1(v_2 - u_2).$$

Thus relabelling $D_{ip+j} \mapsto D_{(i,j)}$ we conclude that

$$|\mathrm{im}(\varphi_K)| = \frac{1}{p^{\omega(D)+1}} \sum_{(D_{(0,0)},\ldots,D_{(p-1,p-1)})} \prod_{v\in\mathbb{F}_p^2} \prod_{l \mid D_v} \chi_l\biggl(\prod_{u\in\mathbb{F}_p^2} D_u^{\Phi(u,v)}\biggr). \tag{4-14}$$

as required. □

5. An expression for the k-th moment

Define

$$S_k(X) = \sum_{K, D_K < X^{p-1}} |\mathrm{im}(\varphi_K)|^k.$$

Computing $S_k(X)$ will allow us to determine the k-th moment

$$M(k) = \lim_{X\to\infty} \frac{S_k(X)}{\sum_{K,D_K<X^{p-1}} 1}$$

of the function $|\mathrm{im}(\varphi_K)|$. We will then show that knowing $M(k)$ for all $k \in \mathbb{Z}_{\geq 1}$ will be enough to determine the distribution of the values of $|\mathrm{im}(\varphi_K)|$. Our goal for the remainder of the paper will thus be computing $S_k(X)$.

We want to use Theorem 4.2 to obtain a formula for $|\mathrm{im}(\varphi_K)|^k$.

Proposition 5.1. *For each degree p cyclic field K let σ_K denote a generator of the Galois group and D_K the discriminant and $\chi = \prod_{l \mid D_K} \chi_l$ a corresponding character. Then for any $k \in \mathbb{Z}_{k\geq 1}$*

$$|\mathrm{im}(\varphi_K)|^k = \frac{1}{p^k \cdot p^{k\omega(\tilde{D}_K)}} \sum_{(D_u)} \prod_{v\in(\mathbb{F}_p^2)^k} \prod_{\substack{l \text{ prime} \\ l \mid D_v}} \chi_l\biggl(\prod_{u\in(\mathbb{F}_p^2)^k} D_u^{\Phi_k(u,v)}\biggr)$$

where the sum is over p^{2k}-tuples of coprime positive integers (D_u) indexed by $u \in (\mathbb{F}_p^2)^k$ satisfying $\prod_{u\in(\mathbb{F}_p^2)^k} D_u = \tilde{D}_K$.

Proof. From Theorem 4.2 we see $|\mathrm{im}(\varphi_K)|^k$ involves a k-fold product of summations over factorizations of $\tilde{D}_K$, so we want to simultaneously consider k different factorizations of $\tilde{D}_K$. We follow the same method as in [Fouvry and Klüners 2007, pages 471–472], and denote any k factorizations of $\tilde{D}_K$ as

$$\tilde{D}_K = \prod_{u_1\in\mathbb{F}_p^2} D_{u_1}^{g} = \cdots = \prod_{u_k\in\mathbb{F}_p^2} D_{u_k}^{(k)}$$

where each index $u_i \in \mathbb{F}_p^2$ (note this differs from the notation in the previous section). Define $D_{u_1,\ldots,u_k} = \gcd(D_{u_1}^{g}, \ldots, D_{u_k}^{(k)})$. From this we obtain a further factorization of each $D_{u_l}^{(l)}$ by

$$D_{u_l}^{(l)} = [\prod_{\substack{1\leq j\leq k \\ j\neq l}} [\prod_{u_j\in\mathbb{F}_p^2} D_{u_1,\ldots,u_k}.$$

Hence taking (4-4) in Theorem 4.2 to the k-th power we get

$$\frac{1}{p^k\cdot p^{k\omega(\tilde{D}_K)}}\sum_{(D^g_{u_1})}\cdots\sum_{(D^{(k)}_{u_k})}\prod_{i=1}^{k}\prod_{v_i\in\mathbb{F}_p^2}\prod_{\substack{l\text{ prime}\\ l\,\mid\, D^{(i)}_{v_i}}}\chi_l\Bigg(\prod_{u_i\in\mathbb{F}_p^2}(D^{(i)}_{u_i})^{\Phi(u_i,v_i)}\Bigg)$$

where the summations are over p^2-tuples of positive integers $(D^{(i)}_{u_i})$ such that $\prod_{u_i\in\mathbb{F}_p^2}D^{(i)}_{u_i}=\tilde{D}_K$. By multiplicativity of the χ_l we can simplify this as

$$\frac{1}{p^k\cdot p^{k\omega(\tilde{D}_K)}}\sum_{(D_{u_1,\dots,u_k})}\prod_{(v_1,\dots,v_k)\in(\mathbb{F}_p^2)^k}\prod_{\substack{l\text{ prime}\\ l\,\mid\, D_{v_1,\dots,v_k}}}\chi_l\Bigg(\prod_{(u_1,\dots,u_k)\in(\mathbb{F}_p^2)^k}D_{u_1,\dots,u_k}^{\Phi(u_1,v_1)+\cdots+\Phi(u_k,v_k)}\Bigg)$$

where the sum is over all p^{2k}-tuples of positive integers $(D_{u_1,\dots,u_k})$ such that

$$\prod_{(u_1,\dots,u_k)\in(\mathbb{F}_p^2)^k}D_{u_1,\dots,u_k}=\tilde{D}_K.$$

To simplify notation we let $u=(u_1,\dots,u_k)$, $v=(v_1,\dots,v_k)$, and $\Phi_k(u,v)=\sum_{i=1}^k\Phi(u_i,v_i)$. Then the expression becomes

$$\frac{1}{p^k\cdot p^{k\omega(\tilde{D}_K)}}\sum_{(D_u)}\prod_{v\in(\mathbb{F}_p^2)^k}\prod_{\substack{l\text{ prime}\\ l\,\mid\, D_v}}\chi_l\Bigg(\prod_{u\in(\mathbb{F}_p^2)^k}D_u^{\Phi_k(u,v)}\Bigg)$$

where the sum is over p^{2k}-tuples of coprime positive integers with $\prod_{u\in(\mathbb{F}_p^2)^k}D_u=\tilde{D}_K$. □

We now sum the expression from Proposition 5.1 over all degree p cyclic fields with discriminant up to X. To this end we define the following notation:

Let $\mathcal{P}(X)$ denote the set of p^{2k}-tuples of coprime positive integers (D_u) indexed by $u=(u_1,\dots,u_k)\in\mathbb{F}_p^{2k}$ (with $u_i\in\mathbb{F}_p^2$) whose prime factors are congruent to 1 mod p or equal to p and $p^{\mathrm{ord}_p(D)}D<X$ where we denote $D=\prod_{u\in(\mathbb{F}_p^2)^k}D_u$.

Let $\mathcal{C}(D)$ be the set of tuples of nontrivial order p characters $(\chi_{l'})_{l'\mid D}$ (see Section 2B).

Theorem 5.2. *For each degree p cyclic field K let σ_K denote a generator of the Galois group and D_K the discriminant. Then for any $k\in\mathbb{Z}_{k\geq 1}$*

$$\sum_{K,D_K<X^{p-1}}|\mathrm{im}(\varphi_K)|^k=\frac{1}{(p-1)\cdot p^k}\sum_{(D_u)\in\mathcal{P}(X)}\sum_{(\chi_{l'})\in\mathcal{C}(D)}\frac{\mu^2(D)}{p^{k\omega(D)}}\times\prod_{v\in\mathbb{F}_p^{2k}}\prod_{\substack{l\textit{ prime}\\ l\,\mid\, D_v}}\chi_l\Bigg(\prod_{u\in(\mathbb{F}_p^2)^k}D_u^{\Phi_k(u,v)}\Bigg)$$

where on the right hand side we denote $D=\prod_{u\in(\mathbb{F}_p^2)^k}D_u$ and $\Phi_k(u,v)=\sum_{i=1}^k\Phi(u_i,v_i)$ with Φ defined in (4-2).

Note that we are summing over cyclic degree p fields satisfying $D_K<X^{p-1}$ but on the right-hand side the condition is $p^{\mathrm{ord}_p(D)}D<X$.

Proof. Fix a degree p cyclic field K. Summing ((4-4) over all tuples of order p characters $(\chi_{l'})_{l' \mid \tilde{D}_K}$ characters corresponds to summing over all degree p cyclic fields of discriminant D_K but overcounts by a factor of $p-1$ since for a fixed discriminant D_K the characters $\prod_{l \mid \tilde{D}_K} \chi_l$ and $\prod_{l \mid \tilde{D}_K} \chi_l^j$ for $1 \le j \le p-1$ correspond to the same field (see Section 2B). Thus for any $D \in \mathbb{Z}$ which is a discriminant of a degree p cyclic field, by Proposition 5.1 we get

$$\sum_{K, D_K = D} |\mathrm{im}(\varphi_K)|^k = \frac{1}{(p-1)} \frac{1}{p^k p^{k\omega(D)}} \sum_{(D_u)} \sum_{(\chi_{l'}) \in \mathcal{C}(D)} \prod_{v \in \mathbb{F}_p^{2k}} \prod_{\substack{l \mid D_v \\ l \text{ prime}}} \chi_l\Bigl(\prod_{u \in \mathbb{F}_p^{2k}} D_u^{\Phi_k(u,v)}\Bigr)$$

where the sum is over p^{2k}-tuples of coprime positive integers (D_u) indexed by $u \in (\mathbb{F}_p^2)^k$ satisfying $\prod_{u \in (\mathbb{F}_p^2)^k} D_u = \tilde{D}$ ($\tilde{D}$ defined as in (3-1)). Since we are interested in computing the average over all degree p Galois fields we sum over these to get

$$\sum_{K, D_K < X^{p-1}} |\mathrm{im}(\varphi_K)|^k = \frac{1}{(p-1)p^k} \sum_{\substack{D \in \mathbb{Z} \\ 0 < p^{\mathrm{ord}_p(D)} D < X}} \mu^2(D) \frac{1}{p^{k\omega(D)}} \sum_{(D_u)} \sum_{(\chi_{l'}) \in \mathcal{C}(D)} \prod_{v \in \mathbb{F}_p^{2k}} \prod_{\substack{l \mid D_v \\ l \text{ prime}}} \chi_l\Bigl(\prod_{u \in \mathbb{F}_p^{2k}} D_u^{\Phi_k(u,v)}\Bigr)$$

and $\prod_{u \in (\mathbb{F}_p^2)^k} D_u = D$. Then by definition of $\mathcal{P}(X)$ we obtain

$$\sum_{K, D_K < X^{p-1}} |\mathrm{im}(\varphi_K)|^k = \frac{1}{(p-1)p^k} \sum_{(D_u) \in \mathcal{P}(X)} \mu^2(D) \frac{1}{p^{k\omega(D)}} \times \sum_{(\chi_{l'}) \in \mathcal{C}(D)} \prod_{v \in \mathbb{F}_p^{2k}} \prod_{\substack{l \mid D_v \\ l \text{ prime}}} \chi_l\Bigl(\prod_{u \in \mathbb{F}_p^{2k}} D_u^{\Phi_k(u,v)}\Bigr).$$

This proves the theorem. □

The goal will now be an asymptotic analysis of this formula.

6. Analytic tools

We list the analytic results that will be needed in the sequel. The first two we take directly from [Fouvry and Klüners 2007].

Lemma 6.1. *There exists an absolute constant B_0, such that for every $X \ge 3$ and every $l \ge 0$ we have*

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

Lemma 6.2. *Let $\gamma \in \mathbb{R}$ with $\gamma > 0$. Then we have*

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

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

Let $\mathcal{O} = \mathbb{Z}[\zeta_3]$, the ring of integers of the quadratic extension $\mathbb{Q}(\zeta_3)$. Let $\left(\frac{x}{y}\right)_3$ denote the cubic residue symbol for $x, y \in \mathcal{O}$ coprime. Let $N(\cdot) : \mathbb{Q}(\zeta_3) \to \mathbb{Q}$ denote the norm function. In the following A, B, Q will denote positive integers.

We will need the following results for estimating bilinear sums. They are all versions of the large sieve inequality. The first two containing the $(AB)^\epsilon$-type factor will be used when A and B are close together, and the latter two which do not contain this factor will be used when A and B are far apart. The first is Theorem 2 from [Heath-Brown 2000].

Proposition 6.3. *Let c_n be a sequence of complex numbers indexed by elements of $\mathcal{O}$. Then for any $\epsilon > 0$*

$$\sum_{\substack{m\in\mathcal{O},\\ N(m)\leq A}} \left| \sum_{\substack{n\in\mathcal{O},\\ N(n)\leq B}} \mu^2(N(n)N(m))c_n\left(\frac{n}{m}\right)_3\right|^2 \ll_\epsilon (A+B+(AB)^{2/3})(AB)^\epsilon \sum_{n\in\mathcal{O}}|c_n|^2$$

where each of the sums are over square-free elements $m, n \in \mathcal{O}$ congruent to $1 \bmod 3$.

Next we have a version for cubic Dirichlet characters and sums over integers, which is Theorem 1.4 from [Baier and Young 2010].

Proposition 6.4. *Let c_m be any sequence of complex numbers indexed by $\mathbb{N}$. Then for any $\epsilon > 0$*

$$\sum_{\substack{q\in\mathbb{N}\\ Q<q<2Q}} \sum_{\chi \bmod q} \left| \sum_{\substack{m\in\mathbb{N}\\ A<m<2A}} c_m\mu^2(m)\chi(m)\right|^2 \ll_\epsilon (Q^{11/9}+Q^{2/3}A)(QA)^\epsilon \sum_{m\in\mathbb{N}} \mu^2(m)|c_m|^2$$

where the second sum is over χ which are primitive Dirichlet characters satisfying $\chi^3 = 1$.

The next version is Theorem 7.13 from [Iwaniec and Kowalski 2004, Section 7.5, page 179] (due to Bombieri and Davenport) and applies to all Dirichlet characters (not necessarily cubic).

Proposition 6.5. *Let c_n be a sequence of complex numbers indexed by $\mathbb{N}$. Then*

$$\sum_{\substack{q\in\mathbb{N}\\ q<Q}} \sum_{\chi \bmod q} \left| \sum_{\substack{m\in\mathbb{N}\\ A<m<2A}} c_m\mu^2(m)\chi(m)\right|^2 \ll (Q^2+A) \sum_{m\in\mathbb{N}} |c_m|^2$$

where the second sum is over χ which are primitive Dirichlet characters.

Finally we will need a generalized version of Siegel–Walfisz for character sums, stated as Main Theorem in [Goldstein 1970]. We state a slightly weaker simplified version here.

Proposition 6.6. *Let $\epsilon > 0$. Let $K/\mathbb{Q}$ be Galois of degree n and let χ be a nontrivial finite Hecke character of K with conductor f_χ. Then there exists a positive constant $c = c(\epsilon)$, not depending on K or χ such that*

$$\sum_{\substack{\mathfrak{p}\subset\mathcal{O}_K\ prime\\ N(\mathfrak{p})\leq x\\ (\mathfrak{p},f_\chi)=1}} \chi(\mathfrak{p}) = O(dx\log^2 x \exp(-cn(\log x)^{1/2}/d))$$

where $d = n^3|D_K N(f_\chi)|^\epsilon c^{-n}$. The implied constant does not depend on K or χ.

We now prove another version of a large sieve bound for cubic characters.

We require a preliminary lemma. The next result is Exercise 2 from [Iwaniec and Kowalski 2004, Section 7.4, page 178]. In their terminology a set of $\alpha_r = (\alpha_{r,1}, \dots, \alpha_{r,k}) \in \mathbb{R}^k$ is δ-spaced if $\max_i |\alpha_{r,i} - \alpha_{r',i}| \geq \delta$ for all $r \neq r'$. The definition extends to elements of $\mathbb{R}^k/\mathbb{Z}^k$ by choosing representatives in $\mathbb{R}^k$ for which $|\alpha_{r,i} - \alpha_{r',i}|$ is minimal, for all i (that is, which make the spacing minimal).

Lemma 6.7. *Let $d \geq 1$ and $\delta > 0$ and let $\alpha_r = (\alpha_{r,1}, \dots, \alpha_{r,d})$ be δ-spaced points in $\mathbb{R}^d/\mathbb{Z}^d$ and a_n a sequence in $\mathbb{C}$ indexed by $n = (n_1, \dots, n_d) \in \mathbb{Z}^d$ with $1 \leq n_i \leq N$. Then*

$$\sum_r \left| \sum_n a_n \exp(2\pi i (n \cdot \alpha_r)) \right|^2 \ll_d (\delta^{-d} + N^d) \sum_n |a_n|^2.$$

Proposition 6.8. *For each $n \in \mathcal{O}$ let ψ_n be a primitive cubic Hecke character of modulus $(n) \subset \mathcal{O}$.*

For any $d \in \mathcal{O}$ let z_d be the smallest positive integer such that $d^{-1} = z/z_d$ for some $z \in \mathcal{O}$ (clearly $z_d \leq |N(d)|$). Let $\mathcal{P}(B) \subset \mathcal{O}$ be a set of elements d satisfying $z_d < B$.

For all $d \in \mathcal{O}$ let $a_d \in \mathbb{C}$ such that $|a_d| \leq 1$. Then

$$\sum_{n \in \mathcal{P}(B)} \left| \sum_{\substack{m \in \mathcal{O} \\ N(m) \leq A}} a_m \psi_n(m) \right|^2 \ll (B^2 + A)A.$$

Proof. For $r, n \in \mathcal{O}$ define the generalized Gauss sum for the character ψ as

$$g(r, n) = \sum_{d \in (\mathcal{O}/n)^\times} \psi_n(d) \check{e}(rd/n)$$

where $\check{e}(z) = \exp(2\pi i (z + \bar{z}))$. It satisfies the property (see [Baier and Young 2010, Section 2.2])

$$g(rs, n) = \overline{\psi_n(s)} g(r, n)$$

for $s \in \mathcal{O}$ coprime to n. We will write $g(n) = g(1, n)$.

We can write the sum in the statement as

$$\sum_n \left| \sum_m a_m \psi_n(m) \right|^2 = \sum_n \frac{1}{|g(n)|^2} \left| \sum_m a_m g(m, n) \right|^2 = \sum_n \frac{1}{|g(n)|^2} \left| \sum_{d \in (\mathcal{O}/n)^\times} \psi_n(d) \sum_m a_m \check{e}(dm/n) \right|^2.$$

Since ψ_n is a primitive character of $(\mathcal{O}/n)^\times$ we can sum over all such characters to get the bound

$$\sum_n \left| \sum_m a_m \psi_n(m) \right|^2 \leq \sum_n \frac{1}{|g(n)|^2} \sum_\chi \left| \sum_{d \in (\mathcal{O}/n)^\times} \chi(d) \sum_m a_m \check{e}(dm/n) \right|^2$$

where the summation $\sum_\chi$ is over primitive characters of $(\mathcal{O}/n)^\times$. Then expanding the square and using orthogonality of characters we obtain, with the notation $b_d = \sum_m a_m \check{e}(dm/n)$,

$$\begin{aligned}\sum_n \left|\sum_m a_m \psi_n(m)\right|^2 &\le \sum_n \frac{1}{|g(n)|^2} \sum_\chi \sum_{d,d' \in (\mathcal{O}/n)^\times} \chi(d)\bar{\chi}(d') b_d \bar{b}_{d'} \\ &\le \sum_n \frac{1}{|g(n)|^2} \sum_{d,d' \in (\mathcal{O}/n)^\times} b_d \bar{b}_{d'} \sum_\chi \chi(dd'^{-1}) \\ &\le \sum_n \frac{|(\mathcal{O}/n)^\times|}{|g(n)|^2} \sum_{d \in (\mathcal{O}/n)^\times} \left|\sum_m a_m \check{e}(dm/n)\right|^2 \\ &\le \sum_n \sum_{d \in (\mathcal{O}/n)^\times} \left|\sum_m a_m \check{e}(dm/n)\right|^2. \end{aligned} \tag{6-1}$$

We now want to apply the multivariable large sieve inequality of Lemma 6.7 so we will rewrite the summation accordingly.

Let $R = \{(n,d) \mid n \in \mathcal{P}(B), d \in (\mathcal{O}/n)^\times\}$. For any $(n,d) \in R$, using $d \in \mathcal{O}$ to also denote any choice of representative of $d \in (\mathcal{O}/n)^\times$ (everything that follows will be independent of such a choice), write $d/n = d_1 + \zeta_3 d_2$ in the basis $\{1, \zeta_3\}$ with $d_i \in \mathbb{Q}$ and similarly write $m = s_1 + \zeta_3 s_2$ with $s_i \in \mathbb{Z}$. Then a computation shows $dm/n = (d_1 s_1 - d_2 s_2, d_1 s_2 + d_2 s_1 - d_2 s_2)$ in coordinates in $\{1, \zeta_3\}$, and

$$\operatorname{tr}(dm/n) = dm/n + \overline{dm}/\bar{n} = s_1(2d_1 - d_2) + s_2(-d_1 - d_2) = (s_1, s_2) \cdot (2d_1 - d_2, d_1 - d_2).$$

So given $r = (n,d) \in R$ define $\alpha_r = (2d_1 - d_2, d_1 - d_2) \in \mathbb{Q}^2$. Then

$$\check{e}(dm/n) = \exp(2\pi i \operatorname{tr}(dm/n)) = \exp(2\pi i (s_1, s_2) \cdot \alpha_r)$$

Hence we can rewrite (6-1) as

$$\sum_n \left|\sum_m a_m \psi_n(m)\right|^2 \le \sum_{r \in R} \left| \sum_{(s_1,s_2) \in \mathbb{Z}^2} a_{s_1 + \zeta s_2} \exp(2\pi i (s_1, s_2) \cdot \alpha_r)\right|^2 \tag{6-2}$$

where $s_1, s_2 \ll A^{1/2}$ since $N(m) \le A$ (see Section 2C).

We claim the sequence $S = \{\alpha_r\}_{r \in R}$ is $1/B$-spaced (see definition before Lemma 6.7). For any $(n,d) \in R$, d is coprime to n so the map $R \to \mathbb{Q}(\zeta_3)$ defined by $(n,d) \mapsto d/n$ is injective. Furthermore $(d_1, d_2) \to (2d_1 - d_2, d_1 - d_2)$ is an invertible linear map, hence the elements of S are all distinct. Note that for any distinct $a/c, b/c \in \mathbb{Q}$ we have $|a/c - b/c| \ge 1/c$. Hence the spacing of a set in $\mathbb{Q}^2$ is bounded below in terms of the denominators of the coordinates of its elements.

Since $n \in \mathcal{P}(B)$ there exists $z \in \mathcal{O}$ such that $nz = z_n \in \mathbb{Z}$ and $z_n < B$. We can write

$$\frac{d}{n} = \frac{dz}{nz} = \frac{a}{z_n} + \frac{b}{z_n}\zeta$$

for some $a, b \in \mathbb{Z}$. Since $z_n \le B$ it follows that S is $1/B$-spaced as required.

Thus by Lemma 6.7 and (6-2) we get

$$\sum_n \Bigl|\sum_m a_m \psi_n(m)\Bigr|^2 \ll \sum_{r\in R}\Bigl|\sum_{s_1,s_2} a_{s_1+\zeta s_2}\exp(2\pi i(s_1,s_2)\cdot\alpha_r)\Bigr|^2 \ll (B^2+A)A. \qquad \square$$

7. Determining the main term

We start with the expression for $\sum_{K,D_K<X^{p-1}}|\mathrm{im}(\varphi_K)|^k$ which we derived in Theorem 5.2,

$$S_k(X)=\frac{1}{(p-1)\cdot p^k}\sum_{(D_u)\in\mathcal{P}(X)}\sum_{(\chi_{l'})\in\mathcal{C}(\prod D_u)}\frac{\mu^2(\prod D_u)}{p^{k\omega(\prod D_u)}}\prod_v\prod_{l\,|\,D_v}\chi_l\Bigl(\prod_u D_u^{\Phi_k(u,v)}\Bigr) \tag{7-1}$$

and recall the notation:

- $\mathcal{P}(X)$ denotes the set of p^{2k}-tuples of coprime positive integers (D_u) indexed by $u=(u_1,\dots,u_k)\in\mathbb{F}_p^{2k}$ whose prime factors are congruent to 1 mod p or equal to p and $p^{\mathrm{ord}_p(D)}D<X$ (with $D=\prod_{u\in\mathbb{F}_p^{2k}}D_u$).
- $\mathcal{C}(D)$ denotes the set of tuples of nontrivial order p characters $(\chi_{l'})_{l'\,|\,D}$.
- $\Phi_k(u,v)=\sum_{i=1}^k\Phi(u_i,v_i)$ with Φ defined in (4-2).

For the remainder of the paper we will use the convention that the implied constants in any big-O notation which appears are allowed to depend on p and k, but not X.

Fix $k\in\mathbb{Z}_{\ge 1}$ and let $\Delta=1+\log^{-(p-1)\cdot p^k}X$. Define $\boldsymbol{A}$ to be a p^{2k}-tuple of variables $(A_u)_{u\in\mathbb{F}_p^{2k}}$ with each $A_u=\Delta^j$ for some $j\ge 0$. We can partition $S_k(X)$ according to the various $\boldsymbol{A}$.

Let $\mathcal{P}(X,\boldsymbol{A})\subset\mathcal{P}(X)$ be the subset of tuples (D_u) satisfying $A_u\le D_u<\Delta A_u$ for all $u\in\mathbb{F}_p^{2k}$. Let $S_k(X,\boldsymbol{A})$ be the above sum (7-1) but now restricted in the first summation to tuples $(D_u)\in\mathcal{P}(X,\boldsymbol{A})$. Thus we have

$$S_k(X)=\sum_{\boldsymbol{A}}S_k(X,\boldsymbol{A})$$

summing over all $\boldsymbol{A}$ with $\prod_{u\in\mathbb{F}_p^{2k}}A_u<X$.

Note that since $\Delta=1+\log^{-(p-1)\cdot p^k}X$ there are $O((\log X)^{p^{2k}(1+(p-1)\cdot p^k)})$ possible $\boldsymbol{A}$ with $S_k(X,\boldsymbol{A})$ not empty. This is since there are $O((\log X)^{(1+(p-1)\cdot p^k)})$ choices for each $1<A_u\le X$.

We now consider certain families of tuples $\boldsymbol{A}$ for which $S_k(X,\boldsymbol{A})$ makes a negligible contribution to the sum. These will be the same as the four families from Section 5.4 in [Fouvry and Klüners 2007]. In the case of the first the argument is identical but for completeness we reproduce it here. The proofs in the remaining three cases must be generalized.

First we reduce the sum $S_k(X)$ to terms where all of the D_u satisfy $\omega(D_u)\le\Omega$, where we define $\Omega=e(p-1)p^{2k}(\log\log X+B_0)$ with B_0 the constant given by Lemma 6.1.

Let $\mathcal{P}(X,\boldsymbol{A},\Omega)\subset\mathcal{P}(X,\boldsymbol{A})$ be the subset of tuples (D_u) additionally satisfying $\omega(D_u)\le\Omega$ for all $u\in\mathbb{F}_p^{2k}$. Let $S_k(X,\boldsymbol{A},\Omega)$ be the above sum (7-1) but now restricted in the first summation to tuples $(D_u)\in\mathcal{P}(X,\boldsymbol{A},\Omega)$.

Lemma 7.1. *With the above notation, for all tuples $\boldsymbol{A}$*

$$\sum_{\boldsymbol{A}} S_k(X,\boldsymbol{A}) = \sum_{\boldsymbol{A}} S_k(X,\boldsymbol{A},\Omega) + O\biggl(\frac{X}{\log X}\biggr).$$

Proof. Let S_0 be the sum of the terms in (7-1) where not all of the D_u satisfy $\omega(D_u) \le \Omega$. We will bound S_0. Let $n = \prod_{u\in\mathbb{F}_p^{2k}} D_u$. We can trivially bound (7-1) by setting all $\chi_l = 1$. For any positive square-free $n \in \mathbb{Z}$ we have $\bigl|\bigl\{(D_u) \in \mathcal{P}(X) \mid \prod_{u\in\mathbb{F}_p^{2k}} D_u = n\bigr\}\bigr| = p^{2k\omega(n)}$ (this is just the number of ways of writing n as a product of p^{2k} positive integers) and $|\mathcal{C}(n)| = (p-1)^{\omega(n)}$. Thus applying the trivial bound to S_0 gives

$$S_0 \ll \sum_{n<X,\omega(n)>\Omega} \mu^2(n)(p^k(p-1))^{\omega(n)}.$$

Then splitting the sum up by the number of prime factors and applying Lemma 6.1 we get the bound

$$\begin{aligned}\sum_{n<X,\omega(n)>\Omega} \mu^2(n)(p^k(p-1))^{\omega(n)} &\ll \sum_{l\ge\Omega} \frac{X}{\log X}(p^k(p-1))^l \frac{(\log\log X + B_0)^l}{l!}\\ &\ll \frac{X}{\log X}\sum_{l\ge\Omega}\biggl(\frac{p^k(p-1)(\log\log X + B_0)}{l/e}\biggr)^l\\ &\ll \frac{X}{\log X}\sum_{l\ge\Omega}\biggl(\frac{1}{p^k}\biggr)^l\\ &\ll \frac{X}{\log X}\end{aligned}$$

where in the second-last inequality we are using $l/e \ge (p-1)p^{2k}(\log\log X + B_0)$ by definition of Ω. □

Thus we can assume in the remainder that all variables D_u satisfy $\omega(D_u) \le \Omega$ (we will only need this fact to bound family 4 in Section 7D).

7A. *The first family.* Note that it is possible there exists an $\boldsymbol{A}$ for which $S_k(X,\boldsymbol{A},\Omega)$ is not empty, but

$$\prod_{u\in\mathbb{F}_p^{2k}} \Delta A_u > X. \tag{7-2}$$

Thus for any tuple $(D_u) \in \mathcal{P}(X,\boldsymbol{A},\Omega)$ the condition $\prod_u D_u < X$ imposes dependencies between the D_u. We wish to remove this dependency to allow application of subsequent analytic results in which we will sum over each D_u independently.

Let $\mathcal{F}_1$ denote the set of $\boldsymbol{A}$ such that (7-2) is satisfied.

Lemma 7.2. *With the above notation*

$$\sum_{\boldsymbol{A}\in\mathcal{F}_1} S_k(X,\boldsymbol{A},\Omega) \ll X/\log X.$$

Proof. Applying the trivial bound as in Lemma 7.1 and applying Lemma 6.2 we have

$$\sum_{\boldsymbol{A}\in\mathcal{F}_1} S_k(X,\boldsymbol{A},\Omega) \ll \sum_{\Delta^{-p^{2k}}X\le D\le X} (p^k(p-1))^{\omega(D)} \ll (1-\Delta^{-p^{2k}})X(\log X)^{(p-1)\cdot p^k-1}.$$

Using that $(1+x)^\alpha = 1+\alpha x+O(x^2)$ for $x\to 0$, setting $\alpha=-p^{2k}$ and $x=\log^{-(p-1)\cdot p^k} X$ we get

$$\Delta^{-p^{2k}} = (1+\log^{-(p-1)\cdot p^k} X)^{-p^{2k}} = 1-p^{2k}\log^{-(p-1)\cdot p^k} X + O(\log^{-2((p-1)\cdot p^k)} X).$$

This gives the bound

$$\sum_{\boldsymbol{A}\in\mathcal{F}_1} S_k(X,\boldsymbol{A},\Omega) \ll (p^{2k}\log^{-(p-1)\cdot p^k} X + O(\log^{-2((p-1)\cdot p^k)} X))X(\log X)^{(p-1)\cdot p^k-1} \ll X/\log X. \quad \square$$

Thus if $\boldsymbol{A}\notin\mathcal{F}_1$ then any $(D_u)\in\mathcal{P}(X,\boldsymbol{A},\Omega)$ automatically satisfies $\prod_{u\in\mathbb{F}_p^{2k}} D_u < X$ so this condition can be dropped from the definition of $\mathcal{P}(X,\boldsymbol{A},\Omega)$ for $\boldsymbol{A}\notin\mathcal{F}_1$.

7B. *The second family.* We now bound the terms in which the range of summation is too short for too many variables D_u. Let $X^\ddagger = \exp(\log^\eta X)$ for some small $\eta>0$ which we will specify later.

Let $\mathcal{F}_2$ be the set of $\boldsymbol{A}$ which satisfy

$$\text{at most } p^{k-1} \text{ variables satisfy } A_u > X^\ddagger. \tag{7-3}$$

Lemma 7.3. *In the above notation*

$$\sum_{\boldsymbol{A}\in\mathcal{F}_2} S_k(X,\boldsymbol{A},\Omega) \ll X(\log X)^{\eta(p-1)\cdot p^k-1/p}.$$

Proof. Let r be the number of variables greater than $X^\ddagger$. We factor the sum $\sum_{\boldsymbol{A}\in\mathcal{F}_2} S_k(X,\boldsymbol{A},\Omega)$ into two parts corresponding to terms with all variables $D_u\le X^\ddagger$ and terms with all $D_u > X^\ddagger$ and then apply the trivial bound as in the proof of Lemma 7.1. This results in

$$\sum_{\boldsymbol{A}\in\mathcal{F}_2} S_k(X,\boldsymbol{A},\Omega) \ll \sum_{r=0}^{p^{k-1}} \sum_{m<(X^\ddagger)^{p^{2k}-r}} \mu^2(m)(p^{2k}-r)^{\omega(m)}\left(\frac{p-1}{p^k}\right)^{\omega(m)} \sum_{n<X/m} \mu^2(n)\left(\frac{(p-1)r}{p^k}\right)^{\omega(n)}.$$

Then applying Lemma 6.2 to the second term above we get

$$\begin{aligned}\sum_{\boldsymbol{A}\in\mathcal{F}_2} S_k(X,\boldsymbol{A},\Omega) &\ll \sum_{r=0}^{p^{k-1}} \sum_{m<(X^\ddagger)^{p^{2k}-r}} \frac{((p-1)\cdot p^k)^{\omega(m)}}{m} X(\log X)^{(p-1)r/p^k-1} \\ &\ll X\left(\sum_{r=0}^{p^{k-1}} (\log X)^{(p-1)r/p^k-1}\right)\left(\sum_{m<(X^\ddagger)^{p^{2k}-r}} \frac{((p-1)\cdot p^k)^{\omega(m)}}{m}\right).\end{aligned}$$

We trivially bound

$$\sum_{r=0}^{p^{k-1}} (\log X)^{(p-1)r/p^k-1} \ll (\log X)^{(p-1)p^{k-1}/p^k-1} = (\log X)^{-1/p}. \tag{7-4}$$

We will apply Mertens' formula $\prod_{q<x}\bigl(1-\frac{1}{q}\bigr)^{-1} \ll \log x$ (where the product is over q prime), to the second term above. Recall $X^{\ddagger} = \exp(\log^{\eta} X)$. We have

$$\begin{aligned}\sum_{m<(X^{\ddagger})^{p^{2k}-r}} \frac{((p-1)\cdot p^k)^{\omega(m)}}{m} &\ll \prod_{q<(X^{\ddagger})^{p^{2k}-r}} \left(1+\frac{(p-1)\cdot p^k}{q}+\cdots\right) \\ &\ll \Biggl[\prod_{q<(X^{\ddagger})^{p^{2k}-r}} \left(1-\frac{1}{q}\right)^{-1}\Biggr]^{(p-1)\cdot p^k} \\ &\ll (\log X^{\ddagger})^{p^k(p-1)} = (\log X)^{\eta p^k(p-1)}. \end{aligned} \tag{7-5}$$

Putting together (7-4) and (7-5) we get

$$\sum_{A\in\mathcal{F}_2} S_k(X, \boldsymbol{A}, \Omega) \ll X(\log X)^{\eta(p-1)\cdot p^k-1/p}. \qquad \square$$

Clearly for any $\epsilon > 0$, for small enough η we get $X(\log X)^{\eta(p-1)\cdot p^k-1/p} \ll X/(\log X)^{1/p-\epsilon}$.

7C. *The third family, the case $p = 3$.* For the third and fourth families we will first let $p = 3$ and bound the error term unconditionally. In Section 7E we will handle the case of general p under the assumption of GRH.

We define some terminology which will be used in the remainder of the paper.

Definition 7.4. We say indices $u, v \in \mathbb{F}_p^{2k}$ are linked if $\Phi_k(u, v) \neq 0$ or $\Phi_k(v, u) \neq 0$. Otherwise we say they are unlinked. We say a set $\mathcal{U} \subset \mathbb{F}_p^{2k}$ is unlinked if u and v are unlinked for all $u, v \in \mathcal{U}$.

Let $X^{\dagger} = \log^{8(1+9^k(1+2\cdot 3^k))} X$.

Let $\mathcal{F}_3$ denote the set of $\boldsymbol{A}$ such that there are two linked indices u and v with

$$A_u, A_v > X^{\dagger}. \tag{7-6}$$

Fix such an $\boldsymbol{A}$ and two linked indices u, v. We split

$$S_k(X, \boldsymbol{A}, \Omega) = \sum_{i=1}^{4} S_{k,i}(X, \boldsymbol{A}, \Omega) \tag{7-7}$$

into four terms depending on whether $3 \mid D_u$, $3 \mid D_v$, $3 \mid D_w$ for some $w \neq u, v$, or $3 \nmid D_w$ for all w. For simplicity we only present the proof of bounding $S_{k,1}(X, \boldsymbol{A}, \Omega)$, the arguments in the other cases being almost identical.

We now consider two cases: case 1 occurs when both $\Phi_k(u, v)$ and $\Phi_k(v, u)$ are nonzero in (7-1) and case 2 occurs when only one of these is nonzero.

Case 1: Both $\Phi_k(u, v)$ and $\Phi_k(v, u)$ are nonzero.

Define $\mathcal{Q}(X) \subset \mathcal{O}$ to be the set of elements $d \in \mathcal{O}$ congruent to 1 mod $3\mathcal{O}$ which are products of split primes, $\mu^2(N(d)) = 1$ and $N(d) \le X$.

Note $\mathcal{Q}(X)$ is closed under conjugation in $\mathbb{Q}(\zeta_3)$.

Lemma 7.5. *If $\Phi_k(u, v)$ and $\Phi_k(v, u)$ are both nonzero then*

$$S_{k,1}(X, \boldsymbol{A}, \Omega) \ll \sum_{(D_w)_{w\neq u,v} \in \mathcal{P}(X,\boldsymbol{A},\Omega)} \Bigg| \sum_{\substack{d_u \in \mathcal{Q}(\Delta A_u/3) \\ d_v \in \mathcal{Q}(\Delta A_v)}} \mu^2(D_{uv}) a(d_u) a(d_v) \left(\frac{d_u}{d_v}\right)_3 \Bigg| \tag{7-8}$$

with $|a(d_u)|, |a(d_u)| \le 1$, and we denote $D_{uv} = N(d_u)N(d_v)$.

Proof. In the following equation to simplify notation we write $D = \prod_{w \in \mathbb{F}_3^{2k}} D_w$ and $D' = D_u D_v$. Let $\boldsymbol{A}_1$ denote the $(3^{2k} - 2)$-tuple $(A_w)_{w\neq u,v}$ and $\boldsymbol{A}_2$ denote the pair (A_u, A_v).

Then from (7-1) we get by splitting up the summations and bounding

$$\begin{aligned} S_{k,1}(X, \boldsymbol{A}, \Omega) &= \frac{1}{2\cdot 3^k} \sum_{\substack{(D_w)\in\mathcal{P}(X,\boldsymbol{A},\Omega) \\ 3\,\mid\, D_u}} \sum_{(\chi_{l'})\in\mathcal{C}(D)} \frac{\mu^2(D)}{3^{k\omega(D)}} \prod_{y\in\mathbb{F}_3^{2k}} \prod_{l\,\mid\, D_y} \chi_l\Bigg(\prod_{z\in\mathbb{F}_3^{2k}} D_z^{\Phi_k(z,y)}\Bigg) \\ &\le \frac{1}{2\cdot 3^k} \sum_{(D_w)\in\mathcal{P}(X,\boldsymbol{A}_1,\Omega)} \sum_{(\chi_{l'})\in\mathcal{C}(D/D')} \frac{1}{3^{k\omega(D/D')}} \\ &\qquad\qquad \times \Bigg| \sum_{\substack{(D_u,D_v)\in\mathcal{P}(X,\boldsymbol{A}_2,\Omega) \\ 3\,\mid\, D_u}} \sum_{(\chi_{l'})\in\mathcal{C}(D')} \frac{\mu^2(D)}{3^{k\omega(D')}} \prod_{\substack{y=u,v \\ l\,\mid\, D_y}} \chi_l\Bigg(\prod_{z\in\mathbb{F}_3^{2k}} D_z^{\Phi_k(z,y)}\Bigg)\Bigg|. \end{aligned}$$

Suppose $z \in \mathbb{Z}$ is square free and a product of primes congruent to 1 mod $3\mathbb{Z}$. This implies z factors into split primes in $\mathbb{Q}(\zeta_3)$, hence there are exactly $2^{\omega(z)}$ ideals $I \subset \mathcal{O}$ such that $N(I) = z$. Furthermore for any ideal $I \subset \mathcal{O}$ such that $N(I) = z$ there exists a unique element with $(d) = I$ such that $d \in \mathcal{Q}(z)$.

Since $p = 3$ there are 2 nontrivial cubic characters χ_l corresponding to any prime l. They are $\left(\frac{n}{\pi}\right)_3$ and $\left(\frac{n}{\bar{\pi}}\right)_3$ where $\pi, \bar{\pi} \equiv 1$ mod $3\mathcal{O}$ and $N(\pi) = N(\bar{\pi}) = l$. Thus there is a bijection between $\mathcal{C}(z)$ and ideals $I \subset \mathcal{O}$ such that $N(I) = z$, defined by

$$(\chi_{l'})_{l'\,\mid\, z} \mapsto \Bigg(\prod_{\substack{l'\,\mid\, z \\ \chi_{l'} = (\frac{\cdot}{\pi})}} \pi\Bigg). \tag{7-9}$$

This map is injective by unique factorization of ideals and hence surjective since the size of both sets is equal. In particular this implies

$$\sum_{(\chi_{l'})\in\mathcal{C}(D/D')} \frac{1}{3^{k\omega(D/D')}} \le \frac{2^{\omega(D/D')}}{3^{k\omega(D/D')}} < 1.$$

So far we have shown

$$S_{k,1}(X,\boldsymbol{A},\Omega)\ll\sum_{(D_w)\in\mathcal{P}(X,\boldsymbol{A}_1,\Omega)}\Biggl|\sum_{\substack{(D_u,D_v)\in\mathcal{P}(X,\boldsymbol{A}_2,\Omega)\\ 3\,|\,D_u}}\ \sum_{(\chi_{l'})\in\mathcal{C}(D')}\frac{\mu^2(D)}{3^{k\omega(D')}}\prod_{\substack{y=u,v\\ l\,|\,D_y}}\chi_l\Biggl(\prod_{z\in\mathbb{F}_3^{2k}}D_z^{\Phi_k(z,y)}\Biggr)\Biggr|.$$

It also follows from the above bijection (7-9) that in the sum over D_u, D_v we can replace the χ_l with cubic residue symbols

$$\sum_{\substack{(D_u,D_v)\in\mathcal{P}(X,\boldsymbol{A}_2,\Omega)\\ 3\,|\,D_u}}\ \sum_{(\chi_{l'})\in\mathcal{C}(D')}\frac{\mu^2(D)}{3^{k\omega(D')}}\prod_{\substack{y=u,v\\ l\,|\,D_y}}\chi_l\Biggl(2\prod_{z\in\mathbb{F}_3^{2k}}D_z^{\Phi_k(z,y)}\Biggr)$$

$$\ll\sum_{\substack{(D_u,D_v)\in\mathcal{P}(X,\boldsymbol{A}_2,\Omega)\\ 3\,|\,D_u}}\ \sum_{\substack{d_u\in\mathcal{Q}(\Delta A_u/3)\\ N(d_u)=D_u/3}}\ \sum_{\substack{d_v\in\mathcal{Q}(\Delta A_v)\\ N(d_v)=D_v}}\frac{\mu^2(D)}{3^{k\omega(D')}}\prod_{y=u,v}\biggl(\frac{\prod_{z\in\mathbb{F}_3^{2k}}D_z^{\Phi_k(z,y)}}{d_y}\biggr)_3\chi_3\Biggl(\prod_{z\in\mathbb{F}_3^{2k}}D_z^{\Phi_k(z,u)}\Biggr)\tag{7-10}$$

For any d_u in the above sum let

$$b(d_u)=\frac{\mu^2(\prod_{w\neq v}D_w)}{3^{k\omega(N(d_u))}}\prod_{y\neq u,v}\prod_{l\,|\,D_y}\chi_l((3N(d_u))^{\Phi_k(u,y)})\biggl(\frac{\prod_{y\neq u,v}D_v^{\Phi_k(y,u)}}{d_u}\biggr)_3$$

and similarly for $b(d_v)$ $\bigl($which will additionally contain the factor of $\chi_3\bigl(\prod_{z\in\mathbb{F}_3^{2k}}D_z^{\Phi_k(z,u)}\bigr)\bigr)$. Note also that $\mu^2(D)=\mu^2\bigl(\prod_{w\neq v}D_w\bigr)\mu^2\bigl(\prod_{w\neq u}D_w\bigr)\mu^2(D_uD_v)$. Changing notation to $D_u=N(d_u)$ and $D_v=N(d_v)$ and plugging in $b(d_u)$ and $b(d_v)$ we can rewrite (7-10) as

$$\sum_{\substack{d_u\in\mathcal{Q}(\Delta A_u/3)\\ A_u/3\le N(d_u)<\Delta A_u/3}}\ \sum_{\substack{d_v\in\mathcal{Q}(\Delta A_v)\\ A_v\le N(d_v)<\Delta A_v}}\mu^2(D_uD_v)b(d_u)b(d_v)\biggl(\frac{D_u}{d_v}\biggr)_3^{\Phi_k(u,v)}\biggl(\frac{D_v}{d_u}\biggr)_3^{\Phi_k(v,u)}.$$

Note Φ_k is either 1 or 2, and squaring a cubic character is the same as conjugating it. Since $\mathcal{Q}(X)$ is closed under conjugation, removing Φ_k from the exponent permutes the coefficients. As a result of this procedure rename $b(d_u)$ to $a(d_u)$ and $b(d_v)$ to $a(d_v)$ if necessary. Letting $a(d_u)=0$ for $N(d_u)<A_u/3$ we can extend the range of summation to $N(d_u)\le\Delta A_u/3$, and similarly to $N(d_v)\le\Delta A_v$.

For any $D_u=N(d_u)=d_u\bar{d}_u$ and $D_v=N(d_v)=d_v\bar{d}_v$ not divisible by 3 by the properties (2-2) and the law of cubic reciprocity (2-3) we have

$$\biggl(\frac{D_u}{d_v}\biggr)_3\biggl(\frac{D_v}{d_u}\biggr)_3=\biggl(\frac{d_u}{d_v}\biggr)_3\biggl(\frac{\bar{d}_u}{d_v}\biggr)_3\biggl(\frac{d_v}{d_u}\biggr)_3\biggl(\frac{\bar{d}_v}{d_u}\biggr)_3=\biggl(\frac{d_u}{d_v}\biggr)_3^2\biggl(\frac{\bar{d}_u}{d_v}\biggr)_3\biggl(\frac{\bar{d}_u}{d_v}\biggr)_3^2=\biggl(\frac{d_u}{\bar{d}_v}\biggr)_3.\tag{7-11}$$

This proves the lemma. □

Proposition 7.6. *For all $\boldsymbol{A}\in\mathcal{F}_3$*

$$S_{k,1}(X,\boldsymbol{A},\Omega)\ll X/\log^{1+9^k(1+2\cdot 3^k)}X.$$

Proof. We will apply the standard strategy of bounding bilinear sums using Cauchy–Schwarz followed by a large sieve type bound. By Cauchy–Schwarz applied to the summand on the right-hand side in Lemma 7.5 we have, for any fixed tuple $(D_w)_{w\neq u,v} \in \mathcal{P}(X, \boldsymbol{A}, \Omega)$,

$$\Bigg|\sum_{\substack{d_u\in\mathcal{Q}(\Delta A_u/3)\\ d_v\in\mathcal{Q}(\Delta A_v)}} \mu^2(D_{uv})a(d_u)a(d_v)\left(\frac{d_u}{d_v}\right)_3\Bigg| \ll A_v^{1/2}\Bigg(\sum_{d_v\in\mathcal{Q}(\Delta A_v)}\Bigg|\sum_{d_u\in\mathcal{Q}(\Delta A_u/3)}\mu^2(D_{uv})a(d_u)\left(\frac{d_u}{d_v}\right)_3\Bigg|^2\Bigg)^{1/2} \tag{7-12}$$

where $D_{uv} = N(d_u)N(d_v)$.

Note that $d_u \mapsto \mu^2(D_{uv})\big(\frac{d_u}{d_v}\big)_3$ is a primitive cubic Hecke character of modulus (d_v). Note also that $\mathcal{Q}(\Delta A_v)$ satisfies the conditions of the set $\mathcal{P}(\Delta A_v)$ in Proposition 6.8 since by definition for any $d \in \mathcal{Q}(\Delta A_v)$, $N(d) = d\bar{d} < \Delta A_v$. Thus by Proposition 6.8

$$\sum_{d_v\in\mathcal{Q}(\Delta A_v)}\Bigg|\sum_{d_u\in\mathcal{Q}(\Delta A_u/3)}\mu^2(D_{uv})a(d_u)\left(\frac{d_u}{d_v}\right)_3\Bigg|^2 \ll (A_v^2 + A_u)A_u.$$

Plugging this into the bound (7-12) we get

$$\Bigg|\sum_{d_v}\sum_{d_u}\mu^2(D_{uv})a(d_u)a(d_v)\left(\frac{d_u}{d_v}\right)_3\Bigg| \ll A_v^{1/2}((A_v^2 + A_u)A_u)^{1/2} = A_vA_u\left(\frac{A_v}{A_u} + \frac{1}{A_v}\right)^{1/2}. \tag{7-13}$$

By symmetry $\big($recall by cubic reciprocity $\big(\frac{d_u}{d_v}\big)_3 = \big(\frac{d_v}{d_u}\big)_3\big)$ we can also bound this by $A_vA_u\big(\big(\frac{A_u}{A_v} + \frac{1}{A_u}\big)\big)^{1/2}$.

Now by symmetry we can assume without loss of generality that $A_v \le A_u$.

First suppose $A_v^2 < A_u$. Recall $\boldsymbol{A} \in \mathcal{F}_3$ implies $A_v, A_u > X^\dagger = \log^{8(1+9^k(1+2\cdot 3^k))} X$. Plugging the bound (7-13) into Lemma 7.5 we get

$$S_{k,1}(X, \boldsymbol{A}, \Omega) \ll \sum_{(D_w)_{w\neq u,v}\in\mathcal{P}(X,\boldsymbol{A},\Omega)} A_vA_u\left(\frac{A_v}{A_u} + \frac{1}{A_v}\right)^{1/2} \ll X\left(\frac{1}{A_u^{1/2}} + \frac{1}{A_v}\right)^{1/2} \ll X/\log^{1+9^k(1+2\cdot 3^k)} X.$$

Now suppose $A_u \le A_v^2$. Then by Proposition 6.3 we directly get the bound, for any $\epsilon > 0$,

$$\sum_{d_v\in\mathcal{Q}(\Delta A_v)}\Bigg|\sum_{d_u\in\mathcal{Q}(\Delta A_u/3)}\mu^2(D_{uv})a(d_u)\left(\frac{d_u}{d_v}\right)_3\Bigg|^2 \ll (A_u + A_v + (A_uA_v)^{2/3})(A_uA_v)^\epsilon A_u. \tag{7-14}$$

Plugging (7-14) into Lemma 7.5 we get

$$\begin{aligned} S_{k,1}(X, \boldsymbol{A}, \Omega) &\ll \sum_{(D_w)_{w\neq u,v}\in\mathcal{P}(X,\boldsymbol{A},\Omega)} A_v^{1/2}((A_u + A_v + (A_uA_v)^{2/3})(A_uA_v)^\epsilon A_u)^{1/2} \\ &\ll X\left(\left(\frac{1}{A_v} + \frac{1}{A_u} + \frac{1}{(A_uA_v)^{1/3}}\right)(A_uA_v)^\epsilon\right)^{1/2} \\ &\ll X\left(\frac{1}{X^{\dagger 1/2}}\right)^{1/2} \\ &\ll X/\log^{1+9^k(1+2\cdot 3^k)} X. \end{aligned}$$

□

This proves the desired bound in Case 1.

Case 2: Only one of $\Phi_k(u,v)$ and $\Phi_k(v,u)$ is nonzero. Without loss of generality assume $\Phi_k(u,v)$ is nonzero.

For any $X>0$ define $\mathcal{R}(X)$ to be the set of positive $d\in\mathbb{Z}$ which are a product of primes congruent to 1 mod $3\mathbb{Z}$ and $d<X$.

Lemma 7.7. *For any linked indices u and v with $\Phi_k(u,v)\neq 0$ and $\Phi_k(v,u)=0$*

$$S_{k,1}(X,\boldsymbol{A},\Omega)\ll\sum_{(D_w)_{w\neq u,v}\in\mathcal{P}(X,\boldsymbol{A},\Omega)}\Biggl|\sum_{\substack{d_v\in\mathcal{Q}(\Delta A_v)\\ D_u\in\mathcal{R}(\Delta A_u/3)}}\mu^2(D_{uv})a(D_u)a(d_v)\biggl(\frac{D_u}{d_v}\biggr)_3\Biggr|,\tag{7-15}$$

with $|a(d_u)|$, $|a(d_u)|\le 1$, $D_{uv}=D_uN(d_v)$.

Proof. This is a simpler version of the proof of Lemma 7.5. Since $\Phi_k(v,u)=0$ the symbol $\bigl(\frac{D_v}{d_u}\bigr)_3$ does not appear and hence we do not apply cubic reciprocity unlike in that proof. Hence we are left with $\bigl(\frac{D_u}{d_v}\bigr)_3$ which is what appears in the statement above. □

Proposition 7.8. *For all $\boldsymbol{A}\in\mathcal{F}_3$*

$$S_{k,1}(X,\boldsymbol{A},\Omega)\ll X/\log^{1+9^k(1+2\cdot 3^k)}X.$$

Proof. The above expression is no longer symmetric in u and v hence we must consider several subcases.

By Cauchy–Schwarz applied to the summand on the right-hand side in Lemma 7.7 we have, for any fixed tuple $(D_w)_{w\neq u,v}\in\mathcal{P}(X,\boldsymbol{A},\Omega)$,

$$\Biggl|\sum_{\substack{d_v\in\mathcal{Q}(\Delta A_v)\\ D_u\in\mathcal{R}(\Delta A_u/3)}}\mu^2(D_{uv})a(D_u)a(d_v)\biggl(\frac{D_u}{d_v}\biggr)_3\Biggr| \ll A_u^{1/2}\Biggl(\sum_{D_u\in\mathcal{R}(\Delta A_u/3)}\Biggl|\sum_{d_v\in\mathcal{Q}(\Delta A_v)}\mu^2(D_{uv})a(d_v)\biggl(\frac{D_u}{d_v}\biggr)_3\Biggr|^2\Biggr)^{1/2}\tag{7-16}$$

where $D_{uv}=D_uN(d_v)$.

For fixed $D_u\in\mathbb{Z}$ the map $d_v\mapsto\mu^2(D_{uv})\bigl(\frac{D_u}{d_v}\bigr)_3$ is a primitive cubic Hecke character on $\mathcal{O}$ with modulus $(9D_u)$ (see [Baier and Young 2010, Section 2.1]). Note also that $\mathcal{R}(\Delta A_u/3)$ satisfies the conditions of the set $\mathcal{P}(\Delta A_u/3)$ in Proposition 6.8, since by definition for any $d\in\mathcal{R}(\Delta A_u/3)$, $d\in\mathbb{Z}$ and $d<\Delta A_u/3$. Thus by Proposition 6.8

$$\sum_{D_u\in\mathcal{R}(\Delta A_u/3)}\Biggl|\sum_{d_v\in\mathcal{Q}(\Delta A_v)}\mu^2(D_{uv})a(d_v)\biggl(\frac{D_u}{d_v}\biggr)_3\Biggr|^2\ll(A_u^2+A_v)A_v.$$

Plugging this into (7-16) we get

$$\Biggl|\sum_{\substack{d_v\in\mathcal{Q}(\Delta A_v)\\ D_u\in\mathcal{R}(\Delta A_u/3)}}\mu^2(D_{uv})a(D_u)a(d_v)\biggl(\frac{D_u}{d_v}\biggr)_3\Biggr|\ll A_u^{1/2}((A_u^2+A_v)A_v)^{1/2}=A_uA_v\biggl(\frac{A_u}{A_v}+\frac{1}{A_u}\biggr)^{1/2}.\tag{7-17}$$

First suppose $A_u^2 < A_v$. Plugging (7-17) into Lemma 7.7 we get

$$\begin{aligned} S_{k,1}(X, \boldsymbol{A}, \Omega) &\ll \sum_{(D_w)_{w\neq u,v}\in\mathcal{P}(X,\boldsymbol{A},\Omega)} A_u A_v \left(\frac{A_u}{A_v} + \frac{1}{A_u}\right)^{1/2} \\ &\ll X\left(\frac{1}{A_v^{1/2}} + \frac{1}{A_u}\right)^{1/2} \\ &\ll X/\log^{1+9^k(1+2\cdot 3^k)} X. \end{aligned}$$

Next suppose $A_v^2 < A_u$. We again apply Cauchy–Schwarz as in (7-16) with summations reversed. Note $\chi(D_u) = \left(\frac{D_u}{d_v}\right)_3$ is a primitive Dirichlet character of modulus $N(d_v)$ for all $d_v \in \mathcal{Q}(\Delta A_v)$. Then by Proposition 6.5 we have

$$\sum_{d_v\in\mathcal{Q}(\Delta A_v)} \left| \sum_{D_u\in\mathcal{R}(\Delta A_u/3)} a(D_u)\mu^2(D_{uv}) \left(\frac{D_u}{d_v}\right)_3 \right|^2 \ll (A_v^2 + A_u)A_u$$

and hence

$$S_{k,1}(X, \boldsymbol{A}, \Omega) \ll \sum_{(D_w)_{w\neq u,v}\in\mathcal{P}(X,\boldsymbol{A},\Omega)} A_v^{1/2}((A_v^2 + A_u)A_u)^{1/2} \ll X\left(\frac{1}{A_u^{1/2}} + \frac{1}{A_v}\right)^{1/2} \ll X/\log^{1+9^k(1+2\cdot 3^k)} X.$$

In the case when the variables A_u and A_v are close together, specifically $A_u < A_v < A_u^2$ or $A_v < A_u < A_v^2$ we again apply Cauchy–Schwarz, followed by Proposition 6.4. We obtain

$$\sum_{d_v\in\mathcal{Q}(\Delta A_v)} \left| \sum_{D_u\in\mathcal{R}(\Delta A_u/3)} a(D_u)\mu^2(D_{uv}) \left(\frac{D_u}{d_v}\right)_3 \right|^2 \ll (A_v^{11/9} + A_v^{2/3}A_u)(A_uA_v)^{\epsilon} A_u.$$

Then

$$S_k(X, \boldsymbol{A}, \Omega) \ll \sum_{(D_w)_{w\neq u,v}\in\mathcal{P}(X,\boldsymbol{A},\Omega)} A_v^{1/2}((A_v^{11/9} + A_v^{2/3}A_u)(A_uA_v)^{\epsilon}A_u)^{1/2} \ll X\left(\left(\frac{A_v^{2/9}}{A_u} + \frac{1}{A_v^{1/3}}\right)(A_uA_v)^{\epsilon}\right)^{1/2}.$$

Then using that $A_u < A_v < A_u^2$ we get

$$S_k(X, \boldsymbol{A}, \Omega) \ll X\left(\left(\frac{1}{A_u^{5/9}} + \frac{1}{A_u^{1/3}}\right)(A_uA_v)^{\epsilon}\right)^{1/2} \ll X\left(\frac{1}{A_u^{1/4}}\right)^{1/2} \ll X/\log^{1+9^k(1+2\cdot 3^k)} X.$$

The case $A_v < A_u < A_v^2$ is similar. □

This proves the desired bound in Case 2.

Finally summing over all $\boldsymbol{A} \in \mathcal{F}_3$ and recalling that there are $O((\log X)^{9^k(1+2\cdot 3^k)})$ possible $\boldsymbol{A}$ with $S_k(X, \boldsymbol{A}, \Omega)$ not empty, we have proven

$$\sum_{\boldsymbol{A}\in\mathcal{F}_3} S_k(X, \boldsymbol{A}, \Omega) \ll X/\log X.$$

7D. ***The fourth family, the case $p = 3$.*** Recall we previously defined$X^{\ddagger} = \exp(\log^{\eta} X)$ and $X^{\dagger} = \log^{8(1+9^k(1+2\cdot 3^k))} X$.

Now consider the fourth family $\mathcal{F}_4$ which consists of those $\boldsymbol{A}$ such that $\boldsymbol{A} \notin \mathcal{F}_3$ and there are two linked indices u, v and

$$A_u > X^{\ddagger}, \quad 2/\Delta \le A_v < X^{\dagger}. \tag{7-18}$$

Note that given $A_u > X^{\ddagger}$ the condition $A_v < X^{\dagger}$ is forced by the assumption that $\boldsymbol{A} \notin \mathcal{F}_3$. For fixed u we in fact consider the collection of all indices $v \in \mathbb{F}_3^{2k}$ which satisfy the above condition. Of the set of $v \in \mathbb{F}_3^{2k}$ linked with u satisfying (7-18), let S_1 be the subset of v such that $\Phi_k(u, v) \ne 0$ and let S_2 be the subset of v such that $\Phi_k(v, u) \ne 0$. We assume $S_1 \cup S_2$ is not empty.

As in Section 7C we split $S_k(X, \boldsymbol{A}, \Omega) = \sum_{i=1}^{4} S_{k,i}(X, \boldsymbol{A}, \Omega)$ into four terms depending on whether $3 \mid D_u$, $3 \mid D_v$ for some $v \in S$, $3 \mid D_w$ for some $w \ne u$ and $w \notin S$, or $3 \nmid D_w$ for all w. In the following we bound $S_{k,1}(X, \boldsymbol{A}, \Omega)$, the arguments in the other cases being almost identical.

Lemma 7.9. *For each prime $l \in \mathbb{Z}$ congruent to $1 \bmod 3$ fix a nontrivial order 3 character of modulus l, denoted χ_l.*

For $\boldsymbol{A}$, u and S_1, S_2 as defined above we have

$$S_{k,1}(X, \boldsymbol{A}, \Omega) \ll \sum_{(D_w)_{w \ne u} \in \mathcal{P}(X, \boldsymbol{A}, \Omega)} \left| \sum_{d_u \in \mathcal{Q}(\Delta A_u/3)} \frac{\mu^2(D)}{3^{\omega(N(d_u))}} \psi(d_u) \right|$$

where we denote $D = \prod_{w \in \mathbb{F}_3^{2k}} D_w$, $D' = \prod_{w \ne u} D_w$, and ψ is a cubic Hecke character defined by

$$\psi(d_u) = \left(\frac{\prod_{v \in S_2} D_v}{d_u} \right)_3 \prod_{v \in S_1} \prod_{l \mid D_v} \chi_l(D_u).$$

Proof. The proof is similar to Lemma 7.5. □

We remark that the $v \in S$ by assumption satisfy $A_v < X^{\dagger}$. The modulus of ψ is $f_{\psi} = 9 \prod_{v \in S_1 \cup S_2} D_v$ and $N(f_{\psi}) \le (X^{\dagger})^{2\cdot 3^{2k}}$. Also note ψ is a nontrivial character since $D_v \ge A_v \ge 2/\Delta > 1$.

Proposition 7.10. *With the above notation we have, for some $t \in \mathbb{R}$*

$$\sum_{d_u \in \mathcal{Q}(\Delta A_u/3)} \frac{\mu^2(D)}{3^{\omega(d_u)}} \psi(d_u) \ll \frac{(\log X)^t}{\exp(c^4 (\log X)^{\eta/4 - 2\cdot 9^k \epsilon}/3^{2+\epsilon})}.$$

Proof. Partitioning the sum according to the number of prime factors in $\mathcal{O}$ we get

$$\sum_{d_u} \frac{\mu^2(D)}{3^{\omega(d_u)}} \psi(d_u) = \sum_{l=0}^{\Omega} \frac{1}{3^l} \sum_{\pi_1, \dots, \pi_l} \mu^2\left(\prod_{v \ne u} D_v N(\pi_1 \cdots \pi_l) \right) \psi(\pi_1 \cdots \pi_l)$$

where (π_i) are prime ideals in $\mathcal{O}$ with $\pi_i \equiv 1 \bmod 3$ and $N(\pi_1 \cdots \pi_l) \le \Delta A_u$. We can relabel the π_i so that $N(\pi_1) \le N(\pi_2) \le \cdots \le N(\pi_l)$ and split the sum on the right-hand side up to get

$$\sum_{l=0}^{\Omega} \frac{1}{3^l} \sum_{\pi_1,\dots,\pi_{l-1}} \psi(\pi_1 \cdots \pi_{l-1}) \sum_{\pi_l} \mu^2\Bigl(\prod_{v \ne u} D_v N(\pi_1 \cdots \pi_l)\Bigr) \psi(\pi_l) \tag{7-19}$$

where $N(\pi_1 \cdots \pi_{l-1}) \le \Delta A_u$ and $A_u^{1/l} \le N(\pi_l) \le \Delta A_u / N(\pi_1 \cdots \pi_{l-1})$.

Note $\omega\bigl(\prod_{v \ne u} D_v N(\pi_1 \cdots \pi_{l-1})\bigr) \le 3^{2k}\Omega$ factoring in $\mathbb{Z}$, hence the number of prime factors in $\mathcal{O}$ is at most $2 \cdot 3^{2k}\Omega$. Hence removing μ^2 from (7-19) adds at most an additional $O(\Omega)$ terms of absolute value 1. Furthermore notice the summations in (7-19) are only over primes split in $\mathcal{O}$. The number of inert primes (π) in $\mathcal{O}$ with $\pi^2 = N\pi < A_u$ is $O(A_u^{1/2})$. Then looking at the inner sum of (7-19) we obtain the bound

$$\sum_{\pi_l} \mu^2\Bigl(\prod_{v \ne u} D_v N(\pi_1 \cdots \pi_l)\Bigr) \psi(\pi_l) \ll \sum_{\pi_l, (\pi_l, f_\psi)=1} \psi(\pi_l) + \Omega + A_u^{1/2}$$

where the summation on the right-hand side is now over all prime ideals in $\mathcal{O}$ with $A_u^{1/l} \le N(\pi_l) \le \Delta A_u / N(\pi_1 \cdots \pi_{l-1})$.

Now we apply Proposition 6.6 with $f_\chi = f_\psi$ and $x = \Delta A_u / N(\pi_1 \cdots \pi_{l-1})$ to get, for some constant c,

$$\sum_{\pi_l, (\pi_l, f_\psi)=1} \psi(\pi_l) \ll \frac{N(f_\psi)^\epsilon x (\log x)^2}{\exp(c^4 (\log x)^{1/2} / 3^{2+\epsilon} N(f_\psi)^\epsilon)}.$$

Using that $N(f_\psi)^\epsilon \le (X^\dagger)^{2 \cdot 9^k \epsilon}$ which implies $\exp(-1/N(f_\psi)^\epsilon) \ll \exp(-1/(X^\dagger)^{2 \cdot 9^k \epsilon})$, we get

$$\sum_{\pi_l, (\pi_l, f_\psi)=1} \psi(\pi_l) \ll \frac{(X^\dagger)^{2 \cdot 9^k \epsilon} x (\log x)^2}{\exp(c^4 (\log x)^{1/2} / 3^{2+\epsilon} (X^\dagger)^{2 \cdot 9^k \epsilon})}. \tag{7-20}$$

Now $x \ge N(\pi_l) \ge A_u^{1/l}$. We claim $A_u^{1/l} \gg \exp(\log^{\eta/2} X)$. Note $\Omega \ll \log\log X$ by definition. Let $\theta = \log X$. Then

$$-\log l + \eta \log\theta \ge -\log\Omega + \eta\log\theta \gg -\log\log\theta + \eta\log\theta \ge \frac{\eta}{2}\log\theta.$$

Taking exp of this inequality gives $(1/l)\log^\eta X \gg \log^{\eta/2} X$. Thus

$$\log A_u^{1/l} \ge \log(X^\ddagger)^{1/l} = (1/l)\log^\eta X \gg \log^{\eta/2} X.$$

Combining these facts we have $(\log x)^{1/2} \gg \log^{\eta/4} X$ so we can write $c_0(\log x)^{1/2} \ge \log^{\eta/4} X$ for some constant c_0. Noting that $X^\dagger$ is some fixed power of $\log X$ we get the bound

$$\sum_{\pi_l, (\pi_l, f_\psi)=1} \psi(\pi_l) \ll X^{\dagger 2 \cdot 9^k \epsilon} \frac{A_u}{N(\pi_1 \cdots \pi_{l-1})} \frac{(\log x)^2}{\exp(c^4 (\log X)^{\eta/4 - 2 \cdot 9^k \epsilon} / c_0 3^{2+\epsilon})}. \tag{7-21}$$

We plug (7-21) back into (7-19) to get

$$\sum_{d_u} \frac{\mu^2(D)}{3^{\omega(d_u)}} \psi(d_u) \ll \sum_{l=0}^{\Omega} \frac{1}{3^l} \sum_{\pi_1,\dots,\pi_{l-1}} X^{\dagger 2 \cdot 9^k \epsilon} \frac{A_u}{N(\pi_1 \cdots \pi_{l-1})} \frac{(\log x)^2}{\exp(c^4 (\log X)^{\eta/4 - 2 \cdot 9^k \epsilon} / c_0 3^{2+\epsilon})}$$
$$\ll A_u \frac{X^{\dagger 2 \cdot 9^k \epsilon} (\log x)^2}{\exp(c^4 (\log X)^{\eta/4 - 2 \cdot 9^k \epsilon} / c_0 3^{2+\epsilon})} \sum_{l=0}^{\Omega} \sum_{\pi_1,\dots,\pi_{l-1}} \frac{1}{N(\pi_1 \cdots \pi_{l-1})}.$$

Noting the bounds

$$\sum_{\substack{\pi_1,\dots,\pi_{l-1} \\ N(\pi_1 \cdots \pi_{l-1}) \le \Delta A_u}} \frac{1}{N(\pi_1 \cdots \pi_{l-1})} \ll \log A_u \Omega \ll \log\log X \log x \ll \log X$$

we get, for some $t \in \mathbb{R}$

$$\sum_{d_u} \frac{\mu^2(D)}{3^{\omega(d_u)}} \psi(d_u) \ll A_u \frac{(\log X)^t}{\exp(c^4 (\log X)^{\eta/4 - 2 \cdot 9^k \epsilon} / c_0 3^{2+\epsilon})}. \qquad \square$$

Combining Lemma 7.9 and Proposition 7.10 we get

$$S_{k,1}(X, \boldsymbol{A}) \ll \sum_{(D_w)_{w \ne u} \in \mathcal{P}(X, \boldsymbol{A}, \Omega)} \left| \sum_{d_u \in \mathcal{Q}(\Delta A_u/3)} \frac{\mu^2(D)}{3^{\omega(N(d_u))}} \psi(d_u) \right| \ll X \frac{(\log X)^t}{\exp(c^4 (\log X)^{\eta/4 - 2 \cdot 9^k \epsilon} / c_0 3^{2+\epsilon})}.$$

Then summing over all $\boldsymbol{A}$ and using that there are $O((\log X)^{9^k(1+2 \cdot 3^k)})$ possible $\boldsymbol{A}$ with $S_k(X, \boldsymbol{A}, \Omega)$ not empty

$$\sum_{\boldsymbol{A} \in \mathcal{F}_4} S_k(X, \boldsymbol{A}) \ll X \frac{(\log X)^{t + 9^k(1+2 \cdot 3^k)}}{\exp(c^4 (\log X)^{\eta/4 - 2 \cdot 9^k \epsilon} / c_0 3^{2+\epsilon})} = o(X).$$

7E. ***The third and fourth families for all p.*** Assume GRH for Artin L-functions. The missing ingredients required to extend our result to general p unconditionally are analogs of Proposition 6.3 and 6.4. That is, we cannot deal with the case in family 3 when A_u and A_v are close together. We will instead give a proof assuming GRH. The following argument replaces the sections containing families 3 and 4 for $p = 3$ above.

Suppose $\boldsymbol{A} \notin \mathcal{F}_1 \cup \mathcal{F}_2$. In particular there are at least $p^{k-1} + 1$ indices $w \in \mathbb{F}_p^{2k}$ which satisfy $A_w > X^{\ddagger}$. Let A_u be the largest of these. Let S_1 be the set of indices v linked with u such that $\Phi_k(u, v) \ne 0$ and let S_2 be the set of v such that $\Phi_k(v, u) \ne 0$ and suppose $S_1 \cup S_2$ is not empty.

Let $\zeta = e^{2\pi i/p}$ and let $\mathcal{O} = \mathbb{Z}[\zeta]$ the ring of integers of $\mathbb{Q}(\zeta)$ which is a degree $p - 1$ extension of $\mathbb{Q}$. For each prime $l \in \mathbb{Z}$ congruent to 1 mod p fix a nontrivial order p character of modulus l, denoted χ_l.

For $A, B \in \mathbb{Z}$ with $(A, B) = 1$ define

$$\left[\frac{A}{B}\right]_p = \prod_{l \mid B} \frac{(\chi_l + \cdots + \chi_l^{p-1})}{p}(A)$$

which does not depend on the choice of χ_l above.

Recall we defined $\mathcal{R}(X)$ to be the set of positive $d \in \mathbb{Z}$ which are a product of primes congruent to 1 mod $p\mathbb{Z}$ and $d < X$. Then we have, by an argument similar to the proof of Lemma 7.5,

$$S_k(X, \boldsymbol{A}) \ll \sum_{(D_w)_{w \neq u} \in \mathcal{P}(X, \boldsymbol{A}, \Omega)} \left| \sum_{D_u \in \mathcal{R}(\Delta A_u)} \mu^2(D) \left(\frac{D_u}{f_u}\right)_p \left[\frac{C_u}{D_u}\right]_p \right| \tag{7-22}$$

where $D = \prod_{w \in \mathbb{F}_p^{2k}} D_w$, $C_u = \prod_{w \in S_2} D_w$, and any choice of $f_u \in \mathcal{O}$ with $N(f_u) = \prod_{w \in S_1} D_w$. Divisibility by p of the D_w is handled as in the $p = 3$ cases in the previous section. For simplicity we will assume $p \nmid D_u C_u N(f_u)$.

Let $K = \mathbb{Q}(\zeta, \sqrt[p]{C_u})$ and let F_q denote the Frobenius of q in K. Let $K_1 = \mathbb{Q}(\zeta)$. Define a character $\rho : \mathrm{Gal}(K/K_1) \to \mu_p$ by $\sigma(\sqrt[p]{C_u}) = \rho(\sigma)\sqrt[p]{C_u}$ (viewing $\mathrm{Gal}(K/K_1)$ as a subgroup of $\mathrm{Gal}(K/\mathbb{Q})$). Denote by $\rho' = \mathrm{Ind}_{\mathrm{Gal}(K/K_1)}^{\mathrm{Gal}(K/\mathbb{Q})} \rho$ the induction to $\mathrm{Gal}(K/\mathbb{Q})$.

Lemma 7.11. *With the above notation, for any $D_u \in \mathcal{R}(\Delta A_u)$ and any prime $q \equiv 1 \bmod p$*

$$\left[\frac{C_u}{D_u}\right]_p = \frac{\prod_{q \mid D_u} \operatorname{tr} \rho'(F_q)}{p^{\omega(D_u)}}.$$

Proof. Note $F_q \in \mathrm{Gal}(K/K_1)$. We claim that for any $q \equiv 1 \bmod p$ we have $\rho(F_q)$ is trivial if and only $\chi_q(C_u)$ is trivial. Indeed $\chi_q : \mathbb{F}_q^\times \to \mu_p$ has kernel equal to the p-th powers in $\mathbb{F}_q^\times$. For any prime $\hat{q}$ in $\mathbb{Q}(\zeta, \sqrt[p]{C_u})$ lying above q we have, using $q = 1 + pn$ for some n, that

$$F_q(\sqrt[p]{C_u}) \bmod \hat{q} = (\sqrt[p]{C_u})^q \bmod \hat{q} = C_u^n \sqrt[p]{C_u} \bmod \hat{q}.$$

Let $\mathbb{F}_{\hat{q}}$ denote the residue field of the prime $\hat{q}$. Note $C_u \in \mathbb{F}_q^\times \subset \mathbb{F}_{\hat{q}}^\times$ and $C_u^n \bmod q$ is trivial exactly when C_u is a p-th power. Thus we have shown $\rho(F_q)$ is trivial if and only $\chi_q(C_u)$ is trivial. In particular $\sum_{i=1}^{p-1} \rho^i(F_q) = \sum_{i=1}^{p-1} \chi_q^i(C_u)$.

By properties of induced representations and since $F_q \in \mathrm{Gal}(K/K_1)$, we have

$$\operatorname{tr} \rho'(F_q) = \sum_{g \in \mathrm{G}(K/\mathbb{Q})/\mathrm{G}(K/K_1)} \hat{\operatorname{tr}}(g^{-1} F_q g) = \sum_{i=1}^{p-1} \rho(F_q^i)$$

where $\hat{\operatorname{tr}}(h) = \rho(h)$ if $h \in \mathrm{Gal}(K/K_1)$ and 0 otherwise. The result follows. □

Let M be the degree p cyclic field corresponding to the character $\chi_{N(f_u)}$ and let $L = KM$. Define the representation $\sigma = \chi_{N(f_u)} \otimes \rho'$ of $\mathrm{Gal}(L/\mathbb{Q})$.

Lemma 7.12. *The L-function*

$$L_0(s, \sigma) = \prod_{\substack{q \equiv 1 (p) \\ q \nmid D}} \det\left(I - \frac{\sigma(F_q)}{q^s}\right)^{-1}$$

is convergent for $\mathrm{Re}\, s > \frac{1}{2}$.

Proof. Consider the regular representation $\chi' = \oplus_{i=0}^{p-2}\chi^i$ where $\chi : \mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q}) \to \mu_{p-1}$ is a fixed character. Let τ be the representation of $\mathrm{Gal}(L/\mathbb{Q})$ given by $\tau = \sigma \otimes \chi'$.

Note the following facts. By standard properties of Artin L-functions we can factor $L(s,\tau) = \prod_{i=0}^{p-1} L(s, \chi_{N(f_u)} \cdot \chi^i \otimes \rho'(F_q))$. If $q \equiv 1 \bmod p$ then $\chi^i(F_q) = 1$ for all i. If $q \not\equiv 1 \bmod p$ then $\mathrm{tr}[\chi_{N(f_u)} \cdot \chi' \otimes \rho'(F_q)] = 0$. Furthermore for any q we have

$$\det\biggl(I - \frac{\chi_{N(f_u)} \cdot \chi' \otimes \rho'(F_q)}{q^s}\biggr) = 1 - \frac{\mathrm{tr}[\chi_{N(f_u)} \cdot \chi' \otimes \rho'(F_q)]}{q^s} + O\biggl(\frac{1}{q^{2s}}\biggr).$$

Putting these together we get

$$\begin{aligned} L(s,\tau) &= \prod_{q \nmid D} \det\biggl(I - \frac{\chi_{N(f_u)} \cdot \chi' \otimes \rho'(F_q)}{q^s}\biggr)^{-1} \\ &= \prod_{\substack{q \equiv 1(p) \\ q \nmid D}} \det\biggl(I - \frac{\chi_{N(f_u)} \otimes \rho'(F_q)}{q^s}\biggr)^{-p} \prod_{\substack{q \not\equiv 1(p) \\ q \nmid D}} \biggl(1 + O\biggl(\frac{1}{q^{2s}}\biggr)\biggr)^{-1}. \end{aligned}$$

The last product above is absolutely convergent for $\operatorname{Re} s > \frac{1}{2}$ and hence has no zeros. Note $L(s,\tau)$ is entire since it can be factored as a product of 1-dimensional L-functions or ones which are induced from 1-dimensional L-functions, which are all known to be entire. By assumption of GRH $L(s,\tau)$ has no zeros to the right of $s = \frac{1}{2}$. Thus there exists a branch of $\log(L_0(s,\sigma)^p)$ (and hence of $(L_0(s,\sigma)^p)^{1/p})$ on $\operatorname{Re} s > \frac{1}{2}$ and the result follows. □

Define the function

$$L_0(s) = \prod_{\substack{q \equiv 1(p) \\ q \nmid D}} \biggl(1 + \frac{\mathrm{tr}\,\sigma(F_q)}{pq^s}\biggr).$$

Lemma 7.13. *There exists a function* $F(s) = \prod_q (1 + O(1/q^{2s}))$ *which is absolutely convergent for* $\operatorname{Re} s > \frac{1}{2}$ *and a branch of* $(L_0(s,\sigma)F(s))^{1/p}$ *defined on* $\operatorname{Re} s > \frac{1}{2}$ *such that*

$$L_0(s) = (L_0(s,\sigma)F(s))^{1/p}.$$

Proof. By definition we have

$$L_0(s,\sigma) = \prod_{\substack{q \equiv 1(p) \\ q \nmid D}} \det(I - \frac{\sigma(F_q)}{q^s})^{-1} = \prod_{\substack{q \equiv 1(p) \\ q \nmid D}} \biggl(1 - \frac{\mathrm{tr}\,\sigma(F_q)}{q^s} + O(1/q^{2s})\biggr)^{-1}.$$

By a similar computation we have

$$L_0(s)^p = \prod_{\substack{q \equiv 1(p) \\ q \nmid D}} \biggl(1 + \frac{\mathrm{tr}\,\sigma(F_q)}{q^s} + O(1/q^{2s})\biggr).$$

Now we have

$$L_0(s,\sigma)L_0(s)^{-p} = \prod_{\substack{q\equiv 1(p)\\ q\nmid D}} (1+O(1/q^{2s}))^{-1}.$$

By assumption of GRH $L(s,\sigma)$ (and hence $L_0(s,\sigma)$) has no zeros to the right of $s=\frac{1}{2}$ and neither does $F(s)$ since it is a convergent product. Thus there exists a branch of $\log L_0(s,\sigma)F(s)$ (and hence of $(L_0(s,\sigma)F(s))^{1/p}$) on $\operatorname{Re} s > \frac{1}{2}$ and the result follows. □

In particular it follows from the above lemma that $L_0(s,\sigma)^{1/p}$ has no poles for $\operatorname{Re} s > \frac{1}{2}$.

Let $\psi(d) = \mu^2(D)\prod_{q\mid d} \operatorname{tr}\sigma(F_q)/p^{\omega(d)}$. Then by Lemma 7.11 we have

$$L_0(s) = \sum_{d\in\mathcal{R}(\infty)} \frac{\psi(d)}{d^s} = \sum_{D_u\in\mathcal{R}(\infty)} \mu^2(D)\left(\frac{D_u}{f_u}\right)_p\left[\frac{C_u}{D_u}\right]_p \cdot d^{-s}.$$

We now apply a standard argument for bounding sums of L-series coefficients (see for instance [Davenport 2000, pages 105–106]). We have

$$\sum_{d\in\mathcal{R}(x)} \psi(d) = \int_{2-iT}^{2+iT} L_0(s)x^s\,\frac{ds}{s} + O\left(\frac{x^2}{T\log x}\right). \tag{7-23}$$

Consider the integral of $L_0(s)x^s/s$ over the rectangle with vertices $\left(\frac{1}{2}+\epsilon, \pm iT\right)$, $(2, \pm iT)$. By Lemma 7.13 we have $|L_0(s)| \ll |L_0(s,\sigma)^{1/p}|$ on this rectangle. Furthermore by the Lindelof conjecture $|L_0(s,\sigma)| \ll (TD)^\epsilon$ on this rectangle. In addition for s on the lines $y\pm iT$ with $y\in\left[\frac{1}{2}+\epsilon, 2\right]$ we have the bounds $|x^s| \ll x^2$ and $1/|s| \ll 1/T$. Thus shifting the above integral to the $\frac{1}{2}+\epsilon$ line we get the bound

$$\begin{aligned}\left|\int_{2-iT}^{2+iT} L_0(s)x^s\,\frac{ds}{s}\right| &\ll \int_{2-iT}^{2+iT} |L_0(s,\sigma)^{1/p}|\left|\frac{x^s}{s}\right| ds\\ &\ll \int_{1/2+\epsilon-iT}^{1/2+\epsilon+iT} |L_0(s,\sigma)^{1/p}||x^s|\,\frac{ds}{|s|} + O\left(\frac{x^2(TD)^\epsilon}{T}\right)\\ &\ll x^{1/2+\epsilon}(TD)^\epsilon \int_{1/2+\epsilon-iT}^{1/2+\epsilon+iT} \frac{1}{|s|}\,ds + O\left(\frac{x^2(TD)^\epsilon}{T}\right)\\ &\ll x^{1/2+\epsilon}(TD)^\epsilon + O\left(\frac{x^2(TD)^\epsilon}{T}\right).\end{aligned} \tag{7-24}$$

Combining (7-23) and (7-24) and setting $T=x^3$ we get

$$\sum_{d\in\mathcal{R}(x)} \psi(d) \ll x^{1/2+\epsilon'}D^\epsilon.$$

Then we bound the inner sum in (7-22) as

$$\sum_{D_u\in\mathcal{R}(\Delta A_u)} \mu^2(D)\left(\frac{D_u}{f_u}\right)_p\left[\frac{C_u}{D_u}\right]_p = \sum_{d\in\mathcal{R}(\Delta A_u)} \psi(d) \ll A_u^{1/2+\epsilon}D^\epsilon.$$

Note that $D^{\epsilon} \leq A_u^{p^{2k}\epsilon}$ since $A_w < A_u$ for all $w \in \mathbb{F}_p^{2k}$. Then summing over all the remaining D_w we get

$$S_k(X, \boldsymbol{A}) \ll X/A_u^{1/4} \ll X/X^{\ddagger 1/4} = o(X).$$

This argument shows that we can remove all $\boldsymbol{A}$ in which there is a variable larger than $X^{\ddagger}$ and linked with any other $A_w > 1$. This is equivalent to removing the $\boldsymbol{A}$ which belong to families 3 or 4.

We summarize the results of this section in the following theorem:

Theorem 7.14. *Let $\sum'_{\boldsymbol{A}} S_k(X, \boldsymbol{A})$ denote a summation over all tuples $\boldsymbol{A}$ which do not belong to any of the* 4 *families*, *that is they do not satisfy any of* (7-2), (7-3), (7-6), (7-18). *Then*

$$S_k(X) = \sum\nolimits'_{\boldsymbol{A}} S_k(X, \boldsymbol{A}) + o(X).$$

8. Computing the k-th moment

We now want to prove Theorem 1.6.

For this section we define the following notation. Let $\mathcal{N}(k) = \mathcal{N}(k, p)$ which we recall is the number of vector subspaces of $\mathbb{F}_p^k$. Let $\mathcal{S}(X)$ be the set of positive square-free integers of the form $n = p_1 \cdots p_r$ and each p_i is either a prime congruent to 1 mod p or equal to p^2, and such that $n < X$.

We will do this by proving the following:

Theorem 8.1. *For any $k \in \mathbb{Z}_{\geq 1}$*

$$S_k(X) = p^{-k}(\mathcal{N}(k+1) - \mathcal{N}(k)) \sum_{n \in \mathcal{S}(X)} (p-1)^{\omega(n)-1} + o(X).$$

Note $S_k(X) = \sum_{K, D_K < X^{p-1}} |\mathrm{im}(\varphi_K)|^k$ is a sum over discriminants up to X^{p-1}. Recall from Section 2 that the number of degree p cyclic fields with discriminant up to X^{p-1} is

$$\sum_{n \in \mathcal{S}(X)} (p-1)^{\omega(n)-1} = cX + o(X).$$

Thus it follows immediately from combining these facts with the above theorem that

$$\lim_{X \to \infty} \frac{\sum_{K, D_K < X} |\mathrm{im}(\varphi_K)|^k}{\sum_{K, D_K < X} 1} = \frac{\mathcal{N}(k+1) - \mathcal{N}(k)}{p^k}.$$

We start by proving some facts about maximal unlinked sets of indices. Recall that for $u, v \in \mathbb{F}_p^2$ written as $u = (u_1, u_2)$ and $v = (v_1, v_2)$ we defined

$$\Phi(u, v) = (u_1)(v_2 - u_2).$$

If we represent each index $u \in \mathbb{F}_p^{2k}$ as $u = (u_{11}, u_{12}, u_{21}, u_{22}, \ldots, u_{k1}, u_{k2})$ then

$$\Phi_k(u, v) = \sum_{i=1}^{k} \Phi((u_{i1}, u_{i2}), (v_{i1}, v_{i2})) = \sum_{i=1}^{k} (u_{i1})(v_{i2} - u_{i2}).$$

We defined $u, v \in \mathbb{F}_p^{2k}$ to be unlinked if $\Phi_k(u, v) = 0$ and $\Phi_k(v, u) = 0$. We say a set $\mathcal{U} \subset \mathbb{F}_p^{2k}$ is unlinked if u and v are unlinked for all $u, v \in \mathcal{U}$.

We will show that for each $\boldsymbol{A}$ in the sum in Theorem 7.14 all the indices $u \in \mathbb{F}_p^{2k}$ with $\Delta A_u \geq 2$ form a maximal unlinked set.

Let $\pi : \mathbb{F}_p^{2k} \to \mathbb{F}_p^{k}$ be the projection onto the even coordinates, that is

$$\pi(u_{11}, u_{12}, u_{21}, u_{22}, \ldots, u_{k1}, u_{k2}) = (u_{12}, u_{22}, \ldots, u_{k2})$$

and let ρ be the projection onto the odd coordinates. Let $V_1 = \ker \pi$ and let $V_2 = \ker \rho$. Then $\mathbb{F}_p^{2k} \cong V_1 \oplus V_2$ given by $\xi : v \mapsto (\rho v, \pi v)$.

For any subset $V \subseteq \mathbb{F}_p^{2k}$ define $\pi(V)^{\perp} = \{v \in \mathbb{F}_p^{k} \mid v \cdot u = 0, \forall u \in \pi(V)\}$.

We will start by classifying the maximal unlinked subspaces of $\mathbb{F}_p^{2k}$.

Lemma 8.2. *Let V be a subspace of $\mathbb{F}_p^{2k}$. Then V is an unlinked set if and only if $\xi(V) \subseteq \pi(V)^{\perp} \oplus \pi(V)$ (here $\pi(V)^{\perp} \oplus \pi(V)$ is viewed as a subspace of $V_1 \oplus V_2$). Equality holds if and only if V is a maximal unlinked subspace.*

Proof. For any $v, w \in V$

$$\Phi_k(v, w) = \sum_{i=1}^{k} v_{i1}(w_{i2} - v_{i2}) = \rho(v) \cdot \pi(w - v).$$

Suppose V is unlinked. Fix any $u \in V$. Let $w = u + v \in V$ in the above equation. Then we get $\Phi_k(v, w) = \rho(v) \cdot \pi(u) = 0$ so $\rho(v) \in \pi(V)^{\perp}$. Since $v \in V$ was arbitrary this implies $\xi(V) \subseteq \pi(V)^{\perp} \oplus \pi(V)$. The converse is clear from the above equation.

For the second part note that $\dim \pi(V)^{\perp} \oplus \pi(V) = k$ for all subspaces V of $\mathbb{F}_p^{2k}$. Thus by the first part any unlinked subspace is contained in one of dimension k. This completes the proof. □

Next we determine when translation preserves the property of being unlinked.

Lemma 8.3. *Suppose $V \subset \mathbb{F}_p^{2k}$ is an unlinked subspace. Let $a \in \mathbb{F}_p^{2k}$. Then $V + a$ is unlinked if and only if $\rho(a) \in \pi(V)^{\perp}$.*

Proof. Let $v, w \in V$. Then we compute

$$\Phi_k(v + a, w + a) = \sum_{i=1}^{k} (v_{i1} + a_{i1})(w_{i2} - v_{i2}) = \sum_{i=1}^{k} (a_{i1})(w_{i2} - v_{i2}) = \rho(a) \cdot \pi(w) - \rho(a) \cdot \pi(v).$$

If $\rho(a) \in \pi(V)^{\perp}$ we see $V + a$ is unlinked. If $V + a$ is unlinked then setting $w = v + u$ for any $u \in V$ we see $\rho(a) \in \pi(V)^{\perp}$. □

Next we show that every maximal unlinked set is a coset of some unlinked subspace.

Lemma 8.4. *Let $\mathcal{U} \subset \mathbb{F}_p^{2k}$ be a maximal unlinked set and let $a \in \mathcal{U}$. Let $V = \mathcal{U} - a$. Then $V \subset \mathbb{F}_p^{2k}$ is an unlinked subspace.*

Proof. First we show V is a subspace. Let $u, v \in \mathcal{U}$. We need to show that $(u-a)+(v-a)+a = u+v-a \in \mathcal{U}$. Since $\mathcal{U}$ is maximal we show $u+v-a$ is unlinked with every element of $\mathcal{U}$. Let $w \in \mathcal{U}$. We have

$$\begin{aligned}\Phi_k(u+v-a, w) &= \sum_{i=1}^{k}(u_{i1}+v_{i1}-a_{i1})(w_{i2}-u_{i2}-v_{i2}+a_{i2})\\ &= \sum_{i=1}^{k}(u_{i1}+v_{i1}-a_{i1})((a_{i2}-u_{i2})+(w_{i2}-v_{i2}))\\ &= \sum_{i=1}^{k} v_{i1}(a_{i2}-u_{i2})+u_{i1}(w_{i2}-v_{i2})-a_{i1}(w_{i2}-v_{i2})\\ &= 0\end{aligned}$$

where the last two equalities follow since u, v, w, a are all unlinked and for instance $v_{i1}(-u_{i2}+a_{i2}) = -v_{i1}(u_{i2}-v_{i2})+v_{i1}(-v_{i2}+a_{i2})$. Similarly

$$\Phi_k(w, u+v-a) = \sum_{i=1}^{k}(w_{i1})(u_{i2}+v_{i2}-a_{i2}-w_{i2}) = 0.$$

Thus V is a subspace. Next we show V is unlinked.

For any $w \in V$ we have $\mathcal{U} = \mathcal{U}+w$. Let $u' \in V$ and let $u = u'+a$, so $u \in \mathcal{U}$. Note $a \in \mathcal{U}$. Then we have

$$\begin{aligned}0 &= \Phi_k(u+w, a+w)\\ &= \sum_{i=1}^{k}(u_{i1}+w_{i1})(a_{i2}+w_{i2}-u_{i2}-w_{i2})\\ &= \Phi_k(u,a)+\sum_{i=1}^{k} w_{i1}(a_{i2}-u_{i2})\\ &= -\sum_{i=1}^{k} w_{i1}u'_{i2} = \rho(w)\cdot\pi(u')\end{aligned}$$

Since $w, u' \in V$ were arbitrary this shows $\rho(V) \in \pi(V)^{\perp}$ so $\xi(V) \subset \pi(V)^{\perp} \oplus \pi(V)$. By Lemma 8.2 V is unlinked. □

With the above results we can classify all the maximal unlinked sets.

Proposition 8.5. *The maximal unlinked sets $\mathcal{U} \subset \mathbb{F}_p^{2k}$ are exactly the sets of the form $\mathcal{U} = V + a$ where V is a subspace which is a maximal unlinked set and $\rho(a) \in \pi(V)^{\perp}$.*

Proof. Suppose $\mathcal{U}$ is a maximal unlinked set. By Lemma 8.4 $\mathcal{U} - a = V$ for some subspace V which is unlinked and some $a \in \mathcal{U}$. Since $\mathcal{U} = V + a$ by Lemma 8.3 we see $\rho(a) \in \pi(V)^{\perp}$.

Let $\mathcal{W} = \xi^{-1}(\pi(V)^{\perp} \oplus \pi(V))$ so that $V \subset \mathcal{W}$. By Lemma 8.2 $\mathcal{W}$ is unlinked. Note $\rho(a) \in \pi(V)^{\perp} = \pi(\mathcal{W})^{\perp}$. Thus $\mathcal{W}+a$ is unlinked and $\mathcal{U} \subset \mathcal{W}+a$. By maximality we get $\mathcal{U} = \mathcal{W}+a$.

The converse is clear from the above lemmas. □

As mentioned in the above proofs $\dim \pi(V)^{\perp} \oplus \pi(V) = k$ for all subspaces V and hence, by Proposition 8.5, every maximal unlinked set has size p^k.

With this we can rewrite $S_k(X)$ in a form closer to Theorem 8.1.

Proposition 8.6. *Let $k \in \mathbb{Z}_{\geq 1}$. Let U be the number of maximal unlinked sets in $\mathbb{F}_p^{2k}$. Then*

$$\sum_{K, D_K < X^{p-1}} |\mathrm{im}(\varphi_K)|^k = \binom{U}{p^k} \sum_{n \in \mathcal{S}(X)} (p-1)^{\omega(n)-1} + o(X)$$

where $\mathcal{S}(X)$ is the set of square-free positive integers of the form $n = p_1 \cdots p_r$ where each p_i is either a prime congruent to 1 mod *p or equal to p^2, such that $n < X$.*

Proof. Given two maximal unlinked sets $\mathcal{U}_i$ for $i = 1, 2$, if $a \in \mathcal{U}_1 \cap \mathcal{U}_2$ then $V_i = \mathcal{U}_i - a$ is a vector space which is also a maximal unlinked set. If the $\mathcal{U}_i$ are distinct then so are the V_i and hence $V_1 \cap V_2$ is at most $k-1$ dimensional. Hence the largest possible intersection of two distinct maximal unlinked sets has size p^{k-1}. Thus a set of $p^{k-1}+1$ unlinked indices determines a unique maximal unlinked set.

Let $\boldsymbol{A}$ be a tuple as in the statement of Theorem 7.14, that is $\boldsymbol{A} \notin \mathcal{F}_i$ for $i = 1, 2, 3, 4$. Let $\mathcal{U}_0$ be the set of indices $u \in \mathbb{F}_p^{2k}$ such that $A_u > X^{\ddagger}$. Since $\boldsymbol{A} \notin \mathcal{F}_2$ this implies there are at least $p^{k-1}+1$ indices $u \in \mathbb{F}_p^{2k}$ with $A_u > X^{\ddagger}$ so $|\mathcal{U}_0| \geq p^{k-1}+1$. Then since $\boldsymbol{A} \notin \mathcal{F}_3$ and $X^{\ddagger} > X^{\dagger}$ the set $\mathcal{U}_0$ is unlinked, so by the above remark determines a unique maximal unlinked set $\mathcal{U} \supset \mathcal{U}_0$.

Hence any $u \in \mathbb{F}_p^{2k}$ such that $u \notin \mathcal{U}$ is linked with some $v \in \mathcal{U}_0$. Since $\boldsymbol{A} \notin \mathcal{F}_3 \cup \mathcal{F}_4$ this implies $A_u < 2/\Delta$ and hence $D_u = 1$.

Let $\boldsymbol{A}_1$ be the tuple consisting of the coordinates of $\boldsymbol{A}$ in $\mathcal{U}$. Thus in the expression $S_k(X, \boldsymbol{A})$ all of the characters χ_l evaluate to 1 so we can write

$$\begin{aligned} S_k(X, \boldsymbol{A}) &= \frac{1}{(p-1)\cdot p^k} \sum_{(D_u) \in \mathcal{P}(X, \boldsymbol{A})} \sum_{(\chi_{l'}) \in \mathcal{C}(D)} \frac{\mu^2(D)}{p^{k\omega(D)}} \prod_{v} \prod_{l \mid D_v} \chi_l\biggl(\prod_u D_u^{\Phi_k(u,v)}\biggr) + o(X) \\ &= \frac{1}{(p-1)\cdot p^k} \sum_{(D_u)_{u \in \mathcal{U}} \in \mathcal{P}(X, \boldsymbol{A}_1)} \sum_{(\chi_{l'}) \in \mathcal{C}(D')} \frac{\mu^2(D')}{p^{k\omega(D')}} + o(X) \\ &= \frac{1}{(p-1)\cdot p^k} \sum_{(D_u)_{u \in \mathcal{U}} \in \mathcal{P}(X, \boldsymbol{A}_1)} \frac{\mu^2(D')}{p^{k\omega(D')}} \cdot (p-1)^{\omega(D')} + o(X) \end{aligned}$$

where we denote $D = \prod_{u \in \mathbb{F}_p^{2k}} D_u$ and $D' = \prod_{u \in \mathcal{U}} D_u$.

Let $\mathcal{A}(\mathcal{U})$ be the set of tuples $\boldsymbol{A}$ which determine $\mathcal{U}$ by the above procedure, that is for which $A_u < 2/\Delta$ for all $u \notin \mathcal{U}$ and $A_u > X^{\ddagger}$ for all $u \in \mathcal{U}$.

We can partition $\sum'_{\boldsymbol{A}} S_k(X, \boldsymbol{A}) = \sum_{\mathcal{U}} S_k(X, \mathcal{U})$ where we define $S_k(X, \mathcal{U}) = \sum'_{\boldsymbol{A} \in \mathcal{A}(\mathcal{U})} \sum S_k(X, \boldsymbol{A})$. Notice that for each $u \in \mathcal{U}$ the range of summation of each D_u in $S_k(X, \mathcal{U})$ is $X^{\ddagger} < D_u < X$. It follows from Section 7B that we can extend this to $1 \leq D_u < X$ since it is proven there that the summation over these terms is contained in the error term.

Thus we have

$$S_k(X,\mathcal{U}) = \frac{1}{p^k(p-1)} \sum_{\prod_{j=1}^{p^k} n_j \in \mathcal{S}(X)} \mu^2\Bigl(\prod_{j=1}^{p^k} n_j\Bigr)\Bigl(\frac{p-1}{p^k}\Bigr)^{\omega\left(\prod_{j=1}^{p^k} n_j\right)} + o(X)$$
$$= \frac{1}{p^k} \sum_{n\in\mathcal{S}(X)} \mu^2(n)(p-1)^{\omega(n)-1} + o(X).$$

The last equality follows since there are $p^{k\omega(n)}$ ways of writing a positive integer n as a product of p^k positive integers. The proposition follows by summing $S_k(X,\mathcal{U})$ over all maximal unlinked sets $\mathcal{U}$. □

The final step of the proof will be the next proposition. Define $n(k,r)$ to be the number of r-dimensional subspaces of $\mathbb{F}_p^k$. We will need two properties of this function which can be found in Lemmas 1 and 3 from [Fouvry and Klüners 2007].

Lemma 8.7. *The function $n(k,r)$ satisfies*

$$n(k,r) = n(k,k-r), \qquad \sum_{r=0}^{k} p^r n(k,r) = \mathcal{N}(k+1) - \mathcal{N}(k).$$

Lemma 8.8. *The number U of maximal unlinked sets $\mathcal{U}\subset\mathbb{F}_p^{2k}$ is*

$$U = \mathcal{N}(k+1) - \mathcal{N}(k).$$

Proof. By Lemma 8.3 if V is a maximal unlinked subspace then $V+a$ is maximal unlinked if and only if $\rho(a)\in\pi(V)^\perp$. Hence given any k-dimensional subspace $V\subset\mathbb{F}_p^{2k}$ which is a maximal unlinked set there are $p^k(p^{\dim\pi(V)^\perp})$ vectors which translate V to a maximal unlinked set. However since translating by a_1 and a_2 gives the same set if and only if a_1 and a_2 are in the same coset of V this implies that there are $p^{\dim\pi(V)^\perp}$ distinct maximal unlinked sets that can be obtained from V.

Now let S be the set of k-dimensional subspaces $V\subset\mathbb{F}_p^{2k}$ which satisfy Lemma 8.2. We compute the size of this set. Fix some subspace $V_0\subset\mathbb{F}_p^k$ with $\dim V_0 = r$ and suppose V satisfies $\pi(V) = V_0$. So $\dim\pi(V)^\perp = k-r$. If $V\in S$ then $V = \pi(V)^\perp\oplus\pi(V) = V_0^\perp\oplus V_0$ and hence there is a unique $V\in S$ with $\pi(V) = V_0$. Hence the number of $V\in S$ with $\dim\pi(V) = r$ is $n(k,r)$.

Thus we have

$$U = \sum_{V\in S} p^{\dim\pi(V)^\perp} = \sum_{r=0}^{k} p^r n(k,r) = \mathcal{N}(k+1) - \mathcal{N}(k)$$

by Lemma 8.7. □

Thus combining Lemma 8.8 with Proposition 8.6 we have shown

$$S_k(X) = p^{-k}(\mathcal{N}(k+1) - \mathcal{N}(k)) \sum_{n<X} (p-1)^{\omega(n)-1} + o(X)$$

which proves Theorem 8.1. As remarked at the beginning of the section it follows that

$$\lim_{X\to\infty} \frac{\sum_{K,D_K<X} |\mathrm{im}(\varphi_K)|^k}{\sum_{K,D_K<X} 1} = p^{-k}(\mathcal{N}(k+1)-\mathcal{N}(k)). \tag{8-1}$$

Thus we have computed all the moments of the function $|\mathrm{im}(\varphi_K)|$ over degree p cyclic fields. We now refer to a result of Fouvry and Klüners which shows that these moments determine a distribution.

The combination of Proposition 1 and Theorem 2 from [Fouvry and Klüners 2006] can be summarized in the following form. This form is slightly more general than the original but follows by the same exact proof (see [Fouvry and Klüners 2006, pages 7–15]) which only uses properties of the function $\mathcal{N}(k, p)$ and $\eta_s(p)$ (defined in the introduction).

Let $\mathcal{F}$ be a family of number fields. Let f be a function on $\mathcal{F}$ valued in $\{1, p, p^2, \dots\}$. Let

$$\mathcal{M}(k) = \lim_{X\to\infty} \frac{\sum_{K\in\mathcal{F},|D_K|<X} f^k(K)}{\sum_{K\in\mathcal{F},|D_K|<X} 1}.$$

Proposition 8.9. *Let p be a prime. Suppose that for every $k \in \mathbb{Z}_{\geq 1}$*

$$\mathcal{M}(k) = p^{-k}(\mathcal{N}(k+1)-\mathcal{N}(k)).$$

Then for every $s \in \mathbb{Z}_{\geq 0}$ the density of the set $\{K \in \mathcal{F} \mid f(K) = p^s\}$ is

$$\frac{\eta_\infty(p)}{\eta_s(p)\eta_{s+1}(p)p^{s(s+1)}}.$$

By letting $\mathcal{F}$ be the set of degree p cyclic fields and $f = |\mathrm{im}(\varphi_K)| = p^{\mathrm{rk}_p \,\mathrm{im}(\varphi_K)}$ we see that Proposition 8.9 combined with (8-1) immediately implies Theorems 1.4 and 1.5.

Acknowledgements

The author extends his thanks to several people for helpful and stimulating discussions while working on this problem: John Friedlander, Jürgen Klüners, Jacob Tsimerman, Asif Zaman. He would also like to thank Leo Goldmakher and Matthew Young for providing useful references on character sums.

References

[Alberts 2016] B. Alberts, "Cohen–Lenstra moments for some nonabelian groups", preprint, 2016. arXiv

[Baier and Young 2010] S. Baier and M. P. Young, "Mean values with cubic characters", *J. Number Theory* **130**:4 (2010), 879–903. MR Zbl

[Bhargava 2014] M. Bhargava, "The geometric sieve and the density of squarefree values of invariant polynomials", preprint, 2014. arXiv

[Boston and Wood 2017] N. Boston and M. M. Wood, "Non-abelian Cohen–Lenstra heuristics over function fields", *Compos. Math.* **153**:7 (2017), 1372–1390. MR Zbl

[Cohen and Lenstra 1984] H. Cohen and H. W. Lenstra, Jr., "Heuristics on class groups of number fields", pp. 33–62 in *Number theory* (Noordwijkerhout, Netherlands, 1983), edited by H. Jager, Lecture Notes in Math. **1068**, Springer, 1984. MR Zbl

[Cohen and Martinet 1987] H. Cohen and J. Martinet, "Class groups of number fields: numerical heuristics", *Math. Comp.* **48**:177 (1987), 123–137. MR Zbl

[Datskovsky and Wright 1988] B. Datskovsky and D. J. Wright, "Density of discriminants of cubic extensions", *J. Reine Angew. Math.* **386** (1988), 116–138. MR Zbl

[Davenport 2000] H. Davenport, *Multiplicative number theory*, 3rd ed., Graduate Texts in Math. **74**, Springer, 2000. MR Zbl

[Davenport and Heilbronn 1971] H. Davenport and H. Heilbronn, "On the density of discriminants of cubic fields, II", *Proc. Roy. Soc. Lond. Ser. A* **322**:1551 (1971), 405–420. MR Zbl

[Ellenberg et al. 2016] J. S. Ellenberg, A. Venkatesh, and C. Westerland, "Homological stability for Hurwitz spaces and the Cohen–Lenstra conjecture over function fields", *Ann. of Math.* (2) **183**:3 (2016), 729–786. MR Zbl

[Fouvry and Klüners 2006] É. Fouvry and J. Klüners, "Cohen–Lenstra heuristics of quadratic number fields", pp. 40–55 in *Algorithmic number theory*, edited by F. Hess et al., Lecture Notes in Comput. Sci. **4076**, Springer, 2006. MR Zbl

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

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

[Gerth 1987] F. Gerth, III, "Densities for ranks of certain parts of p-class groups", *Proc. Amer. Math. Soc.* **99**:1 (1987), 1–8. MR Zbl

[Goldstein 1970] L. J. Goldstein, "A generalization of the Siegel–Walfisz theorem", *Trans. Amer. Math. Soc.* **149** (1970), 417–429. MR Zbl

[Hall 1938] P. Hall, "A partition formula connected with Abelian groups", *Comment. Math. Helv.* **11**:1 (1938), 126–129. MR Zbl

[Heath-Brown 2000] D. R. Heath-Brown, "Kummer's conjecture for cubic Gauss sums", *Israel J. Math.* **120**:part A (2000), 97–124. MR Zbl

[Iwaniec and Kowalski 2004] H. Iwaniec and E. Kowalski, *Analytic number theory*, Amer. Math. Soc. Colloq. Publ. **53**, Amer. Math. Soc., Providence, RI, 2004. MR Zbl

[Koymans and Pagano 2018] P. Koymans and C. Pagano, "On the distribution of $\mathrm{Cl}(K)[l^\infty]$ for degree l cyclic fields", preprint, 2018. arXiv

[Lemmermeyer 2013] F. Lemmermeyer, "The ambiguous class number formula revisited", *J. Ramanujan Math. Soc.* **28**:4 (2013), 415–421. MR Zbl

[Mayer 1992] D. C. Mayer, "Multiplicities of dihedral discriminants", *Math. Comp.* **58**:198 (1992), 831–847. MR Zbl

[Milovic 2017] D. Milovic, "On the 16-rank of class groups of $\mathbb{Q}(\sqrt{-8p})$ for $p \equiv -1 \mod 4$", *Geom. Funct. Anal.* **27**:4 (2017), 973–1016. MR Zbl

[Milovic 2018] D. Z. Milovic, "On the 8-rank of narrow class groups of $\mathbb{Q}(\sqrt{-4pq})$, $\mathbb{Q}(\sqrt{-8pq})$, and $\mathbb{Q}(\sqrt{8pq})$", *Int. J. Number Theory* **14**:8 (2018), 2165–2193.

[Neukirch 1999] J. Neukirch, *Algebraic number theory*, Grundlehren der Math. Wissenschaften **322**, Springer, 1999. MR Zbl

[Smith 2017] A. Smith, "2^∞-Selmer groups, 2^∞-class groups, and Goldfeld's conjecture", preprint, 2017. arXiv

[Stevenhagen 1995] P. Stevenhagen, "Rédei-matrices and applications", pp. 245–259 in *Number theory* (Paris, 1992-93), edited by S. David, Lond. Math. Soc. Lecture Note Ser. **215**, Cambridge Univ. Press, 1995. MR Zbl

[Wood 2019] M. M. Wood, "Nonabelian Cohen–Lenstra moments", *Duke Math. J.* **168**:3 (2019), 377–427. MR Zbl

[Wright 1989] D. J. Wright, "Distribution of discriminants of abelian extensions", *Proc. Lond. Math. Soc.* (3) **58**:1 (1989), 17–50. MR Zbl

Communicated by Bjorn Poonen
Received 2017-06-30 Revised 2019-10-30 Accepted 2019-11-27

jack.klys@ucalgary.ca *Department of Mathematics and Statistics, University of Calgary, Canada*

Algebra & Number Theory

msp.org/ant

PRODUCTION

production@msp.org

Silvio Levy, Scientific Editor

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2020 is US $415/year for the electronic version, and $620/year (+$60, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

ANT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

Algebra & Number Theory

Volume 14 No. 4 2020

The distribution of p-torsion in degree p cyclic fields 815
JACK KLYS

On the motivic class of an algebraic group 855
FEDERICO SCAVIA

A representation theory approach to integral moments of L-functions over function fields 867
WILL SAWIN

Deformations of smooth complete toric varieties: obstructions and the cup product 907
NATHAN ILTEN and CHARLES TURO

Mass equidistribution on the torus in the depth aspect 927
YUEKE HU

The basepoint-freeness threshold and syzygies of abelian varieties 947
FEDERICO CAUCCI

On the Ekedahl–Oort stratification of Shimura curves 961
BENJAMIN HOWARD

A moving lemma for relative 0-cycles 991
AMALENDU KRISHNA and JINHYUN PARK