ПРЕПРИНТЫ ПОМИ РАН

Учредитель: Санкт-Петербургское отделение Математического института им. В. А. Стеклова Российской академии наук

Свидетельство о регистрации средства массовой информации: ЭЛ №ФС 77-33560 от 16 октября 2008 г. Выдано Федеральной службой по надзору в сфере связи и массовых коммуникаций

Контактные данные: 191023, г. Санкт-Петербург, наб. реки Фонтанки, дом 27

телефоны:(812)312-40-58; (812) 571-57-54

e-mail: admin@pdmi.ras.ru

http://www. pdmi.ras.ru /preprint/

Заведующая информационно-издательским сектором Симонова В.Н

PDMI Preprint 15/2010

QUASIMORPHISMS AND RANDOM WALKS

A. V. Malyutin

St.Petersburg Department Steklov Mathematical Institute RAN,
Fontanka 27, 191011 St.Petersburg, Russia

malyutin@pdmi.ras.ru

December 15, 2010

Abstract

We study interrelations between the theory of quasimorphisms and theory of random walks on groups, and establish the following criterion of transience for subsets of countable groups: if a subset of a countable group has bounded images under any three linearly independent homogeneous quasimorphisms on the group, then this subset is transient (with respect to all nondegenerate random walks on the group). From this it follows by results of M. Bestvina, K. Fujiwara, J. Birman, W. Menasco, and others that generic elements in mapping class groups of surfaces are pseudo-Anosov, generic braids in Artin's braid groups represent prime links and knots, generic elements in the commutant of the free group have large stable commutator length, etc.

Key words: Quasimorphism, random walk, transience, mapping class group, pseudo-Anosov, braid, knot, commutator.

Supported by the grant RFBR 09-01-12175-ofi-m.

Introduction

Let G be a countable group and let μ be a probability measure on G. Recall that the *right random walk* $(X_k)_{k\geq 0}$ *on* G *with distribution* μ (or, briefly, μ*-walk*) is the time-homogeneous Markov chain whose state space is G, the transition probabilities are given by $P(g,h) = \mu(g^{-1}h)$, and the initial distribution is concentrated at the identity of the group. Realizations of this process are called *paths* of the random walk. We say that μ is *nondegenerate* if its support generates G as a semigroup; a random walk is *nondegenerate* if its distribution is nondegenerate.

Given a group G and a mapping[1] $f : G \to R$ to some space R, what is the behaviour of f with respect to random walks on G? (By the "behaviour" of f with respect to a random walk $(X_k)_{k\geq 0}$ on G we mean various properties of random variables $f(X_k)$ and sequences $f(\tau_k)$, where $(\tau_k)_{k\geq 0}$ are paths of $(X_k)_{k\geq 0}$.) In certain cases, the study of this question turns out to be productive and helps to obtain a new insight on the structure of G and properties of f.

In the present paper, we study the case where G is a countable group, $R = \mathbb{R}^{\mathsf{d}}$ ($\mathsf{d} \in \mathbb{N}$), and f is an $\mathbb{R}^{\mathsf{d}}$*-quasimorphism*, i. e., a map $G \to \mathbb{R}^{\mathsf{d}}$ such that the set

$$D_f := \{f(g_1g_2) - f(g_1) - f(g_2)\}_{g_1,g_2\in G}$$

is bounded in $\mathbb{R}^{\mathsf{d}}$. (See Sec. 1 for definitions.) We say that an $\mathbb{R}^{\mathsf{d}}$-quasimorphism is *nondegenerate* if its image is not contained in a bounded neighborhood of a hyperplane in $\mathbb{R}^{\mathsf{d}}$.

Observe that quasimorphisms are similar to homomorphisms, while the "behaviour" of homomorphisms $G \to \mathbb{R}^{\mathsf{d}}$ with respect to random walks on G is described by the classical theory of random walks on Euclidean spaces (due to the obvious fact that homomorphisms $G \to \mathbb{R}^{\mathsf{d}}$ send random walks on G to random walks on $\mathbb{R}^{\mathsf{d}}$). Therefore, one could expect that some theorems about random walks on Euclidean spaces may be generalized to describe behaviour of quasimorphisms with respect to random walks. M. Björklund and T. Hartnick [4] proved that the (analogues of the) central limit theorem and the law of the iterated logarithm are valid for quasimorphisms. In this paper, we show that several other properties of random walks on Euclidean spaces (kindred to the property of transience for random walks in $\mathbb{R}^{\mathsf{d}}$ with $\mathsf{d} \geq 3$) still hold true for quasimorphisms. We will also present corollaries for mapping class groups, braid groups, commutants.

Our basic result is the following theorem, which is the direct analogue of a well-known fact about random walks on $\mathbb{R}^{\mathsf{d}}$ ($\mathbb{Z}^{\mathsf{d}}$).

0.1. Theorem. *Let* G *be a countable group and let* $\Phi : G \to \mathbb{R}^{\mathsf{d}}$, $\mathsf{d} \in \mathbb{N}$, *be a nondegenerate* $\mathbb{R}^{\mathsf{d}}$*-quasimorphism. Then for each nondegenerate probability measure* μ *on* G *and for every bounded subset* $Q \subset \mathbb{R}^{\mathsf{d}}$ *there exists a constant* $C := C(G, \Phi, \mu, Q)$ *such that for any* $k \in \mathbb{N}$ *and* $\mathbf{x} \in \mathbb{R}^n$ *we have*

$$\mu^{*k}\left(\Phi^{-1}(\mathbf{x} + Q)\right) \;<\; Ck^{-\mathsf{d}/2},$$

where μ^{*k} *denotes the* k*-fold convolution of* μ.

(In the case where $\mathsf{d} = 1$ and Φ is square-integrable with respect to μ, the statement of Theorem 0.1 trivially follows from the above-mentioned central limit theorem of M. Björklund and T. Hartnick.)

[1] The "mapping" might stand here for a characteristic of group elements, an invariant, norm, homomorphism, etc.

Since the series $\sum_{k=1}^{\infty} k^{-\mathsf{d}/2}$ is convergent whenever $\mathsf{d} \geq 3$, Theorem 0.1 and Borel–Cantelli lemma readily imply the following corollary, which generalizes the well-known fact that every nondegenerate random walk on $\mathbb{Z}^{\mathsf{d}}$ with $\mathsf{d} \geq 3$ is transient.

0.2. Corollary. *Let G be a countable group. Assume that $\mathsf{d} \geq 3$. Then each nondegenerate $\mathbb{R}^{\mathsf{d}}$-quasimorphism on G sends a. e. path of every nondegenerate random walk on G to a sequence tending to infinity.*

Theorem 0.1 immediately implies the following corollaries.

0.3. Corollary. *If a subset S of a countable group G has bounded image under a nondegenerate $\mathbb{R}^{\mathsf{d}}$-quasimorphism $G \to \mathbb{R}^{\mathsf{d}}$, $\mathsf{d} \in \mathbb{N}$, then for every nondegenerate probability measure μ on G there exists a constant $C := C(\mu)$ such that for each $k \in \mathbb{N}$ we have*

$$\mu^{*k}(S) \ < \ Ck^{-\mathsf{d}/2}.$$

0.4. Corollary. *If a subset S of a countable group G has bounded image under a nondegenerate quasimorphism $G \to \mathbb{R}$, then, for every nondegenerate probability measure μ on G, the probability that the random μ-walk on G hits S at the k-th step, tends to 0 as k tends to infinity.*

Recall that a subset of a group is said to be *transient* with respect to a random walk on the group if a. e. path of the random walk visits the subset only finitely many times. Corollary 0.2 yields the following criterion of transience for subsets of countable groups.

0.5. Corollary. *If a subset of a countable group has bounded image under a nondegenerate $\mathbb{R}^3$-quasimorphism of the group, then the subset is transient with respect to each nondegenerate random walk on the group.*

Pseudo-Anosov elements in mapping class groups. M. Bestvina and K. Fujiwara [2, 3] showed that if the mapping class group $\mathrm{MCG}(M)$ of a compact surface M is not virtually Abelian, then there exists an infinite number of linearly independent homogeneous quasimorphisms $\mathrm{MCG}(M) \to \mathbb{R}$ each of which sends all non-pseudo-Anosov elements in $\mathrm{MCG}(M)$ to 0. This means that (if $\mathrm{MCG}(M)$ is not virtually Abelian) for each $\mathsf{d} \in \mathbb{N}$ there exists a nondegenerate $\mathbb{R}^{\mathsf{d}}$-quasimorphism $\mathrm{MCG}(M) \to \mathbb{R}^{\mathsf{d}}$ that sends the set of non-pseudo-Anosov elements in $\mathrm{MCG}(M)$ to (the bounded subset) $\{\mathbf{0}\} \subset \mathbb{R}^{\mathsf{d}}$. By Corollary 0.5, this implies the following result.

0.6. Corollary. *If the mapping class group $\mathrm{MCG}(M)$ of a compact surface M is not virtually Abelian, then the subset of non-pseudo-Anosov elements in $\mathrm{MCG}(M)$ is transient for each nondegenerate random walk on $\mathrm{MCG}(M)$.*

I. Rivin [13, 14] proved that simple random walks on mapping class groups of closed orientable surfaces give rise to non-pseudo-Anosov elements with asymptotic probability zero. J. Maher [11] proved that random walks on mapping class groups of orientable surfaces with punctures give rise to non-pseudo-Anosov elements with asymptotic probability zero. (Methods of J. Maher apply to many subgroups of mapping class groups.) E. Kowalski [10] showed that the subset T of non-pseudo-Anosov elements is transient for the case of closed orientable surfaces and simple symmetric random walks. (The approach developed by E. Kowalski shows that $\mu^{*k}(T)$ decays exponentially fast with k, while the approach via quasimorphisms gives only the superpolinomiality.)

We can strengthen Corollary 0.6 in the following way. Denote by $T := T(M)$ be the set of non-pseudo-Anosov elements in the mapping class group $\mathrm{MCG}(M)$ of a compact surface M. It is well known that $\bigcup_{k\in\mathbb{N}\cup\{0\}} T^k = \mathrm{MCG}(M)$. For an element $g \in \mathrm{MCG}(M)$ we set $\|g\|_T := \min\{k \in \mathbb{N}\cup\{0\} \,|\, g \in T^k\}$.

0.7. Corollary. *If the mapping class group* $\mathrm{MCG}(M)$ *of a compact surface* M *is not virtually Abelian, then for a. e. path* $(\tau_k)_{k\geq 0}$ *of every nondegenerate random walk, the sequence* $(\|\tau_k\|_T)_{k\geq 0}$ *tends to infinity.*

Proof. The definition of $\mathbb{R}^{\mathsf{d}}$-quasimorphism implies that if a subset S of a group G has bounded image under a nondegenerate $\mathbb{R}^{\mathsf{d}}$-quasimorphism Φ, then $\Phi(S^m)$ is bounded for each m. By the mentioned above result of M. Bestvina and K. Fujiwara, there exists a nondegenerate $\mathbb{R}^3$-quasimorphism $\Phi : \mathrm{MCG}(M) \to \mathbb{R}^3$ such that $\Phi(T(M)) = \{\mathbf{0}\}$. Therefore, $\Phi(T^m)$ is bounded in $\mathbb{R}^3$ for each m. In particular, if $(x_k)_{k\geq 0}$ is a sequence in $\mathrm{MCG}(M)$, then the sequence $(\|x_k\|_T)_{k\geq 0}$ tends to infinity whenever the sequence $(\Phi(\tau_k))_{k\geq 0}$ tends to infinity in $\mathbb{R}^3$. It remains to notice that, by Corollary 0.2, for a. e. path $(\tau_k)_{k\geq 0}$ of every nondegenerate random walk, the sequence $(\Phi(\tau_k))_{k\geq 0}$ tends to infinity in $\mathbb{R}^3$. □

Braid groups and knots. The above-mentioned result of M. Bestvina and K. Fujiwara about mapping class groups trivially implies that, for the Artin braid group B_n with $n \geq 3$, there exists an infinite number of linearly independent homogeneous quasimorphisms $B_n \to \mathbb{R}$ each of which sends all non-pseudo-Anosov braids in B_n to 0. (In order to see this, one can use the natural homomorphism $B_n \to \mathrm{MCG}(S_{n+1})$ to the mapping class group $\mathrm{MCG}(S_{n+1})$ of $(n+1)$-punctured sphere. This homomorphism sends (non-)pseudo-Anosov braids to (non-)pseudo-Anosov elements in $\mathrm{MCG}(S_{n+1})$; its image is a finite index subgroup in $\mathrm{MCG}(S_{n+1})$.) By Corollary 0.5, this implies that Corollaries 0.6, 0.7 are valid in the case of braid groups.

0.8. Corollary. *Let* B_n *be the braid group of index* $n \geq 3$ *and let* $T_n \subset B_n$ *be the subset of all non-pseudo-Anosov braids in* B_n*. Then, for each* $m \in \mathbb{N}$*, the subset* T_n^m *is transient for each nondegenerate random walk on* B_n*.*

It is deduced in [12, Proposition 6.1] from results of J. S. Birman, W. W. Menasco, and I. A. Dynnikov that all braids in $B_n \smallsetminus T_n^2$ represent[2] *prime* knots and links (that is, every braid representing *composite*, *split*, or *trivial* link is the product of two non-pseudo-Anosov braids). By Corollary 0.8, this fact implies the following result.

0.9. Corollary. *In the braid group* B_n *with* $n \geq 3$*, the set of those braids that represent non-prime (i.e., composite, split, or trivial) knots and links is transient for each nondegenerate random walk on* B_n*.*

We conjecture that techniques developed by T. Ito in [8, 9] may be used to establish new relations between quasimorphisms of braid groups and properties of links represented by braids. In particular, we conjecture that these techniques allow to (prove the existence and) construct a function $\phi : \mathbb{N} \to \mathbb{N}$ such that, for each $k \in \mathbb{N}$, the set $S(n,k) \subset B_n$ of those braids in B_n that represent knots and links of genus $\leq k$ is contained in $T_n^{\phi(k)}$. In view of results presented above, this would imply that, for any $n \geq 3$ and $k \in \mathbb{N}$, $S(n,k)$ is transient for each nondegenerate random walk on B_n.

[2] In the classical sense of J. W. Alexander and A. A. Markov

Commutants and (stable) commutator length. Another direction where we can apply Theorem 0.1 is the study of (stable) commutator length (see [5] for references and definitions). In particular, Corollary 0.2 implies the following result (cf. [5]).

0.10. Corollary. *Let G be a countable group. Assume that the vector space $Q(G)/H^1(G)$, where $Q(G)$ denotes the vector space of homogeneous quasimorphisms of G and $H^1(G) \subset Q(G)$ is the vector space of real-valued homomorphisms of G, has dimension at least 3. Then for a. e. path $(\tau_k)_{k\geq 0}$ of every nondegenerate random walk on the commutant $[G,G]$, the sequence $(\mathrm{scl}(\tau_k))_{k\geq 0}$ tends to infinity.*

Acknowledgements. The author is cordially grateful to N. Yu. Netsvetaev and T. Smirnova-Nagnibeda for useful discussions.

1 Preliminaries on $\mathbb{R}^n$-quasimorphisms

A function $\varphi : G \to \mathbb{R}$ on a group G is called a *quasimorphism* with *defect* d if the following condition is fulfilled:

$$\sup_{g_1,g_2\in G} |\varphi(g_1g_2) - \varphi(g_1) - \varphi(g_2)| = d < \infty.$$

In order to study n-tuples of quasimorphisms, it is convenient to use the following notion of $\mathbb{R}^n$-*quasimorphisms*. We say that a map $\Phi : G \to \mathbb{R}^n$ $(n \in \mathbb{N})$ is $\mathbb{R}^n$-*quasimorphism* if the set

$$D_\Phi := \{\Phi(g_1g_2) - \Phi(g_1) - \Phi(g_2)\}_{g_1,g_2\in G}$$

is bounded in $\mathbb{R}^n$.

It is clear that a map $F : G \to \mathbb{R}^n$ is an $\mathbb{R}^n$-quasimorphism if and only if all its coordinate functions[3] are quasimorphisms. This yields a natural bijection between the set of all n-tuples of quasimorphisms of G and the set of all $\mathbb{R}^n$-quasimorphisms of G.

In order to perform basic estimates and characterize properties of $\mathbb{R}^n$-quasimorphisms (e. g., to define the notion of defect for $\mathbb{R}^n$-quasimorphism), we need to pick a reference norm in $\mathbb{R}^n$. In what follows, we will use the 1-norm[4] defined by

$$\|(x_1,\dots,x_n)\|_1 := \sum_{i=1}^n |x_i|.$$

We define the *defect* d_Φ of an $\mathbb{R}^n$-quasimorphism Φ to be

$$d_\Phi := \sup_{\mathbf{v}\in D_\Phi} \|\mathbf{v}\|_1.$$

A quasimorphism $\phi : G \to \mathbb{R}$ is said to be *homogeneous* if $\phi(g^k) = k\phi(g)$ for all $g \in G$, $k \in \mathbb{Z}$. We say that an $\mathbb{R}^n$-quasimorphism $\Phi : G \to \mathbb{R}^n$ is *homogeneous* if $\Phi(g^k) = k \cdot \Phi(g)$ for all $g \in G$, $k \in \mathbb{Z}$. The definition obviously implies that an $\mathbb{R}^n$-quasimorphism is homogeneous if and only if all its coordinate quasimorphisms are homogeneous.

[3] By the coordinate functions of F we mean the functions $f_1 : G \to \mathbb{R}, \dots, f_n : G \to \mathbb{R}$ such that $F(g) = (f_1(g),\dots,f_n(g))$.

[4] We chose the 1-norm only because it simplifies some of our formulas below.

1.1. Lemma. *Each $\mathbb{R}^n$-quasimorphism $\Phi : G \to \mathbb{R}^n$ has a unique homogeneous $\mathbb{R}^n$-quasimorphism $\overline{\Phi} : G \to \mathbb{R}^n$ such that the map $\overline{\Phi} - \Phi$ is bounded. For every $g \in G$ we have $\overline{\Phi}(g) = \lim_{k\to\infty} \frac{\Phi(g^k)}{k}$ and $\left\|\overline{\Phi}(g) - \Phi(g)\right\|_1 \leq d_\Phi$.*

Proof. Since for any $h \in G$, $k \in \mathbb{N}$ we obviously have

$$\|\Phi(h^k) - k \cdot \Phi(h)\|_1 \ \leq \ (k-1)d_\Phi,$$

it follows that for any $k, m \in \mathbb{N}$, $g \in G$ we have

$$\begin{aligned}\left\|\frac{\Phi(g^k)}{k} - \frac{\Phi(g^m)}{m}\right\|_1 &= \frac{\left\|m \cdot \Phi(g^k) - \Phi(g^{km}) + \Phi(g^{km}) - k \cdot \Phi(g^m)\right\|_1}{km} \\ &\leq \frac{(m-1)d_\Phi + (k-1)d_\Phi}{km} = \left(\frac{1}{k} + \frac{1}{m} - \frac{2}{km}\right) d_\Phi. \quad (1)\end{aligned}$$

Therefore, for every $g \in G$ the sequence $\{\Phi(g^k)/k\}_{k\in\mathbb{N}}$ converges being a Cauchy sequence. Consequently, the function

$$\overline{\Phi}(g) := \lim_{k\to\infty} \frac{\Phi(g^k)}{k} \quad (2)$$

is well-defined. From (1) it follows that for every $g \in G$ we have

$$\left\|\overline{\Phi}(g) - \Phi(g)\right\|_1 \ \leq \ d_\Phi.$$

This means that the map $\overline{\Phi} - \Phi$ is bounded and hence $\overline{\Phi}$ is an $\mathbb{R}^n$-quasimorphism. Moreover, from (2) it easily follows that $\overline{\Phi}$ is homogeneous. Observe that the difference of any two distinct homogeneous $\mathbb{R}^n$-quasimorphisms is unbounded. It follows that $\overline{\Phi}$ is a unique homogeneous $\mathbb{R}^n$-quasimorphism of G such that the map $\overline{\Phi} - \Phi$ is bounded. □

We say that an $\mathbb{R}^n$-quasimorphism is *degenerate* if its image is contained in a bounded neighborhood of a hyperplane in $\mathbb{R}^n$. Obviously, an $\mathbb{R}^n$-quasimorphism is degenerate if and only if a nontrivial linear combination of its coordinate quasimorphisms is a bounded function.

1.2. Lemma. *Let G be a group and let $\Phi : G \to \mathbb{R}^n$ ($n \in \mathbb{N}$) be a nondegenerate $\mathbb{R}^n$-quasimorphism. Then the image $\Phi(G)$ is cobounded (i. e., it forms an ε-net) in $\mathbb{R}^n$.*

Proof. Let $\overline{\Phi} : G \to \mathbb{R}^n$ be the homogeneous $\mathbb{R}^n$-quasimorphism with bounded difference $\overline{\Phi} - \Phi$. Since Φ is nondegenerate, then obviously so is $\overline{\Phi}$. Therefore, there is an n-tuple $(g_1, \ldots, g_n)$ of elements of G such that the vectors

$$\mathbf{v}_1 := \overline{\Phi}(g_1), \ \ \ldots, \ \ \mathbf{v}_n := \overline{\Phi}(g_n)$$

are linearly independent. Since $\overline{\Phi}$ is homogeneous, the following inequality holds for any integers $k_1, \ldots, k_n$:

$$\|\overline{\Phi}(g_1^{k_1} \cdots g_n^{k_n}) - (k_1 \cdot \mathbf{v}_1 + \cdots + k_n \cdot \mathbf{v}_n)\|_1 \ \leq \ (n-1)d_{\overline{\Phi}},$$

where $d_{\overline{\Phi}}$ is the defect of $\overline{\Phi}$. This means that for each point $\mathbf{v}$ of the lattice generated by the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$, there is a point $\mathbf{w} \in \overline{\Phi}(G)$ such that $|\mathbf{v} - \mathbf{w}| \leq (n-1)d_{\overline{\Phi}}$. The image $\overline{\Phi}(G)$ is thus cobounded. Since $\overline{\Phi} - \Phi$ is bounded it follows that $\Phi(G)$ is also cobounded. □

2 Proof of Theorem 0.1 and special pairs of sequences

Our proof of Theorem 0.1 is based on the notion of *special* pairs of sequences. The definition is as follows.

2.1. Definition. Let A be a set, and let $\mathcal{Y}$ be a family of two-element subsets of A. (In this paper, we mostly interested in the case where A is finite or countable, and $\mathcal{Y}$ is finite and consists of pairwise disjoint subsets of A.) Let $V = (v_1, \dots, v_n)$ and $W = (w_1, \dots, w_n)$, $n \in \mathbb{N}$, be two distinct finite sequences over A. We say that the pair $\{V, W\}$ is $\mathcal{Y}$-*special* (*special* when $\mathcal{Y}$ is fixed) if the following two conditions hold:

i) for each $i \in \{1, \dots, n\}$ we have either $v_i = w_i$ or $\{v_i, w_i\} \in \mathcal{Y}$,

ii) there are no $i, j \in \{1, \dots, n\}$ such that $v_i = w_j \neq w_i = v_j$.

We will study sets of sequences that do not contain special pairs.

2.2. Example. Let Ab be a torsion-free Abelian group (say, the additive group of real numbers). Let $a, b, c \in Ab$ and assume that $b \neq c$. Let $\mathcal{Y}$ be the family consisting of the unique element $\{b, c\} \subset Ab$. Let L_a be a set of finite sequences over Ab such that the sum of elements in each sequence from L_a equals a. Then L_a has no $\mathcal{Y}$-special pairs.

2.3. Notation. If A is a set and $n \in \mathbb{N}$, we will denote by A^n the set of all sequences of length n over A. If ν is a measure on A, we will denote by ν^n the corresponding product measure on A^n.

We deduce Theorem 0.1 from the following two propositions.

2.4. Proposition. *Let A be a countable or finite set, let $\mathsf{d} \in \mathbb{N}$, and let $\mathcal{Y} = \{Y_1, \dots, Y_{\mathsf{d}}\}$ be a family consisting of d pairwise disjoint*[5] *two-element subsets of A. Let ν be a probability measure on A with $\operatorname{supp}(\nu) \supset Y := Y_1 \cup \dots \cup Y_{\mathsf{d}}$. Then there exists a constant $C(\nu)$ such that for each $n \in \mathbb{N}$ and for every subset $L \subset A^n$ without $\mathcal{Y}$-special pairs we have*

$$\nu^n(L) \ < \ C(\nu) n^{-\mathsf{d}/2}.$$

2.5. Proposition. *Let G be a group and let $\Phi : G \to \mathbb{R}^{\mathsf{d}}$, $\mathsf{d} \in \mathbb{N}$, be a nondegenerate $\mathbb{R}^{\mathsf{d}}$-quasimorphism. Let $R > 0$ be a positive real number. Then there exists a family $\mathcal{Y}$ consisting of d pairwise disjoint two-element subsets of G such that for each $n \in \mathbb{N}$ and for every $\mathcal{Y}$-special pair $\{(g_1, \dots, g_n), (h_1, \dots, h_n)\} \subset G^n$ we have $\|\Phi(g_1 \cdots g_n) - \Phi(h_1 \cdots h_n)\|_1 \geq R$. Moreover, if $S \subset G$ is a subset generating G as a semigroup, then there exist $p \in \mathbb{N}$ and a family $\mathcal{Y}$ that satisfies all the above properties and consists of subsets of S^p.*

The proofs of Propositions 2.4 and 2.5 may be found in Sections 4 and 5 below. Now, we deduce Theorem 0.1 from these propositions.

Proof of Theorem 0.1. Recall that we consider a countable group G and a nondegenerate $\mathbb{R}^{\mathsf{d}}$-quasimorphism $\Phi : G \to \mathbb{R}^{\mathsf{d}}$ ($\mathsf{d} \in \mathbb{N}$). Our aim is to show that for any

[5] In fact, the statement of the proposition holds true in the (more general) case of pairwise distinct (not necessarily disjoint) two-element subsets; see Rem. 4.1.

nondegenerate probability measure μ on G and bounded subset $Q \subset \mathbb{R}^{\mathtt{d}}$ there exists a constant $C := C(G, \Phi, \mu, Q)$ such that for any $k \in \mathbb{N}$ and $\mathbf{x} \in \mathbb{R}^n$ we have

$$\mu^{*k}\left(\Phi^{-1}\left(\mathbf{x}+Q\right)\right) \;<\; Ck^{-\mathtt{d}/2}.$$

Let μ be a nondegenerate probability measure on G and let $Q \subset \mathbb{R}^{\mathtt{d}}$ be a bounded subset. Set $S_\mu := \mathrm{supp}(\mu)$. ($S_\mu$ generates G as a semigroup since μ is nondegenerate.) Set $Q' := Q + D_\Phi$, where $D_\Phi := \{\Phi(g_1 g_2) - \Phi(g_1) - \Phi(g_2)\}_{g_1, g_2 \in G}$. (The set Q' is bounded in $\mathbb{R}^{\mathtt{d}}$ because D_Φ is bounded by the definition of $\mathbb{R}^{\mathtt{d}}$-quasimorphism.) Let $R_{Q'}$ denote the diameter of Q' with respect to our reference norm $\|\cdot\|_1$ (defined by $\|(x_1, \ldots, x_{\mathtt{d}})\|_1 = \sum_{i=1}^{\mathtt{d}} |x_i|$). Let pr_k, $k \in \mathbb{N}$, denote the natural projection from G^k to G (this projection sends the sequence $(g_1, \ldots, g_k) \in G^k$ to the element $g_1 \cdots g_k \in G$).

By Proposition 2.5, there exist $p \in \mathbb{N}$ and a family $\mathcal{Y} = \{Y_1, \ldots, Y_{\mathtt{d}}\}$ consisting of $\mathtt{d}$ pairwise disjoint two-element subsets of $S_\mu^p = \mathrm{supp}(\mu^{*p})$ such that for each $k \in \mathbb{N}$ and for every $\mathcal{Y}$-special pair $\{(g_1, \ldots, g_k), (h_1, \ldots, h_k)\} \subset G^k$ of k-sequences over G we have $\|\Phi(g_1 \cdots g_k) - \Phi(h_1 \cdots h_k)\|_1 \geq R_{Q'} + 1$.

This means that, for any $k \in \mathbb{N}$ and $\mathbf{x} \in \mathbb{R}^{\mathtt{d}}$, the set $\mathrm{pr}_k^{-1}\left(\Phi^{-1}\left(\mathbf{x}+Q'\right)\right)$ has no $\mathcal{Y}$-special pairs.

Since $\mathrm{supp}(\mu^{*p}) \supset Y := Y_1 \cup \cdots \cup Y_{\mathtt{d}}$ and $Y_1, \ldots, Y_{\mathtt{d}}$ are pairwise disjoint in G, it then follows by Proposition 2.4 that there exists a constant $C(\mu^{*p}, Q')$ such that for any $k \in \mathbb{N}$, $\mathbf{x} \in \mathbb{R}^{\mathtt{d}}$ we have

$$(\mu^{*p})^k\left(\mathrm{pr}_k^{-1}\left(\Phi^{-1}\left(\mathbf{x}+Q'\right)\right)\right) \;<\; C(\mu^{*p}, Q')k^{-\mathtt{d}/2},$$

which is equivalent to

$$\mu^{*pk}\left(\Phi^{-1}\left(\mathbf{x}+Q'\right)\right) \;<\; C(\mu^{*p}, Q')k^{-\mathtt{d}/2} \tag{3}$$

(since for any measure ν on G and $k \in \mathbb{N}$ we obviously have $\mathrm{pr}_k(\nu^k) = \nu^{*k}$).

Let us show that for any $k \in \mathbb{N}$ and $\mathbf{x} \in \mathbb{R}^n$ we have

$$\mu^{*k}\left(\Phi^{-1}\left(\mathbf{x}+Q\right)\right) \;<\; C_\star k^{-\mathtt{d}/2}, \qquad \text{where} \quad C_\star := \max\{p^{d/2}, (2p)^{d/2} C(\mu^{*p}, Q')\}. \tag{4}$$

Indeed, if $k < p$, then

$$\mu^{*k}\left(\Phi^{-1}\left(\mathbf{x}+Q\right)\right) \;\leq\; 1 \;<\; p^{d/2}k^{-\mathtt{d}/2} \;\leq\; C_\star k^{-\mathtt{d}/2}.$$

Suppose that $k \geq p$. Let $k = mp + r$, where $m \in \mathbb{N}$, $r \in \{0, 1, \ldots, p-1\}$. Then we have

$$\begin{aligned}\mu^{*k}\left(\Phi^{-1}\left(\mathbf{x}+Q\right)\right) \;&=\; \mu^{*(mp+r)}\left(\Phi^{-1}\left(\mathbf{x}+Q\right)\right) \\ &=\; \sum_{g \in G} \mu^{*r}(g)\mu^{*mp}\left(g^{-1}\Phi^{-1}\left(\mathbf{x}+Q\right)\right).\end{aligned} \tag{5}$$

Observe that

$$\begin{aligned}g^{-1}\Phi^{-1}\left(\mathbf{x}+Q\right) \;&\subset\; \Phi^{-1}\left(\Phi\left(g^{-1}\Phi^{-1}\left(\mathbf{x}+Q\right)\right)\right) \\ &\subset\; \Phi^{-1}\left(\Phi(g^{-1}) + \mathbf{x} + Q + D_\Phi\right) \;=\; \Phi^{-1}\left(\Phi(g^{-1}) + \mathbf{x} + Q'\right) \\ &=\; \Phi^{-1}\left(\mathbf{x}_g + Q'\right), \quad \text{where} \quad \mathbf{x}_g := \mathbf{x} + \Phi(g^{-1}).\end{aligned} \tag{6}$$

By (3) and (6) we have

$$\mu^{*mp}\left(g^{-1}\Phi^{-1}\left(\mathbf{x}+Q\right)\right) \;\leq\; \Phi^{-1}\left(\mathbf{x}_g+Q'\right) \;<\; C(\mu^{*p},Q')m^{-\mathsf{d}/2}. \tag{7}$$

Since $\sum_{g\in G}\mu^{*r}(g)=1$, (5) and (7) yield

$$\mu^{*k}\left(\Phi^{-1}\left(\mathbf{x}+Q\right)\right) \;=\; \mu^{*(mp+r)}\left(\Phi^{-1}\left(\mathbf{x}+Q\right)\right) \;<\; C(\mu^{*p},Q')m^{-\mathsf{d}/2}. \tag{8}$$

Since $k=mp+r$, $m\in\mathbb{N}$, $p\in\mathbb{N}$, and $r\in\{0,\dots,p-1\}$, we have $k<2mp$, whence

$$C(\mu^{*p},Q')m^{-\mathsf{d}/2} \;<\; C(\mu^{*p},Q')(2p)^{\mathsf{d}/2}k^{-\mathsf{d}/2} \;\leq\; C_\star k^{-\mathsf{d}/2}. \tag{9}$$

Inequality (4) and the theorem are thus proved. □

3 A result from Sperner theory

In this section, we prove the following proposition, which will be used in the proof of Proposition 2.4.

3.1. Proposition. *Let $X=X_1\cup\dots\cup X_{\mathsf{d}}$ ($\mathsf{d}\in\mathbb{N}$) be the union of finite (may be empty) pairwise disjoint sets X_1, …, X_{d}. Let $\mathcal{A}\subset 2^X$ be a family of subsets of X such that for any $S,T\in\mathcal{A}$ there exists $j\in\{1,\dots,\mathsf{d}\}$ for which $(S\cap X_j)\not\subset(T\cap X_j)$ and $(T\cap X_j)\not\subset(S\cap X_j)$. Then*

$$|\mathcal{A}| \;\leq\; \prod_{i=1}^{\mathsf{d}}\binom{|X_i|}{\left\lfloor\frac{|X_i|+1}{2}\right\rfloor}.$$

As it is common in combinatorics, Proposition 3.1 has a number of interpretations[6] and relates to a number of deep theorems[7]. It is related to the *Sperner theory* (see, e. g., [7]), to the *theory of perfect graphs* (see, e. g., [1]), etc. As a consequence, there are various ways to prove the proposition. Despite the Sperner theory seems to be the most relevant one, we will prove our proposition in the more general settings of graph theory.

Recall that a *simple graph* (i. e., an undirected graph without loops or multiple edges) is a pair $\Gamma=(V,E)$, where $V=V(\Gamma)$ is the set of *vertices* of Γ and $E=E(\Gamma)$ (the set of *edges*) is a subset of the set of all unordered pairs of elements of V. Two vertices x, y of Γ are *adjacent* if $\{x,y\}\in E(\Gamma)$. Throughout this section, by a graph we mean a simple graph. A graph Γ is said to be a *comparability graph* if there exists a partial order on $V(\Gamma)$ such that the vertices of each edge in $E(\Gamma)$ are comparable with respect to this order.

A *clique* in a graph Γ is a set $U\subset V(\Gamma)$ of pairwise adjacent vertices, and an *independent set* (*stable set*, *anticlique*) is a set of pairwise non-adjacent vertices. The set of all cliques (resp., anticliques) of a graph Γ is denoted by $\mathfrak{C}(\Gamma)$ (resp., $\mathfrak{A}(\Gamma)$). For a finite graph Γ, let $\alpha(\Gamma)$ be the number of vertices in the largest independent set of Γ ($\alpha(\Gamma)$ is called the *independence number* or *stability number*). Let $\theta(\Gamma)$ denote the *clique covering number* of Γ, i. e., $\theta(\Gamma)$ is the least number of cliques which cover all the vertices of Γ. Clearly, for any finite graph Γ we have

$$\alpha(\Gamma) \;\leq\; \theta(\Gamma) \tag{10}$$

since every clique of Γ has at most one vertex in each independent set of Γ.

3.2. Theorem (Dilworth). *For each comparability graph Γ we have*

$$\alpha(\Gamma) \;=\; \theta(\Gamma).$$

[6] See, e. g., the formula for random walks, which appears in the proof of Proposition 2.4.

[7] For example, in the case where $\mathsf{d}=1$, the proposition is the Sperner theorem.

3.3. Remark. The equality $\alpha(\Gamma) = \theta(\Gamma)$ holds true for every perfect graph. (This follows from the *Perfect Graph Theorem* (Lovász 1972), which states that a graph is perfect if and only if its complement is perfect.) Every comparability graph is perfect (Mirsky's theorem).

Recall that the *normal* (or *strong*) *product* $\Gamma \boxtimes \Delta$ of two graphs Γ and Δ is a graph with vertex set $V(\Gamma) \times V(\Delta)$; two distinct pairs (x_1, y_1) and (x_2, y_2), where $x_1, x_2 \in V(\Gamma)$ and $y_1, y_2 \in V(\Delta)$, are adjacent in $\Gamma \boxtimes \Delta$ if and only if x_1 is equal or adjacent to x_2, and y_1 is equal or adjacent to y_2. Note that the normal product is an associative operation.

3.4. Lemma. *Let Γ and Δ be finite simple graphs. Then*

$$\alpha(\Gamma)\alpha(\Delta) \overset{E1}{\leq} \alpha(\Gamma \boxtimes \Delta) \overset{E2}{\leq} \theta(\Gamma \boxtimes \Delta) \overset{E3}{\leq} \theta(\Gamma)\theta(\Delta).$$

Proof. Observe that the product $A \times B$ of independent sets $A \in \mathfrak{A}(\Gamma)$ and $B \in \mathfrak{A}(\Delta)$ is an independent set in $\Gamma \boxtimes \Delta$. This obviously implies inequality E1. Inequality E2 is a particular case of (10). In order to check E3, we observe that the product $C_1 \times C_2$ of cliques $C_1 \in \mathfrak{C}(\Gamma)$, $C_2 \in \mathfrak{C}(\Delta)$ is a clique in $\Gamma \boxtimes \Delta$. Consequently, if $\mathcal{C}_1 \subset \mathfrak{C}(\Gamma)$ and $\mathcal{C}_2 \subset \mathfrak{C}(\Delta)$ are minimal "covering" families of cliques (such that $|\mathcal{C}_1| = \theta(\Gamma)$ and $|\mathcal{C}_2| = \theta(\Delta)$), then $\{C_1 \times C_2 : C_1 \in \mathcal{C}_1, C_2 \in \mathcal{C}_2\}$ is a family of cliques that covers all the vertices of $\Gamma \boxtimes \Delta$ and contains $\theta(\Gamma)\theta(\Delta)$ cliques, whence $\theta(\Gamma \boxtimes \Delta) \leq \theta(\Gamma)\theta(\Delta)$. □

3.5. Corollary. *Let $\mathsf{d} \in \mathbb{N}$ and let $\Gamma_1, \ldots, \Gamma_{\mathsf{d}}$ be finite simple graphs with $\alpha(\Gamma_i) = \theta(\Gamma_i)$ for each i. Then*

$$\prod_{i=1}^{\mathsf{d}} \alpha(\Gamma_i) \;=\; \alpha(\Gamma_1 \boxtimes \ldots \boxtimes \Gamma_{\mathsf{d}}) \;=\; \theta(\Gamma_1 \boxtimes \ldots \boxtimes \Gamma_{\mathsf{d}}) \;=\; \prod_{i=1}^{\mathsf{d}} \theta(\Gamma_i).$$

Proof. This follows from Lemma 3.4 by induction on d. □

For a finite set Z, let B_Z denote the graph with the set of vertices $V(B_Z) := 2^Z$ and the set of edges $E(B_Z)$ consisting of all the pairs $\{S, T\}$, $S \neq T \in 2^Z$, for which $S \subset T$ or $T \subset S$.

3.6. Theorem (Sperner). *For every finite set Z we have*

$$\alpha(B_Z) = \binom{|Z|}{\left\lfloor \frac{|Z|+1}{2} \right\rfloor}.$$

Proof of Proposition 3.1. Consider the graphs $B_{X_1}, \ldots, B_{X_{\mathsf{d}}}$ and their normal product $B_{X_1} \boxtimes \ldots \boxtimes B_{X_{\mathsf{d}}}$. We have a natural one-to-one correspondence between sets 2^X and $V(B_{X_1} \boxtimes \ldots \boxtimes B_{X_{\mathsf{d}}})$. Observe that, under this correspondence, the family $\mathcal{A}$ of the proposition is an anticlique in the graph $B_{X_1} \boxtimes \ldots \boxtimes B_{X_{\mathsf{d}}}$. Consequently, we have

$$|\mathcal{A}| \;\leq\; \alpha(B_{X_1} \boxtimes \ldots \boxtimes B_{X_{\mathsf{d}}}). \tag{11}$$

Since $B_{X_1}, \ldots, B_{X_{\mathsf{d}}}$ are comparability graphs, it follows by Dilworth's theorem (Theorem 3.2) and Corollary 3.5 that

$$\alpha(B_{X_1} \boxtimes \ldots \boxtimes B_{X_{\mathsf{d}}}) \;=\; \prod_{i=1}^{\mathsf{d}} \alpha(B_{X_i}). \tag{12}$$

By Sperner's theorem (Theorem 3.6) we have

$$\prod_{i=1}^{\mathsf{d}} \alpha(B_{X_i}) \;=\; \prod_{i=1}^{\mathsf{d}} \binom{|X_i|}{\lfloor \frac{|X_i|+1}{2} \rfloor}. \tag{13}$$

The desired inequality follows from (11)–(13). □

4 Proof of Proposition 2.4

Proof. Our proof consists of two parts.

Part 1. First, we prove that the proposition holds true in the case where the measure ν is uniform on the set $Y = Y_1 \cup \cdots \cup Y_{\mathsf{d}}$ (i. e., $\nu(a) = \nu(Y)/|Y|$ for each $a \in Y$).

Part 2. We show that the general case reduces to the case where ν is uniform on Y.

Part 1. Assume that ν is uniform on Y and adopt the following notation:

$$\lambda := \frac{\nu(Y)}{|Y|} = \frac{\nu(Y)}{2\mathsf{d}},$$
$$Y_0 := A \smallsetminus (Y),$$
$$\nu_0 := \nu(Y_0) = 1 - 2\mathsf{d}\lambda.$$

Let us show that for each $n \in \mathbb{N}$ and for every set $L \subset A^n$ without $\mathcal{Y}$-special pairs the following inequality holds:

$$\nu^n(L) \;\le\; \mathcal{C}(\mathsf{d}, n, \nu_0), \qquad \text{where} \quad \mathcal{C}(\mathsf{d}, n, \nu_0) := \sum_{\substack{n_0, n_1, \ldots, n_{\mathsf{d}} \\ n_0 + n_1 + \cdots + n_{\mathsf{d}} = n \\ n_0, n_1, \ldots, n_{\mathsf{d}} \ge 0}} \left(\nu_0^{n_0} \lambda^{n - n_0} \cdot \binom{n}{n_0, n_1, \ldots, n_{\mathsf{d}}} \cdot \prod_{i=1}^{\mathsf{d}} \binom{n_i}{\lfloor \frac{n_i+1}{2} \rfloor} \right). \tag{14}$$

It is sufficient to consider the case where $Y_0 \neq \varnothing$ (i. e., $A \neq Y$) because the case $A = Y$ transforms to the former one via passing to the set $A' := A \cup \{w\}$ with a new element $w \notin A$ and assigning $\nu(w) = 0$. In order to prove (14), we split A^n into $(\mathsf{d}+1)^n$ classes of the form

$$Y_{i_1} \times Y_{i_2} \times \cdots \times Y_{i_n}, \quad \text{where} \quad i_j \in \{0, 1, \ldots, \mathsf{d}\}.$$

(This is possible because we assume that Y_i's are pairwise disjoint, while $A = Y_0 \cup Y_1 \cup \cdots \cup Y_{\mathsf{d}}$ by definition of Y_0.) If $K = Y_{i_1} \times Y_{i_2} \times \cdots \times Y_{i_n}$ is a class of this partition and $\ell \in \{0, 1, \ldots, \mathsf{d}\}$, we set

$$I_\ell(K) := \{j \in \{1, \ldots, n\} \,:\, i_j = \ell\},$$
$$n_\ell(K) := |I_\ell(K)|.$$

In order to prove (14), let us show that for an arbitrary class K of the above partition of A^n we have

$$\nu^n(L \cap K) \;\le\; \nu_0^{n_0(K)} \lambda^{n - n_0(K)} \prod_{i=1}^{\mathsf{d}} \binom{n_i(K)}{\lfloor \frac{n_i(K)+1}{2} \rfloor}. \tag{15}$$

In order to prove (15), we split K into subclasses in the following way: we let two sequences $(w_1,\dots,w_n)$ and $(w'_1,\dots,w'_n)$ from K be in one and the same subclass if and only if $w_i = w'_i$ for each $i \in I_0(K)$. (Thus, each subclass of K consists of $2^{n-n_0(K)}$ elements. If A is finite, then K splits into $|Y_0|^{n_0(K)}$ subclasses.) If J is a subclass of K and $W = (w_1,\dots,w_n) \in J$, then the elements w_i, $i \in I_0(K)$, are determined by J and do not depend on $W \in J$. It follows that the value $\prod_{i\in I_0(K)} \nu(w_i)$ is determined by J. We set $\nu_0^*(J) := \prod_{i\in I_0(K)} \nu(w_i)$. Then we have

$$\nu^n(L\cap J) \;=\; \nu_0^*(J)\cdot\lambda^{n-n_0(K)}\cdot|L\cap J|. \tag{16}$$

Proposition 3.1 implies that for each subclass J of the class K we have

$$|L\cap J| \;\le\; \prod_{i=1}^{\mathsf{d}} \binom{n_i(K)}{\left\lfloor\frac{n_i(K)+1}{2}\right\rfloor}. \tag{17}$$

(In order to see this in terms of Proposition 3.1, assign $X_i := I_i(K)$ for each $i \in \{1,\dots,\mathsf{d}\}$. Then, for each $i \in \{1,\dots,\mathsf{d}\}$, choose an element y_i in the pair Y_i, and let $\mathscr{F} : J \to 2^X$ be the bijection defined by

$$\mathscr{F}\big(w_1,\dots,w_n\big) = \{i \in I_1(K)\cup\dots\cup I_{\mathsf{d}}(K) \,:\, w_i \in \{y_1,\dots,y_{\mathsf{d}}\}\}\,.$$

Since L has no $\mathcal{Y}$-special pairs, it follows that the image $\mathcal{A} := \mathscr{F}(L\cap J)$ satisfies the requirement of Proposition 3.1 (cf. the definition of $\mathcal{Y}$-special pairs with this requirement). Therefore, the inequality of Proposition 3.1 gives us inequality (17).)

Let $\widehat{K}$ denote the set of all subclasses of the class K. Then (16) and (17) yield

$$\nu^n(L\cap K) \;=\; \sum_{J\in\widehat{K}} \nu^n(L\cap J) \le \left(\sum_{J\in\widehat{K}} \nu_0^*(J)\right)\cdot\lambda^{n-n_0(K)}\cdot\prod_{i=1}^{\mathsf{d}} \binom{n_i(K)}{\left\lfloor\frac{n_i(K)+1}{2}\right\rfloor}. \tag{18}$$

At the same time, we clearly have

$$\sum_{J\in\widehat{K}} \nu_0^*(J) \;=\; \prod_{i\in I_0(K)} \sum_{w\in Y_0} \nu(w) \;=\; \nu_0^{n_0(K)}. \tag{19}$$

Inequality (15) directly follows from (18) and (19). The required estimate (14) readily follows from (15) by summing over all classes of the partition.

In order to complete the proof of Part 1, it suffices to show that there exists a constant $C := C(\nu)$ such that for each $n \in \mathbb{N}$ we have $\mathcal{C}(\mathsf{d},n,\nu_0) < Cn^{-\mathsf{d}/2}$. This property may be easily derived from well-known facts of the theory of random walks on integer lattices. Let $\theta = \theta(\mathsf{d},\nu_0)$ be the probability measure on $\mathbb{Z}^{\mathsf{d}}$ defined by

$$\begin{gathered}\theta(\mathbf{e}_i) = \theta(-\mathbf{e}_i) = \lambda, \qquad i = 1,\dots,\mathsf{d};\\ \theta(\mathbf{0}) = \nu_0 (= 1-2\mathsf{d}\lambda),\end{gathered}$$

and let

$$D := \{(z_1,\dots,z_{\mathsf{d}}) \in \mathbb{Z}^{\mathsf{d}} \,:\, z_i \in \{0,1\} \text{ for each } i \in \{1,\dots,\mathsf{d}\}\}.$$

Then it is obvious that for each $n \in \mathbb{N}$ we have $\theta^{*n}(D) = \mathcal{C}(\mathsf{d},n,\nu_0)$, i. e., the value $\mathcal{C}(\mathsf{d},n,\nu_0)$ is equal to the probability that the random walk in $\mathbb{Z}^{\mathsf{d}}$ with distribution θ will hit D at the nth step. Since θ is nondegenerate in $\mathbb{Z}^{\mathsf{d}}$, there exists a constant

$N > 0$ such that $\theta^{*n}(z) < Nn^{-\mathsf{d}/2}$ for all $z \in \mathbb{Z}^{\mathsf{d}}$, $n \in \mathbb{N}$ (see, e. g., [15, p. 72]). Since D consists of 2^{d} elements of $\mathbb{Z}^{\mathsf{d}}$, we have $\mathcal{C}(\mathsf{d}, n, \nu_0) < 2^{\mathsf{d}} N n^{-\mathsf{d}/2}$ for each $n \in \mathbb{N}$.

Part 2. Let $\mathcal{Y} =: \{\{a_1, b_1\}, \ldots, \{a_{\mathsf{d}}, b_{\mathsf{d}}\}\}$ so that $Y = \{a_1, b_1, \ldots, a_{\mathsf{d}}, b_{\mathsf{d}}\}$. Let $Y' := \{a'_1, b'_1, \ldots, a'_{\mathsf{d}}, b'_{\mathsf{d}}\}$ be a set of $2\mathsf{d}$ elements not in A, and let $A' := A \cup Y'$. Let $\nu_m := \min_{x \in Y} \nu(x)$. (Note that $\nu_m > 0$ since $\mathrm{supp}(\nu) \supset Y$.) Let ν' be the probability measure on A' defined by

$\nu'(x) := \nu(x)$ if $x \in A \setminus Y$,

$\nu'(x) := \nu_m$ if $x \in Y$, and

$\nu'(x') := \nu(x) - \nu_m$ if $x' \in Y'$, where x is the element in Y that corresponds to x'.

Let f denote the map $A' \to A$ of "forgetting the primes" for elements (i. e., $f(w) = w$ if $w \in A$ and $f(w') = w$ if $w' \in Y'$ and w is the element in Y that corresponds to w'). Let $F : (A')^n \to A^n$ denote the map of "forgetting the primes" for sequences (i. e., $F(w_1, \ldots, w_n) = (f(w_1), \ldots, f(w_n))$). (If $W := (w_1, \ldots, w_n)$ is a sequence over A with precisely $t \in \mathbb{N}_0$ occurrences of elements from Y, then $F^{-1}(W)$ is a set of 2^t sequences over A'.)

The following two claims are obvious. (In order to check the second one, it is enough to notice that F sends each $\mathcal{Y}$-special pair to a $\mathcal{Y}$-special pair.)

Claim 2.1. *For any $n \in \mathbb{N}$, $S \subset A^n$, we have $(\nu')^n(F^{-1}(S)) = \nu^n(S)$.*

Claim 2.2. *If $L \subset A^n$ is a subset without $\mathcal{Y}$-special pairs, then $F^{-1}(L) \subset (A')^n$ is also a subset without $\mathcal{Y}$-special[8] pairs.*

Since ν' is homogeneous on Y, it follows by Part 1 of this proof that there exists a constant $C' := C'(\nu')$ such that for each $n \in \mathbb{N}$ and for any subset $L' \subset (A')^n$ without $\mathcal{Y}$-special pairs we have

$$(\nu')^n(L') \;<\; C' n^{-\mathsf{d}/2}.$$

Therefore, by Claims 2.1 and 2.2, for each $n \in \mathbb{N}$ and for any subset $L \subset A^n$ without $\mathcal{Y}$-special pairs we have

$$\nu^n(L) \;=\; (\nu')^n(F^{-1}(L)) \;<\; C' n^{-\mathsf{d}/2}.$$

Proposition 2.4 is thus proved. □

4.1. Remark. The statement of Proposition 2.4 holds true in the (more general) case where the family $\mathcal{Y} = \{Y_1, \ldots, Y_{\mathsf{d}}\}$ consists of d pairwise distinct (not necessarily disjoint) subsets. This may be proved by an argument similar to the one from Part 2 above.

Let A, d, $\mathcal{Y} = \{Y_1, \ldots, Y_{\mathsf{d}}\}$, ν be as in Proposition 2.4, and assume that Y_1, $\ldots$, Y_{d} are pairwise distinct but not disjoint ($|Y| < 2\mathsf{d}$). Let Z_Y be the set of $\mathsf{d}|Y|$ elements not in A that contains the elements $w_{[1]}, \ldots, w_{[\mathsf{d}]}$ for each $w \in Y$. Set $A^\dagger := A \setminus Y \cup Z_Y$. Let $\nu^\dagger$ be the probability measure on $A^\dagger$ defined by

$\nu^\dagger(x) := \nu(x)$ if $x \in A^\dagger \setminus Z_Y = A \setminus Y$,

$\nu^\dagger(x_{[i]}) := \nu(x)/\mathsf{d}$ if $x_{[i]} \in Z_Y$, where x is the element in Y that corresponds to $x_{[i]}$.

[8] In $(A')^n$, we consider special pairs with respect to Y, not with respect to $Y \cup Y'$.

Let $\mathcal{Y}^\dagger = \{Y_1^\dagger, \dots, Y_{\mathsf{d}}^\dagger\}$ be the family of d two-element subsets of Z_Y defined as follows: we set $Y_i^\dagger := \{v_{[i]}, w_{[i]}\}$ if $Y_i = \{v, w\}$. It is obvious that the set $Y^\dagger := Y_1^\dagger \cup \dots \cup Y_{\mathsf{d}}^\dagger$ consists of $2\mathsf{d}$ elements and $\mathcal{Y}^\dagger$ consists of d pairwise disjoint subsets. (Observe also that $\operatorname{supp}(\nu^\dagger) \supset Z_Y \supset Y^\dagger$.)

Let f denote the map $A^\dagger \to A$ of "forgetting the indexes" for elements in Z_Y (i. e., $f(w_{[j]}) = w$ if $w_{[j]} \in Z_Y$ and w is the element in Y that corresponds to $w_{[j]}$; $f(w) = w$ if $w \in A^\dagger \smallsetminus Z_Y = A \smallsetminus Y$). Let $F : (A^\dagger)^* \to A^*$ denote the corresponding map for sequences (i. e., $F(w_1, \dots, w_k) = (f(w_1), \dots, f(w_k))$).

Here, we can use claims, similar to Claims 2.1 and 2.2:

Claim 3.1. *For any $n \in \mathbb{N}$, $S \subset A^n$, we have $(\nu^\dagger)^n(F^{-1}(S)) = \nu^n(S)$.*

Claim 3.2. *If $L \subset A^n$ is a subset without $\mathcal{Y}$-special pairs, then $F^{-1}(L) \subset (A^\dagger)^n$ is a subset without $\mathcal{Y}^\dagger$-special pairs.*

Claim 3.1 is obvious. Let us prove Claim 3.2. It is enough to show that F sends each $\mathcal{Y}^\dagger$-special pair to a $\mathcal{Y}$-special pair. If $V = (v_1, \dots, v_k)$, $W = (w_1, \dots, w_k)$ is a $\mathcal{Y}^\dagger$-special pair in $(A^\dagger)^n$, then by definition we have

(a) $v_i = w_i$ whenever $\{v_i, w_i\} \notin \mathcal{Y}^\dagger$,

(b) there are no $i, j \in \{1, \dots, k\}$ such that $v_i = w_j$, $w_i = v_j$, and $\{v_i, w_i\} = \{v_j, w_j\} \in \mathcal{Y}^\dagger$.

The definition of $\mathcal{Y}^\dagger$ yields that f sends $\mathcal{Y}^\dagger$ to $\mathcal{Y}$ (bijectively). In view of (a) this implies that

(A) $f(v_i) = f(w_i)$ whenever $\{f(v_i), f(w_i)\} \notin \mathcal{Y}$.

Furthermore, condition (a) implies that $\{v_i, w_i\}$ is in $\mathcal{Y}^\dagger$ if and only if $\{f(v_i), f(w_i)\}$ is in $\mathcal{Y}$. (Indeed, if $\{v_i, w_i\}$ is in $\mathcal{Y}^\dagger$ then $\{f(v_i), f(w_i)\}$ is in $\mathcal{Y}$ by definitions of $\mathcal{Y}^\dagger$ and f; if $\{f(v_i), f(w_i)\}$ is in $\mathcal{Y}$ then $\{v_i, w_i\}$ is in $\mathcal{Y}^\dagger$ because otherwise we have $v_i = w_i$ by condition (a) whence $f(v_i) = f(w_i)$.) This fact shows that condition (b) directly implies the following

(B) There is no $i \neq j \in \{1, \dots, k\}$ such that $f(v_i) = f(w_j)$, $f(w_i) = f(v_j)$, and $\{f(v_i), f(w_i)\} = \{f(v_j), f(w_j)\} \in \mathcal{Y}$.

Conditions (A) and (B) mean exactly that $F(V) = f(v_1) \dots f(v_k)$, $F(W) = f(w_1) \dots f(w_k)$ is a $\mathcal{Y}$-special pair. Claim 3.2 is thus proved.

Since $\mathcal{Y}^\dagger$ consists of d pairwise disjoint subsets and $\operatorname{supp}(\nu^\dagger) \supset Z_Y \supset Y^\dagger$, it follows by

Since the quadruple $(\mathsf{d}, A^\dagger, \mathcal{Y}^\dagger, \nu^\dagger)$ satisfies condition $(\mathrm{C}_{\mathcal{Y}})$, it follows by Proposition 2.4 that there exists a constant $C^\dagger := C^\dagger(\nu^\dagger)$ such that for each $n \in \mathbb{N}$ and for any subset $L^\dagger \subset (A^\dagger)^n$ without $\mathcal{Y}^\dagger$-special pairs we have

$$(\nu^\dagger)^n(L^\dagger) \;<\; C^\dagger n^{-\mathsf{d}/2}.$$

Therefore, by Claims 3.1 and 3.2, for each $n \in \mathbb{N}$ and for any subset $L \subset A^n$ without $\mathcal{Y}$-special pairs we have

$$\nu^n(L) \;=\; (\nu^\dagger)^n(F^{-1}(L)) \;<\; C^\dagger n^{-\mathsf{d}/2}.$$

5 Proof of Proposition 2.5

In order to prove Proposition 2.5, we introduce the following auxiliary notion.

5.1. Definition. Let $n, m \in \mathbb{N}$, and let $(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ be an m-tuple of vectors in $\mathbb{R}^n$. We define the characteristic $\mathcal{V}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ as follows:

$$\mathcal{V}(\mathbf{v}_1, \ldots, \mathbf{v}_m) := \inf_{(t_1,\ldots,t_m)\in\mathbb{R}^m\setminus\{\mathbf{0}\}} \frac{\left\|\sum_{i=1}^m t_i \cdot \mathbf{v}_i\right\|_1}{\|(t_1, \ldots, t_m)\|_1}.$$

5.2. Properties of $\mathcal{V}$. We give several simple properties of the characteristic $\mathcal{V}$.

1. For any $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$ and $t_1, \ldots, t_m \in \mathbb{R}$ we have, by definition,

$$\left\|\sum_{i=1}^m t_i \cdot \mathbf{v}_i\right\|_1 \;\geq\; \mathcal{V}(\mathbf{v}_1, \ldots, \mathbf{v}_m)\,\|(t_1, \ldots, t_m)\|_1 \;=\; \mathcal{V}(\mathbf{v}_1, \ldots, \mathbf{v}_m)\sum_{i=1}^m |t_i|. \tag{20}$$

2. $\mathcal{V}(\mathbf{v}_1, \ldots, \mathbf{v}_m) = 0$ if and only if the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly dependent.
3. For any $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$ and $\delta_1, \ldots, \delta_m \in \{+1, -1\}$ we have

$$\mathcal{V}(\delta_1 \cdot \mathbf{v}_1, \ldots, \delta_m \cdot \mathbf{v}_m) \;=\; \mathcal{V}(\mathbf{v}_1, \ldots, \mathbf{v}_m).$$

5.3. Claim. *Let $n, m \in \mathbb{N}$, let $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, and let $(\mathbf{x}_1, \ldots, \mathbf{x}_k)$, $k \in \mathbb{N}$, be a sequence of vectors in $\mathbb{R}^n$ such that $\mathbf{x}_i \in \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ for each $i \in \{1, \ldots, k\}$. Then we have*

$$\left\|\sum_{i=1}^k \mathbf{x}_i\right\|_1 \;\geq\; k \cdot \mathcal{V}(\mathbf{v}_1, \ldots, \mathbf{v}_m) \;\geq\; \mathcal{V}(\mathbf{v}_1, \ldots, \mathbf{v}_m).$$

Proof. This readily follows from (20). □

5.4. Lemma. *Let G be a group and let $\Phi : G \to \mathbb{R}^n$ be an $\mathbb{R}^n$-quasimorphism with defect d_Φ. Let $k \in \mathbb{N}$ and let g_0 g_1, …, g_k, h_1, …, h_k be elements of G. Then*

$$\left\|\Phi(g_0 h_1 g_1 \cdots h_k g_k) - \Phi(g_0 g_1 \cdots g_k)\right\|_1 \;\geq\; \left\|\sum_{i=1}^k \Phi(h_i)\right\|_1 - 3k d_\Phi.$$

Proof. From the definition of defect, it follows by induction that we have

$$\left\|\Phi(g_0 h_1 g_1 \cdots h_k g_k) - \sum_{i=0}^k \Phi(g_i) - \sum_{i=1}^k \Phi(h_i)\right\|_1 \;\leq\; 2k d_\Phi,$$

$$\left\|\Phi(g_0 g_1 \cdots g_k) - \sum_{i=0}^k \Phi(g_i)\right\|_1 \;\leq\; k d_\Phi.$$

Consequently, we have

$$\left\|\Phi(g_0 h_1 g_1 \cdots h_k g_k) - \Phi(g_0 g_1 \cdots g_k) - \sum_{i=1}^k \Phi(h_i)\right\|_1 \;\leq\; 3k d_\Phi,$$

which obviously implies the statement of the lemma. □

5.5. Lemma. *Let G be a group and let $\Phi : G \to \mathbb{R}^n$, $n \in \mathbb{N}$, be an $\mathbb{R}^n$-quasimorphism with defect d_Φ. Let c_1, …, c_m be elements of G. Let $k \in \mathbb{N}$ and let $(g_0, g_1, \ldots, g_k)$, $(h_1, \ldots, h_k)$ be sequences of elements of G such that $h_i \in \{c_1, \ldots, c_m\}$ for each $i \in \{1, \ldots, k\}$. Then we have*

$$\left\|\Phi(g_0 h_1 g_1 \cdots h_k g_k) - \Phi(g_0 g_1 \cdots g_k)\right\|_1 \;\geq\; k \cdot \left(\mathcal{V}\left(\Phi(c_1), \ldots, \Phi(c_m)\right) - 3d_\Phi\right), \tag{21}$$

whence it follows that

$$\left\|\Phi(g_0 h_1 g_1 \cdots h_k g_k) - \Phi(g_0 g_1 \cdots g_k)\right\|_1 \;\geq\; \mathcal{V}\left(\Phi(c_1), \ldots, \Phi(c_m)\right) - 3d_\Phi. \tag{22}$$

Proof. By Lemma 5.4, we have

$$\left\|\Phi(g_0h_1g_1\cdots h_kg_k)-\Phi(g_0g_1\cdots g_k)\right\|_1 \;\geq\; \left\|\sum_{i=1}^{k}\Phi(h_i)\right\|_1 - 3kd_\Phi. \tag{23}$$

By Claim 5.3 we have

$$\left\|\sum_{i=1}^{k}\Phi(h_i)\right\|_1 \;\geq\; k\cdot\mathcal{V}\left(\Phi(c_1),\ldots,\Phi(c_m)\right). \tag{24}$$

Inequality (21) follows from (23) and (24). Inequality (22) follows from (21) because $\|\Phi(g_0h_1g_1\cdots h_kg_k)-\Phi(g_0g_1\cdots g_k)\|_1\geq 0$ and $k\in\mathbb{N}$. □

5.6. Lemma. *Let G be a group and let $\overline{\Phi}: G\to\mathbb{R}^n$, $n\in\mathbb{N}$, be a homogeneous $\mathbb{R}^n$-quasimorphism with defect $d_{\overline{\Phi}}$. Let $\mathcal{Y}=\{\{a_1,b_1\},\ldots,\{a_m,b_m\}\}$, $m\in\mathbb{N}$, be a family of two-element subsets of G. Set*

$$\mathcal{V}_{\overline{\Phi}}(\mathcal{Y}) := \mathcal{V}\left(\overline{\Phi}(a_1^{-1}b_1),\ldots,\overline{\Phi}(a_m^{-1}b_m)\right).$$

Then for each $t\in\mathbb{N}$ and for every $\mathcal{Y}$-special pair $\{(g_1,\ldots,g_t),(h_1,\ldots,h_t)\}\subset G^t$ we have

$$\|\Phi(g_1\cdots g_t)-\Phi(h_1\cdots h_t)\|_1 \;\geq\; \mathcal{V}_{\overline{\Phi}}(\mathcal{Y})-3d_{\overline{\Phi}}.$$

Proof. Set $g:=g_1\cdots g_t$ and $h:=h_1\cdots h_t$. Since $\{(g_1,\ldots,g_t),(h_1,\ldots,h_t)\}$ is $\mathcal{Y}$-special, there exists a family $(x_1,\ldots,x_m)$ with $x_i\in\{a_i^{-1}b_i,b_i^{-1}a_i\}$ for each i such that for some $k\in\mathbb{N}$ and $g_0',\ldots,g_k'\in G$ we have

$$g=g_0'g_1'\cdots g_k'$$

and

$$h=g_0'z_1g_1'\cdots z_kg_k',$$

where $z_j\in\{x_1,\ldots,x_m\}$ for each $j\in\{1,\ldots,k\}$.

By Lemma 5.5 we have

$$\left\|\overline{\Phi}(h)-\overline{\Phi}(g)\right\|_1 \;\geq\; \mathcal{V}\left(\overline{\Phi}(x_1),\ldots,\overline{\Phi}(x_m)\right)-3d_{\overline{\Phi}}.$$

It remains to observe that, since $\overline{\Phi}$ is homogeneous and $x_i\in\{a_i^{-1}b_i,b_i^{-1}a_i\}$ for each i, we have $\mathcal{V}\left(\overline{\Phi}(x_1),\ldots,\overline{\Phi}(x_m)\right)=\mathcal{V}_{\overline{\Phi}}(\mathcal{Y})$ (see properties of $\mathcal{V}$ in 5.2). □

5.7. Lemma. *Let G be a group and let $\Phi: G\to\mathbb{R}^{\mathsf{d}}$, $\mathsf{d}\in\mathbb{N}$, be a nondegenerate $\mathbb{R}^{\mathsf{d}}$-quasimorphism. Then for any $r>0$ there exists a d-set $\{g_1,\ldots,g_{\mathsf{d}}\}\subset G$ such that*

$$\mathcal{V}_\Phi(g_1,\ldots,g_{\mathsf{d}})\geq r. \tag{25}$$

Moreover, if $S\subset G$ is a subset generating G as a semigroup, then for any $r>0$ there exist $p\in\mathbb{N}$ and a d-set $\{g_1,\ldots,g_{\mathsf{d}}\}$ of elements from S^p such that (25) holds.

Proof. By Lemma 1.2, the image $\Phi(G)$ is cobounded in $\mathbb{R}^{\mathsf{d}}$. This means that there exists $\varepsilon>0$ such that for each $\mathbf{w}\in\mathbb{R}^{\mathsf{d}}$ there is $\mathbf{v}\in\Phi(G)$ with $\|\mathbf{w}-\mathbf{v}\|_1\leq\varepsilon$. Let $\mathbf{v}_1,\ldots,\mathbf{v}_{\mathsf{d}}$ be points from $\Phi(G)$ such that

$$\|\mathbf{v}_i-(r+\varepsilon)\cdot\mathbf{e}_i\|_1\leq\varepsilon\quad(i=1,\ldots,\mathsf{d}).$$

Then for any $t_1, \dots, t_{\mathsf{d}} \in \mathbb{R}$ we have

$$\left\| \sum_{i=1}^{\mathsf{d}} t_i \cdot \mathbf{v}_i \right\|_1 \;\geq\; \left\| \sum_{i=1}^{\mathsf{d}} t_i (r+\varepsilon) \cdot \mathbf{e}_i \right\|_1 - \left\| \sum_{i=1}^{\mathsf{d}} \left(t_i (r+\varepsilon) \cdot \mathbf{e}_i - t_i \cdot \mathbf{v}_i \right) \right\|_1$$
$$\geq\; (r+\varepsilon) \sum_{i=1}^{\mathsf{d}} |t_i| - \varepsilon \sum_{i=1}^{\mathsf{d}} |t_i| \;=\; r \sum_{i=1}^{\mathsf{d}} |t_i|.$$

This means that $\mathcal{V}(\mathbf{v}_1, \dots, \mathbf{v}_{\mathsf{d}}) \geq r$. It remains to choose $g_i \in \Phi^{-1}(\mathbf{v}_i)$ for each $i \in \{1, \dots, \mathsf{d}\}$.

Now, let $S \subset G$ be a subset generating G as a semigroup. Let us show that for each $m \in \mathbb{N}$ the image $\Phi\left(\bigcup_{k \in \mathbb{N}} S^{mk}\right)$ is cobounded in $\mathbb{R}^{\mathsf{d}}$. Observe that the set $m \cdot \Phi(G)$ is cobounded in $\mathbb{R}^{\mathsf{d}}$ since $\Phi(G)$ is cobounded in $\mathbb{R}^{\mathsf{d}}$. Recall that for any $g \in G$ we have

$$\left\| \Phi(g^m) - m \cdot \Phi(g) \right\|_1 \;\leq\; (m-1) d_\Phi,$$

where d_Φ is the defect of Φ. Therefore, the set $\bigcup_{g \in G} \Phi(g^m)$ is also cobounded in $\mathbb{R}^{\mathsf{d}}$. Since S generates G as a semigroup, we have

$$\bigcup_{k \in \mathbb{N}} S^k = \bigcup_{k \in \mathbb{N} \cup \{0\}} S^k = G.$$

Therefore, we have

$$\bigcup_{g \in G} g^m \;=\; \bigcup_{g \in \bigcup_{k \in \mathbb{N}} S^k} g^m \;\subset\; \bigcup_{k \in \mathbb{N}} \left(S^k\right)^m \;=\; \bigcup_{k \in \mathbb{N}} \left(S^m\right)^k.$$

It is thus shown that for each $m \in \mathbb{N}$ the image $\Phi\left(\bigcup_{k \in \mathbb{N}} S^{mk}\right)$ is cobounded in $\mathbb{R}^{\mathsf{d}}$.

Obviously, there exists $m_0 \in \mathbb{N}$ such that S^{m_0} contains the identity of the group. By the above, the image $\Phi\left(\bigcup_{k \in \mathbb{N}} S^{m_0 k}\right)$ is cobounded in $\mathbb{R}^{\mathsf{d}}$. Then the argument from the first part of the proof shows that there exists a d-set $\{g_1, \dots, g_{\mathsf{d}}\}$ of elements from $\bigcup_{k \in \mathbb{N}} S^{m_0 k}$ such that (25) holds. At the same time, we have $S^{m_0} \subset S^{2m_0} \subset S^{3m_0} \subset \cdots$ because $S^{m_0} \ni e$. Therefore, there exists $q \in \mathbb{N}$ such that $S^{q m_0}$ contains all the elements $g_1, \dots, g_{\mathsf{d}}$. It remains to set $p := q m_0$. □

Proof of Proposition 2.5. It is clearly enough to prove the second statement of the proposition, which is stronger than the first one. Let $S \subset G$ be a subset generating G as a semigroup. Let $\overline{\Phi}$ be the homogeneous quasimorphism corresponding to Φ (see Lemma 1.1). By Lemma 5.7, there exist $p \in \mathbb{N}$ and a d-set $\{g_1, \dots, g_{\mathsf{d}}\} \subset S^p$ such that

$$\mathcal{V}\left(\overline{\Phi}(g_1), \dots, \overline{\Phi}(g_{\mathsf{d}})\right) \;\geq\; R + 3 d_{\overline{\Phi}} + 2 d_\Phi, \tag{26}$$

where $d_{\overline{\Phi}}$ and d_Φ are the defects of $\overline{\Phi}$ and Φ, respectively. From the proof of Lemma 5.7 it is clear that we may assume without loss of generality that S^p contains the identity e of G. Then S^{2p} contains the set $Y := \{g_1, \dots, g_{\mathsf{d}}, g_1^2, \dots, g_{\mathsf{d}}^2\}$. We set $\mathcal{Y} := \{\{g_1, g_1^2\}, \dots, \{g_{\mathsf{d}}, g_{\mathsf{d}}^2\}\}$ and show that $2p$ and $\mathcal{Y}$ meet the requirements of the proposition. Indeed, suppose that $k \in \mathbb{N}$ and $\{(x_1, \dots, x_k), (y_1, \dots, y_k)\} \subset G^k$ is a $\mathcal{Y}$-special pair. Set $x := x_1 \cdots x_k$ and $y := y_1 \cdots y_k$. Then by Lemma 5.6 we have

$$\left\| \overline{\Phi}(x) - \overline{\Phi}(y) \right\|_1 \;\geq\; \mathcal{V}\left(\overline{\Phi}(g_1^2 g_1^{-1}), \dots, \overline{\Phi}(g_{\mathsf{d}}^2 g_{\mathsf{d}}^{-1})\right) - 3 d_{\overline{\Phi}} \overset{(26)}{\geq} R + 2 d_\Phi. \tag{27}$$

Since $\overline{\Phi}$ is the homogeneous quasimorphism corresponding to Φ, we have $\left\|\overline{\Phi}(x) - \Phi(x)\right\|_1 \leq d_\Phi$ and $\left\|\overline{\Phi}(y) - \Phi(y)\right\|_1 \leq d_\Phi$ (see Lemma 1.1). Consequently,

$$\|\Phi(x) - \Phi(y)\|_1 \;\geq\; \left\|\overline{\Phi}(x) - \overline{\Phi}(y)\right\|_1 - 2d_\Phi \overset{(27)}{\geq} R.$$

It remains to show that the elements $g_1, \ldots, g_{\mathsf{d}}, g_1^2, \ldots, g_{\mathsf{d}}^2$ are pairwise distinct. In order to see this, observe that for any $i, j \in \{1, \ldots, \mathsf{d}\}$ and $r, s \in \mathbb{Z}$ such that $(i, r) \neq (j, s)$ we have

$$\begin{aligned}\left\|\overline{\Phi}\left(g_i^r g_j^{-s}\right)\right\|_1 \;&\geq\; \left\|r \cdot \overline{\Phi}(g_i) - s \cdot \overline{\Phi}(g_j)\right\|_1 - d_{\overline{\Phi}} \\ &\geq\; \mathcal{V}\left(\overline{\Phi}(g_1), \ldots, \overline{\Phi}(g_{\mathsf{d}})\right) - d_{\overline{\Phi}} \overset{(26)}{\geq} R \;>\; 0,\end{aligned}$$

whence it follows that $g_i^r \neq g_j^s$ (because, due to the homogeneity of $\overline{\Phi}$, we have $\overline{\Phi}(e) = \mathbf{0}$ and $\left\|\overline{\Phi}(e)\right\|_1 = 0$). □

References

[1] C. Berge, V. Chvátal (eds.), *Topics on Perfect Graphs*, Math. Stud. **88** (Annals of Discrete Math. **21**), North-Holland, Amsterdam etc., 1984.

[2] M. Bestvina, K. Fujiwara, *Bounded cohomology of subgroups of mapping class groups*, Geom. Topol. **6** (2002), 69–89.

[3] M. Bestvina, K. Fujiwara, *Quasi-homomorphisms on mapping class groups*, Glasnik Matematicki **42**:1 (2007), 213–236.

[4] M. Björklund, T. Hartnick, *Biharmonic functions on groups and limit theorems for quasimorphisms along random walks*, Geom. Topol. **15** (2011), 123–143.

[5] D. Calegari, J. Maher, *Statistics and compression of scl*, preprint 2010. arXiv:1008.4952.

[6] R. P. Dilworth, *A decomposition theorem for partially ordered sets*, Ann. Math. **51**:1 (1950), 161–166.

[7] K. Engel, *Sperner Theory*, Cambridge University Press, Cambridge, New York, 1997.

[8] T. Ito, *Braid ordering and the geometry of closed braid*, preprint 2008. arXiv:0805.1447.

[9] T. Ito, *Braid ordering and knot genus*, preprint 2008. arXiv:0805.2042.

[10] E. Kowalski, *The Large Sieve and its Applications: Arithmetic Geometry, Random Walks and Discrete Groups*, Cambridge Tracts in Math. **175**, Cambridge University Press, Cambridge, 2008.

[11] J. Maher, *Random walks on the mapping class group*, preprint 2006. arXiv:math/0604433.

[12] A. V. Malyutin, N. Yu. Netsvetaev, *Dehornoy's ordering on the braid group and braid moves*, Algebra i Analiz **15**:3 (2003), 170–187; Engl. transl., St. Petersburg Math. J. **15**:3 (2004), 437–448.

[13] I. Rivin, *Counting reducible matrices, polynomials, and surface and free group automorphisms*, preprint 2006. arXiv:math/0604489.

[14] I. Rivin, *Walks on groups, counting reducible matrices, polynomials, and surface and free group automorphisms*, preprint 2007. arXiv:math.NT/0703532.

[15] F. Spitzer, *Principles of Random Walk*. Graduate Texts in Mathematics, vol. **34**, Springer-Verlag, 1976.