ON THE DISTRIBUTION OF EIGENVALUES IN FAMILIES OF CAYLEY GRAPHS

MATILDE LALIN AND ANWESH RAY

Abstract. We consider the family of undirected Cayley graphs associated with odd cyclic groups, and study statistics for the eigenvalues in their spectra. Our results are motivated by analogies between arithmetic geometry and graph theory.

1. Introduction

It is well known that there are deep analogies between the arithmetic of curves in positive characteristic and the combinatorial properties of graphs. Ihara [Iha66a] and Sunada [Sun86] defined the zeta function associated to a locally finite connected graph. The definition closely resembles the *Selberg zeta function* and encodes asymptotics for closed walks in the graph. There is a close analogy between the zeta functions of curves in positive characteristic and the Ihara zeta functions associated to graphs. Further analogies have been realized by Baker and Norine [BN09], who etablished an analogue of the Abel-Jacobi map and the Riemann-Roch theorem in graph theory. Recently, statistical questions have been framed and studied for curves varying in certain naturally occurring ensembles, see for instance, [Rud08, KR09, Xio10, BDFL11, Ent12, TX14, CWZ15, BDF+16, BDFL16, San19, Ray22]. There has also been a growing interest in graph theoretic questions that are motivated by arithmetic statistics. The sandpile group (also known as the Jacobian group, the critical group, or the Picard group) is an abelian group naturally associated to a graph that codifies structural information of the graph. The statistics of such groups have been studied by various authors [CKL+15, Woo17]. More recently, the arithmetic statistics in the context of the Iwasawa theory of graphs has been studied in [DRLV22].

It is natural to study statistical questions for the poles of the Ihara zeta function [Iha66b, Bas92, Ter11]. For a finite regular graph, Ihara related the poles of the zeta function to the spectrum of the adjacency matrix. The spacing between the eigenvalues of regular graphs has been studied via numerical computation, and related to the Gaussian orthogonal ensemble, cf. [JMRR99, New05]. Cayley graphs are central objects in the combinatorial and geometric group theory. They allow us to visualize a group with respect to a generating set. Its spectrum can be described very concretely [Val94]. The main focus of this work is to study distribution questions for the family of eigenvalues that arise from Cayley graphs of finite odd cyclic groups.

Date: March 30, 2023.

2020 *Mathematics Subject Classification.* 05C25, 11H06, 11N60, 11M38.

Key words and phrases. arithmetic statistics, graph theory, eigenvalue distribution.

1.1. **Main result.** We now state our main result. Let $r \in \mathbb{Z}_{\geq 1}$ and let $\mathcal{F}$ be an infinite family of r-regular graphs with real eigenvalues. A *counting function* $h : \mathcal{F} \to \mathbb{Z}_{\geq 1}$ is a function such that for any $x \in \mathbb{R}_{\geq 0}$, there are only finitely many graphs $X \in \mathcal{F}$ that satisfy $h(X) \leq x$. Note that any eigenvalue of an r-regular graph in $\mathcal{F}$ lies in the interval $[-r, r]$. Given an interval $J = [a, b]$ contained in $[-r, r]$, we wish to evaluate the probability that a random eigenvalue of a graph in $\mathcal{F}$ does lie in J. We consider families of Cayley graphs associated with cyclic groups with odd order. Given an odd natural number n, let C_n denote the cyclic group with n elements, which we identify with $\mathbb{Z}/n\mathbb{Z}$. For a pair $\gamma = (G, S)$ such that $S = S^{-1}$, recall that X_γ is the associated Cayley graph. The association $\gamma \mapsto X_\gamma$ gives us a way to parametrize Cayley graphs. In greater detail, we set

$$\mathcal{F}_r := \left\{ X_\gamma \mid \gamma = (G, S); G = C_n \text{ for some odd number } n > 1, \#S = r, S = S^{-1} \right\}.$$

Thus, $\mathcal{F}_r$ is an infinite family of r-regular graphs; define a height function by setting $h(X_\gamma) := n$, where $\gamma = (C_n, S)$. There are only finitely many Cayley graphs for a given group, and hence, h is a counting function.

For $x \geq 0$, let $\sqrt{x}$ be the non-negative square root of x. Set δ denote the normalized arcsine distribution

$$\delta(u) := \begin{cases} \frac{1}{\pi\sqrt{1-u^2}} & \text{if } u \in [-1, 1]; \\ 0 & \text{otherwise}; \end{cases}$$

and let $\delta^{(k)}(u)$ be the k-fold convolution

$$\delta^{(k)}(u) := \int_{-1}^{1} \cdots \int_{-1}^{1} \delta(u_1)\delta(u_2) \cdots \delta(u_{k-1})\delta\left(u - \sum_{i=1}^{k-1} u_i\right) du_1 \cdots du_{k-1}.$$

Given a graph X, let $\mathrm{Sp}(X)$ denote its spectrum, i.e., the set of eigenvalues for its adjacency matrix.

Theorem 1.1. *Let $r \in \mathbb{Z}_{\geq 2}$ and set $k := \lfloor \frac{r}{2} \rfloor$. Let $J = [a, b]$ be an interval contained in $[-r, r]$. Set $I = [c, d] \subseteq [-k, k]$ to denote the interval $[a/2, b/2]$ (resp. $[(a-1)/2, (b-1)/2]$) if r is even (resp. r is odd). Then, we have that*

$$\lim_{n \to \infty} \frac{\#\{(X, \alpha) \mid X \in \mathcal{F}_r, \alpha \in \mathrm{Sp}(X), \text{ such that } h(X) = n, \alpha \in J\}}{\#\{(X, \alpha) \mid X \in \mathcal{F}_r, \alpha \in \mathrm{Sp}(X), \text{ such that } h(X) = n\}} = \int_c^d \delta^{(k)}(u) du.$$

More precisely, as $n \to \infty$,

$$\frac{\#\{(X, \alpha) \mid X \in \mathcal{F}_r, \alpha \in \mathrm{Sp}(X), \text{ such that } h(X) = n, \alpha \in J\}}{\#\{(X, \alpha) \mid X \in \mathcal{F}_r, \alpha \in \mathrm{Sp}(X), \text{ such that } h(X) = n\}}$$
$$= \int_c^d \delta^{(k)}(u) du + O(n^{-1+\epsilon}),$$

where the implicit constant in the error term only depends on r.

The above theorem is a direct consequence of Theorem 5.5. We note that the eigenvalues in the spectrum of a Cayley graph in $\mathcal{F}_1$ are all equal to 1, hence, we only consider the case when $r \geq 2$. The above probabilities are reinterpreted in terms of certain lattice point counts in certain regions in euclidean space. These lattice point counts are estimated via a combination combinatorial and analytic techniques.

The above result motivates the study of eigenvalue distributions in more general families of r-regular Cayley graphs, ordered by the size of the underlying group. The authors expect that the study of such questions shall lead to many interesting developments in the future.

Acknowledgements. The authors would like to thank Daniel Vallières for helpful discussions. This work is supported by the Natural Sciences and Engineering Research Council of Canada, Discovery Grant 355412-2022, the Fonds de recherche du Québec - Nature et technologies, Projet de recherche en équipe 300951, the Simons Foundation, and the Centre de recherches mathématiques.

2. Preliminaries

2.1. **Basic definitions and properties of graphs.** In this section, we recall a few preliminary notions in graph theory. For a more detailed exposition, we refer to [Sun13]. The data of a directed graph X consists of a tuple $(V_X, \mathbf{E}_X)$, where V_X is the set of vertices and $\mathbf{E}_X$ is the set of directed edges. The map $\operatorname{inc} : \mathbf{E}_X \to V_X \times V_X$ is the incidence map. The source and target maps $o : \mathbf{E}_X \to V_X$ and $t : \mathbf{E}_X \to V_X$ send a directed edge to the source and target vertices respectively, and $\operatorname{inc}(e) = (o(e), t(e))$. Given $e \in \mathbf{E}_X$, we say that e joins $o(e)$ to $t(e)$. We shall make the following assumptions on our graphs X.

(1) The sets V_X and $\mathbf{E}_X$ are finite.
(2) Given vertices $v_1 \in V_X$ and $v_2 \in V_X$ (not necessarily distinct), there is at most one edge joining v_1 to v_2.
(3) Given any edge $e \in \mathbf{E}_X$, there is an edge $\bar{e}$ such that $o(\bar{e}) = t(e)$ and $t(\bar{e}) = o(e)$. Note that if $o(e) = t(e)$ (i.e., e is a loop), then $\bar{e} = e$.

Consider an equivalence relation on the set $\mathbf{E}_X$ of directed edges, identifying e with $\bar{e}$. The equivalence classes consist of undirected edges of the graph and are denoted by E_X. Given $v \in V_X$, we consider the set of directed edges that orginate at v,

$$\mathbf{E}_{X,v} := \{e \in \mathbf{E}_X \mid o(e) = v\}.$$

The valency of v is defined as follows $\operatorname{val}_X(v) := \#\mathbf{E}_{X,v}$. Alternatively, the valency is also the number of undirected edges which contain v as one of its vertices. The graph is r-regular if $\operatorname{val}_X(v) = r$ for all $v \in V_X$. Let $g := \#V_X$ and choose a labelling on the vertices $V_X = \{v_1, \ldots, v_g\}$. Consider the $g \times g$ diagonal matrix defined by

$$d_{i,j} = \begin{cases} 0 \text{ if } i \neq j; \\ \operatorname{val}_X(v_i) \text{ if } i = j. \end{cases}$$

The adjacency matrix $A = (a_{i,j})$ is defined as follows

$$a_{i,j} = \begin{cases} \text{number of undirected edges between } v_i \text{ and } v_j \text{ if } i \neq j; \\ \text{twice the number of undirected loops at } v_i \text{ if } i = j. \end{cases}$$

The spectrum $\operatorname{Sp}(X)$ is the set of eigenvalues α of the adjacency matrix A. The matrix $Q := D - A$ is called the *Laplacian matrix* of the graph X. Let us explain the explicit relationship between the eigenvalues of the adjacency matrix and the poles of *Ihara zeta function* associated to X. A path in X is a sequence $C = a_1, \ldots, a_k$, where $(a_1, \ldots, a_k)$ is a tuple of directed edges of X such that $t(a_i) = o(a_{i+1})$. The length of C is the number of edges $\nu(C) := k$. A closed path is a path for which $o(a_1) = t(a_k)$, i.e. a path starting and ending at the same vertex. The path is said to have a *backtrack* if $a_{j+1} = \overline{a_j}$ for some j in the range $1 \leq j \leq k-1$. The path is said to have a *tail* if $a_k = \overline{a_1}$. Given a closed path D, and an integer $m > 0$, D^m is the path obtained by moving around D m-times. A closed path C is said to be *primitive* if C has no backtracks or tails and $C \neq D^m$ for any closed path D and $m > 1$. Two closed paths are equivalent if one can be obtained from the other by changing the starting vertex. Note that if C is primitive, there are exactly $\nu(C)$ paths that are equivalent to C. A *prime* path is an equivalence class of primitive paths. We shall sometimes denote the equivalence class of a primitive path C by $[C]$, however, when there is no cause for confusion, we simply use C itself to denote the prime path with primitive representative C.

The *Ihara polynomial* $h_X(u)$ denotes the determinant

$$h_X(u) := \det(I - Au + Qu^2).$$

Suppose that X is connected with no vertex of degree 1. With these definitions in mind, the Ihara zeta function is defined to be the product

$$\zeta_X(u) = \prod_C (1 - u^{\nu(C)})^{-1},$$

where the product ranges over all primes in X, ordered according to length. Then the Ihara 3-term determinant formula states that

$$\zeta_X(u)^{-1} = (1-u^2)^{r(X)-1} h_X(u),$$

where $r(X)$ is rank of the fundamental group, $r(X) = |E| - |V| + 1$. Note that when X is r-regular, where $r \geq 2$. Then $Q = (r-1)\,\mathrm{Id}$, and

$$h_X(u) = \prod_{\alpha \in \mathrm{Sp}(X)} \left(1 - \alpha u + (r-1)u^2\right).$$

Therefore, the roots of $h_X(u)$ are given by

$$r_\alpha^\pm = \frac{\alpha \pm \sqrt{\alpha^2 - 4(r-1)}}{2(r-1)},$$

where $\alpha \in \mathrm{Sp}(X)$.

Let G be a finite group and S be a subset of G such that $S^{-1} := \{s^{-1} \mid s \in S\}$ is equal to S. Associated to the pair $\gamma = (G, S)$ is the undirected Cayley graph X_γ. The vertices $V_\gamma := V_{X_\gamma}$ is the set $\{v_g \mid g \in G\}$, and there is an undirected edge joining v_g to v_h if and only if gh^{-1} is contained in S. Consider the special case when G is a finite abelian group and $\widehat{G} := \mathrm{Hom}(G, \mathbb{C}^\times)$ be the group of characters defined on G. A Cayley graph X_γ is r-regular, where, $r := \#S$. Given a character, $\chi \in \widehat{G}$, set $\lambda_\chi(X_\gamma) := \sum_{s \in S} \chi(s)$. The spectrum $\mathrm{Sp}(X_\gamma)$ is the set of eigenvalues of the adjacency matrix and is equal to the set $\{\lambda_\chi(X_\gamma) \mid \chi \in \widehat{G}\}$. Since $S = S^{-1}$, we find that the eigenvalues $\lambda_\chi(X_\gamma)$ are all real and contained in $[-r, r]$.

3. Eigenvalues of Cayley graphs

Let $r \in \mathbb{Z}_{\geq 2}$, and recall that $\mathcal{F}_r$ is the family of Cayley graphs

$$\mathcal{F}_r := \left\{X_\gamma \mid \gamma = (G, S); G = C_n \text{ for some odd number } n > 1, \#S = r, S = S^{-1}\right\}.$$

Since n is assumed to be odd, we find that $r = \#S$ is odd precisely when S contains 0. In this case, we write

$$\begin{cases} r = 2k & \text{if } r \text{ is even,} \\ r = 2k+1 & \text{if } r \text{ is odd.} \end{cases}$$

In either case, we may write $S \backslash \{0\} = T \cup (-T)$, where $T = \{a_1, \dots, a_k\}$ and $1 \leq a_i \leq \left(\frac{n-1}{2}\right)$ for $i = 1, \dots, k$. We shall assume without loss of generality that $1 \leq a_1 < a_2 < \dots < a_k \leq \left(\frac{n-1}{2}\right)$. Let S be a subset of $G = C_n$ such that $\#S = r$ and $S = -S$. We set $\gamma := (G, S)$. As χ ranges over $\widehat{G}$, the eigenvalues $\lambda_\chi(X_\gamma)$ are all distinct. Recall that the spectrum of X_γ is given by $\mathrm{Sp}(X_\gamma) = \{\lambda_\chi(X_\gamma) \mid \chi \in \widehat{G}\}$. For an integer $m \in [0, n-1]$, let $\chi_m \in \widehat{C_n}$ be the character defined by $\chi_m(x) := \exp\left(\frac{2\pi i m x}{n}\right)$. We set $\lambda_m(\gamma) := \lambda_\chi(X_\gamma)$ and we find that

$$\lambda_m(\gamma) = \lambda_{\chi_m}(\gamma) = \begin{cases} 1 + \sum_{j=1}^{k} 2\cos\left(\frac{2\pi m a_j}{n}\right) & \text{if } r \text{ is odd,} \\ \sum_{j=1}^{k} 2\cos\left(\frac{2\pi m a_j}{n}\right) & \text{if } r \text{ is even.} \end{cases} \tag{3.1}$$

In particular, we note that $\lambda_0(\gamma) = r$ is the largest eigenvalue.

We set $A_k(n)$ to denote the set of all vectors $\vec{a} = (a_1, \dots, a_k) \in \mathbb{Z}^k$ such that $1 \leq a_1 < a_2 < \dots < a_k \leq \left(\frac{n-1}{2}\right)$. Given $\vec{a} \in A_k(n)$, let $\gamma_{\vec{a}} := (C_n, S_{\vec{a}})$, where $S_{\vec{a}}$ is the set with r elements such that $S \backslash \{0\} = \{a_1, -a_1, a_2, -a_2, \dots, a_r, -a_r\}$. We set

$$\tau_m(\vec{a}) = \tau_m(\gamma_{\vec{a}}) := \sum_{i=1}^{k} \cos\left(\frac{2\pi m a_i}{n}\right).$$

Therefore, from (3.1), we find that

$$\lambda_m(\vec{a}) = \lambda_m(\gamma_{\vec{a}}) = \begin{cases} 1 + 2\tau_m(\vec{a}) & \text{if } r \text{ is odd,} \\ 2\tau_m(\vec{a}) & \text{if } r \text{ is even.} \end{cases}$$

We find that $\tau_m(\vec{a}) \in [-k, k]$, and that $\lambda_m(\vec{a}) \in [-r, r]$. Let $\mathcal{S}(n, k)$ be the set of pairs $\big(\gamma_{\vec{a}}, \tau_m(\vec{a})\big)$, where $\vec{a} \in A_k(n)$, and $m \in [0, n-1]$. Let $\mathcal{S}(n, k, m)$ be the subset of $\mathcal{S}(n, k)$ where m is fixed. Given and interval $[e, f] \subset \mathbb{R}$, we set $[e, f]_{\mathbb{Z}} := \mathbb{Z} \cap [e, f]$. We identify $\mathcal{S}(n, k)$ with $A_k(n) \times [0, n-1]_{\mathbb{Z}}$. Moreover, we identify $\mathcal{S}(n, k, m)$ with $A_k(n)$, upon identifying $\vec{a} \in A_k(n)$ with $(\vec{a}, m) \in A_k(n) \times [0, n-1]_{\mathbb{Z}}$. We refer to $\mathcal{S}(n, k, m)$ as the *m-slice* in $\mathcal{S}(n, k)$. Let $J = [a, b] \subseteq [-r, r]$ be a closed interval and $I = [c, d] \subseteq [-k, k]$ defined by

$$[c, d] = \begin{cases} [a/2, b/2] & \text{if } r \text{ is even;} \\ [(a-1)/2, (b-1)/2] & \text{if } r \text{ is odd.} \end{cases}$$

Note that $\lambda_m(\vec{a}) \in J$ if and only if $\tau_m(\vec{a}) \in I$. Set $\mathcal{S}_I(n, k) \subseteq \mathcal{S}(n, k)$ (resp. $\mathcal{S}_I(n, k, m) \subseteq \mathcal{S}(n, k, m)$) to denote the subset for which $\tau_m(\vec{a}) \in I$.

We have the following identifications

$$\begin{aligned} &\mathcal{S}_I(n, k, m) = \{(X, \alpha) \mid X \in \mathcal{F}_r, \alpha \in \mathrm{Sp}(X), \text{ such that } h(X) = n\}, \\ &\mathcal{S}_I(n, k, m) = \{(X, \alpha) \mid X \in \mathcal{F}_r, \alpha \in \mathrm{Sp}(X), \text{ such that } h(X) = n, \alpha \in J\}. \end{aligned} \tag{3.2}$$

It is easy to see that

$$\#\mathcal{S}(n,k,m) = \binom{\frac{n-1}{2}}{k},$$
$$\#\mathcal{S}(n,k) = n\binom{\frac{n-1}{2}}{k}.$$

For $n,k \in \mathbb{Z}_{\geq 1}$ such that n is odd and $k \leq \frac{n-1}{2}$, set

$$\mathrm{Prob}_I\,(n,k) := \frac{\#\mathcal{S}_I(n,k)}{\#\mathcal{S}(n,k)}.$$

Note that after making the identifications (3.2), we have that

$$\mathrm{Prob}_I\,(n,k) = \frac{\#\{(X,\alpha) \mid X \in \mathcal{F}_r, \alpha \in \mathrm{Sp}(X),\ \text{such that } h(X)=n \text{ and } \alpha \in J\}}{\#\{(X,\alpha) \mid X \in \mathcal{F}_r, \alpha \in \mathrm{Sp}(X),\ \text{such that } h(X)=n\}}. \tag{3.3}$$

For a fixed integer $k \in \mathbb{Z}_{\geq 1}$, we wish to compute the limit

$$\begin{aligned}\mathrm{Prob}_I(k) &:= \lim_{n\to\infty} \mathrm{Prob}_I\,(n,k)\,,\\ &= \lim_{n\to\infty} \left(\frac{(2^k k!)\#\mathcal{S}_I(n,k)}{n^{k+1}}\right).\end{aligned}$$

We express $\mathrm{Prob}_I(k)$ as an average

$$\mathrm{Prob}_I(n,k) = \frac{1}{n}\sum_{m=0}^{n-1} \mathrm{Prob}_I(n,k,m), \tag{3.4}$$

where

$$\mathrm{Prob}_I(n,k,m) := \frac{\#\mathcal{S}_I(n,k,m)}{\#\mathcal{S}(n,k,m)} = \frac{\#\mathcal{S}_I(n,k,m)}{\binom{\frac{n-1}{2}}{k}}.$$

In the next section, we shall set up a geometric interpretation for the probability $\mathrm{Prob}_I(n,k,m)$. Let B_k denote the k-dimensional box

$$B_k := \{(x_1,\ldots,x_k) \mid 0 \leq x_i < 1/2 \text{ for all } i\},$$

and note that $\mathrm{Vol}(B_k) = 2^{-k}$. Let $B_k(I)$ be the subset of B_k defined by

$$B_k(I) := \{(x_1,\ldots,x_k) \in B_k \mid \sum_{i=1}^{k} \cos(2\pi x_i) \in I\}. \tag{3.5}$$

Let us compute $\mathrm{Vol}\big(B_k(I)\big)$ for an interval $I=[c,d]$. Let $D_k(I) := \{(u_1,\ldots,u_k) \in [-1,1]^k \mid \sum_i u_i \in I\}$, and Φ be the function taking $D_k(I)$ to $B_k(I)$ defined by $\Phi(u_1,\ldots,u_k) = \Big(\frac{1}{2\pi}\cos^{-1}(u_1),\ldots,\frac{1}{2\pi}\cos^{-1}(u_k)\Big)$. For $x \geq 0$, let $\sqrt{x}$ be the non-negative square root of x. Set

$$\delta(u) := \begin{cases} \frac{1}{\pi\sqrt{1-u^2}} & \text{if } u \in [-1,1];\\ 0 & \text{otherwise;}\end{cases}$$

and let $\delta^{(k)}(u)$ be the k-fold convolution

$$\delta^{(k)}(u) := \int_{-1}^{1} \cdots \int_{-1}^{1} \delta(u_1)\delta(u_2)\cdots\delta(u_{k-1})\delta\left(u - \sum_{i=1}^{k-1} u_i\right) du_1 \cdots du_{k-1}.$$

Lemma 3.1. *Let $I = [c, d] \subseteq [-k, k]$ be an interval, $k \in \mathbb{Z}_{\geq 1}$, and let $B_k(I) \subset \mathbb{R}^k$ be the region defined by (3.5). Then, its volume is given by*

$$\mathrm{Vol}\left(B_k(I)\right) = 2^{-k} \int_c^d \delta^{(k)}(u) du.$$

Proof. By the change of variables theorem, we find that

$$\begin{aligned}\mathrm{Vol}\left(B_k(I)\right) &= \int_{D_k(I)} |\det D\Phi| \\ &= (2\pi)^{-k} \int_{D_k(I)} \frac{1}{\left(\sqrt{1-u_1^2}\sqrt{1-u_2^2}\ldots\sqrt{1-u_r^2}\right)} du_1 \ldots du_r \\ &= 2^{-k} \int_c^d \delta^{(k)}(u) du.\end{aligned} \tag{3.6}$$

□

4. A geometric interpretation

We wish to provide a geometric reinterpretation for the quantities $\#\mathcal{S}_I(n, k, m)$ and $\#\mathcal{S}(n, k, m)$ introduced in the previous section. Given an integer $n \geq 1$, let $L_n \subset \mathbb{R}^k$ be the lattice consisting of all vectors of the form $(\frac{a_1}{n}, \ldots, \frac{a_k}{n})$, where $(a_1, \ldots, a_k) \in \mathbb{Z}^k$. Let L'_n be the subset of L_n consisting of such vectors $(\frac{a_1}{n}, \ldots, \frac{a_k}{n})$ for which the coordinates are all mutually distinct. Set $\Omega_n := B_k \cap L_n$, $\Omega_n(I) := B_k(I) \cap L_n$, $\Omega'_n := B_k \cap L'_n$ and $\Omega'_n(I) := B_k(I) \cap L'_n$ and let $\kappa : \mathbb{R} \to [0, 1/2]$ be the function defined as follows

$$\kappa(x) := \begin{cases} \{x\} & \text{if } \{x\} \leq 1/2; \\ 1 - \{x\} & \text{if } \{x\} > 1/2. \end{cases} \tag{4.1}$$

Here, $\{x\}$ denotes the fractional part of x. We note that $\cos(2\pi x) = \cos\left(2\pi\kappa(x)\right)$. For a pair $\gamma = (G, S)$, such that $G = C_n$ and S a subset of G with $S = S^{-1}$, we recall from the previous section that $S\backslash\{0\} = T \cup (-T)$, where $T = \{a_1, \ldots, a_k\}$, and $\vec{a} = (a_1, \ldots, a_k) \in A_k(n)$. Let σ be a permutation of $\{1, \ldots, k\}$ and $m \in [1, n-1]_{\mathbb{Z}}$. Write $m = dm_1$, $n = dn_1$, where $d := (m, n)$. Consider a point $(\vec{a}, m) \in \mathcal{S}(n, k, m)$. Associate to this point and to the permutation σ, the point

$$\begin{aligned}P(\vec{a}, \sigma, m, n) &:= \left(\kappa\left(\frac{ma_{\sigma(1)}}{n}\right), \kappa\left(\frac{ma_{\sigma(2)}}{n}\right), \ldots, \kappa\left(\frac{ma_{\sigma(k)}}{n}\right)\right), \\ &= \left(\kappa\left(\frac{m_1 a_{\sigma(1)}}{n_1}\right), \kappa\left(\frac{m_1 a_{\sigma(2)}}{n_1}\right), \ldots, \kappa\left(\frac{m_1 a_{\sigma(k)}}{n_1}\right)\right).\end{aligned} \tag{4.2}$$

It is clear that $P(\vec{a}, \sigma, m, n)$ is contained in Ω_{n_1}.

For ease of notation, identify the m-slice $\mathcal{S}(n, k, m)$ with $A_k(n)$, by identifying $\vec{a}$ with $(\vec{a}, m)$. For $m \in [0, n-1]_{\mathbb{Z}}$, it conveniences us to further subdivide $\mathcal{S}(n, k, m)$ into two disjoint subsets $\mathcal{S}'(n,k,m)$ and $\mathcal{S}''(n,k,m)$ as follows. Set $\mathcal{S}'(n,k,0) := \emptyset$ and $\mathcal{S}''(n,k,0) := \mathcal{S}(n,k,0)$. For $m \in [1, n-1]_{\mathbb{Z}}$, let $\mathcal{S}'(n,k,m)$ consist of all $\vec{a} \in A_k(n)$ such that for $i < j$,

$$a_i \not\equiv \pm a_j \mod n_1,$$

and set $\mathcal{S}''(n,k,m)$ to be the complement $\mathcal{S}(n,k,m)\backslash\mathcal{S}'(n,k,m)$. We shall set $\mathcal{S}'(n,k)$ (resp. $\mathcal{S}''(n,k)$) to denote the disjoint union of slices $\mathcal{S}'(n,k,m)$ (resp. $\mathcal{S}''(n,k,m)$) as m ranges through $[0, n-1]_{\mathbb{Z}}$. It shall turn out that the slices $\mathcal{S}'(n,k,m)$ are better suited to combinatorial arguments than $\mathcal{S}(n,k,m)$. Lemma 4.2 will show that when k is fixed, $\#\mathcal{S}''(n,k)$ is small in comparison to $\#\mathcal{S}(n,k)$, as n goes to ∞. Given an interval $I \subseteq [-k,k]$, we set $\mathcal{S}'_I(n,k,m)$ (resp. $\mathcal{S}''_I(n,k,m)$) to be the subset of $\mathcal{S}'(n,k,m)$ (resp. $\mathcal{S}''(n,k,m)$) for which $\tau_m(\vec{a}) \in I$.

Lemma 4.1. *There is a constant $C_k > 0$ (depending on k), independent of m and n such that*

$$\#\mathcal{S}''(n,k,m) < \frac{C_k n^k}{n_1}.$$

The constant C_k can be taken to be a polynomial in k.

Proof. Given a pair $\tau = (i,j)$ such that $i < j$, let $\mathcal{S}''_\tau(n,k,m)$ be the subset of $\mathcal{S}''(n,k,m)$ such that

$$a_i \equiv \pm a_j \mod n_1.$$

The total number of subsets $\{a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_k\}$ is at most $\binom{\frac{n-1}{2}}{k-1}$. Given one such choice,

$$\begin{aligned} a_j \in &\{a_i + rn_1 \mid -\lfloor n/n_1 \rfloor - 1 \le r \le \lfloor n/n_1 \rfloor + 1\} \\ &\cup \{-a_i + rn_1 \mid -\lfloor n/n_1 \rfloor - 1 \le r \le \lfloor n/n_1 \rfloor + 1\}. \end{aligned}$$

Therefore, we find that

$$\#\mathcal{S}''_\tau(n,k,m) \le 4\left(\lfloor n/n_1 \rfloor + 1\right)\binom{\frac{n-1}{2}}{k-1}.$$

Thus, we have that

$$\#\mathcal{S}''(n,k,m) \le 4\binom{k}{2}\left(\lfloor n/n_1 \rfloor + 1\right)\binom{\frac{n-1}{2}}{k-1} < \frac{C_k n^k}{n_1}.$$

$\square$

Given $n \in \mathbb{Z}_{\ge 1}$, let $d(n)$ denote the number of divisors of n. Note that for any $\epsilon > 0$, $d(n) = O(n^\epsilon)$. Given non-negative functions f, g and h, we write

$$f = g + O_k(h)$$

if there is a constant $C_k > 0$ which depends only on k such that

$$|f - g| \le C_k h.$$

Lemma 4.2. *There is a constant $C_k > 0$, independent of n, and depending on k, such that*

$$\#\mathcal{S}''(n,k) < C_k n^k \left(\sum_{d|n} \frac{\varphi(d)}{d}\right) < C_k d(n) n^k,$$

where $d(n)$ is the number of divisors of n. As a result, we find that for any $\epsilon > 0$,

$$\#\mathcal{S}''(n,k) = O_k(n^{k+\epsilon}),$$

and that

$$\frac{\#\mathcal{S}''(n,k)}{\#\mathcal{S}(n,k)} = O_k(n^{-1+\epsilon}).$$

Proof. Let $m \in [1, n-1]_{\mathbb{Z}}$, set $d = (m, n)$. Then, Lemma 4.1 implies that

$$\#\mathcal{S}''(n,k,m) < C_k \frac{n^k}{n_1} = C_k n^{k-1} d.$$

We find that

$$\begin{aligned}
\#\mathcal{S}''(n,k) &= \#\mathcal{S}''(n,k,0) + \sum_{m=1}^{n-1} \#\mathcal{S}''(n,k,m), \\
&= \#\mathcal{S}''(n,k,0) + \sum_{d|n} \sum_{(m,n)=d} \#\mathcal{S}''(n,k,m), \\
&< n^k + C_k \sum_{d|n} n^{k-1} d \#\{m \in [1, n-1]_{\mathbb{Z}} \mid (m,n) = d\}, \\
&= n^k + C_k \sum_{d|n} n^{k-1} d\varphi\left(\frac{n}{d}\right), \\
&= n^k + C_k n^k \left(\sum_{d|n} \frac{\varphi(d)}{d}\right).
\end{aligned}$$

Replace C_k above with $(C_k + 1)$ to obtain the result. □

Remark 4.3. The function $g(n) := \sum_{d|n} \varphi(d)\left(\frac{n}{d}\right)$ is known as Pillai's arithmetical function [Pil33]. Set $\omega(n)$ to denote the number of prime divisors of n. We note that a bound of

$$g(n) \leq 27n \left(\frac{\log n}{\omega(n)}\right)^{\omega(n)}$$

is explicitly proven by Broughan [Bro01, Theorem 3.1].

Lemma 4.4. *Let $\sigma \in S_k$, $m \in [1, n-1]_{\mathbb{Z}}$, and $\vec{a} \in \mathcal{S}'(n,k,m)$. Let I be an interval contained in $[-k, k]$. Then, the point $P(\vec{a}, \sigma, m, n)$ defined above lies in Ω'_{n_1}. Furthermore, the following are equivalent*

1. *$\tau_m(\vec{a}) \in I$,*
2. *$P(\vec{a}, \sigma, m, n) \in \Omega'_{n_1}(I)$.*

Proof. Suppose that

$$\kappa\left(\frac{m_1 a_{\sigma(i)}}{n_1}\right) = \kappa\left(\frac{m_1 a_{\sigma(j)}}{n_1}\right),$$

where we recall that $\kappa(x)$ is given by (4.1). Then,

$$m_1 a_{\sigma(i)} \equiv \pm m_1 a_{\sigma(j)} \mod n_1.$$

Since m_1 is coprime to n_1, we find that

$$a_{\sigma(i)} \equiv \pm a_{\sigma(j)} \mod n_1.$$

However, since $\vec{a} \in \mathcal{S}'(n,k,m)$,

$$a_{\sigma(i)} \not\equiv \pm a_{\sigma(j)} \mod n_1$$

unless $i = j$. This proves that all the coordinates of the point $P(\vec{a},\sigma,m,n)$ are distinct. Furthermore, all coordinates are of the form b_i/n_1, where b_i is an integer, and these numbers lie in the range $[0,1/2]$. Therefore, the point $P(\vec{a},\sigma,m,n)$ is contained in Ω'_{n_1}.

Write $P(\vec{a},\sigma,m,n) = (y_1,\dots,y_k)$, where, $y_i = \kappa\left(\frac{ma_{\sigma(i)}}{n}\right)$. Recall that $B_k(I)$ consists of all tuples $(x_1,\dots,x_k) \in B_k$ such that $\sum_{i=1}^k \cos(2\pi x_i) \in I$. The point $(y_1,\dots,y_k) \in \Omega'_{n_1}$ lies in $\Omega'_{n_1}(I)$ if and only if $\sum_{i=1}^k \cos(2\pi y_i) \in I$. Observe that

$$\begin{aligned}
&\sum_{i=1}^k \cos(2\pi y_i)\\
=&\sum_{i=1}^k \cos\left(2\pi\kappa\left(\frac{m_1 a_{\sigma(i)}}{n_1}\right)\right)\\
=&\sum_{i=1}^k \cos\left(\frac{2\pi m_1 a_{\sigma(i)}}{n_1}\right)\\
=&\sum_{i=1}^k \cos\left(\frac{2\pi m_1 a_i}{n_1}\right) = \tau_m(\vec{a}).
\end{aligned}$$

This means that the conditions are equivalent. □

Lemma 4.5. *With respect to notation above,*

$$\#\mathcal{S}'(n,k,m) = \left(\frac{(n,m)+1}{2}\right)^k \#\mathcal{S}'(n_1,k,m_1),$$
$$\#\mathcal{S}'_I(n,k,m) = \left(\frac{(n,m)+1}{2}\right)^k \#\mathcal{S}'_I(n_1,k,m_1).$$

Proof. Given a tuple $(a_1,\dots,a_k) \in \mathcal{S}'(n_1,k,m_1)$ and $(a'_1,\dots,a'_k) \in \left[0,\left(\frac{(n,m)-1}{2}\right)\right]^k$, set $(b_1,\dots,b_k)$ be given by $b_i := a_i + a'_i m_1$. Note that since $a_i \in [1,\frac{m_1-1}{2}]$,

$$b_i < \left(\frac{m_1-1}{2}\right) + \left(\frac{(n,m)-1}{2}\right) m_1 = \frac{m-1}{2}.$$

Let $(c_1,\dots,c_k)$ be the permutation of the coordinates of $(b_1,\dots,b_k)$ such that $1 \le c_1 < c_2 < \dots < c_k \le \frac{m-1}{2}$. Note that $a_i \not\equiv \pm a_j \mod m_1$ for $i \neq j$, and therefore, $c_i \not\equiv \pm c_j \mod m_1$ for $i \neq j$. This implies that $(c_1,\dots,c_k) \in \mathcal{S}'(n,k,m)$. The association

$$\big((a_1,\dots,a_k),(a'_1,\dots,a'_k)\big) \mapsto (c_1,\dots,c_k)$$

sets up a bijection

$$\mathcal{S}'(n_1,k,m_1)\times\left[0,\left(\frac{(n,m)-1}{2}\right)\right]^k\to\mathcal{S}'(n,k,m),$$

which restricts to a bijection

$$\mathcal{S}'_I(n_1,k,m_1)\times\left[0,\left(\frac{(n,m)-1}{2}\right)\right]^k\to\mathcal{S}'_I(n,k,m).$$

Therefore, we conclude that

$$\#\mathcal{S}'(n,k,m)=\left(\frac{(n,m)+1}{2}\right)^k\#\mathcal{S}'(n_1,k,m_1),$$
$$\#\mathcal{S}'_I(n,k,m)=\left(\frac{(n,m)+1}{2}\right)^k\#\mathcal{S}'_I(n_1,k,m_1).$$

□

Proposition 4.6. *For $m\in[1,n-1]_{\mathbb{Z}}$, and $I\subseteq[-k,k]$ be a subset. Then the following assertions hold.*

(1) *There is a surjection*

$$\Phi:\Omega'_{n_1}\to\mathcal{S}'(n_1,k,m_1)$$

whose fibres all have cardinality $k!$.

(2) *The map Φ restricts to a map*

$$\Phi_I:\Omega'_{n_1}(I)\to\mathcal{S}'_I(n_1,k,m_1),$$

whose fibers also have cardinality $k!$.

Proof. We set $d:=(m,n)$, $m_1:=m/d$ and $n_1:=n/d$. Consider a point $x=(x_1,\dots,x_k)\in\Omega'_{n_1}$. By definition, there exist distinct integers $b_i\in[0,\frac{n_1-1}{2}]$ such that $x_i=b_i/n_1$. Let $c_i\in[0,n_1-1]$ be such that $m_1c_i\equiv b_i \mod n_1$ for all i. Then, set

$$a'_i=\begin{cases}c_i & \text{if } c_i\in[0,\frac{n_1-1}{2}]\\ (n_1-c_i) & \text{if } c_i\in[\frac{n_1+1}{2},n_1-1].\end{cases}$$

Note that $a'_i\in[0,\frac{n_1-1}{2}]$ and that $\kappa\left(\frac{m_1a'_i}{n_1}\right)=\frac{m_1c_i}{n_1}=\frac{b_i}{n_1}$. There is a tuple $\vec{a}=(a_1,\dots,a_k)\in\mathcal{S}'(n_1,k,m_1)$ and a permutation $\sigma\in S_k$ such that $(a_{\sigma(1)},\dots,a_{\sigma(k)})=(a'_1,\dots,a'_k)$. Thus, there is a unique permutation σ such that

$$(x_1,\dots,x_k)=\left(\kappa\left(\frac{m_1a_{\sigma(1)}}{n_1}\right),\dots,\kappa\left(\frac{m_1a_{\sigma(i)}}{n_1}\right),\dots,\kappa\left(\frac{m_1a_{\sigma(k)}}{n_1}\right)\right).$$

We map the tuple $(x_1,\dots,x_k)$ to $(a_1,\dots,a_k)$. Consider the action of S_k on Ω'_{n_1} via permutation of coordinates. The fibres of this map have cardinality $k!=\#S_k$. That $\Phi^{-1}(\mathcal{S}'_I(n_1,k,m_1))=\Omega'_{n_1}(I)$ is easy to see, and the second assertion follows. □

Lemma 4.7. *Let $m \in [1, n-1]_{\mathbb{Z}}$ and I be an interval contained in $[-k, k]$. Then we have that*

$$\#\mathcal{S}'(n,k,m) = \left(\frac{(n,m)+1}{2}\right)^k \frac{\#\Omega'_{n_1}}{k!},$$
$$\#\mathcal{S}'_I(n,k,m) = \left(\frac{(n,m)+1}{2}\right)^k \frac{\#\Omega'_{n_1}(I)}{k!}.$$

Proof. Lemma 4.5 asserts that the following relations hold

$$\#\mathcal{S}'(n,k,m) = \left(\frac{(n,m)+1}{2}\right)^k \#\mathcal{S}'(n_1,k,m_1),$$
$$\#\mathcal{S}'_I(n,k,m) = \left(\frac{(n,m)+1}{2}\right)^k \#\mathcal{S}'_I(n_1,k,m_1).$$

By Proposition 4.6, we have that

$$\#\mathcal{S}'(n_1,k,m_1) = \frac{\#\Omega_{n_1}}{k!},$$
$$\#\mathcal{S}'_I(n_1,k,m_1) = \frac{\#\Omega_{n_1}(I)}{k!},$$

and combining the above relations, the result follows. □

Corollary 4.8. *Let $m \in [1, n-1]$, we have that*

$$\#\mathcal{S}'(n,k,m) = \left(\frac{(n,m)+1}{2}\right)^k \frac{\#\Omega_{n_1}}{k!} + O_k((n,m)n^{k-1}),$$
$$\#\mathcal{S}'_I(n,k,m) = \left(\frac{(n,m)+1}{2}\right)^k \frac{\#\Omega_{n_1}(I)}{k!} + O_k((n,m)n^{k-1}).$$

Proof. Let $\Omega''_{n_1} := \Omega_{n_1} \backslash \Omega'_{n_1}$ and $\Omega''_{n_1}(I) := \Omega_{n_1}(I) \backslash \Omega'_{n_1}(I)$. The set Ω''_{n_1} consists of all tuples $(x_1, \ldots, x_k)$ such that for some $i < j$, $x_i = x_j$. Clearly, the cardinality of this set can be bounded as follows

$$\#\Omega''_{n_1} \leq \binom{k}{2} \left(\frac{n_1 - 1}{2}\right)^{k-1}.$$

It follows from Lemma 4.7 that

$$\#\mathcal{S}'(n,k,m) = \left(\frac{(n,m)+1}{2}\right)^k \frac{\#\Omega_{n_1}}{k!} - \left(\frac{(n,m)+1}{2}\right)^k \frac{\#\Omega''_{n_1}}{k!},$$
$$\#\mathcal{S}'_I(n,k,m) = \left(\frac{(n,m)+1}{2}\right)^k \frac{\#\Omega_{n_1}(I)}{k!} - \left(\frac{(n,m)+1}{2}\right)^k \frac{\#\Omega''_{n_1}(I)}{k!}.$$

Therefore, the error term can be bounded as follows

$$\left(\frac{(n,m)+1}{2}\right)^k \frac{\#\Omega''_{n_1}}{k!} \leq \frac{1}{k!}\binom{k}{2} \left(\frac{n_1 - 1}{2}\right)^{k-1} \left(\frac{(n,m)+1}{2}\right)^k = O_k((n,m)n^{k-1}),$$

and the result follows. □

5. Proof of the main theorem

In the previous section, we obtained a geometric interpretation for the quantities $\#\mathcal{S}'(n,k,m)$ and $\#\mathcal{S}'_I(n,k,m)$, in terms of the quantities $\#\Omega'_{n_1}$ and $\#\Omega'_{n_1}(I)$ respectively. In order to effectively bound $\#\Omega'_n(I)$, it suffices bound $\#\Omega_n(I)$.

Lemma 5.1. *Suppose that $I=[c,d]$ is an interval contained in $[-k,k]$, and suppose that $(\frac{a_1}{n},\dots,\frac{a_k}{n})\in B_k(I)$. Then, $(\frac{a_1+1}{n},\frac{a_2+1}{n},\dots,\frac{a_k+1}{n})\in B_k\left(\left[c-\frac{2\pi k}{n},d+\frac{2\pi k}{n}\right]\right)$.*

Proof. It suffices to prove that

$$\sum_{i=1}^{k}\cos\left(\frac{2\pi a_i}{n}\right)\in[c,d]\Longrightarrow\sum_{i=1}^{k}\cos\left(\frac{2\pi(a_i+1)}{n}\right)\in\left[c-\frac{2\pi k}{n},d+\frac{2\pi k}{n}\right].$$

To see the above, notice that for $0\le x\le\pi$ and $\epsilon>0$ such that $x+\epsilon\le\pi$, it follows from the mean value theorem that

$$\cos(x)-\epsilon\le\cos(x+\epsilon)\le\cos(x)+\epsilon.$$

□

Proposition 5.2. *Let $I=[c,d]$ be an interval contained in $[-k,k]$. With respect to notation above,*

$$\#\Omega_n(I)=n^k\operatorname{Vol}(B_k(I))+O_k(n^{k-1}).$$

Proof. For a given element $(\frac{a_1}{n},\dots,\frac{a_n}{n})\in\Omega_n(I)$, let $\mathcal{B}_{(a_1,\dots,a_k)}$ denote the box

$$\mathcal{B}_{(a_1,\dots,a_k)}=\left\{(x_1,\dots,x_k)\,\middle|\,\frac{a_j}{n}\le x_j\le\frac{a_j+1}{n}\right\}.$$

It is clear that $\operatorname{Vol}(\mathcal{B}_{(a_1,\dots,a_k)})=\frac{1}{n^k}$. Now we consider the set $U_{k,n}$ which is the union of $\mathcal{B}_{(a_1,\dots,a_k)}$, where $(\frac{a_1}{n},\dots,\frac{a_n}{n})\in\Omega_n(I)$. This is a set containing $B_k(I)$. We have

$$\operatorname{Vol}(B_k(I))\le\frac{\#\Omega_n(I)}{n^k}=\operatorname{Vol}(U_{k,n}).$$

From Lemma 5.1, it follows that

$$U_{k,n}\subseteq B_k\left(\left[c-\frac{2\pi k}{n},d+\frac{2\pi k}{n}\right]\right).$$

Thus we conclude that

$$\begin{aligned}
n^k\operatorname{Vol}(B_k(I))&\le\#\Omega_n(I)\\
&\le n^k\operatorname{Vol}\left(B_k\left(\left[c-\frac{2\pi k}{n},d+\frac{2\pi k}{n}\right]\right)\right)\\
&=n^k\operatorname{Vol}(B_k(I))\\
&\quad+n^k\operatorname{Vol}\left(B_k\left(\left[c-\frac{2\pi k}{n},c\right]\right)\right)+n^k\operatorname{Vol}\left(B_k\left(\left[d,d+\frac{2\pi k}{n}\right]\right)\right).
\end{aligned}$$

To estimate the error term, it suffices to bound

$$\left(\frac{n}{2}\right)^k \int_{c-\frac{2\pi k}{n}}^{c} \delta^{(k)}(u)du \text{ and } \left(\frac{n}{2}\right)^k \int_{d}^{d+\frac{2\pi k}{n}} \delta^{(k)}(u)du.$$

Note that the support of δ is contained in $[-1,1]$, and hence the support of $\delta^{(k)}$ is contained in $[-k,k]$. Therefore, if $c=-k$, then, $\int_{c-\frac{2\pi k}{n}}^{c} \delta^{(k)}(u)du = 0$. Without loss of generality, assume that $c \in (-k,k)$, and let $c' \in (-k,c)$. Set $M := \sup\{|\delta^{(k)}(u)| \mid u \in [c',c]\}$. Note that for $n > \frac{2\pi k}{(c-c')}$,

$$\int_{c-\frac{2\pi k}{n}}^{c} \delta^{(k)}(u)du \leq \frac{2\pi k M}{n}.$$

Therefore, we have shown that

$$\left(\frac{n}{2}\right)^k \int_{c-\frac{2\pi k}{n}}^{c} \delta^{(k)}(u)du = O_k(n^{k-1}).$$

The same reasoning shows that

$$\left(\frac{n}{2}\right)^k \int_{d}^{d+\frac{2\pi k}{n}} \delta^{(k)}(u)du = O_k(n^{k-1}).$$

Therefore, we have shown that

$$\#\Omega_n(I) = n^k \operatorname{Vol}(B_k(I)) + O_k\left(n^{k-1}\right).$$

□

Proposition 5.3. *With respect to notation above,*

$$\#\mathcal{S}'_I(n,k,m) = \#\mathcal{S}'(n,k,m)2^k \operatorname{Vol}(B_k(I)) + O_k((n,m)n^{k-1}).$$

Proof. We have

$$\begin{aligned}\#\mathcal{S}'_I(n,k,m) &= \left(\frac{(n,m)+1}{2}\right)^k \frac{\#\Omega_{n_1}(I)}{k!} + O_k((n,m)n^{k-1}),\\ &= \left(\frac{(n,m)+1}{2}\right)^k \frac{n_1^k}{k!} \operatorname{Vol}(B_k(I)) + O_k((n,m)n^{k-1})\\ &= \left(\frac{(n,m)+1}{2}\right)^k \#\Omega_{n_1} \frac{2^k}{k!} \operatorname{Vol}(B_k(I)) + O_k((n,m)n^{k-1})\\ &= \#\mathcal{S}'(n,k,m)2^k \operatorname{Vol}(B_k(I)) + O_k((n,m)n^{k-1}).\end{aligned}$$

For the above set of inequalities, we use Corollary 4.8 and Proposition 5.2. □

Proposition 5.4. *With respect to notation above,*

$$\#\mathcal{S}'_I(n,k) = \#\mathcal{S}'(n,k)2^k \operatorname{Vol}(B_k(I)) + O_k(n^{k+\epsilon}).$$

Proof. Note that $\mathcal{S}'(n,k,0)$ is empty. By an argument similar to the one employed in Lemma 4.2,

$$\begin{aligned}
\#\mathcal{S}'_I(n,k) &= \sum_{m=1}^{n-1} \#\mathcal{S}'_I(n,k,m), \\
&= \sum_{m=1}^{n-1} \#\mathcal{S}'(n,k,m) 2^k \operatorname{Vol}(B_k(I)) + O_k\left(n^{k-1}\sum_{m=1}^{n-1}(n,m)\right), \\
&= \#\mathcal{S}'(n,k) 2^k \operatorname{Vol}(B_k(I)) + O_k\left(n^{k-1}\sum_{d|n} d\varphi\left(\frac{n}{d}\right)\right), \\
&= \#\mathcal{S}'(n,k) 2^k \operatorname{Vol}(B_k(I)) + O_k(n^{k+\epsilon}),
\end{aligned}$$

where in the second line, we have applied Proposition 5.3. □

Theorem 5.5. *For $k \geq 1$ fixed, and $I = [c,d] \subseteq [-k,k]$,*

$$\operatorname{Prob}_I(n,k) = \int_c^d \delta^{(k)}(u)du + O_k(n^{-1+\epsilon}).$$

Proof. Recall from (3.4) that

$$\operatorname{Prob}_I(n,k) = \frac{1}{n}\sum_{m=0}^{n-1} \operatorname{Prob}_I(n,k,m),$$

where

$$\operatorname{Prob}_I(n,k,m) = \frac{\#\mathcal{S}_I(n,r,m)}{\#\mathcal{S}(n,r,m)} = \frac{\#\mathcal{S}_I(n,k,m)}{\binom{\frac{n-1}{2}}{k}}.$$

Note that $\#\mathcal{S}_I(n,k,m) = \#\mathcal{S}'_I(n,k,m) + \#\mathcal{S}''_I(n,k,m)$, and therefore,

$$\begin{aligned}
\operatorname{Prob}_I(n,k) =& \frac{1}{n}\sum_{m=0}^{n-1} \operatorname{Prob}_I(n,k,m) \\
=& \frac{1}{n}\sum_{m=0}^{n-1} \frac{\mathcal{S}'_I(n,k,m)}{\binom{\frac{n-1}{2}}{k}} + \frac{\mathcal{S}''_I(n,k)}{n\binom{\frac{n-1}{2}}{k}}.
\end{aligned}$$

By Lemma 4.2, we find that

$$\mathcal{S}''_I(n,k) \leq \mathcal{S}''(n,k) = O_k(d(n)n^k) = O_k(n^{k+\epsilon}),$$

and therefore,

$$\operatorname{Prob}_I(n,k) = \frac{1}{n}\frac{\mathcal{S}'_I(n,k)}{\binom{\frac{n-1}{2}}{k}} + O_k(n^{-1+\epsilon}).$$

By Proposition 5.4,

$$\begin{aligned}\#\mathcal{S}'_I(n,k) =&\#\mathcal{S}'(n,k)2^k \operatorname{Vol}(B_k(I)) + O_k(n^{k+\epsilon}),\\ =&\#\mathcal{S}(n,k)2^k \operatorname{Vol}(B_k(I)) + O_k(n^{k+\epsilon}),\\ =&n\binom{\frac{n-1}{2}}{k}2^k \operatorname{Vol}(B_k(I)) + O_k(n^{k+\epsilon}).\end{aligned}$$

Therefore,

$$\begin{aligned}\operatorname{Prob}_I(n,k) =&2^k \operatorname{Vol}(B_k(I)) + O_k(n^{-1+\epsilon}),\\ =&\int_c^d \delta^{(k)}(u)du + O_k(n^{-1+\epsilon}).\end{aligned}$$

This completes the proof. □

Proof of Theorem 1.1. It follows from (3.3) that

$$\operatorname{Prob}_I(n,k) = \frac{\#\{(X,\alpha) \mid X \in \mathcal{F}_r, \alpha \in \operatorname{Sp}(X),\ \text{such that } h(X) = n, \alpha \in J\}}{\#\{(X,\alpha) \mid X \in \mathcal{F}_r, \alpha \in \operatorname{Sp}(X),\ \text{such that } h(X) = n\}}.$$

By theorem 5.5,

$$\operatorname{Prob}_I(n,k) = \int_c^d \delta^{(k)}(u)du + O_k(n^{-1+\epsilon}),$$

which proves the result. □

References

[Bas92] Hyman Bass. The Ihara-Selberg zeta function of a tree lattice. *Internat. J. Math.*, 3(6):717–797, 1992.

[BDF+16] Alina Bucur, Chantal David, Brooke Feigon, Nathan Kaplan, Matilde Lalín, Ekin Ozman, and Melanie Matchett Wood. The distribution of $\mathbb{F}_q$-points on cyclic ℓ-covers of genus g. *International Mathematics Research Notices*, 2016(14):4297–4340, 2016.

[BDFL11] Alina Bucur, Chantal David, Brooke Feigon, and Matilde Lalín. Biased statistics for traces of cyclic p-fold covers over finite fields., 2011.

[BDFL16] Alina Bucur, Chantal David, Brooke Feigon, and Matilde Lalin. Statistics for ordinary Artin-Schreier covers and other p-rank strata. *Transactions of the American Mathematical Society*, 368(4):2371–2413, 2016.

[BN09] Matthew Baker and Serguei Norine. Harmonic morphisms and hyperelliptic graphs. *Int. Math. Res. Not. IMRN*, (15):2914–2955, 2009.

[Bro01] Kevin A. Broughan. The gcd-sum function. *J. Integer Seq.*, 4(2):Article 01.2.2, 19, 2001.

[CKL+15] Julien Clancy, Nathan Kaplan, Timothy Leake, Sam Payne, and Melanie Matchett Wood. On a Cohen-Lenstra heuristic for Jacobians of random graphs. *J. Algebraic Combin.*, 42(3):701–723, 2015.

[CWZ15] GilYoung Cheong, Melanie Wood, and Azeem Zaman. The distribution of points on superelliptic curves over finite fields. *Proceedings of the American Mathematical Society*, 143(4):1365–1375, 2015.

[DRLV22] Cédric Dion, Anwesh Ray, Antonio Lei, and Daniel Vallières. On the distribution of Iwasawa invariants associated to multigraphs. *Preprint, arXiv:2207.07213*, 2022.

[Ent12] Alexei Entin. On the distribution of zeroes of Artin–Schreier l-functions. *Geometric and Functional Analysis*, 22(5):1322–1360, 2012.

[Iha66a] Yasutaka Ihara. On discrete subgroups of the two by two projective linear group over $\mathfrak{p}$-adic fields. *J. Math. Soc. Japan*, 18:219–235, 1966.

[Iha66b] Yasutaka Ihara. On discrete subgroups of the two by two projective linear group over $\wp$-adic fields. *J. Math. Soc. Japan*, 18:219–235, 1966.

[JMRR99] Dmitry Jakobson, Stephen D. Miller, Igor Rivin, and Zeév Rudnick. Eigenvalue spacings for regular graphs. In *Emerging applications of number theory (Minneapolis, MN, 1996)*, volume 109 of *IMA Vol. Math. Appl.*, pages 317–327. Springer, New York, 1999.

[KR09] Pär Kurlberg and Zeév Rudnick. The fluctuations in the number of points on a hyperelliptic curve over a finite field. *Journal of Number Theory*, 129(3):580–587, 2009.

[New05] Derek B. Newland. *Kernels in the Selberg trace formula on the k-regular tree and zeros of the Ihara zeta function*. ProQuest LLC, Ann Arbor, MI, 2005. Thesis (Ph.D.)–University of California, San Diego.

[Pil33] S. S. Pillai. On an arithmetic function. *J. Annamalai Univ.*, 2:243–248, 1933.

[Ray22] Anwesh Ray. Statistics for p-ranks of artin–schreier covers. *Acta Arithmetica*, 205:211–226, 2022.

[Rud08] Zeév Rudnick. Traces of high powers of the Frobenius class in the hyperelliptic ensemble. *arXiv preprint arXiv:0811.3649*, 2008.

[San19] Soumya Sankar. Proportion of ordinarity in some families of curves over finite fields. *arXiv preprint arXiv:1904.12173*, 2019.

[Sun86] Toshikazu Sunada. L-functions in geometry and some applications. In *Curvature and topology of Riemannian manifolds (Katata, 1985)*, volume 1201 of *Lecture Notes in Math.*, pages 266–284. Springer, Berlin, 1986.

[Sun13] Toshikazu Sunada. *Topological crystallography*, volume 6 of *Surveys and Tutorials in the Applied Mathematical Sciences*. Springer, Tokyo, 2013. With a view towards discrete geometric analysis.

[Ter11] Audrey Terras. *Zeta functions of graphs*, volume 128 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2011. A stroll through the garden.

[TX14] Frank Thorne and Maosheng Xiong. Distribution of zeta zeroes for cyclic trigonal curves over a finite field. *preprint*, 2014.

[Val94] Alain Valette. Can one hear the shape of a group? *Rend. Sem. Mat. Fis. Milano*, 64:31–44 (1996), 1994.

[Woo17] Melanie Matchett Wood. The distribution of sandpile groups of random graphs. *J. Amer. Math. Soc.*, 30(4):915–958, 2017.

[Xio10] Maosheng Xiong. The fluctuations in the number of points on a family of curves over a finite field. *Journal de théorie des nombres de Bordeaux*, 22(3):755–769, 2010.

(Lalin) Université de Montréal, Pavillon André-Aisenstadt, Dépt. de mathématiques et de statistique, CP 6128, succ. Centre-ville, Montréal, Québec, H3C 3J7, Canada
Email address: `matilde.lalin@umontreal.ca`

(Ray) Centre de recherches mathématiques, Université de Montréal, Pavillon André-Aisenstadt, 2920 Chemin de la tour, Montréal (Québec) H3T 1J4, Canada
Email address: `anwesh.ray@umontreal.ca`