EQUIDISTRIBUTION OF GEODESICS ON HOMOLOGY CLASSES AND ANALOGUES FOR FREE GROUPS

YIANNIS N. PETRIDIS AND MORTEN S. RISAGER

Abstract. We investigate how often geodesics have homology in a fixed set of the homology lattice of a compact Riemann surface. We prove that closed geodesics are equidistributed on a random set of homology classes and certain arithmetic sets. We explain the analogues for free groups, conjugacy classes and discrete logarithms, in particular, we investigate the density of conjugacy classes with relatively prime discrete logarithms.

1. Introduction

Let M be a compact Riemann surface of genus $g>1$ and let $\pi(T)$ denote the number of prime closed geodesics γ on M whose length l_γ is at most T. Huber [9] and Selberg [23] proved the prime geodesic theorem

$$\pi(T)\sim\frac{e^T}{T},\qquad \text{as } T\to\infty. \tag{1.1}$$

In this paper we investigate how the prime geodesics are distributed among the homology classes $\beta\in\mathbb{Z}^{2g}\overset{\psi}{\simeq}H_1(M,\mathbb{Z})$. If $\tilde\psi:\Gamma\to H_1(M,\mathbb{Z})$ is the map of the fundamental group to the first homology group, we let $\phi=\psi^{-1}\circ\tilde\psi$. For a set $A\subseteq\mathbb{Z}^{2g}$ we will consider to what extend

$$\pi_A(T)=\#\{\{\gamma\}|\gamma \text{ prime } l_\gamma\leq T,\phi(\gamma)\in A\} \tag{1.2}$$

depends on the set A. We recall that to every conjugacy class $\{\gamma\}\subset\Gamma$ corresponds a unique closed oriented geodesic on M of lenght l_γ. We will say that *the prime geodesics are equidistributed on a set* $A\subseteq\mathbb{Z}^{2g}$ if

$$\frac{\pi_A(T)}{\pi(T)}\to d(A),\quad \text{as } T\to\infty, \tag{1.3}$$

where $d(A)$ is the natural density of A in $\mathbb{Z}^{2g}$. This only makes sense if the natural density $d(A)$

$$d(A)=\lim_{T\to\infty}\frac{\#\{\alpha\in A,|\alpha_i|\leq T,i=1,\dots,2g\}}{\#\{\alpha\in\mathbb{Z}^{2g},|\alpha_i|\leq T,i=1,\dots,2g\}} \tag{1.4}$$

exists. Our main result is the following theorem:

Theorem 1.1. *The prime geodesics on a compact Riemann surface of genus $g>1$ are equidistributed on a random set.*

Date: September 27, 2005.

2000 *Mathematics Subject Classification.* Primary 05C25; Secondary 20F69, 37D40, 11M36.

The first author was partially supported by a Humboldt Foundation Research Fellowship, PSC CUNY Research Award, No. 66520-00-35, and NSF grant DMS 0401318 while the second author was supported by a grant from Carlsberg.

For a precise definition in terms of probability of the notion of random set see the introduction to Theorem 2.14. We note that a random set has natural density 1/2. To prove Theorem 1.1 we need to exhibit cancellation in certain exponential sums related to the random set (See Definition 2.6). This kind of cancellation can be verified also for certain individual sets. We get the following result:

Theorem 1.2. *The prime geodesics on a compact Riemann surface of genus $g > 1$ are equidistributed on*

(i) *Finite sets.*
(ii) *Shifted sublattices $\vec{a} + L$ of $\mathbb{Z}^{2g}$, where $\vec{a} \in \mathbb{Z}^{2g}$ and $L \subset \mathbb{Z}^{2g}$ is a lattice.*
(iii) *The set of lattice points with coprime coordinates.*

Remark 1.3. The natural density is zero for the first type of sets (finite), $1/\mathrm{vol}(\mathbb{Z}^{2g}/L)$ for the second type of sets, and $\zeta(2g)^{-1}$ for (iii) by Cesaro's classic result [5]. In all cases where we can prove equidistribution we can get error terms for the rate of convergence in 1.3 but to improve readability we have chosen not make this point explicit in the proofs.

Remark 1.4. The proof of Theorems 1.1 and 1.2 uses the Selberg trace formula with characters as used in [18]. We combine this approach with ideas from [25], where the stationary phase argument used in [18] is simplified to make more transparent the dependence on the homology class. This idea seems to go a back at least to [20]. As an intermediate step toward proving Theorems 1.1 and 1.2 we get strong improvements on average of the local limit theorem of Sharp [24] (see Theorem 2.10). We need also one new ingredient (Corollary 4.5, Lemma 4.4), which tells us that certain averages over A of appropriate functions converge to the density of A.

Remark 1.5. For sets containing exactly one element α the counting function $\pi_\alpha(T)$ was studied by Adachi and Sunada [2, 1] and Phillips and Sarnak [18], as well as many others. Phillips and Sarnak found the full asymptotic expansion with leading term

$$\pi_\alpha(T) \sim (g-1)^g \frac{e^T}{T^{g+1}}, \qquad \text{as } T \to \infty. \tag{1.5}$$

In particular the leading term, in contrast to the lower order terms, does not depend on α. The dependence on α in the lower order terms has been considered in [12, 25], but the results are not strong enough to handle equidistribution for sets of positive density by simply summing up asymptotics. Equation (1.5) is a much stronger statement than Theorem 1.2 (i), since one cannot recover the main term in (1.5) from Theorem 1.2 (i). For sets of positive natural density Theorems 1.1 and 1.2 gives precise information about the asymptotic behaviour of $\pi_A(T)$. Theorem 1.2 (ii) follows also from the Chebotarev density theorem for closed geodesics (see [22, 15, 26]) in the case of abelian covers. The error terms obtained in this way do not seem to be good enough to take linear combinations (inclusion exclusion) to get Theorem 1.2 (iii). The general set A in Theorem 1.1 carries no group like structure and seems out of reach using variants of the Chebotarev density theorem.

Theorems 1.1 and 1.2 have analogues for free groups (and other hyperbolic groups). Let now $\Gamma = F(A_1, \ldots, A_k)$, $k \geq 2$ be the free group on k generators. The words $\gamma \in \Gamma$ can be counted according to their word length $\mathrm{wl}(\gamma)$ and one finds (see [16, 19]) that the function $\Pi(m)$ counting conjugacy classes $\{\gamma\}$ in Γ with length at most m satisfies

$$\Pi(m) \sim \frac{q}{q-1}\frac{q^m}{m}, \qquad \text{as } m \to \infty, \tag{1.6}$$

which is the analogue of (1.1). Here $q = 2k-1$. We define discrete logarithms on the generators

$$\begin{array}{cccc} \log_j : & \Gamma & \to & \mathbb{Z} \\ & A_i & \mapsto & \delta_{ij}. \end{array} \tag{1.7}$$

The above definition extends to Γ by requiring that $\log_j$ is an additive homomorphism. Hence $\log_j$ counts the number of occurrences (with signs) of the generator A_j. We let

$$\begin{array}{cccc} \Phi : & \Gamma & \to & \mathbb{Z}^k \\ & \gamma & \mapsto & (\log_1(\gamma), \ldots, \log_k(\gamma)). \end{array} \tag{1.8}$$

This is the map to the abelianization of Γ exactly as ϕ, and it is well-defined on conjugacy classes. We therefore think of the images of Φ as analogous to homology classes in M (they are homology classes for a certain graph constructed in 3.1). We investigate how conjugacy classes of the free group are distributed in the lattice $\mathbb{Z}^k$. For $B \subseteq \mathbb{Z}^k$ we consider

$$\Pi_B(m) = \#\{\{\gamma\} \in \{\Gamma\} |\, \mathrm{wl}\,(\{\gamma\}) \leq m, \Phi(\{\gamma\}) \in B\}, \tag{1.9}$$

where $\{\Gamma\}$ is the set of conjugacy classes of Γ. We will say that the conjugacy classes *are equidistributed on a set $B \subseteq \mathbb{Z}^k$* if

$$\frac{1}{2}\left(\frac{\Pi_B(m)}{\Pi(m)} + \frac{\Pi_B(m+1)}{\Pi(m+1)}\right) \to d(B), \qquad \text{as } m \to \infty. \tag{1.10}$$

As in 1.3 this only makes sense if the natural density $d(B)$ exist. The fact that we look at averages over m and $m+1$ turns out to be natural. See Remark 1.8 below.

Theorem 1.6. *The free group elements are equidistributed on random sets and*

(i) *Finite sets.*
(ii) *Shifted sublattices $\vec{a} + L$ of $\mathbb{Z}^k$, where $\vec{a} \in \mathbb{Z}^k$ and $L \subset \mathbb{Z}^k$ is a lattice.*
(iii) *The set of lattice points with coprime coordinates.*

Remark 1.7. The main idea in the proof of Theorem 1.6 is to analyze the relevant counting functions

$$\sum_{\substack{\gamma \in \Gamma \\ \mathrm{wl}(\gamma) \leq m}} \chi(\gamma), \tag{1.11}$$

(the sum only runs over cyclically reduced words) where χ is a character on Γ, using an identity due to Ihara. This identity gives an expression for the generating function for $\chi(\gamma)$ as a rational function. This enables us to give asymptotic expansions with an error term for (1.11). We integrate over the character variety to pick up a specific homology class. The identity for the Ihara zeta function is analogous to the Selberg trace formula as encoded in the Selberg zeta function.

We obtain a new proof of the local limit theorem for free groups of Sharp [24] using the spectral theory of a simple graph, rather than the thermodynamic formalism and subshifts of finite type. We also obtain strong improvements on average. (See Theorems 3.7 and 3.10.) The comment about error terms from Remark 1.3 applies also in the case of free groups.

Remark 1.8. In Theorem 1.6 we cannot in general get a limit without averaging for m and $m+1$. If $B = \{\vec{v}, v_i \equiv a_i \,(\text{mod}\, l_i)\,, i = 1, \ldots, k\}$, where all the moduli $l_1, \cdots l_k$ are even the limits over the subsequence with m even and the subsequence with m odd exist and are computed in Section 3.5 and they do *not* coincide. If at least one modulus is odd we do not need to average, i.e. in that case

$$\lim_{m\to\infty} \frac{\Pi_B(m)}{\Pi(m)} = \frac{1}{l_1 \cdots l_k}.$$

Remark 1.9. Theorem 1.6 (ii) for moduli l_1 prime and $l_2 = \ldots = l_k = 1$ was first proved (in a slightly different formulation) by I. Rivin, [19], using graphs, and Theorem 1.6 (i) follows also from [24]. Our proofs are more elementary than [19] in the following sense: (a) we use a simpler graph, in fact one with a single vertex, (b) the analysis is simpler, since we have the Ihara zeta function identity, and we do not use asymptotics of special functions, like Chebychev polynomials, used in [19].

Remark 1.10. An element $\gamma_0 \in \Gamma$ is called a test element if every endomorphism of Γ fixing γ_0 is an automorphism of Γ. The property of being a test element has been studied extensively. We refer to [11] for further explanations and references. The property of being a test element can be characterized by relative primality of discrete logarithms. Recently Kapovich, Schupp, and Shpilrain [11] used Theorem 1.6 (iii) to prove that the property of being a test element in the free group on two generators is neither generic nor negligible in the sense of Gromov ([6], [7]). In fact, this was the application that initiated our interest in the present work. This seems to be the first known non-trivial example of an interesting property in the free group on two generators which is neither generic nor negligible.

2. Counting closed geodesics on Riemann surfaces

Let M be a smooth compact Riemann surface of genus $g > 1$ without boundary. Any such Riemann surface may be realized as $\Gamma\backslash\mathbb{H}$ where $\mathbb{H}$ is the upper half-plane and the fundamental group Γ is isomorphic to a discrete subgroup of $\mathrm{PSL}_2(\mathbb{R})$. There exists a fundamental set of generators $\{a_1, \ldots a_g, b_1, \cdots, b_g\} = \{C_1, \ldots, C_{2g}\} \subset \Gamma$ satisfying the relation

$$[a_1, b_1] \cdots [a_g, b_g] = 1. \tag{2.1}$$

There exists a basis $\omega_1, \ldots \omega_{2g}$ of harmonic 1-forms, dual to $C_1, \ldots C_{2g}$, i.e.

$$\int_{C_i} \omega_j = \delta_{ij}. \tag{2.2}$$

The first homology group $H_1(M, \mathbb{Z})$ can be identified as

$$H_1(M, \mathbb{Z}) \cong \left\{ \sum_{j=1}^{2g} n_j C_j,\, n_j \in \mathbb{Z} \right\} \cong \mathbb{Z}^{2g}. \tag{2.3}$$

For $\gamma \in \Gamma$ with homology $\sum_j n_j C_j$ we write $\phi(\gamma) = (n_1, \ldots, n_{2g}) \in \mathbb{Z}^{2g}$. For $\gamma \in \Gamma$ and $\epsilon \in \mathbb{R}^{2g}/\mathbb{Z}^{2g}$ we consider unitary characters

$$\begin{array}{cccc} \chi_\epsilon(\cdot) & : & \Gamma & \rightarrow & S^1 \\ & & \gamma & \mapsto & e^{2\pi i \langle \phi(\gamma), \epsilon \rangle} \end{array}. \tag{2.4}$$

We consider the set of square-integrable χ_ϵ-automorphic functions, i.e. the set of $f : \mathbb{H} \to \mathbb{C}$ such that

$$f(\gamma z) = \chi_\epsilon(\gamma) f(z) \tag{2.5}$$

and

$$\int_F |f(z)|^2 \, d\mu(z) < \infty, \tag{2.6}$$

where F is a fundamental domain for $\Gamma\backslash\mathbb{H}$. Let L_ϵ denote the Laplacian defined as the closure of $-y^2(\partial_x^2 + \partial_y^2)$ defined on smooth compactly supported functions satisfying (2.5) and (2.6). The Laplacian is self-adjoint and its spectrum consists of a countable set of eigenvalues $0 \leq \lambda_0(\epsilon) \leq \lambda_1(\epsilon) \leq \ldots$. By the maximum principle 0 is an eigenvalue if and only if $\epsilon = 0$. The behaviour of $\lambda_0(\epsilon)$ for ϵ small is of fundamental importance to our investigation.

Proposition 2.1. [18, Lemma 2.1, 2.2] *Let $\lambda_0(\epsilon)$ be the first eigenvalue of L_ϵ of a surface M with $g > 1$. Then*

(i) *$\lambda_0(\epsilon)$ is real analytic in ϵ near $\epsilon = 0$.*
(ii) *$\epsilon = 0$ is a critical point for $\lambda_0(\epsilon)$.*
(iii) *at $\epsilon = 0$ the Hessian $H = \{a_{ij}\}$ is positive definite and satisfies*

$$a_{ij} = \left.\frac{\partial^2 \lambda_0(\epsilon)}{\partial\epsilon_i \partial\epsilon_j}\right|_{\epsilon=0} = \frac{2\pi}{g-1} \langle \omega_i, \omega_j \rangle ,$$

and $\det(\langle \omega_i, \omega_j \rangle) = 1$.

We use this information about the smallest eigenvalue to count closed primitive geodesics on M with certain homological restrictions. The prime geodesics on M are in 1-1 correspondence with the primitive conjugacy classes of Γ. Hence by an abuse of notation we want to count geodesics $\{\gamma\}$ with a given homology class $\phi(\{\gamma\}) = \alpha$. Here $\{\gamma\}$ is the conjugacy class of γ in Γ. The main tool is the Selberg trace formula for $L(\epsilon)$ (see [23, 8, 27]). This relates the eigenvalues $\{\lambda_i(\epsilon)\}_{i=0}^\infty$ to the length spectrum of the surface, i.e. the set of lengths of the closed geodesics. Here l_γ is the length of the corresponding geodesic. We define – following [18, (2.26), (2.29)] –

$$R_\alpha(T) = \sideset{}{'}\sum_{\substack{\{\gamma\}, l_\gamma \leq T \\ \phi(\gamma)=\alpha}} \frac{l_\gamma}{\sinh(l_\gamma/2)}.$$

The $'$ on the sum means that we only sum over prime geodesics.

It is custumary to introduce $s_j(\epsilon)$ subject to $\lambda_j(\epsilon) = s_j(\epsilon)(1 - s_j(\epsilon))$, $\Re(s_j(\epsilon)) \geq 1/2$, $\Im(s_j(\epsilon)) \geq 0$. Hence $\lambda_0(\epsilon)$ close to zero corresponds to $s_0(\epsilon)$ close to 1. It is straightforward to translate Proposition 2.1 into statements about $s_0(\epsilon)$. The trace formula gives estimates for

$$R_\chi(T) = \sideset{}{'}\sum_{\{\gamma\}, l_\gamma \leq T} \frac{\chi(\gamma) l_\gamma}{\sinh(l_\gamma/2)}.$$

Let $\chi_\epsilon^\alpha = \exp(2\pi i \langle \alpha, \epsilon \rangle)$. The orthogonality of characters, i.e.

$$\int_{\mathbb{R}^{2g}/\mathbb{Z}^{2g}} \chi_\epsilon(\gamma) \overline{\chi_\epsilon^\alpha} \, d\epsilon = \delta_{\phi(\gamma)=\alpha}$$

allows to integrate the trace formula over $\mathbb{R}^{2g}/\mathbb{Z}^{2g}$ to get the following result:

Lemma 2.2. [18, (2.37)] *For all ε sufficiently small there exists a $\nu < 1/2$ such that for all $\alpha \in \mathbb{Z}^{2g}$*

$$R_\alpha(T) = 2e^{T/2} \int_{B(\varepsilon)} \frac{e^{(s_0(\epsilon)-1)T}}{s_0(\epsilon) - 1/2} \overline{\chi_\epsilon^\alpha} d\epsilon + O(e^{\nu T}). \tag{2.7}$$

Here $B(\varepsilon)$ is the open ball at zero with radius ε and the implied constant depends only on M.

Remark 2.3. We remark that there is a factor 2 missing in the formula [18, (2.37)]. This is due to the fact that in the trace formula [18, (2.27)] one should take the eigenvalue parameters $\pm r_j(\theta)$, and the contribution of the smallest $\lambda_0(\theta)$ should be counted twice. A small typo in [18, (2.44)] gives an extra factor 1/2 so [18] still get the correct asymptotics (1.5).

Phillips and Sarnak used a stationary phase argument on the integral (2.7) to find the asymptotic behaviour of $R_\alpha(T)$. The asymptotic formula 1.5 follows. Since we want to consider closed geodesics whose homology lies in more general sets than singletons, we consider

$$R_A(T) = \sum_{\substack{\{\gamma\}, l_\gamma \leq T \\ \phi(\gamma) \in A}} \frac{l_\gamma}{\sinh(l_\gamma/2)}.$$

where A is any subset of $\mathbb{Z}^{2g}$. The following lemma shows that in a certain sense a geodesic of 'small' lenght cannot have 'large' homology:

Lemma 2.4. *There exist a constant $c > 0$ such that for all $\gamma \in \Gamma$*

$$|n_i| \leq c l_\gamma$$

where $\phi(\gamma) = (n_1, \dots, n_{2g})$

Proof. This follows from e.g. Lemma 2.1 in [17] where in the present case the relevant modular symbol is formed using the cohomology class ω_i. □

It follows from Lemma 2.4 that

$$R_A(T) = \sum_{\substack{\alpha \in A \\ |\alpha_i| \leq cT}} R_\alpha(T).$$

We can therefore conclude from Lemma 2.2 that

$$R_A(T) = 2e^{T/2} \int_{B(\varepsilon)} \frac{e^{(s_0(\epsilon)-1)T}}{s_0(\epsilon) - 1/2} \sum_{\substack{\alpha \in A \\ |\alpha_i| \leq cT}} \overline{\chi_\epsilon^\alpha} d\epsilon + O(T^{2g} e^{\nu T}). \tag{2.8}$$

Equation (2.8) shows already the main exponential sum supported on the set A. To find the asymptotic behaviour of $R_A(T)$ we shall use a technique based on change of variable as in [24, 20]. This has the advantage over the stationary phase argument used in [18] that it allows us to look at several homology classes simultaneously. For this argument to work we need the following lemma:

Lemma 2.5. *Let*

$$\rho^2 = \frac{2\pi}{(g-1)} \text{ and } M = \{\langle \omega_i, \omega_j \rangle\}.$$

(i) *For every* $\epsilon_0 \in \mathbb{R}^{2g}$

$$e^{(s_0(\epsilon_0/\rho\sqrt{T})-1)T} \to e^{-\langle \epsilon_0, M\epsilon_0\rangle/2}$$

as $T \to \infty$.

(ii) *There exists* $\delta > 0$ *such that for all* $\|\epsilon\| < \delta\rho\sqrt{T}$.

$$\left| e^{(s_0(\epsilon/\rho\sqrt{T})-1)T} - e^{-\langle \epsilon, M\epsilon\rangle/2} \right| \le 2e^{-\langle \epsilon, M\epsilon\rangle/4}.$$

(iii) *There exist constants* $\delta, C > 0$ *such that for all* $T > 0$, $\|\epsilon\| < \delta T^{1/16}$,

$$\left| e^{(s_0(\epsilon/\rho\sqrt{T})-1)T} - e^{-\langle \epsilon, M\epsilon\rangle/2} \right| \le C\frac{1}{T^{1/2}}.$$

Proof. Consider the function $f(\epsilon) = e^{s_0(\epsilon)-1}$. Since $\lambda_0(\epsilon) = s_0(\epsilon)(1 - s_0(\epsilon))$ it is easy to derive from Lemma 2.1 that at $\epsilon = 0$ we have $\nabla f = 0$ and that the Hessian of f at $\epsilon = 0$ is $-\rho^2 M$. Since $s_0(\epsilon)$ is even, any odd number of derivatives of $s_0(\epsilon)$ at $\epsilon = 0$ must vanish. Hence by Taylor's theorem we have

$$f(\epsilon) = 1 - \frac{\langle \epsilon, \rho^2 M\epsilon\rangle}{2} + O(\|\epsilon\|^4).$$

All the claims now follow from Proposition 4.1 in the Appendix if we put $D = \rho^2 M$, $r = \rho$, $\nu = 2$ and $b = 1/2$. □

We shall now see how to use this to find an expansion for the integral in (2.8) By a change of variable we get

(2.9)
$$\frac{\rho^{2g} T^g R_A(T)}{4e^{T/2}} = \int_{B(\varepsilon\rho\sqrt{T})} \frac{e^{(s_0(\epsilon/\rho\sqrt{T})-1)T}}{2s_0(\epsilon/\rho\sqrt{T}) - 1} \sum_{\substack{\alpha\in A \\ |\alpha_i|\le cT}} \overline{\chi^{\alpha}_{\epsilon/\rho\sqrt{T}}} d\epsilon + O(T^{3g} e^{(\nu-1/2)T}).$$

The identity

$$\int_{\mathbb{R}^{2g}} e^{-\langle \epsilon, M\epsilon\rangle/2} \overline{\chi^{\alpha}_{\epsilon/\rho\sqrt{T}}} d\epsilon = (2\pi)^g e^{-\frac{4\pi^2}{\rho^2}\langle \alpha, M^{-1}\alpha\rangle/2T} \tag{2.10}$$

can be easily checked using the Fourier transform. Since $-\langle \epsilon, M\epsilon\rangle/2 + (\varepsilon\rho\sqrt{T})^2/4 \ll -\langle \epsilon, M\epsilon\rangle/4$ when $\epsilon \in B(\varepsilon\rho\sqrt{T})^c$ we conclude

$$\left| e^{\rho^2\varepsilon^2 T/4} \int_{\mathbb{R}^{2g}\setminus B(\varepsilon\rho\sqrt{T})} e^{-\langle \epsilon, M\epsilon\rangle/2} \overline{\chi^{\alpha}_{\epsilon/\rho\sqrt{T}}} d\epsilon \right| \ll \int_{\mathbb{R}^{2g}\setminus B(\varepsilon\rho\sqrt{T})} e^{-\langle \epsilon, M\epsilon\rangle/4} d\epsilon \le C.$$

Therefore the part of the integral (2.10) outside $B(\varepsilon\rho\sqrt{T})$ is of exponential decay in T. It follows that up to an error term of exponential decay

$$\frac{\rho^{2g} T^g R_A(T)}{4e^{T/2}} - \sum_{\substack{\alpha\in A \\ |\alpha_i|\le cT}} (2\pi)^g e^{-\frac{4\pi^2}{\rho^2}\langle \alpha, M^{-1}\alpha\rangle/2T} \tag{2.11}$$

equals

$$\int_{B(\varepsilon\rho\sqrt{T})} \left(\frac{e^{(s_0(\epsilon/\rho\sqrt{T})-1)T}}{2s_0(\epsilon/\rho\sqrt{T}) - 1} - e^{-\langle \epsilon, M\epsilon\rangle/2} \right) \sum_{\substack{\alpha\in A \\ |\alpha_i|\le cT}} \overline{\chi^{\alpha}_{\epsilon/\rho\sqrt{T}}} d\epsilon. \tag{2.12}$$

In order to be able to say something meaningful about this expression we need to see some cancellation in the character sum. We therefore make the following definition:

Definition 2.6. A set $A \subset \mathbb{Z}^{2g}$ is said to have property [C] if, for some $0 \leq \varepsilon < g$ there exist constants $c, \iota > 0$ such that for $T^{-2g} < |\epsilon_1|, \dots, |\epsilon_{2g}| < cT^{\iota}$

$$\sum_{\substack{\alpha \in A \\ |\alpha_i| \leq cT}} \chi^{\alpha}_{\epsilon/\rho\sqrt{T}} = O(f(\epsilon) T^{g+\varepsilon}), \tag{2.13}$$

where the implied constant may depend on ε, A, c, and M. If furthermore $f(\epsilon)$ satisfies

$$\int_{T^{-2g} < |\epsilon_1|, \dots, |\epsilon_{2g}| < cT^{\iota}} |f(\epsilon)| \left| e^{(s_0(\epsilon/\rho\sqrt{T})-1)T} - e^{-\langle \epsilon, M\epsilon \rangle/2} \right| d\epsilon = o(T^{-\varepsilon}) \tag{2.14}$$

and

$$\int_{T^{-2g} < |\epsilon_1|, \dots, |\epsilon_{2g}| < cT^{\iota}} |f(\epsilon)| \left\|\epsilon\right\|^2 \left| e^{(s_0(\epsilon/\rho\sqrt{T})-1)T} \right| d\epsilon = o(T^{1-\varepsilon}) \tag{2.15}$$

as $T \to \infty$ the set A is said to have property [GC].

Remark 2.7. We think of '[C]' as 'cancellation' ($\varepsilon = g$ is the trivial estimate on the exponential sum), and '[GC]' as 'good cancellation'. Note that when a set has positive natural density the sum contains asymptotically a constant times T^{2g} terms, so property [C] requires some bound towards 'square root' cancellation. The requirement that we stay away from a (shrinking) neighborhood of the hypersurfaces $\epsilon_i = 0$ seems inevitable since we cannot expect good cancellation at any hypersurface $\epsilon_i = 0$. Below we shall verify this condition for various sets, one of these being the full set $A = \mathbb{Z}^{2g}$. Using this it is easy to see that the set of sets of property [C] resp. [GC] is an algebra i.e. it is closed under addition and complements. Whether or not it is a σ-algebra (reducing the conditions to tautologies) remains open.

Before giving examples of sets with property [C] resp. [GC] we state and prove the main reason why we care about such sets.

Theorem 2.8. *Assume that $A \subseteq \mathbb{Z}^{2g}$ has property [GC], and that A has asymptotic density $d(A)$. Then*

$$\frac{\pi_A(T)}{\pi(T)} \to d(A) \quad \textit{as } T \to \infty.$$

Lemma 2.9. *Let $\sigma^2 = (2\pi(g-1))^{-1}$. Assume that $A \subseteq \mathbb{Z}^g$ has property [GC]. Then*

$$\frac{T^g R_A(T)}{4e^{T/2}} - \sum_{\substack{\alpha \in A \\ |\alpha_i| \leq cT}} \frac{1}{(2\pi\sigma^2)^g} e^{-\left\langle \alpha, M^{-1}\alpha \right\rangle / 2\sigma^2 T} = o(T^g).$$

Proof. The claim follows from showing that (2.12) is $o(T^g)$. Using the definition of property [GC] we may safely assume $\iota \leq 1/2$. If $\iota < 1/2$ we start by splitting the integral (2.12) in two:

$$\int_{\|\epsilon\| \leq cT^{\iota}} + \int_{cT^{\iota} \leq \|\epsilon\| \leq \varepsilon\rho T^{1/2}}. \tag{2.16}$$

We start by estimating the second integral. Since $s_0(\epsilon)$ is even with $s_0(0) = 1$ we have

$$\left|(2s_0(\epsilon) - 1)^{-1} - 1\right| \le C \|\epsilon\|^2 \tag{2.17}$$

when $\|\epsilon\| \le \varepsilon$. Hence the part of the integrand of (2.12) outside the sum is bounded absolutely by

$$\left|e^{(s_0(\epsilon/\rho\sqrt{T})-1)T} - e^{-\langle \epsilon, M\epsilon\rangle/2}\right| + C\left|e^{(s_0(\epsilon/\rho\sqrt{T})-1)T}\right| \|\epsilon\|^2 T^{-1}.$$

By Lemma 2.5 (ii) this is $O(e^{-\langle \epsilon, M\epsilon\rangle/4}\|\epsilon\|^2 T^{-1})$, and, since $\|\epsilon\| \ge cT^{\iota}$, the integrand (outside the sum) is $O(e^{-c(M)T^{2\iota}})$ where $c(M)$ is some constant depending on M. Hence, using the trivial estimate on the sum $O(T^{2g})$, the integral is $O(T^{2g}e^{-c(M)T^{2\iota}}T^g)$, which decays exponentially.

We now address the first integral in (2.16). This we also split in two integrals:

$$\int_{\substack{\|\epsilon\|\le cT^{\iota} \\ \exists i, |\epsilon_i| \le T^{-2g}}} + \int_{\substack{\|\epsilon\|\le cT^{\iota} \\ T^{-2g} < |\epsilon_i| \le cT^{\iota}}}.$$

To estimate the first we bound the integrand by an absolute constant times T^{2g} coming from the trivial estimate on the exponential sum. We then use the fact that the volume of the integration domain is $O(T^{(2g-1)/2}T^{-2g})$. Hence this integral is $O(T^{(2g-1)/2})$. Hence also $o(T^g)$.

To estimate the second we use (2.17) to conclude that

$$\left|\int_{\substack{\|\epsilon\|\le cT^{\iota} \\ T^{-2g} < |\epsilon_i| \le cT^{\iota}}}\right| \le CT^{-1}\int_{T^{-2g}<|\epsilon_i|\le cT^{\iota}} \|\epsilon\|^2 \left|e^{(s_0(\epsilon/\rho\sqrt{T})-1)T}\right| \Bigg|\sum_{\substack{\alpha\in A \\ |\alpha_i|\le cT}} \overline{\chi^{\alpha}_{\epsilon/\rho\sqrt{T}}}\Bigg| d\epsilon$$

$$+ \int_{T^{-2g}<|\epsilon_i|\le cT^{\iota}} \left|e^{(s_0(\epsilon/\rho\sqrt{T})-1)T} - e^{-\langle \epsilon, M\epsilon\rangle/2}\right| \Bigg|\sum_{\substack{\alpha\in A \\ |\alpha_i|\le cT}} \overline{\chi^{\alpha}_{\epsilon/\rho\sqrt{T}}}\Bigg| d\epsilon.$$

The assumption that A has property [GC] implies that both these terms are $o(T^g)$. This completes the proof. □

To prove Theorem 2.8 we will prove a result which is closely related to the main result in [25] (i.e. a local limit theorem). The following theorem improves the error terms of [25, Theorem 1] on average by a square root.

Theorem 2.10. *Let $\sigma^2 = (2\pi(g-1))^{-1}$. Assume that $A \subseteq \mathbb{Z}^g$ has property [GC] and that A has natural density. Then*

$$\frac{T^g \pi_A(T)}{e^T/T} - \sum_{\substack{\alpha\in A \\ |\alpha_i|\le cT}} \frac{1}{(2\pi\sigma^2)^g} e^{-\langle \alpha, M^{-1}\alpha\rangle/2\sigma^2 T} = o(T^g).$$

We notice that by Gauß-Bonnet the variance σ^2 equals half the volume of the surface.

Proof. We have

$$\pi_A(T) = \int_0^T \frac{\sinh(s/2)}{s} dR_A(s) = \int_0^T \frac{e(s/2)}{2s} dR_A(s) + O(1).$$

Integrating the integral by parts we find

$$\pi_A(T)=\frac{e^{T/2}}{2T}R_A(T)-\int_0^T\frac{1}{4s}e^{s/2}R_A(s)\,ds-\int_0^T\frac{1}{2s^2}e^{s/2}R_A(s)ds+O(1). \tag{2.18}$$

Using $R_A(s)=O(e^{s/2})$, which follows from Lemma 2.9, we easily find that the last integral is $O(e^T/T^2)$. We claim that

$$\int_0^T\frac{1}{s}e^{s/2}R_A(s)\,ds=\frac{e^{T/2}}{T}R_A(T)+o(e^T/T) \tag{2.19}$$

from which it follows that

$$\pi_A(T)=\frac{e^{T/2}}{4T}R_A(T)+o(e^T/T).$$

Substituting this into Lemma 2.9 we get exactly the statement of Theorem 2.10.

To prove the claim we notice that by Lemma 2.9 and Corollary 4.5 we have $R_A(T)=4d(A)e^{T/2}+o(e^{T/2})$, so there exist a positive function $g(T)$ decreasing to zero as $T\to\infty$ such that

$$\left|R_A(T)-4d(A)e^{T/2}\right|\leq g(T)e^{T/2}. \tag{2.20}$$

Consider now

$$\begin{aligned}\int_1^T&\frac{e^{s/2}}{s}R_A(s)ds-\frac{e^{T/2}}{T}R_A(T)\\ &=\int_1^T\frac{e^{s/2}}{s}\left(R_A(s)-4d(A)e^{s/2}\right)ds+4d(A)\int_1^T\frac{e^s}{s}ds-\frac{e^{T/2}}{T}R_A(T)\end{aligned}$$

The second term is $4d(A)e^T/T+O(e^T/T^2)$ and using (2.20) we find

$$=\int_1^T\frac{e^{s/2}}{s}\left(R_A(s)-4d(A)e^{s/2}\right)ds+o(e^T/T).$$

We split the integral into an integral from 1 to $T/2$ and from $T/2$ to T and use the bound (2.20):

$$\left|\int_1^{T/2}\frac{e^{s/2}}{s}\left(R_A(s)-4d(A)e^{s/2}\right)ds\right|\leq g(1)\int_1^{T/2}\frac{e^s}{s}ds=O(e^{T/2}/T)=o(e^T/T),$$

$$\left|\int_{T/2}^T\frac{e^{s/2}}{s}\left(R_A(s)-4d(A)e^{s/2}\right)ds\right|\leq g(T/2)\int_{T/2}^T\frac{e^s}{s}ds=O(g(T/2)e^T/T)=o(e^T/T).$$

This concludes the proof of the claim (2.19). □

Proof of Theorem 2.8. It follows easily from Theorem 2.10, (1.1) and Corollary 4.5. □

We have now shown why the properties in Definition 2.6 are relevant: prime geodesics are equidistributed on sets satisfying property [GC]. But we still need to see that there are sets with this property. There are three cases we consider: Finite sets (which from this point of view should be considered the trivial case: Phillips and Sarnak [18] get much stronger results), highly arithmetic sets (essentially arithmetic progressions) and random sets.

Proposition 2.11. *Finite sets have property [GC].*

Proof. This follows directly from the definition with $\varepsilon = 0$ quoting Lemma 2.5 (ii). $\square$

To handle the arithmetic sets we need the following simple lemma:

Lemma 2.12. *There exist a $c > 0$ such that for every $\varepsilon > 0$*

$$\int_{T^{-2g} \leq |\epsilon_1|,\ldots,|\epsilon_{2g}| \leq cT^{1/16}} \frac{\left| e^{(s_0(\epsilon/\rho\sqrt{T})-1)T} - e^{-\langle \epsilon, M\epsilon\rangle/2} \right|}{|\epsilon_1|\cdots|\epsilon_{2g}|} d\epsilon = o(T^{-1/2+\varepsilon})$$

and

$$\int_{T^{-2g} < |\epsilon_1|,\ldots,|\epsilon_{2g}| < cT^{1/16}} \frac{\|\epsilon\|^2 \left| e^{(s_0(\epsilon/\rho\sqrt{T})-1)T} \right|}{|\epsilon_1|\cdots|\epsilon_{2g}|} d\epsilon = o(T^{\varepsilon})$$

as $T \to \infty$.

Proof. To evaluate the first integral we use Lemma 2.5 (iii) to conclude that there exist $c = \delta$ such that the integrand is

$$O\left(\frac{T^{-1/2}}{|\epsilon_1|\cdots|\epsilon_{2g}|}\right).$$

The integral is therefore $O(T^{-1/2}(\log T)^{2g})$ which is $o(T^{-1/2+\varepsilon})$.

To evaluate the second integral we use Lemma 2.5 (ii) from which we easily find that the numerator of the integral is bounded, and the whole integral is therefore $O((\log T)^{2g})$ which is $o(T^{\varepsilon})$.

$\square$

This allows us to prove the following (where $\alpha \equiv \beta \,(\mathrm{mod}\ l)$ denotes that for all $i = 1, \ldots, 2g$ we have $\alpha_i \equiv \beta_i \,(\mathrm{mod}\ l_i)$):

Theorem 2.13. *Let $l \in \mathbb{N}^{2g}$, $\beta \in \mathbb{Z}^{2g}$. The sets*

$$A_l(\beta) = \{\alpha \in \mathbb{Z}^{2g} | \alpha \equiv \beta \,(\mathrm{mod}\ l)\}$$

$$A_{rp} = \{\alpha \in \mathbb{Z}^{2g} | \gcd(\alpha_1, \ldots \alpha_{2g}) = 1\}$$

have property [GC].

Proof. Consider first *any* arithmetic progression $B = \{b + kl | l \in \mathbb{Z}\}$. We use $|\sin(\pi x)| \geq 2\{x\}$, where $\{x\}$ is the distance between x and the closest integer, to conclude that for any $T', T'' > 0$

$$\Bigg| \sum_{\substack{a \in B \\ -T' \leq a \leq T''}} e^{2\pi i a x/\rho\sqrt{T}} \Bigg| \leq \frac{2}{\left| e^{2\pi i l x/\rho\sqrt{T}} - 1 \right|} \leq \frac{1}{\left| \sin(\pi l x/\rho\sqrt{T}) \right|} \leq \frac{\rho\sqrt{T}}{|2lx|},$$

where we have assumed that $|x|/\rho\sqrt{T} < 1/(2l)$.

We may safely assume that in Lemma 2.12 the constant $c < 1/2$. Then

$$\Bigg| \sum_{\substack{\alpha \in A_l(\beta) \\ |\alpha_i| \leq cT}} \chi^{\alpha}_{\epsilon/\rho\sqrt{T}} \Bigg| = \Bigg| \prod_{j=1}^{2g} \sum_{\substack{\alpha_j \equiv \beta_j (l_j) \\ |a_j| \leq cT}} e^{2\pi i \alpha_j \epsilon_j/\rho\sqrt{T}} \Bigg| \leq 2^{-2g} \frac{\rho^{2g} T^g}{|l_1 \cdots l_{2g}|\, |\epsilon_1 \ldots \epsilon_{2g}|} \tag{2.21}$$

when $\|\epsilon\| \leq \delta\rho\sqrt{T}$. Hence $A_l(\beta)$ has property [C] and it follows from Lemma 2.12 that for some $c > 0$ and $\iota = 1/16$ it also has property [GC].

To see that A_{rp} has the claimed properties we use the inclusion-exclusion principle: relative primality means we take all $2g$-tuples, take off the ones with common divisor of the entries a prime number, add the ones with common divisor a product of distinct primes and so on. This gives

$$\sum_{\substack{\alpha\in A_{rp}\\ |\alpha_i|\leq cT}} \chi^{\alpha}_{\epsilon/\rho\sqrt{T}} = \sum_{l=0}^{\infty}(-1)^l \sum_{p_1,\ldots,p_l}\sum_{\substack{\alpha\in A_{p_1\cdots p_l}\\ |\alpha_i|\leq cT}} \chi^{\alpha}_{\epsilon/\rho\sqrt{T}}. \tag{2.22}$$

Here the second sum on the right is over different primes and

$$A_{p_1\cdots p_l} = A_{(p_1\cdots p_l,\ldots,p_1\cdots p_l)}(\vec{0})\backslash\{\vec{0}\}.$$

We note that for every T all sums are finite. Hence by (2.22) – noticing that we have the same bound when we exclude zero – we have

$$\begin{aligned}\left|\sum_{\substack{\alpha\in A_{rp},\\ l_i|\alpha_i|\leq T}} \chi^{\alpha}_{\epsilon/\rho\sqrt{T}}\right| &\leq \sum_{l=0}^{\infty}\sum_{p_1,\ldots,p_l}\frac{2^{-2g}\rho^{2g}T^g}{(p_1\cdots p_l)^{2g}\left|\epsilon_1\ldots\epsilon_{2g}\right|}\\ &= \frac{4^{-g}\rho^{2g}T^g}{\left|\epsilon_1\ldots\epsilon_{2g}\right|}\prod_p(1+p^{-2g}) = \frac{4^{-g}\rho^{2g}T^g}{\left|\epsilon_1\ldots\epsilon_{2g}\right|}\prod_p\frac{1-p^{-4g}}{1-p^{-2g}}\\ &= \frac{4^{-g}\rho^{2g}T^g\zeta(2g)}{\left|\epsilon_1\ldots\epsilon_{2g}\right|\zeta(4g)}.\end{aligned}$$

Hence A_{rp} has property [C] with any c,ι and Lemma 2.12 gives property [GC] for some $c>0$ and $\iota=1/16$. $\square$

We now turn to the last type of sets that we can prove have property [GC]: random sets. We set up just enough notation for our result to be intelligible. For further details we refer to [10].

Let $(\Omega,\mathscr{A},\mathscr{P})$ be the product (over $\alpha\in\mathbb{Z}^{2g}$) probability space of $(\Omega_\alpha,\mathscr{A}_\alpha,\mathscr{P}_\alpha)$, where Ω_α is the unit interval $[0,1]$, $\mathscr{A}_\alpha$ is the σ-field of Lebesgue measurable sets, and $\mathscr{P}_\alpha$ is the Lebesgue measure. For $\omega\in\Omega$ we define a set $A(\omega)\subseteq\mathbb{Z}^{2g}$ by

$$\alpha\in A(\omega) \text{ if and only if } \omega_\alpha\in[0,1/2[.$$

Let (W_α) be a sequence of independent random variables equidistributed on $[0,1]$ (a Steinhaus sequence). We say that a random set $\subset\mathbb{Z}^{2g}$ has property P, if $\{W\in\Omega|A(W)$ has property $P\}\subseteq\mathscr{A}$ and

$$\mathscr{P}(\{W\in\Omega|A(W)\text{ has property }P\})=1,$$

i.e. if a set whose coefficients are included or excluded at random will have property P with probability 1.

Theorem 2.14. *A random set has property [GC].*

Proof. Consider the following random trigonometric polynomial in $2g$ variables of degree at most $2gcT$

$$f_\alpha(t)=\sum_{|\alpha_i|\leq cT}\epsilon_\alpha(W)e^{2\pi i\langle\alpha,t\rangle},$$

where

$$\epsilon_\alpha(W)=\begin{cases}1, & \text{if } W_\alpha\in[0,1/2[,\\ -1, & \text{if } W_\alpha\in[1/2,1[\end{cases}$$

is the Rademacher sequence derived from W. This is a subnormal sequence [10, p. 67], and we may therefore apply [10, Theorem 3, p. 70] to conclude that for some absolute constant C

$$\mathscr{P}\left(\sup_t |f_\alpha(t)| \geq C\Big(2g \sum_{|\alpha_i|\leq cT} \log(2gcT)\Big)^{1/2}\right) \leq (2gcT)^{-2}e^{-2g}.$$

Hence by the Borel-Cantelli lemma ([10, p. 7])

$$\mathscr{P}\left(\sup_t |f_\alpha(t)| = O\left(\Big(\sum_{|\alpha_i|\leq cT} \log(2gcT)\Big)^{1/2}\right) \text{ as } T \to \infty\right) = 1$$

and hence for every $\varepsilon > 0$

$$\mathscr{P}\left(\sup_t \left|\sum_{|\alpha_i|\leq cT} \epsilon_\alpha(W) e^{2\pi i\langle \alpha, t\rangle}\right| = O(T^{g+\varepsilon}) \text{ as } T \to \infty\right) = 1. \tag{2.23}$$

Since for a random set $A(W)$

$$\sum_{\substack{\alpha\in A(W)\\ |\alpha_i|\leq cT}} \chi^\alpha_{\epsilon/\rho\sqrt{T}} = \sum_{\substack{\alpha\in\mathbb{Z}^{2g}\\ |\alpha_i|\leq cT}} \left(\frac{\epsilon_\alpha(W)}{2} + \frac{1}{2}\right) e^{2\pi i\left\langle \alpha, \epsilon/\rho\sqrt{T}\right\rangle}, \tag{2.24}$$

we conclude from (2.23), Lemma 2.5 (iii) that, with probability 1, the set $A(W)$ has property [GC]. □

We note that from (2.23) and (2.24) with $t = 0$ it follows that a random set has natural density equal to 1/2 (in particular the natural density exists). Hence Theorem 1.1 follows from Theorem 2.14 and Theorem 2.8.

3. Densities in free groups

Let $\Gamma = F(A_1, \ldots, A_k)$, $k \geq 2$ be the free group on k generators and set $q = 2k-1$. We consider the set Γ_c of cyclically reduced words in Γ, i.e. words such that the first letter multiplied with the last letter is not the identity. These words γ can be counted according to their word length $\mathrm{wl}\,(\gamma)$ and one finds (see [16, 19]) that the number of cyclically reduced words of word length m equals

$$\#\{\gamma\in\Gamma_c |\, \mathrm{wl}\,(\gamma) = m\} = q^m + 1 + (k-1)\,(1 + (-1)^m)\,. \tag{3.1}$$

We note that an element $\gamma \in \Gamma$ and the corresponding cyclically reduced element has the same value for any discrete logarithm and therefore for the vector of discrete logarithms $\Phi(g)$, as in (1.8).

We want to consider conjugacy classes of Γ of length $l(\{\gamma\}) \leq m$ instead of cyclically reduced words of word length less than m. The length of a conjugacy class is the cyclically reduced length of any representative of the conjugacy class, which is also the minimal length of the representatives of the conjugacy class. There is a m to 1 correspondence between the set of cyclically reduced words of word length m and the set of conjugacy classes of Γ of length m, taking a cyclically reduced word to its conjugacy class in Γ. The map Φ factorizes though this correspondence and it follows that for any set $B \subset \mathbb{Z}^k$

(3.2)

$$\#\{\gamma\in\Gamma_c |\, \mathrm{wl}\,(\gamma) = m, \Phi(\gamma)\in B\} = m\#\{\{\gamma\}\in\{\Gamma\} | l(\{\gamma\}) = m, \Phi(\{\gamma\})\in B\}.$$

Using partial summation we find

$$\begin{aligned}\#\{\{\gamma\}\in\{\Gamma\}\,|\,l(\{\gamma\})\leq m,\Phi(\{\gamma\})\in B\} \;&=\; m^{-1}\#\{\gamma\in\Gamma_c|\,\mathrm{wl}\,(\gamma)\leq m,\Phi(\gamma)\in B\}\\ &\quad+\int_1^m t^{-2}\#\{\gamma\in\Gamma_c|\,\mathrm{wl}\,(\gamma)\leq t,\Phi(\gamma)\in B\}dt.\end{aligned}\tag{3.3}$$

We use (3.1) to bound the integral by

$$\int_1^m t^{-2}\frac{q^{t+1}}{q-1}\,dt,$$

which is easily seen to be $O(m^{-2}q^m)$ by partial integration.
Hence

$$\begin{aligned}&\#\{\{\gamma\}\in\{\Gamma\}|l(\{\gamma\})\leq m,\Phi(\{\gamma\})\in B\}\\ &\qquad=m^{-1}\#\{\gamma\in\Gamma_c|\,\mathrm{wl}\,(\gamma)\leq m,\Phi(\gamma)\in B\}+O(m^{-2}q^m).\end{aligned}\tag{3.4}$$

We can, therefore, freely move back and forth between counting problems for conjugacy classes and counting problems for cyclically reduced words. Using (3.4) and (3.1) we get

$$\Pi(m)\sim\frac{q}{q-1}\frac{q^m}{m}\qquad\text{as } m\to\infty.$$

3.1. **A graph identity.** We can now explain how to estimate counting functions related to cyclically reduced words using spectral perturbations of the adjacency operator of a graph: for any unitary character χ on Γ we have the following identity (see [16])

$$\sum_{m=1}^\infty n_{\Gamma,\chi}(m)u^m=\frac{2(k-1)u^2}{(1-u^2)}+\frac{uA(\Gamma,\chi)-2(2k-1)u^2}{1-uA(\Gamma,\chi)+(2k-1)u^2},\tag{3.5}$$

where

$$n_{\Gamma,\chi}(m)=\sum_{\substack{\gamma\in\Gamma_c\\ \mathrm{wl}(\gamma)=m}}\chi(\gamma)\tag{3.6}$$

and

$$A(\Gamma,\chi)=\sum_{i=1}^k(\chi(A_i)+\chi(A_i)^{-1})\tag{3.7}$$

is the twisted adjacency operator of the graph to the right of Figure 2. The power series (3.5) is convergent up to the first pole of the left-hand side.

This identity is the main analytic tool we use to prove Theorems 1.6. It is a particular case of the Ihara trace formula which relates geometric data (lengths of paths) to spectral data (eigenvalues of the adjacency operator) for a finite regular graph. In [16] we showed how one can interpret additive characters on free groups as multiplicative characters on a singleton graph and it is this identification that gives (3.5). We refer to [16] for further details. We have (assuming for a moment that $\lambda_1\neq\lambda_2$)

$$\frac{1}{1-uA(\Gamma,\chi)+(2k-1)u^2}=\frac{1}{2k-1}\frac{1}{\lambda_1-\lambda_2}\left(\frac{1}{u-\lambda_1}-\frac{1}{u-\lambda_2}\right),$$

where $\lambda_i=\lambda_i(\chi,\Gamma)$ are the roots of $1-uA(\Gamma,\chi)+(2k-1)u^2$. We note that

$$\lambda_1+\lambda_2=A(\Gamma,\chi)/(2k-1),\quad\lambda_1\lambda_2=1/(2k-1).\tag{3.8}$$

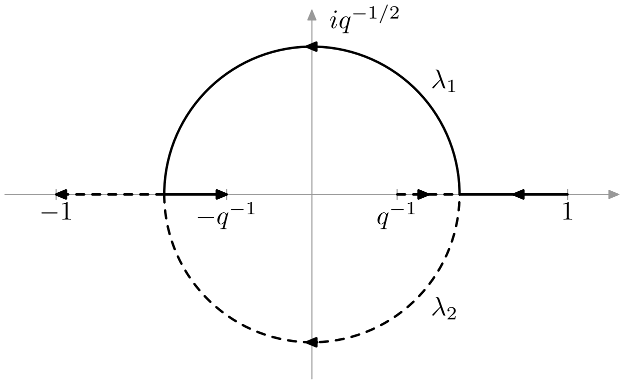

FIGURE 1. The trajectories of the eigenvalues as A moves away from $2k$.

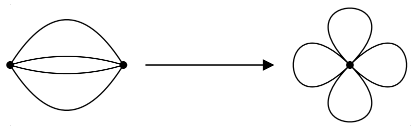

FIGURE 2. The graph and its two-cover, $n = 4$.

We have

$$\begin{aligned} \lambda_1 &= \frac{A(\Gamma,\chi) + \sqrt{A(\Gamma,\chi)^2 - 4(2k-1)}}{2(2k-1)}, \\ \lambda_2 &= \frac{A(\Gamma,\chi) - \sqrt{A(\Gamma,\chi)^2 - 4(2k-1)}}{2(2k-1)}. \end{aligned} \tag{3.9}$$

Remark 3.1. We note that if $A(\Gamma,\chi)^2 - 4(2k-1) > 0$ and $A > 0$ then λ_1 is a strictly increasing function of $A(\Gamma,\chi)$, while λ_2 is a strictly decreasing function of $A(\Gamma,\chi)$. As $A(\Gamma,\chi)$ varies in $[2\sqrt{2k-1}, 2k]$ and attains its maximal value $2k$ we have

$$\frac{1}{\sqrt{2k-1}} \leq \lambda_1 \leq 1, \quad \frac{1}{\sqrt{2k-1}} \geq \lambda_2 \geq \frac{1}{(2k-1)},$$

with the numbers on the right achieved for the trivial character.

When $|A(\Gamma,\chi)| < 2\sqrt{2k-1}$ we have $|\lambda_1| = |\lambda_2| = 1/\sqrt{2k-1}$.

When $A(\Gamma,\chi)^2 - 4(2k-1) > 0$ and $A < 0$ then λ_2 is a strictly increasing function of $A(\Gamma,\chi)$, while λ_1 is a strictly decreasing function of $A(\Gamma,\chi)$. As $A(\Gamma,\chi)$ varies in $[-2k, -2\sqrt{2k-1}]$ and attains its minimal value $-2k$ we have

$$-\frac{1}{2k-1} \geq \lambda_1 \geq -\frac{1}{\sqrt{2k-1}}, \quad -1 \leq \lambda_2 \leq -\frac{1}{\sqrt{2k-1}},$$

with the numbers on the left achieved at the infimum of $A = -2k$.

Remark 3.2. The λ_j, $j = 1, 2$ are not the eigenvalues of the Laplace operator $\Delta(\chi) = A(\chi) - (q+1)I$. The relation is as follows: The resolvent of Δ is $(\Delta(\chi) - \mu)^{-1}$ and has poles at the eigenvalues of $\Delta(\chi)$. Simple algebra shows that $1 - uA(\chi) + qu^2 = -u(\Delta - (u-1)(qu-1)/u)$. When $\chi = 1$, we have $A = q+1$, $\Delta = 0$ and the corresponding u's in the resolvent are 1 and $1/q$. When $\chi = -1$ (i.e. $\chi(A_i) = -1$), we have $A = -(q+1)$, $\Delta = -2(q+1)$ and the corresponding u's are -1 and $-1/q$. Now recall that for $\chi = 1$ and a general finite graph the eigenvalue $q+1$ of A occurs and the eigenvalue $-(q+1)$ of A occurs iff the graph is bipartite, see [21, p. 67]. In our case the eigenvalue $-(q+1)$ occurs when $\chi = -1$. In this case the character has order 2 and gives a double covering of the graph in Fig. 2, which is bipartite. It consists of two vertices, joined by $2k$ edges, see Fig.2. Its spectrum contains $\mathrm{Spec}(A(\chi))$ for $\chi = -1$. The adjacency operator is

$$\begin{pmatrix} 0 & 2k \\ 2k & 0 \end{pmatrix}$$

with eigenvectors $(1,1)$ and $(1,-1)$ with eigenvalues $2k$, $-2k$ respectively.

3.2. **Detecting words with a given homology.** We now explain how to use the orthogonality relations to count words with a given homology. Using

$$\frac{2(k-1)u^2}{(1-u^2)} = (k-1)\sum_{m=1}^{\infty}\left(1+(-1)^k\right)u^k,$$

and

$$\frac{1}{u-\lambda} = -\sum_{m=0}^{\infty}\lambda^{-(m+1)}u^m,$$

we find from (3.5) the following generalization of (3.1)

$$n_{\Gamma,\chi}(m) = \lambda_2^{-m} + \lambda_1^{-m} + (k-1)\left(1+(-1)^{m+1}\right). \tag{3.10}$$

The same expression holds when $\lambda_1 = \lambda_2$, which can be seen by plugging $A = 2\sqrt{2k-1}$ into (3.5).

Consider now $\Phi : \Gamma_c \to \mathbb{Z}^k$ with $\Phi(\gamma) = (\log_1(\gamma), \dots, \log_k(\gamma))$. For $\beta \in \mathbb{Z}^k$ we let

$$n_{\Gamma,\beta}(m) = \#\{\gamma \in \Gamma_c |\, \mathrm{wl}\,(\gamma) = m, \Phi(\gamma) = \beta\}. \tag{3.11}$$

Consider the unitary character

$$\chi_\epsilon(\gamma) = e^{2\pi i\langle \Phi(\gamma), \epsilon\rangle},$$

where $\epsilon \in \mathbb{R}^k/\mathbb{Z}^k$ and $\langle\cdot,\cdot\rangle$ is the inner product between $\mathbb{Z}^k$ and its dual $\mathbb{R}^k/\mathbb{Z}^k$. For $\beta \in \mathbb{Z}^k$ we define the unitary character

$$\chi_\epsilon^\beta = e^{2\pi i\langle \beta, \epsilon\rangle}.$$

Then by the orthogonality relation for abelian groups we have:

$$\int_{\mathbb{R}^k/\mathbb{Z}^k} \chi_\epsilon(\gamma)\overline{\chi_\epsilon^\beta}\,d\epsilon = \delta_{\Phi(\gamma)=\beta}. \tag{3.12}$$

It follows that

$$n_{\Gamma,\beta}(m) = \int_{\mathbb{R}^k/\mathbb{Z}^k} n_{\Gamma,\epsilon}(m)\overline{\chi_\epsilon^\beta}\,d\epsilon, \tag{3.13}$$

where we use the notation $n_{\Gamma,\epsilon} := n_{\Gamma,\chi_\epsilon}$. We shall also write $A(\epsilon) := A(\Gamma, \chi_\epsilon)$, $\lambda_i(\epsilon) := \lambda_i(\chi_\epsilon)$, and $q_i(\epsilon) := \lambda_i(\epsilon)^{-1}$, and $q = q_2(0) = (2k-1)$. The equations (3.10) and (3.8) give

$$\frac{n_{\Gamma,\beta}(m)}{q^m} = \int_{\mathbb{R}^k/\mathbb{Z}^k} (\lambda_1(\epsilon)^m + \lambda_2(\epsilon)^m)\overline{\chi_\epsilon^\beta}\,d\epsilon + O(q^{-m}). \tag{3.14}$$

It is easy to check that

$$A(\epsilon) = 2\sum_{j=1}^{k} \cos(2\pi\epsilon_j).$$

Clearly there is a symmetry $A(\epsilon + (1/2,\dots,1/2)) = -A(\epsilon)$ from which we conclude that

$$\lambda_2(\epsilon + (1/2,\dots,1/2)) = -\lambda_1(\epsilon). \tag{3.15}$$

Using this and $\chi^{\beta}_{\epsilon+(1/2,\dots,1/2)} = (-1)^{\beta_1+\dots+\beta_k}\chi^{\beta}_{\epsilon}$ we see that

$$\frac{n_{\Gamma,\beta}(m)}{q^m} = (1 + (-1)^{m+\beta_1+\dots+\beta_k}) \int_{\mathbb{R}^k/\mathbb{Z}^k} \lambda_1(\epsilon)^m \overline{\chi^{\beta}_{\epsilon}} d\epsilon + O(q^{-m}). \tag{3.16}$$

We now have the following analogue of Lemma 2.5:

Lemma 3.3. *Let*

$$\rho^2 = \frac{4\pi^2}{k-1}.$$

(i) *For every $\epsilon_0 \in \mathbb{R}^k$*

$$\lambda_1(\epsilon_0/\rho\sqrt{m})^m \to e^{-\langle\epsilon_0,\epsilon_0\rangle/2}$$

as $m \to \infty$.

(ii) *There exists $\delta > 0$ such that for all $\|\epsilon\| < \delta\rho\sqrt{m}$.*

$$\left|\lambda_1(\epsilon/\rho\sqrt{m})^m - e^{-\langle\epsilon,\epsilon\rangle/2}\right| \le 2e^{-\langle\epsilon,\epsilon\rangle/4}.$$

(iii) *There exist constants $\delta, C > 0$ such that for all $m \in \mathbb{N}$, $\|\epsilon\| < \delta m^{1/16}$,*

$$\left|\lambda_1(\epsilon/\rho\sqrt{m})^m - e^{-\langle\epsilon,\epsilon\rangle/2}\right| \le C\frac{1}{m^{1/2}}.$$

Proof. The function $A(\epsilon)$ is clearly even in each ϵ_j and symmetric in the ϵ_j's. It follows that $\lambda_1(\epsilon)$ is even (compare (3.9)). Therefore all odd derivatives of the smooth function $\lambda_1(\epsilon)$ vanish at zero, as do all mixed second derivatives. Also $\partial_{ii}\lambda_1|_{\epsilon=0} = -\rho^2$ as can easily be checked by differentiating (3.9). Hence Taylor's theorem implies that Proposition 4.1 can be used and setting $D = \rho^2 I$, $\nu = 2$, $b = 1/2$ and $r = \rho$ gives the desired conclusion. □

3.2.1. *Elements with a given word length.* We let $I(v) = [-v/2, v/2]^k$. Using (3.16) and performing the change of variables $\epsilon \to \epsilon/\rho\sqrt{m}$ in (3.16) we find that

$$\rho^k m^{k/2} \frac{n_{\Gamma,\alpha}(m)}{q^m} = s_{\beta,m} \int_{I(\rho\sqrt{m})} \overline{\chi^{\beta}_{\epsilon/\rho\sqrt{m}}} \lambda_1(\epsilon/\rho\sqrt{m})^m d\epsilon + O(q^{-m}m^{k/2}),$$

where $s_{\beta,m} = 1 + (-1)^{m+\beta_1+\cdots+\beta_k}$. Using the Fourier transform of the Gaussian density function

$$(2\pi)^{k/2} e^{-2\pi^2\langle\beta,\beta\rangle/\rho^2 m} = \int_{\mathbb{R}^k} \overline{\chi^{\beta}_{\epsilon/\rho\sqrt{m}}} e^{-\langle\epsilon,\epsilon\rangle/2} d\epsilon,$$

we can split the relevant integral into three parts to conclude that

$$\begin{aligned}
\rho^k m^{k/2}\frac{n_{\Gamma,\beta}(m)}{q^m} &- s_{\beta,m}(2\pi)^{k/2}e^{-2\pi^2\langle\beta,\beta\rangle/\rho^2 m}\\
=&s_{\beta,m}\int_{B(\delta\rho\sqrt{m})}\overline{\chi^{\beta}_{\epsilon/\rho\sqrt{m}}}\left(\lambda_1(\epsilon/\rho\sqrt{m})^m - e^{-\langle\epsilon,\epsilon\rangle/2}\right)d\epsilon\\
&+ s_{\beta,m}\int_{I(\rho\sqrt{m})\backslash B(\delta\rho\sqrt{m})}\overline{\chi^{\beta}_{\epsilon/\rho\sqrt{m}}}\lambda_1(\epsilon/\rho\sqrt{m})^m d\epsilon\\
&- s_{\beta,m}\int_{\mathbb{R}^k\backslash B(\delta\rho\sqrt{m})}\overline{\chi^{\beta}_{\epsilon/\rho\sqrt{m}}}e^{-\langle\epsilon,\epsilon\rangle/2}d\epsilon + O(q^{-m/2}m^{k/2})\\
=&s_{\beta,m}(A_1(m,\beta)+A_2(m,\beta)+A_3(m,\beta)) + O(q^{-m}m^{k/2}).
\end{aligned}\tag{3.17}$$

Lemma 3.4. *There exists a $d>0$, depending only on k and δ, such that*

$$\begin{aligned}
A_2(m,\beta) &= O(q^{-dm})\\
A_3(m,\beta) &= O(q^{-dm}).
\end{aligned}$$

The implied constants are independent of β.

Proof. For ϵ bounded away from the identity in $\mathbb{R}^k/\mathbb{Z}^k$, $\lambda_1(\epsilon)$ is bounded away from 1, which is the maximum of λ_1. Hence there exists $d_1>0$ (depending on δ) such that $\lambda_1(\epsilon) < q^{-d_1}$ for $\epsilon \in I(1)\backslash B(\delta)$. We, therefore, have $|A_2(m,\beta)| \le Cq^{-d_1 m}m^{k/2}$. Choosing $d = d_1/2$ does the job.

Since $-\langle\epsilon,\epsilon\rangle/2 + (\delta\rho\sqrt{m})^2/4 \le -\langle\epsilon,\epsilon\rangle/4$ when $\epsilon\in B(\delta\rho\sqrt{m})^c$, we conclude

$$\left|e^{\rho^2\delta^2 m/4}A_3(m,\beta)\right| \le 4\int_{\mathbb{R}^k\backslash B(\delta\rho\sqrt{m})} e^{-\langle\epsilon,\epsilon\rangle/4} \le C,$$

from which the result easily follows. □

We have the following lemma.

Lemma 3.5. *There exist $d>0$ which depends only on k such that*

$$\rho^k m^{k/2}\frac{n_{\Gamma,\beta}(m)}{q^m} - s_{\beta,m}(2\pi)^{k/2}e^{-\langle\beta,\beta\rangle(k-1)/2m} = s_{\beta,m}A_1(m,\beta) + O(q^{-dm}),$$

where the implied constants is independent on β.

Proof. This follows directly from (3.17) and Lemma 3.4. □

3.2.2. *Elements with word length less than a given length.* We now let

$$\begin{aligned}
N_\Gamma(m) &= \#\{\gamma\in\Gamma_c|\,\mathrm{wl}\,(\gamma)\le m\},\\
N_{\Gamma,\beta}(m) &= \#\{\gamma\in\Gamma_c|\,\mathrm{wl}\,(\gamma)\le m, \Phi(\gamma)=\beta\}.
\end{aligned}$$

We aim at proving a result for $N_{\Gamma,\beta}(m)$ analogous to Lemma 3.5. We note that from (3.1) we get

$$N_\Gamma(m) = \frac{q^{m+1}}{q-1} + O(m). \tag{3.18}$$

We shall write $\beta\sim m$ if $\beta\in\mathbb{Z}^k$ and $m\in\mathbb{N}$ has the same parity, i.e. if $m+\beta_1+\ldots+\beta_k$ is even. Using (3.16) we find that

$$N_{\Gamma,\beta}(m) = 2\int_{\mathbb{R}^k/\mathbb{Z}^k}\sum_{\substack{n\le m\\ n\sim\beta}} q^n\lambda_1(\epsilon)^n\overline{\chi^{\beta}_{\epsilon}}d\epsilon + O(m). \tag{3.19}$$

Writing

$$\delta_\beta = \begin{cases} 1, & \text{if } \beta_1 + \ldots + \beta_k \text{ is odd,} \\ 0, & \text{otherwise,} \end{cases}$$

we have

$$\sum_{\substack{n \le m \\ n \sim \beta}} q^n \lambda_1(\epsilon)^n = \frac{(q\lambda_1(\epsilon))^{2\left[\frac{m-\delta_\beta}{2}\right]+2+\delta_\beta} - (q\lambda_1(\epsilon))^{2-\delta_\beta}}{(q\lambda_1(\epsilon))^2 - 1}. \tag{3.20}$$

Inserting this in (3.19) we find that

$$N_{\Gamma,\beta}(m) = 2\frac{q^{2\left[\frac{m-\delta_\beta}{2}\right]+2+\delta_\beta}}{q^2-1} \int_{\mathbb{R}^k/\mathbb{Z}^k} \overline{\chi_\epsilon^\beta} g_\beta(\epsilon, m) \lambda_1(\epsilon)^m d\epsilon + O(m),$$

where

$$g_\beta(\epsilon, m) = \frac{q^2-1}{(q\lambda_1(\epsilon))^2 - 1} \lambda_1(\epsilon)^{2\left[\frac{m-\delta_\beta}{2}\right]+2+\delta_\beta - m}.$$

Clearly $g_\beta(\epsilon, m)$ is uniformly bounded in $\mathbb{R}^k/\mathbb{Z}^k$, independently of β, it satisfies $g_\beta(0, m) = 1$, and close to zero we have $g_\beta(\epsilon, m) - 1 = O(\langle \epsilon, \epsilon \rangle)$, where the implied constant does not depend on m or β.

We simplify by taking average over two successive m. It is easy to check that

$$\frac{1}{2}\left(\frac{N_{\Gamma,\beta}(m)}{q^{m+1}/(q-1)} + \frac{N_{\Gamma,\beta}(m+1)}{q^{m+2}/(q-1)}\right) = \int_{\mathbb{R}^k/\mathbb{Z}^k} \overline{\chi_\epsilon^\beta} h_\beta(\epsilon, m) \lambda_1(\epsilon)^m d\epsilon + O(q^{-m} m), \tag{3.21}$$

where

$$h_\beta(\epsilon, m) = \begin{cases} \dfrac{q g_\beta(m,\epsilon) + \lambda_1(\epsilon) g_\beta(m+1,\epsilon)}{q+1}, & \text{if } m \sim \beta, \\ \dfrac{g_\beta(m,\epsilon) + q\lambda_1(\epsilon) g_\beta(m+1,\epsilon)}{q+1}, & \text{otherwise.} \end{cases}$$

The function $h_\beta(m, \epsilon)$ inherits its properties from those of $g_\beta(m, \epsilon)$: It is uniformly bounded in $\mathbb{R}^k/\mathbb{Z}^k$ independent of β, it satisfies $h_\beta(0, m) = 1$, and close to zero $h_\beta(\epsilon, m) - 1 = O(\langle \epsilon, \epsilon \rangle)$ where the implied constant does not depend on m or β.

We now use the same techniques that lead to Lemma 3.5. We start by doing the change of variables $\epsilon \to \epsilon/\rho\sqrt{m}$ to get (up to an error $O(m^{k/2+1} q^{-m})$)

$$\rho^k m^{k/2} \frac{1}{2}\left(\frac{N_{\Gamma,\beta}(m)}{q^{m+1}/(q-1)} + \frac{N_{\Gamma,\beta}(m+1)}{q^{m+2}/(q-1)}\right) = \int_{I(\rho\sqrt{m})} \overline{\chi_{\epsilon/\rho\sqrt{m}}^\beta} h_\beta(\epsilon/\rho\sqrt{m}, m) \lambda_1(\epsilon/\rho\sqrt{m})^m d\epsilon.$$

In analogy with (3.17) we get

$$\begin{aligned} &\rho^k m^{k/2} \frac{1}{2}\left(\frac{N_{\Gamma,\beta}(m)}{q^{m+1}/(q-1)} + \frac{N_{\Gamma,\beta}(m+1)}{q^{m+2}/(q-1)}\right) - (2\pi)^{k/2} e^{-2\pi^2 \langle \beta, \beta \rangle / \rho^2 m} \\ &= \int_{B(\delta\rho\sqrt{m})} \overline{\chi_{\epsilon/\rho\sqrt{m}}^\beta} \left(h_\beta(\epsilon/\rho\sqrt{m}, m) \lambda_1(\epsilon/\rho\sqrt{m})^m - e^{-\langle \epsilon, \epsilon \rangle/2}\right) d\epsilon \\ &\quad + \int_{I(\rho\sqrt{m}) \setminus B(\delta\rho\sqrt{m})} \overline{\chi_{\epsilon/\rho\sqrt{m}}^\beta} h_\beta(\epsilon/\rho\sqrt{m}, m) \lambda_1(\epsilon/\rho\sqrt{m})^m d\epsilon \\ &\quad - \int_{\mathbb{R}^k \setminus B(\delta\rho\sqrt{m})} \overline{\chi_{\epsilon/\rho\sqrt{m}}^\beta} e^{-\langle \epsilon, \epsilon \rangle/2} d\epsilon + O(q^{-m/2} m^{k/2+1}) \\ &= B_1(m, \beta) + B_2(m, \beta) + B_3(m, \beta) + O(q^{-m} m^{k/2+1}). \end{aligned} \tag{3.22}$$

With this notation we have

Lemma 3.6. *There exist $d > 0$ which depends only on k such that*

$$\rho^k m^{k/2}\frac{1}{2}\left(\frac{N_{\Gamma,\beta}(m)}{q^{m+1}/(q-1)}+\frac{N_{\Gamma,\beta}(m+1)}{q^{m+2}/(q-1)}\right)-(2\pi)^{k/2}e^{-\langle\beta,\beta\rangle(k-1)/2m} = B_1(m,\beta)+O(q^{-dm}),$$

where the implied constant is independent on β.

Proof. Using that $h(\epsilon/\rho\sqrt{m})$ is uniformly bounded the proof of Lemma 3.4 can be copied almost word by word to prove $B_2(m,\beta), B_3(m,\beta) = O(q^{-dm})$. □

3.3. **A local limit theorem.** We can now state and prove a local limit theorem, i.e. a theorem that gives information (uniform in β) about the asymptotic probability for an element to satisfy $\Phi(\gamma) = \beta$. To be more precise we have the following theorem:

Theorem 3.7. *Let $\sigma^2 = (k-1)^{-1}$. Then*

$$\sup_{\beta\in\mathbb{Z}^k}\left|m^{k/2}\frac{n_{\Gamma,\beta}(m)}{q^m}-\frac{s_{m,\beta}}{(2\pi\sigma^2)^{k/2}}e^{-\langle\beta,\beta\rangle/2\sigma^2 m}\right| = o(1)$$

and

$$\sup_{\beta\in\mathbb{Z}^k}\left|\frac{m^{k/2}}{2}\left(\frac{N_{\Gamma,\beta}(m)}{q^{m+1}/(q-1)}+\frac{N_{\Gamma,\beta}(m+1)}{q^{m+2}/(q-1)}\right)-\frac{e^{-\langle\beta,\beta\rangle/2\sigma^2 m}}{(2\pi\sigma^2)^{k/2}}\right| = o(1).$$

Proof. We ignore the oscillation and possible cancellation due to χ_ϵ^β. Using

$$\sup_\beta |A_1(m,\beta)| \leq \int_{B(\delta\rho\sqrt{m})}\left|\lambda_1(\epsilon/\sigma\sqrt{m})^m - e^{-\langle\epsilon,\epsilon\rangle/2}\right| d\epsilon$$

the first claim follows from Lemma 3.5, Lemma 3.3 (i) and (ii) and the dominated convergence theorem.

By Lemma 3.3 (i) and the decay properties of $h_\beta(\epsilon, m)$ close to zero we have (using the triangle inequality)

$$\left|h_\beta(\epsilon/\rho\sqrt{m})\lambda_1(\epsilon/\rho\sqrt{m})^m - e^{-\langle\epsilon,\epsilon\rangle/2}\right| \leq C\frac{\|\epsilon\|^2}{\rho^2 m}e^{-\langle\epsilon,\epsilon\rangle/4}+\left|\lambda_1(\epsilon/\sigma\sqrt{m})^m - e^{-\langle\epsilon,\epsilon\rangle/2}\right|,$$

when $\|\epsilon\| < \delta\rho\sqrt{m}$. The right-hand-side is independent of β. Hence

$$\sup_\beta |B_1(m,\beta)| \leq \int_{B(\delta\rho\sqrt{m})}\left(C\frac{\|\epsilon\|^2}{\rho^2 m}e^{-\langle\epsilon,\epsilon\rangle/4}+\left|\lambda_1(\epsilon/\sigma\sqrt{m})^m - e^{-\langle\epsilon,\epsilon\rangle/2}\right|\right) d\epsilon.$$

The integrand on the right converges pointwise to zero by Lemma 3.3 (i). Using Lemma 3.3 (ii) we see that it can be bounded from above by $C'\|\epsilon\|^2 e^{-\langle\epsilon,\epsilon\rangle/4} + 2e^{-\langle\epsilon,\epsilon\rangle/4}$ which is integrable on $\mathbb{R}^k$. The bounded convergence theorem now gives $\sup_\beta |B_1(m,\beta)| \to 0$ and quoting Lemma 3.6 we conclude the theorem.

□

Remark 3.8. The statement in the Theorem 3.7 concerning $n_{\Gamma,\beta}(m)$ was also proved by R. Sharp [24, proposition 3]. A related but weaker result was proved by I. Rivin [19, Theorem 5.1]. We emphasize that these papers have a different value for σ^2.

This is due to an erroneous calculation in [19]. The left hand side of [19, Eq. (22)] should read

$$1-\frac{1}{2n(c+\sqrt{c^2-1})}\left(\frac{c}{k}+\frac{c^2}{(c^2-1)^{1/2}k}\right)\langle\theta,\theta\rangle+o\left(\frac{1}{n}\right).$$

Once this is corrected the values of the variances agree.

3.4. **Densities of discrete logarithms in a given set.** In this section we define a class of sets B for which we can study the density of discrete logarithms in B. We need to consider sets for which we can see cancellation in an exponential sum supported in B. To be precise:

Definition 3.9. A set $B\subset\mathbb{Z}^k$ is said to have property (C) if, for some $0\leq\varepsilon<k/2$ there exist constants $c,\iota>0$ such that for $m^{-k}<|\epsilon_1|,\dots,|\epsilon_k|<cm^\iota$:

$$\sum_{\substack{\beta\in B,\delta_a=\delta\\|\beta_i|\leq m}}\chi^{\beta}_{\epsilon/\rho\sqrt{m}}=O(f_\delta(\epsilon)m^{k/2+\varepsilon}),\qquad\delta=0,1,\tag{3.23}$$

where the implied constant may depend on ε, B, c, δ and k. If furthermore $f_\delta(\epsilon)$ satisfies

$$\int_{m^{-k}<|\epsilon_1|,\dots,|\epsilon_k|<cm^\iota}|f_\delta(\epsilon)|\left|\lambda_1(\epsilon/\rho\sqrt{m})^m-e^{\langle\epsilon,\epsilon\rangle/2}\right|d\epsilon=o(m^{-\varepsilon})$$

and

$$\int_{m^{-k}<|\epsilon_1|,\dots,|\epsilon_k|<cm^\iota}|f_\delta(\epsilon)|\left\|\epsilon\right\|^2\left|\lambda_1(\epsilon/\rho\sqrt{m})^m\right|d\epsilon=o(m^{1-\varepsilon})$$

as $m\to\infty$ the set B is said to have property (GC).

Obviously Definition 3.9 is analogous to Definition 2.6, as it also quantifies a certain form of cancellation in an exponential sum. The sets with Property (C) resp. Property (GC) also form an algebra. In fact Remark 2.7 applies here also.

Theorem 3.10. *Let $\sigma^2=(k-1)^{-1}$. Assume that $B\subset\mathbb{Z}^k$ has property (GC). Then*

$$\sum_{\substack{\beta\in B\\|\beta_i|\leq m}}\left(\frac{m^{k/2}}{2}\left(\frac{N_{\Gamma,\beta}(m)}{q^{m+1}/(q-1)}+\frac{N_{\Gamma,\beta}(m+1)}{q^{m+2}/(q-1)}\right)-\frac{1}{(2\pi\sigma^2)^{k/2}}e^{-\langle\beta,\beta\rangle/2\sigma^2m}\right)=o(m^{k/2}).$$

Before proving it we state and prove a corollary which is the main reason why property (GC) is interesting:

Corollary 3.11. *Assume that $B\subset\mathbb{Z}^k$ has property (GC) and assume that B has natural density $d(B)$. Then*

$$\frac{1}{2}\left(\frac{N_{\Gamma,B}(m)}{N_\Gamma(m)}+\frac{N_{\Gamma,B}(m+1)}{N_\Gamma(m+1)}\right)\to d(B)$$

as $m\to\infty$.

Proof. Theorem 3.10, (3.18) and Corollary 4.5. We also notice that $|\beta_i|\leq m$ for cyclically reduced words of length m, as all discrete logarithms are less than the length. So

$$N_{\Gamma,B}(m)=\sum_{\substack{\beta\in B\\|\beta_i|\leq m}}N_{\Gamma,\beta}(m)$$

and

$$\frac{1}{2}\left(\frac{N_{\Gamma,B}(m)}{N_\Gamma(m)}+\frac{N_{\Gamma,B}(m+1)}{N_\Gamma(m+1)}\right)-\sum_{\substack{\beta\in B\\|\beta_i|\leq m}}\frac{1}{2}\left(\frac{N_{\Gamma,\beta}(m)}{N_\Gamma(m)}+\frac{N_{\Gamma,\beta}(m+1)}{N_\Gamma(m+1)}\right)\to 0,$$

as $m\to\infty$, because it is bounded by

$$\sum_{\|\beta\|=m+1}\frac{n_{\Gamma,\beta}(m+1)}{N_\Gamma(m+1)},$$

which tends to 0 by Theorem 3.7. □

Proof of Theorem 3.10: Using the definition of Property (GC) we may safely assume $\iota\leq 1/2$. Quoting Lemma 3.6 we see that the theorem would follow from

$$\sum_{\substack{\beta\in B\\|\beta_i|\leq m}}B_1(m,\beta)+O(q^{-dm}m^k)=o(m^{k/2}).$$

Hence we need the following estimate:

$$\sum_{\substack{\beta\in B\\|\beta_i|\leq m}}\int_{B(\delta\rho\sqrt{m})}\overline{\chi^{\beta}_{\epsilon/\rho\sqrt{m}}}\left(h_\beta(\epsilon/\rho\sqrt{m},m)\lambda_1(\epsilon/\rho\sqrt{m})^m-e^{-\langle\epsilon,\epsilon\rangle/2}\right)d\epsilon=o(m^{k/2}).$$

After this point the proof is, mutatis mutandis, a repetition of the proof of Lemma 2.9. The only new issue is that we need to split the sum into two sums, according to the value of δ_β. We shall not repeat the details. □

Theorem 1.6 follows now from Corollary 3.11, Corollary 4.5, the discussion about partial summation leading up to Section 3.1 and the following result:

Theorem 3.12. *The following sets have Property (GC):*

(i) *Random sets.*
(ii) *Finite sets.*
(iii) *Sets whose coordinates are arithmetic progressions.*
(iv) *The set k-tuples whose coordinates are coprime integers.*

Proof. The proof of these claims is, mutatis mutandis, identical to the proofs of Proposition 2.11, Theorem 2.13 and Theorem 2.14. Again the only real difference is that we have to take into account splitting according to the value of δ_β. The only place where this is non-trivial is in the case of arithmetic progressions. Here we notice that the set $\{\gamma\in B_l(\alpha)|\delta_\beta=1\}$ (resp. $\delta_\beta=0$) is the disjoint union of the 2^{k-1} sets $\{\beta\in B_{l_1,\ldots,l_k}(\nu)|\nu_i\equiv\eta_i\,(\mathrm{mod}\ 2)\}$ with $\eta_i\in\{0,1\}$ and $\eta_1+\ldots+\eta_k$ odd (resp. even). But these sets are all either empty or sets of progressions with moduli l_i (when $(2,l_i)\neq 1$) or $2l_i$ (when $(2,l_i)=1$) using the chinese remainder theorem. Using this observation the rest is straightforward, and we omit the details. □

3.5. **A more direct proof for arithmetic progressions.** In this section we prove a slightly more precise version of Theorem 1.6 (ii).

Theorem 3.13. *Let*

$$N_{\Gamma,a_1,\ldots,a_k}(m)=\#\left\{\gamma\in\Gamma_c|\,\mathrm{wl}(\gamma)\leq m,\log_i(\gamma)\equiv a_i\,(\mathrm{mod}\ l_i)\,,\,i=1,\ldots,k\right\}$$

(a) If $2 \not| (l_1, l_2, \ldots, l_k)$ *we have*

$$\frac{N_{\Gamma,a_1,\ldots a_k}(m)}{\#\{\gamma \in \Gamma_c | \operatorname{wl}(\gamma) \leq m\}} \to \frac{1}{l_1 l_2 \cdots l_k}$$

as $m \to \infty$.

(b) If the l_j, $j = 1, \ldots, k$ *are all even, then*

$$\frac{1}{2}\left(\frac{N_{\Gamma,a_1,\ldots,a_k}(m)}{\#\{\gamma \in \Gamma_c | \operatorname{wl}(\gamma) \leq m\}} + \frac{N_{\Gamma,a_1,\ldots,a_k}(m+1)}{\#\{\gamma \in \Gamma_c | \operatorname{wl}(\gamma) \leq m+1\}}\right) \to \frac{1}{l_1 l_2 \cdots l_k}$$

as $m \to \infty$.

For notational simplicity we restrict ourselves to the case $k = 2$. The generalization to $k > 2$ is straightforward. Consider the abelian group $\mathbb{Z}/l_j\mathbb{Z}$. Consider the set of additive unitary characters on $\mathbb{Z}/l_j\mathbb{Z}$. These are parametrized by $g \in \mathbb{Z}/l_j\mathbb{Z}$ writing

$$\chi_{g,l_j}(a) = \exp\left(\frac{2\pi i g a}{l_j}\right).$$

The orthogonality relation for representations of finite groups (which in this simple example is easy to verify directly) gives

$$\frac{1}{l_j} \sum_{g \in \mathbb{Z}/l_j\mathbb{Z}} \chi_{g,l_j}(a)\overline{\chi_{g,l_j}(a_j)} = \begin{cases} 1, & \text{if } a \equiv a_j \pmod{l_j} \\ 0, & \text{otherwise.} \end{cases} \tag{3.24}$$

Putting $a = \log_1(\gamma)$ enables us to see - using characters - if $\log_1(\gamma)$ lies in a specific arithmetic progression. Multiplying two such identities (or using the orthogonality relation for $\mathbb{Z}/l_1\mathbb{Z} \times \mathbb{Z}/l_2\mathbb{Z}$) we find

(3.25)

$$\frac{1}{l_1 l_2} \sum_{\substack{g \in \mathbb{Z}/l_1\mathbb{Z} \\ g' \in \mathbb{Z}/l_2\mathbb{Z}}} \overline{\chi_{g,l_1}(a_1)\chi_{g',l_2}(a_2)}\chi_{g,g',l_1,l_2}(\gamma) = \begin{cases} 1, & \text{if } \log_j(\gamma) \equiv a_j \pmod{l_j}, j = 1, 2 \\ 0, & \text{otherwise.} \end{cases}$$

Here

$$\begin{aligned} \chi_{g,g',l_1,l_2}(\gamma) &= \chi_{g,l_1}(\log_1(\gamma))\chi_{g',l_2}(\log_2(\gamma))) \\ &= \exp\left(2\pi i \left(\frac{g \log_1(\gamma)}{l_1} + \frac{g' \log_2(\gamma)}{l_2}\right)\right), \end{aligned}$$

which is a unitary character on Γ. We note that

$$A(\Gamma, \chi_{g,g',l_1.l_2}) = 2\cos\left(\frac{2\pi g}{l_1}\right) + 2\cos\left(\frac{2\pi g'}{l_2}\right),$$

which is clearly less than or equal to $2k$. We sum over $\operatorname{wl}(\gamma) \leq m$ in (3.10) to get

$$\sum_{\substack{\gamma \in \Gamma_c \\ \operatorname{wl}(\gamma) \leq m}} \chi(\gamma) = \frac{\lambda_2^{-(m+1)} - \lambda_2^{-1}}{\lambda_2^{-1} - 1} + \frac{\lambda_1^{-(m+1)} - \lambda_1^{-1}}{\lambda_1^{-1} - 1} + (k-1)\left(m - \left(1 + (-1)^{m+1}\right)/2\right).$$

As $m \to \infty$ we have

$$\sum_{\substack{\gamma \in \Gamma_c \\ \operatorname{wl}(\gamma) \leq m}} \chi(\gamma) = \frac{\lambda_1^{-(m+1)}}{\lambda_1^{-1} - 1} + \frac{\lambda_2^{-(m+1)}}{\lambda_2^{-1} - 1} + O(m), \tag{3.26}$$

as long as 1 is not an eigenvalue. By Remark 3.1, when $\chi^2 \neq 1$,

$$\lim_{m\to\infty} \lambda_j^{-m}/(2k-1)^m = 0.$$

We now distinguish two cases:

(a) The only character with $\chi^2 = 1$ is the trivial character 1. We conclude from (3.25) that

$$\frac{\#\left\{\gamma \in \Gamma_c \,\middle|\, \mathrm{wl}\,(\gamma) \leq m,\ \begin{matrix} \log_1(\gamma) \equiv a_1 \,(\mathrm{mod}\ l_1) \\ \log_2(\gamma) \equiv a_2 \,(\mathrm{mod}\ l_2) \end{matrix}\right\}}{\#\{\gamma \in \Gamma_c |\, \mathrm{wl}\,(\gamma) \leq m\}} \to \frac{1}{l_1 l_2} \tag{3.27}$$

as $m \to \infty$.

(b) There exist another real character χ. This happens if both l_j are even and $g = l_1/2$, $g' = l_2/2$. In particular

$$\chi_{g,g',l_1,l_2}(a_1,a_2) = e^{\pi i(a_1+a_2)} = \begin{cases} 1, & \text{if } a_1 + a_2 \text{ is even,} \\ -1, & \text{if } a_1 + a_2 \text{ is odd.} \end{cases}$$

In this case we sum the contribution from the real characters and recall that from (3.7) and Remark 3.1 we have that the second real character gives eigenvalues $-1/(2k-1)$ and -1. Using (3.6) we get

$$n_{\Gamma,1}(m) = (2k-1)^m + 1^m + O(1), \quad n_{\Gamma,\chi} = (-(2k-1))^m + (-1)^m + O(1).$$

Using (3.11) and (3.25) we get

$$n_{\Gamma,a_1,a_2}(m) = \frac{1}{l_1 l_2}(2k-1)^m \left(1 + (-1)^m \overline{\chi(a_1,a_2)}\right) + O(d^m),$$

where $d = \sup(|\lambda_1|^{-1}, |\lambda_2|^{-1}) < q$ for the nonreal characters. We sum for $m = 1,\dots,l$. Depending of the value of $\chi(a_1,a_2)$ we sum either over the odd or the even exponents of $(2k-1)^j$. For instance, assuming that $\chi(a_1,a_2) = 1$, we get for $l = 2s$

$$N_{\Gamma,a_1,a_2}(l) = \frac{2}{l_1 l_2} \sum_{m=2m' \leq 2s} q^m + O(d^l) = \frac{2}{l_1 l_2} q^2 \frac{q^l - 1}{q^2 - 1} + O(d^l),$$

while for $l = 2s+1$ we get (up to an error of type $O(d^l)$)

$$N_{\Gamma,a_1,a_2}(l) = \frac{2}{l_1 l_2} \sum_{2m' \leq 2s+1} q^{2m'} = \frac{2}{l_1 l_2} \sum_{m' \leq s} q^{2m'} = \frac{2}{l_1 l_2} q^2 \frac{q^{2s} - 1}{q^2 - 1} = \frac{2}{l_1 l_2} \frac{q}{q^2 - 1}(q^l - 1).$$

We note that

$$\#\{\gamma \in \Gamma_c |\, \mathrm{wl}\,(\gamma) \leq l\} = \sum_{m \leq l} q^m + O(l) = \frac{q}{q-1} q^l + O(l).$$

Finally, as $l \to \infty$,

$$\begin{aligned} &\frac{N_{\Gamma,a_1,a_2}(l)}{\#\{\gamma \in \Gamma_c |\, \mathrm{wl}\,(\gamma) \leq l\}} + \frac{N_{\Gamma,a_1,a_2}(l+1)}{\#\{\gamma \in \Gamma_c |\, \mathrm{wl}\,(\gamma) \leq l+1\}} \\ &\qquad \to \frac{2}{l_1 l_2} \left(\frac{q^2/(q^2-1)}{q/(q-1)} + \frac{q/(q^2-1)}{q/(q-1)} \right) = \frac{2}{l_1 l_2}. \end{aligned}$$

The case $\chi(a_1 + a_2) = -1$ is similar. This proves the second part of Theorem 3.13. We note that the subsequence of odd and even m's do not have the same limit.

4. Appendix

In this appendix we state and prove some general results which we have been unable to find in the existing literature.

The first two parts of the following proposition have, however, appeared previously in e.g [24, 25] but we recall them for convenience.

Proposition 4.1. *Let $f : \mathbb{R}^k \to \mathbb{R}$ be a function for which there exists $\nu > 0$ such that*

$$f(\epsilon) = 1 - \frac{\langle \epsilon, D\epsilon \rangle}{2} + O(\|\epsilon\|^{2+\nu}),$$

as $\|\epsilon\| \to 0$, where D is a positive definite $k \times k$ matrix. Let $r > 0$ be a constant.

(i) *For every $\epsilon_0 \in \mathbb{R}^k$ we have*

$$f(\epsilon_0/r\sqrt{m})^m \to \exp(-\langle \epsilon, D\epsilon \rangle /2r^2),$$

as $m \to \infty$.

(ii) *There exist a $\delta > 0$ such that for all $m \in \mathbb{N}$*

$$\left| f(\epsilon/r\sqrt{m})^m - e^{-\langle \epsilon, D\epsilon \rangle /2r^2} \right| < 2e^{-\langle \epsilon, D\epsilon \rangle /4r^2}$$

whenever $\|\epsilon\| < \delta r\sqrt{m}$.

(iii) *For every $0 \le b \le \nu/2$ there exist constants $\delta, C > 0$ such that for all $m \in \mathbb{N}$, $\|\epsilon\| \le \delta m^{b/(4+2\nu)}$,*

$$\left| f(\epsilon/r\sqrt{m})^m - e^{-\langle \epsilon, D\epsilon \rangle /2r^2} \right| \le C/m^{\min(1-\varepsilon,\nu/2-b)}.$$

Proof. We have

$$f(\epsilon_0/r\sqrt{m}) = 1 - \frac{\langle \epsilon_0, D\epsilon_0 \rangle}{2mr^2} + O(\|\epsilon_0/\sqrt{m}\|^{2+\nu}).$$

and therefore for m sufficiently large

$$f(\epsilon_0/r\sqrt{m})^m = \left(1 - \frac{\langle \epsilon_0, D\epsilon_0 \rangle}{2mr^2}\right)^m + R(\epsilon_0, m), \tag{4.1}$$

where

$$|R(\epsilon_0, m)| \le \sum_{k=1}^{m} \binom{m}{k} \frac{C^k \|\epsilon_0\|^{(2+\nu)k}}{m^{(1+\nu/2)k}} = \left(1 + \frac{C\|\epsilon_0\|^{2+\nu}}{m^{1+\nu/2}}\right)^m - 1 \tag{4.2}$$

The first result now follows from

$$\lim_{m\to\infty} (1 - x/m^c)^m = \begin{cases} e^{-x} & \text{if } c = 1 \\ 1 & \text{if } c > 1 \end{cases}. \tag{4.3}$$

For the second result we can choose δ sufficiently small such that for $\|\epsilon\| < \delta$

$$f(\epsilon) - 1 \le -\frac{1}{4} \langle \epsilon, D\epsilon \rangle .$$

Using $(1 - x/m)^m < e^{-x}$ we find that for $\|\epsilon\| < \delta r\sqrt{m}$ we have $|f(\epsilon/r\sqrt{m})^m| \le e^{-\langle \epsilon, D\epsilon \rangle /4r^2}$ from which (ii) easily follows.

To prove (iii) we need to consider the rate af convergence in (4.3). We first consider $c = 1$. We use the Taylor series of $\log(1 - u)$ to see that

$$x + m\log(1 - x/m) \to 0$$

as $m \to \infty$. In fact it is $O(x^2/m)$:

$$x - m\sum_{j=1}^{\infty} \frac{x^j}{m^j j} = -\sum_{j=2}^{\infty} \frac{x^j}{m^{j-1} j} = O(x\sum_{1}^{\infty}(x/m)^j) = O\left(x\frac{|x/m|}{1-|x/m|}\right).$$

Since $(e^u - 1)/u \to 1$ as $u \to 0$, we have $e^u - 1 = O(u)$ for u going to zero. We assume that $|x| \leq \delta' m^{1/2}$. Hence $|x|^2/m$ can be made small by making δ' small, and we have:

$$e^{x+m\log(1-x/m)} - 1 = O(x + m\log(1-|x/m|)) = O(x^2/m).$$

By multiplying with e^{-x} we get

$$(1-x/m)^m - e^{-x} = O(e^{-x}x^2/m)$$

which holds for all $|x| \leq \delta' m^{1/2}$. We note that $e^{-x}x^2/m \leq m^{-1+\varepsilon}$ when $0 \leq x \leq m^{\varepsilon/2}$ and $e^{-x}x^2/m = O(m^{-L})$ for any positive L when $m^{\varepsilon/2} \leq x \leq m^{1/2}$.

Hence for any $\varepsilon > 0$ there exist C_ε such that when $0 \leq x \leq \delta' m^{1/2}$

$$\left|(1-x/m)^m - e^{-x}\right| \leq C_\varepsilon m^{-1+\varepsilon}.$$

For the case $c > 1$ we have $m\log(1-x/m^c) \to 0$ and, in fact, $m\log(1-x/m^c) = O(x/m^{c-1})$ by the same argument as before. So when $|x| < \delta' m^{c-1}$

$$(1-x/m^c)^m - 1 = e^{m\log(1-x/m^c)} - 1 = O(m\log(1-x/m^c)) = O(x/m^{c-1}).$$

Hence there exist a constant $B > 0$ such that if we fix $b \leq c-1$ and restrict x in the set $|x| \leq \delta' m^b$ we have

$$(1-x/m^c)^m - 1 \leq Bm^{1+b-c}.$$

Using (4.1) we have

$$f(\epsilon/r\sqrt{m})^m = \left(1 - \frac{\langle \epsilon, D\epsilon\rangle}{2mr^2}\right)^m + R(\epsilon, m)$$

whenever $\|\epsilon\| \leq \delta r\sqrt{m}$. We take $c = 1 + \nu/2$ in (4.2). Let $b \leq c - 1 = \nu/2$. Hence there exist a constant $C > 0$ such that if we let

$$\langle \epsilon, \epsilon\rangle \leq \delta'' m^{1/2} \quad \text{ and } \quad \langle \epsilon, \epsilon\rangle^{2+\nu} \leq \delta'' m^b$$

then

$$\left|f(\epsilon/r\sqrt{m})^m - e^{-\langle \epsilon, D\epsilon\rangle/2r^2}\right| \leq C\max(m^{\varepsilon-1}, m^{-\nu/2+b}).$$

This completes the proof. □

Lemma 4.2. *Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable with continuous, bounded derivative, non-increasing for $x > 0$, even and integrable. Assume that*

$$\lim_{x\to\pm\infty} xf(x) = 0.$$

Let $B \subset \mathbb{R}$ have density $d(B)$. Then

$$\lim_{x\to\infty} \sum_{t\in B, |t|\leq x} \frac{1}{\sqrt{x}} f\left(\frac{t}{\sqrt{x}}\right) = d(B)\int_{\mathbb{R}} f(u)\,du.$$

Proof. Fix $\epsilon > 0$. Choose x_0 such that for $t > x_0$ we have

$$2(d(B) - \epsilon)t < |B \cap [-t, t]| < 2(d(B) + \epsilon)t.$$

To simplify notation b will always denote an element of B. We use summation by parts to get:

$$\begin{aligned}
\sum_{|b| \le x} \frac{1}{\sqrt{x}} f\left(\frac{b}{\sqrt{x}}\right) &= \sum_{0 \le b \le x} \frac{1}{\sqrt{x}} f\left(\frac{b}{\sqrt{x}}\right) + \sum_{0 \ge b \ge -x} \frac{1}{\sqrt{x}} f\left(\frac{b}{\sqrt{x}}\right) \\
&= |B \cap [0, x]| \frac{1}{\sqrt{x}} f(\sqrt{x}) - \int_0^x |B \cap [0, t]| \frac{1}{x} f'(t/\sqrt{x})\, dt \\
&+ |B \cap [-x, 0]| \frac{1}{\sqrt{x}} f(-\sqrt{x}) - \int_0^x |B \cap [-t, 0]| \frac{-1}{x} f'(-t/\sqrt{x})\, dt \\
&= |B \cap [-x, x]| \frac{1}{\sqrt{x}} f(\sqrt{x}) - \int_0^x |B \cap [-t, t]| \frac{1}{x} f'(t/\sqrt{x})\, dt,
\end{aligned}$$

where we also used that f is even and f' is odd.

$$\begin{aligned}
\limsup \sum_{|b| \le x} \frac{1}{\sqrt{x}} f(b/\sqrt{x}) \le{}& \limsup \frac{|B \cap [-x, x]|}{x} \sqrt{x} f(\sqrt{x}) \\
&+ \limsup \left(-\int_{x_0}^x 2t(d(B) + \epsilon) \frac{1}{x} f'(t/\sqrt{x})\, dt \right) \\
&+ \limsup \int_0^{x_0} |B \cap [-t, t]| \frac{1}{x} f'(t/\sqrt{x})\, dt.
\end{aligned}$$

where we have used the fact that $f' \le 0$. The first term on the right tends to 0, by the existence of the density and the conditions on f. The third term tends to 0, as f' is bounded, B is discrete and $[0, x_0]$ is a finite interval. The second term can be analyzed as follows: We do integration by parts.

$$\begin{aligned}
-\int_{x_0}^x \frac{2t}{x} f'(t/\sqrt{x})\, dt &= -\left[-f\left(\frac{t}{\sqrt{x}}\right) \frac{2t}{\sqrt{x}} \right]_{x_0}^x + \int_{x_0}^x f\left(\frac{t}{\sqrt{x}}\right) \frac{2}{\sqrt{x}}\, dt \\
&= \left(2\sqrt{x} f(\sqrt{x}) - f(x_0/\sqrt{x}) \frac{2x_0}{\sqrt{x}} \right) + \int_{x_0/\sqrt{x}}^{\sqrt{x}} 2f(u)\, du.
\end{aligned}$$

The first term tends to 0 and the second to the integral $\int_0^\infty 2f(u)\, du = \int_{-\infty}^\infty f(u)\, du$. This proves that

$$\limsup \sum_{|b| \le x} \frac{1}{\sqrt{x}} f(b/\sqrt{x}) \le (d(B) + \epsilon) \int_{\mathbb{R}} f(u)\, du$$

for every ϵ. We let ϵ tend to 0 and we work similarly with $\liminf$. □

Remark 4.3. A similar result holds for any $0 < c < 1$:

$$\frac{1}{x^c} \sum_{0 \le a \le x} f(a/x^c) \to d(B) \int_0^\infty f(u)\, du.$$

Lemma 4.2 is a special case of Lemma 4.4 below. We included it only because it is easier to follow the proof in one dimension.

Lemma 4.4. *Let $f:\mathbb{R}^k\to\mathbb{R}$ be integrable, differentiable, with continuous partial derivatives, even in each variable, with $f_i\leq 0$, $f_{ij}\geq 0$ for $i\neq j$, $f_{ijk}\leq 0$ for distinct i,j,k etc on the higher partial derivatives in the set $\{t\in\mathbb{R}^k, t_i\geq 0\}$. Let B have natural density $d(B)$. Then*

$$\lim_{x\to\infty}\sum_{t\in B,|t_i|\leq x}\frac{1}{x^{k/2}}f(t_1/\sqrt{x},\dots,t_k/\sqrt{x})=d(B)\int_{\mathbb{R}^k}f(u_1,\dots,u_k)\,du_1\dots du_k.$$

Proof. As usual we denote $[s]$ the greatest integer part of s and by $\{s\}$ its fractional part. Fix $\epsilon>0$. We use the norm $|t|=\max|t_j|$. Choose x_0 such that for $|t|>x_0$ we have

$$2^k(d(B)-\epsilon)t_1\dots t_k<\left|B\cap\prod_i[-t_i,t_i]\right|<2^k(d(B)+\epsilon)t_1\dots t_k.$$

To simplify notation b will denote an element of B. We use summation by parts [13, Th. 1.6, p 24] to get

$$\sum_{|b_i|\leq x}\frac{1}{x^{k/2}}f(b/\sqrt{x})=\sum_{\vec{a}\in\{0,1\}^k}(-1)^{\sum a_i}\int_{|t_i|\leq x}\left|B\cap\prod_i[-t_i,t_i]\right|\prod_{i=1}^k\frac{\partial}{\partial t_i}^{a_i}\left(\frac{1}{x^{k/2}}f(t/\sqrt{x})\right)dt.$$

In this equation it is understood that for $\vec{a}=\vec{0}$ there is no integration and we substitute $t_i=x$. Also $\partial^{a_i}/\partial t_i$ means no derivative in t_i if $a_i=0$. We split the integration in $|t|\leq x_0$ and $|t|>x_0$. The conditions on the partial derivatives of f imply that

$$\begin{aligned}
&\sum_{|b_j|\leq x}\frac{1}{x^{k/2}}f(b/\sqrt{x})\\
&=\sum_{\vec{a}\in\{0,1\}^k}(-1)^{\sum a_i}\int_{|t_j|\leq x_0}\left|B\cap\prod_i[-t_i,t_i]\right|\prod_{i=1}^k\frac{\partial}{\partial t_i}^{a_i}\left(\frac{1}{x^{k/2}}f(t/\sqrt{x})\right)dt\\
&\quad+\sum_{\vec{a}\in\{0,1\}^k}(-1)^{\sum a_i}\int_{x_0\leq|t_j|\leq x}\left|B\cap\prod_i[-t_i,t_i]\right|\prod_{i=1}^k\frac{\partial}{\partial t_i}^{a_i}\left(\frac{1}{x^{k/2}}f(t/\sqrt{x})\right)dt\\
&\leq\sum_{\vec{a}\in\{0,1\}^k}(-1)^{\sum a_i}\int_{|t_j|\leq x_0}\left|B\cap\prod_i[-t_i,t_i]\right|\prod_{i=1}^k\frac{\partial}{\partial t_i}^{a_i}\left(\frac{1}{x^{k/2}}f(t/\sqrt{x})\right)dt\\
&\quad+(d(B)+\epsilon)\sum_{\vec{a}\in\{0,1\}^k}(-1)^{\sum a_i}\int_{x>|t|>x_0}\prod_i(2[t_i])\prod_{i=1}^k\frac{\partial}{\partial t_i}^{a_i}\left(\frac{1}{x^{k/2}}f(t/\sqrt{x})\right)dt\\
&\quad+(d(B)+\epsilon)\sum_{\vec{a}\in\{0,1\}^k}(-1)^{\sum a_i}\int_{x>|t|>x_0}\prod_i(2\{t_i\})\prod_{i=1}^k\frac{\partial}{\partial t_i}^{a_i}\left(\frac{1}{x^{k/2}}f(t/\sqrt{x})\right)dt
\end{aligned}$$

We note that the third term can be written by the change of variables $u=t/\sqrt{x}$ as

$$(d(B)+\epsilon)\sum_{\vec{a}\in\{0,1\}^k}(-1)^{\sum a_i}\int_{\sqrt{x}>|u|>x_0/\sqrt{x}}\prod_i(2\{u_i\})\frac{1}{x^{\sum a_j/2}}\prod_{i=1}^k\frac{\partial}{\partial u_i}^{a_i}(f(u))\,du$$

and this is the integral of an integrable function. If $\sum a_j>0$ it tends to 0. The term with $a_i=0$, $i=1,\dots,k$ together with the second term can be combined so that we use the summation by parts formula backwards.

$$
\begin{aligned}
&\sum_{|b_j|\le x} \frac{1}{x^{k/2}} f(b/\sqrt{x}) \\
&= \sum_{\vec{a}\in\{0,1\}^k} (-1)^{\sum a_i} \int_{|t_j|\le x_0} \left| B \cap \prod_i [-t_i, t_i] \right| \prod_{i=1}^{k} \frac{\partial}{\partial t_i}^{a_i} \left(\frac{1}{x^{k/2}} f(t/\sqrt{x}) \right) dt \\
&\quad + (d(B)+\epsilon) \sum_{x_0<|b|\le x} \frac{1}{x^{k/2}} f(b/\sqrt{x}) + O(1/x^{1/2}).
\end{aligned}
$$

The second term is a Riemann sum for the integral

$$\int_{x_0<|t|\le x} \frac{1}{x^{k/2}} f(t/\sqrt{x})\, dt.$$

The conditions on the function f guarantee that the Riemann sum approximates the integral with an error $O(x^{-k/2} \max |f(t)|)$. We take $\limsup$ on the inequality, as $x \to \infty$. The first term on the right tends to 0, as it is an integral over a compact set. We end up with

$$
\begin{aligned}
\limsup \sum_{|b_j|\le x} \frac{1}{x^{k/2}} f(b/\sqrt{x}) \quad &\le \quad (d(B)+\epsilon) \limsup \sum_{x_0<|b|\le x} \frac{1}{x^{k/2}} f(b/\sqrt{x}) \\
&= (d(B)+\epsilon) \lim \int_{x_0<|t|\le x} \frac{1}{x^{k/2}} f(t/\sqrt{x})\, dt \\
&= (d(B)+\epsilon) \lim \int_{[x_0/\sqrt{x}, \sqrt{x}]^k} 2^k f(u)\, du \\
&= (d(B)+\epsilon) \int_{\mathbb{R}^k} f(u)\, du.
\end{aligned}
$$

We let ϵ tend to 0 and we work similarly with $\liminf$.

□

We have the following corollary:

Corollary 4.5. *Let M be a positive definite matrix of determinant 1. Assume that $B \subset \mathbb{Z}^k$ has natural density $d(B)$. Then*

$$\sum_{\substack{\beta\in B \\ |\beta_i|\le m}} \frac{1}{(2\pi\sigma^2 m)^{k/2}} e^{-\langle \beta, M^{-1}\beta\rangle/2\sigma^2 m} \to d(B),$$

as $m \to \infty$.

Acknowledgments:
We would like to thank I. Kapovich for valuable comments and for initiating our interest in this problem. The authors are grateful to P. Sarnak for useful comments and suggestions that lead us to consider random sets. The first author will like to thank the Max-Planck-Institut für Mathematik, where he was a visitor for the year 2005, and the second author gratefully acknowledges the hospitality of the Institute for Advanced Study in Princeton.

References

[1] T. Adachi, Distribution of closed geodesics with a preassigned homology class in a negatively curved manifold. Nagoya Math. J. **110** (1988), 1–14.

[2] T. Adachi, T. Sunada, Homology of closed geodesics in a negatively curved manifold. J. Differential Geom. **26** (1987), no. 1, 81–99.

[3] M. Babillot, F. Ledrappier, Lalley's theorem on periodic orbits of hyperbolic flows. Ergodic Theory Dynam. Systems 18 (1998), no. 1, 17–39.

[4] A. V. Borovik, A. G. Myasnikov, V. Shpilrain, Measuring sets in infinite groups, in *Computational and statistical group theory (Las Vegas, NV/Hoboken, NJ, 2001)*, 21–42, Contemp. Math., 298, Amer. Math. Soc., Providence, RI, 2002.

[5] E. Cesàro, Démonstration élémentaire et généralisation de quelques théorèmes de M. Berger, Mathesis 1,1881, 99–102.

[6] M. Gromov, Hyperbolic Groups in *Essays in group theory*, 75–263, Springer, New York, 1987.

[7] M. Gromov, Asymptotic invariants of infinite groups in *Geometric group theory, Vol. 2 (Sussex, 1991)*, 1–295, Cambridge Univ. Press, Cambridge, 1993.

[8] D. Hejhal, The Selberg trace formula for PSL(2, R). Vol. 1. Lecture Notes in Mathematics, 1001. Springer-Verlag, Berlin, 1983. viii+806pp.

[9] H. Huber, Zur analytischen Theorie hyperbolischen Raumformen und Bewegungsgruppen I, Math. Ann. **138** (1959), 1–26; II Math. Ann. **142** (1960/1961), 385–398; Nachtrag zu II, Math. Ann. **143** (1961), 463—464.

[10] J.-P. Kahane, *Some random series of functions*, Second edition, Cambridge Univ. Press, Cambridge, 1985

[11] I. Kapovich, P. Schupp, V. Shpilrain, On the density of test elements in a free group of rank two, In preparation.

[12] M. Kotani, A note on asymptotic expansions for closed geodesics in homology classes. Math. Ann. **320** (2001), no. 3, 507–529.

[13] E. Krätzel, Lattice points. Mathematics and its Applications (East European Series), 33. Kluwer Academic Publishers Group, Dordrecht, 1988. 320 pp. ISBN: 90-277-2733-3

[14] S. Lalley, Closed geodesics in homology classes on surfaces of variable negative curvature. Duke Math. J. 58 (1989), no. 3, 795–821.

[15] W. Parry, M. Pollicott, The Chebotarov theorem for Galois coverings of Axiom A flows. Ergodic Theory Dynam. Systems **6** (1986), no. 1, 133–148.

[16] Y. N. Petridis, M. S. Risager, Discrete logarithms in free groups, to appear in Proc. Amer. Math. Soc.

[17] Y. N. Petridis, M. S. Risager, The distribution of values of the Poincaré pairing for hyperbolic Riemann surfaces, J. für die Reine und Angew. Mathematik, 579, **2005** 159–173.

[18] R. Phillips, P. Sarnak, Geodesics in homology classes. Duke Math. J. 55 (1987), no. 2, 287–297.

[19] I. Rivin, Growth in free groups (and other stories), arXiv:math.CO/9911076.

[20] J. Rousseau-Egele, Un théorème de la limite locale pour une classe de transformations dilatantes et monotones par morceaux. Ann. Probab. **11** (1983), no. 3, 772–788.

[21] P. Sarnak, *Some applications of modular forms*, Cambridge Tracts in Mathematics, 99. Cambridge University Press, Cambridge, 1990. x+111 pp. ISBN 0-521-40245-6.

[22] P. Sarnak, Class numbers of indefinite binary quadratic forms, J. Number Theory **15** (1982), no. 2, 229–247.

[23] A. Selberg, Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series, J. Indian Math. Soc. (N.S.) **20** 1956, 47–87.

[24] R. Sharp, Local limit theorems for free groups. Math. Ann. 321 (2001), no. 4, 889–904.

[25] R. Sharp, A local limit theorem for closed geodesics and homology. Trans. Amer. Math. Soc. 356 (2004), no. 12, 4897–4908.

[26] T. Sunada, Geodesic flows and geodesic random walks. Geometry of geodesics and related topics (Tokyo, 1982), 47–85, Adv. Stud. Pure Math., 3, North-Holland, Amsterdam, 1984.

[27] A. Venkov, Spectral theory of automorphic functions. A translation of Trudy Mat. Inst. Steklov. **153** (1981). Proc. Steklov Inst. Math. 1982, no. 4 (153), ix+163 pp. 1983.

Department of Mathematics and Computer Science,, City University of New York, Lehman College,, 250 Bedford Park Boulevard West, Bronx, NY 10468-1589

The Graduate Center, Mathematics Ph.D. Program, 365 Fifth Avenue, Room 4208, New York, NY 10016-4309

E-mail address: `petridis@comet.lehman.cuny.edu`

Department of Mathematical Sciences, University of Aarhus, Ny Munkegade Building 530, 8000 Aarhus C, Denmark

E-mail address: `risager@imf.au.dk`