THE GEOMETRY OF SPHERES IN FREE ABELIAN GROUPS

MOON DUCHIN, SAMUEL LELIÈVRE, AND CHRISTOPHER MOONEY

Abstract. We study word metrics on $\mathbb{Z}^d$ by developing tools that are fine enough to measure dependence on the generating set. We obtain counting and distribution results for the words of length n. With this, we show that counting measure on spheres always converges to cone measure on a polyhedron (strongly, in an appropriate sense). Using the limit measure, we can reduce probabilistic questions about word metrics to problems in convex geometry of Euclidean space. We give several applications to the statistics of "size-like" functions.

1. Introduction

Suppose one wants to study the density of group elements that have a certain property, or the average value of some statistic in a group. If the group is finitely generated by a set S, then there is an associated word metric on the group that measures how far an element is from the identity, and the ball of radius n is a finite set. Arguably the most natural approach to a density question is to put counting measure on the ball of radius n, measure the proportion of those points with the desired property, and let n tend to infinity.

Given a group G with a fixed finite generating set S (say symmetrized, so that $S = S^{-1}$), let S_n denote the sphere of radius n in the Cayley graph, which is just the set of group elements whose distance from the identity in the word metric is exactly n; that is, they are group elements whose minimal spelling has n letters, or which are reached by geodesics of length n based at the identity. Likewise, B_n is the (closed) ball of radius n. Then a reasonable way to consider the density in G of a property with respect to a generating set is to find the expectation over large balls B_n. Furthermore, one might be interested in understanding the frequency of a property among long words, which amounts to finding the expectation over large spheres S_n—a strictly harder problem, as we will demonstrate.

More generally, we will study the averages $\frac{1}{|B_n|}\sum_{\mathsf{x}\in B_n} f(\mathsf{x})$ and $\frac{1}{|S_n|}\sum_{\mathsf{x}\in S_n} f(\mathsf{x})$, not just for characteristic functions. For some functions we will find that these averages grow on the order of n^k, in which case we normalize the average and seek a limit, or the growth coefficient of n^k.

We will show that averages for "size-like" functions over spheres in $(\mathbb{Z}^d, S)$ must exist; further, the averages can be reduced to integrals on convex polyhedra in Euclidean space, with respect to an appropriate geometrically defined measure. One of the themes will be to study statistics that are a priori dependent on the choice of generating set. In some cases, we will be able to quantify the extent of the dependence; in other cases, we will find that there is no dependence.

Recall that a function $g : \mathbb{R}^d \to \mathbb{R}$ is called *homogeneous (of order k)* if $g(a\mathsf{x}) = a^k g(\mathsf{x})$ for $a \geq 0$. Let us call a function $f : \mathbb{Z}^d \to \mathbb{R}$ *coarsely homogeneous* if there is

Date: August 8, 2011.

some homogeneous function $g : \mathbb{R}^d \to \mathbb{R}$ such that $f \overset{+}{\asymp} g$, meaning that $|g(\mathsf{x}) - f(\mathsf{x})|$ is uniformly bounded over $\mathsf{x} \in \mathbb{Z}^d$. We say that f is *asymptotically homogeneous* (or *size-like*) if there is some homogeneous function such that $f \sim g$, meaning that the ratio $f(\mathsf{x})/g(\mathsf{x}) \to 1$ as $\mathsf{x} \to \infty$. (Here the notation $\mathsf{x} \to \infty$ means that sequences leave all compact sets.) In particular, coarsely homogeneous implies asymptotically homogeneous when $k \geq 1$ and g is nonzero.

For any free abelian group $\mathbb{Z}^d$ with any finite generating set S, let Q be the convex hull of the points corresponding to S in $\mathbb{R}^d$, and let L be its boundary polytope, and let $\hat{A}$ denote the cone from $A \subseteq L$ to the origin, so that $Q = \hat{L}$. For $A \subseteq L$ define the *cone measure* by

$$\mu(A) = \mu_L(A) = \frac{\mathrm{Vol}(\hat{A})}{\mathrm{Vol}(Q)}.$$

This is the Euclidean volume of the cone from A to the origin normalized by the volume of Q, so that $\mu = \mu_L$ is a Borel probability measure on L (and we will suppress L from the notation when it is understood). As we will discuss below, it is not hard to show that S_n looks more and more like the the dilate nL, or in other terms, that $\frac{1}{n}S_n \to L$ as a Gromov-Hausdorff limit. Our main contribution is to establish that counting measure on spheres limits to the cone measure on L in an appropriate sense to carry out averaging operations.[1]

Theorem 1.1 (Sphere averages)**.** *For any finite presentation* $(\mathbb{Z}^d, S)$ *and any function* $f : \mathbb{Z}^d \to \mathbb{R}$ *asymptotic to* $g : \mathbb{R}^d \to \mathbb{R}$ *with* g *homogeneous of order* k*, let* $L = \partial\,\mathrm{CHull}(S)$ *and let* μ_L *be cone measure on* L*. Then*

$$\frac{1}{|S_n|} \sum_{\mathsf{x} \in S_n} f(\mathsf{x}) = (v_{g,S}) \cdot n^k + O(n^{k-1}),$$

with coefficient $v_{g,S} := \int_L g(\mathsf{x})\ d\mu_L(\mathsf{x})$*.*

As a consequence of the main theorem, we get the leading term for the average value over balls in the word metric.

Remark 1.2 (Ball averages)**.** *With the same assumptions as above, let* $Q = \mathrm{CHull}(S)$*. Then*

$$\frac{1}{|B_n|} \sum_{\mathsf{x} \in B_n} f(\mathsf{x}) = (V_{g,S}) \cdot n^k + O(n^{k-1}).$$

with coefficient $V_{g,S} := \int_Q g(\mathsf{x})\ d\,\mathrm{Vol}(\mathsf{x})$*, normalized so that* $\mathrm{Vol}(Q) = 1$*.*

Crucially, this statement about ball averages can be observed much more easily than the sphere averages in the main theorem. (Just use the fact that $\frac{1}{n}B_n$ becomes uniformly distributed in Q, and that the error term is counted by points in a region of lower-order volume.) Going further, one can deduce from ball asymptotics that if the counting measures on $\frac{1}{n}S_n$ do converge to a measure on L, then it must be cone measure: that is the unique measure on L which, considering the necessary scaling properties, is compatible with Lebesgue measure on Q. However, there is no guarantee that the limit in Theorem 1.1 exists at all, even given the limit statement in Remark 1.2.

[1]We note that working in the asymptotic cone, $\mathbb{R}^d$, would be a substitute for the rescaling by dilations; furthermore, this suggests a natural generalization of these questions to other groups. To do this, one would need a theory of ultralimits of measures; in logic, this goes by the name of Loeb measure. These ideas have not yet been imported into geometric group theory, a translation that seems as though it would be quite fruitful.

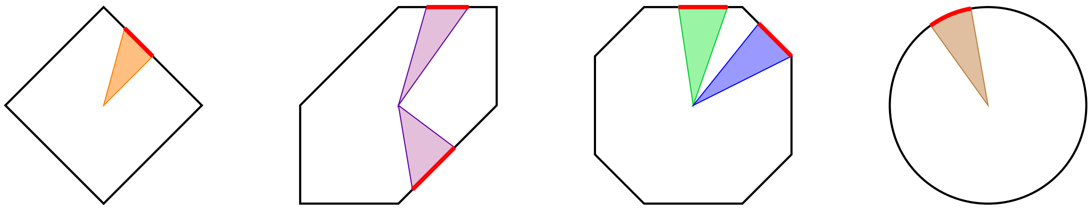


FIGURE 1. Six arcs are shown in red in this figure, each having cone measure 1/14; in other words, all of the colored regions have 1/14 as much area as the convex body they are in. In the square and the hexagon, all sides have equal measure because in each example the triangles subtended by the sides are mutually congruent. On the other hand, for this octagon generated by the *chess-knight* moves $\{(\pm 2, \pm 1), (\pm 1, \pm 2)\}$, the measure of its two types of sides (shown with green and blue) is in the ratio $4:3$. Cone measure is defined on any convex, centrally symmetric figure, and in particular it is uniform on the circle.

In the next section, we will give further examples to illustrate that the sphere problem is strictly harder than the ball problem.

As a consequence of Theorem 1.1, we can deduce a distribution result: the ball average for size-like functions compares to the sphere average by a simple multiplicative factor which is independent of the generating set. This factor depends only on the dimension d and the growth order k.

Proposition 1.3 (Spheres versus balls)**.** *For any function $f : \mathbb{Z}^d \to \mathbb{R}$ that is asymptotically homogeneous of order k,*

$$\lim_{n\to\infty} \frac{1}{|B_n|} \sum_{\mathsf{x}\in B_n} \frac{1}{n^k} f(\mathsf{x}) = \left(\frac{d}{d+k}\right) \lim_{n\to\infty} \frac{1}{|S_n|} \sum_{\mathsf{x}\in S_n} \frac{1}{n^k} f(\mathsf{x}).$$

That is, the coefficients of growth for sphere averages and ball averages are related by the simple expression $V_{g,S} = \left(\frac{d}{d+k}\right) v_{g,S}$.

Next, we can apply Theorem 1.1 to reduce problems about the asymptotic study of the geometry of Cayley graphs for $\mathbb{Z}^d$ to problems in convex geometry. A collection of such applications is given in §5.

The central examples of asymptotically homogeneous functions come from considering distance in the word metric. Using the standard embedding of $\mathbb{Z}^d$ in $\mathbb{R}^d$ as the integer lattice, the word metric on the Cayley graph for $(\mathbb{Z}^d, S)$ is within bounded distance of a norm on $\mathbb{R}^d$, namely the norm induced by the convex, centrally symmetric polyhedron L. This norm, denoted $\|\mathsf{x}\|_L$, is defined as the unique norm for which L is the unit sphere. Then it is a basic fact (Lemma 3.5 below) that there is a uniform bound K such that

$$\|\mathsf{x}\|_L \;\le\; |\mathsf{x}| \;\le\; \|\mathsf{x}\|_L + K$$

for all $\mathsf{x} \in \mathbb{Z}^d$, where $|\cdot|$ is the word length in the Cayley graph and K is the largest distance in the word metric from the identity to any lattice point in Q. Burago proved more generally in [4] that periodic metrics on $\mathbb{R}^d$ are at bounded distance from norms; we give a hands-on proof with the optimal constant for word metrics here. This ensures that distance in the Cayley graph for any finite generating set can be regarded as a coarsely homogeneous function. It is also immediate that $f(x) = |\mathsf{x}|^p$ is asymptotically homogeneous of order p, and it follows from this that the pth moments of a word metric—expected position in a large ball, variance and standard deviation, skewness, and so on—are in a sense independent of the choice of finite generating set (discussed below). As a special case of the main theorem, we get asymptotics for the growth function $\beta(n)$ and spherical growth function $\sigma(n)$—the number of words of length up to n and exactly n, respectively—for $\mathbb{Z}^d$ with an arbitrary generating set.

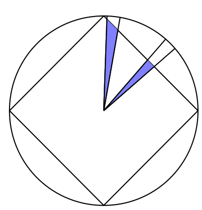
Corollary 1.4 (Growth functions)**.** *Fixing $(\mathbb{Z}^d, S)$ as above, let $\beta(n) = \#B_n$ and $\sigma(n) = \#S_n$ be the growth function and the spherical growth function, respectively. Then*

$$\beta(n) = Vn^d + O(n^{d-1}) \ ; \qquad \sigma(n) = (d\cdot V)n^{d-1} + O(n^{d-2}) \ ;$$

where $V = \mathrm{Vol}(Q)$ and d is the dimension.

The much stronger fact that these growth functions are always eventually polynomial for $\mathbb{Z}^d$ was shown by Benson in [2] (that is, for sufficiently large n, each function gives values equal to a well-defined polynomial in n, with coefficients depending on S); it would be interesting to study whether each of the coefficients has a geometric interpretation, as the leading term does. We note that Benson's results dictate that the growth functions of virtually abelian groups need not be polynomial, but are in general *quasipolynomial*, having coefficients that oscillate with finite period. We cite an example in the next section in which even the leading coefficient oscillates, showing that a result in the form of our Theorem 1.1 does not naively generalize to virtually abelian groups.

Notice that the limit measure μ_L is uniform with respect to surface area on each face of L, in contrast to the hitting measure for random walks in the word metric, which has a Gaussian distribution. This is of interest because randomness problems in groups are often approached by studying asymptotics of random walks, rather than probabilities with respect to the word metric. This shows that they are in general quite different. For instance, with $\mathbb{Z}^2$ and its standard generators, consider the distribution of points on S_n with respect to direction viewed radially from the origin. The random walk is most likely to hit a point near the diagonal directions; by contrast, the counting measure gives greater weight to the axial directions.



Acknowledgments. We thank Alex Eskin and Ralf Spatzier. The first author is partially supported by NSF grant DMS-0906086, the second author is partially supported by ANR projects Teichmüller (ANR-06-BLAN-0038) and Modunombres, and the third author is partially supported by NSF grant RTG-0602191.

2. Counting in balls and spheres

It is a problem of extremely classical interest to count the lattice points in balls in metric spaces. In one well-known form, this is the Gauss circle problem: where $R(n)$

is the number of standard lattice points in the round disk of radius n centered at the origin in the plane, Gauss first proved that

$$R(n) = \pi n^2 + O(n).$$

Given strong enough estimates for functions on balls, one can derive estimates for the annular regions $\Delta_n = B_n \setminus B_{n-1}$ defined as differences of successive balls. From Gauss's estimate in the circle problem we get $R(n) - R(n-1) = 2\pi n + O(n)$, which is vacuous: the estimate tells us nothing about lattice points in the annulus. Historically, the next progress on the circle problem was by Sierpinski, who proved in 1906 that

$$R(n) = \pi n^2 + O(n^{2/3}),$$

from which we get $R(n) - R(n-1) = 2\pi n + O(n^{2/3})$, giving a nontrivial estimate for annuli. Finding the optimal error term for the circle problem (conjecturally $n^{\frac{1}{2}+\epsilon}$) is a deep problem with ties to the Riemann hypothesis. It is natural to consider analogs of the Gauss circle problem in other metrics on $\mathbb{R}^d$, or other symmetric spaces, by counting lattice points in metric balls. There is a particularly nice answer in the case of lattice polytopes in $\mathbb{R}^d$, where the number of lattice points in integer dilates nP of a polytope P is exactly given by a polynomial in n of degree d, by a beautiful theorem of Ehrhart. These are called *Ehrhart polynomials*, and there are several excellent treatments in the literature, for instance in [1]. In general, when counting lattice points in dilates, one expects the leading term to come from volume, and the error term to come from the boundary effects. In polytopes, the second-order coefficient is given by the surface area of the boundary (suitably normalized); in general, the presence of flat faces makes the error order easy to compute, in contrast to the rounded boundary in the Gauss problem.

Sphere asymptotics for word metrics can be regarded as a group-theoretic version of the Gauss circle problem, because for integer-valued metrics such as word metrics, the annular region $B_n \setminus B_{n-1}$ is precisely the sphere S_n. This work studies counting and distribution problems in all finitely generated word metrics on free abelian groups by relating the algebraic counting problem to a geometric counting problem for polytopes. This gives us first-order asymptotics and a sharp error term.

There are several issues that should be clarified at the outset. First, functions that can be averaged over balls do not necessarily admit well-defined averages over spheres and annuli. Furthermore, the (directional) spherical estimates that we obtain in this paper are strictly better than the difference of (directional) ball estimates, even though the error order is optimal in the ball estimates. In this section, we present simple illustrations of some of the subtleties.

Here is a straightforward example in lattice-point counting to explicitly illustrate this issue. Take Q to be the unit square $[-\frac{1}{2}, \frac{1}{2}] \times [-\frac{1}{2}, \frac{1}{2}]$, L its boundary, and $\|\cdot\|_L$ the corresponding norm on $\mathbb{R}^2$ (that is, the norm whose unit circle is L, which in this case is half of the sup norm). The ball of radius n with respect to $\|\cdot\|_L$ is the dilate nQ of Q by n. Setting up a straightforward Riemann sum lets us count the number of lattice points lying inside nQ to first order:

$$\#\mathbb{Z}^2 \cap nQ = n^2 \operatorname{Area}(Q) + O(n) = n^2 + O(n).$$

However, it does not follow that the number of lattice points in the annuli between successive dilates equals $2n + O(1)$; instead, the number oscillates with the parity of n, between no points and $4n$ points. Correspondingly, there is no well-defined coefficient of n in the lattice-point count $n^2 + O(n)$ given above. Thus, the indicator function of $\mathbb{Z}^2$

is well-defined to first order over the ball nQ but not over the annular region which is the difference of two balls. (Indeed, for rational polytopes one finds that lattice points in dilates are always counted by quasipolynomials.) The same phenomenon, that ball averages are well-defined but sphere averages are not, can also be observed in groups. Consider Cannon's example (see [5, Example VI.A.9]): the orientation-preserving part of the Euclidean reflection group in the equilateral triangle, which is virtually abelian. With appropriate generators, the spherical growth function oscillates between $4n+2$ and $5n-2$, and so never even becomes monotone.

In this paper, we will deal not only with geometric indicator functions, but also with averaging of more general functions. Let us briefly recall an arithmetic example to underscore the principle that ball averages are often better-behaved than sphere averages. One can study random properties of the integers by choosing uniformly over $\{1, 2, \ldots, n\}$ and letting $n \to \infty$. This is what is meant by classical statements of analytic number theory (see [7]) such as

The probability that two integers are relatively prime is $6/\pi^2$.

We can express the probability of relative primality in terms of the Euler phi function $\phi(n)$; here again, the order oscillates (that is, $\frac{1}{n}\phi(n)$ has multiple accumulation points—one for the primes and another for the powers of two, for instance), while the average order $\frac{1}{n^2}[\phi(1)+\cdots+\phi(n)]$ converges. Let f be the function on $\mathbb{Z}^2$ that is the indicator for the relative primality of the coordinates. With respect to an appropriate Cayley graph for $\mathbb{Z}^2$, this means that sphere averages for f do not exist, whereas ball averages tend to $6/\pi^2$.

We note that this function is coarsely homogeneous of order zero, but not asymptotically homogeneous (i.e., relative primality is not size-like), so Theorem 1.1 does not apply. For size-like functions on free abelian groups, we will find that sphere averages must exist for all Cayley graphs; we get distribution results which count points in every direction, as illustrated in Figure 2.

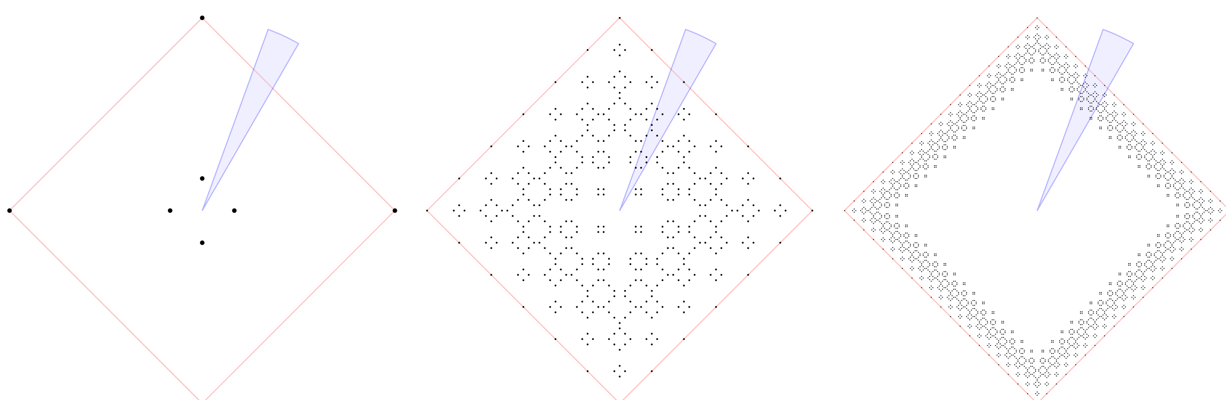


FIGURE 2. For the generating set $S = \pm\{6\mathsf{e}_1, \mathsf{e}_1, 6\mathsf{e}_2, \mathsf{e}_2\}$, spheres of radius $n = 1$, 6, and 20 are shown relative to the dilated limit shape nL. As $n \to \infty$, the proportion of points in S_n that lie in any fixed direction is converging.

We will show below that

$$\#S_n \sim \#(\mathbb{Z}^d \cap \Delta_n L) \sim \mathrm{Vol}(\Delta_n L) = dn^{d-1} \cdot \mathrm{Vol}(Q),$$

where $\Delta_n L = nQ \setminus (n-1)Q$ is the annular region between dilates of L. The first comparison here reduces the sphere counting problem in the group to the lattice point counting problem in a geometric annulus, and the second comparison is a solution to the lattice point counting problem. Counting problems like this can be studied in many other kinds of metric spaces. In [9], for instance, the lattice point counting problem is solved in the Heisenberg group, and partial results are offered on the sphere counting problem.

3. The limit metric

When considering presentations $(\mathbb{Z}^d, S)$, we will assume throughout that S is symmetric, so that $S = -S = S_1$ is also the sphere of radius 1 in the group.

First let us consider $G = \mathbb{Z}$. With any finite generating set, the spheres of large radius are divided into a positive part and a negative part, each of uniformly bounded diameter in $\mathbb{R}$. In particular, if a is the largest positive element in the generating set, then the most efficient spelling of a very large integer uses almost exclusively the letter a; in the language we will develop below, $\pm a$ are the only *significant generators.* Let K be the smallest value such that the ball of radius K in the word metric contains all of the integers $-a \le m \le a$ (so that K is a constant depending on S). The ball of radius K includes all positive integers up to a. The ball of radius $K+1$ then includes all positive integers up to $2a$, since integers between a and $2a$ can be obtained by adding the generator a. Continuing, we see that, for any $n > K$, the ball of radius $n-1$ includes all positive integers up to $(n-K)a$. Thus the sphere of radius n is totally contained in the interval $\big((n-K)a,\ na\big]$. That means that $\frac{1}{n}S_n \cap \mathbb{R}_+$ is contained in the interval $\big(a(1-K/n),\ a\big]$, and therefore $\frac{1}{n}S_n \to \{\pm a\}$. This depends on the choice of generating set, but in fact only on its convex hull in $\mathbb{R}$.

For $G = \mathbb{Z}^d$ the situation is similar. We will study word-length from a geometric point of view. Let $S \subset \mathbb{Z}^d$ be a fixed finite set of generators. We adopt additive notation, so that every element of $\mathbb{Z}^d$ has a representative in the form $\mathsf{w} = \alpha_1\mathsf{a}_1 + \alpha_2\mathsf{a}_2 + \cdots + \alpha_r\mathsf{a}_r$ where $S = \{\pm\mathsf{a}_1, \ldots, \pm\mathsf{a}_r\}$ and $\alpha_i \in \mathbb{Z}$. Let $|\mathsf{w}|$ denote the length of w in the word metric, or the minimal $\sum|\alpha_i|$ over all representatives as above. A spelling is called a *geodesic representative* (or a *geodesic spelling*) if it realizes this minimum, since these spellings correspond to geodesic paths in the Cayley graph.

Let Q be the convex hull in $\mathbb{R}^d$ of the generating set S, and let L denote its boundary. By construction, L is a centrally symmetric convex polyhedron. Let $\|\cdot\|_L$ denote the norm on $\mathbb{R}^d$ induced by L: this is the unique norm for which L is the unit sphere. Namely, for $\mathsf{x} \in \mathbb{R}^d$, $\|\mathsf{x}\|_L$ equals the unique $\lambda \ge 0$ such that $\mathsf{x} \in \lambda L$.

For any subset $M \subset \mathbb{R}^d$, let $\Delta M = \{t\mathsf{x} : t \in \mathbb{R}, t \ge 0, \mathsf{x} \in M\}$ be the infinite cone on M from the origin with respect to dilation. Then let $\Delta_k M = \{t\mathsf{x} : k-1 < t \le k, \mathsf{x} \in M\}$ be the annular region from the $(k-1)$st to the kth dilation, so that the cone $\hat{M}$ from M to the origin is equal to $\Delta_1 M \cup \{0\}$. For σ a codimension-1 face of L, we will call $\Delta\sigma$ the *sector* associated to σ. Vol denotes the Lebesgue measure on $\mathbb{R}^d$ (or Area if $d = 2$). The extreme points of Q are called *significant generators.* These are necessarily elements of the generating set, and it will turn out that many of the properties we study in this paper depend only on this subset of S. In particular, we will see shortly that the significant generators completely determine the averages of size-like functions in the word metric. The first basic observation is that L encodes the large-scale geometry of the group with this generating set.

Lemma 3.1. *An element $\mathsf{a} \in S$ is on the polytope L if and only if $n\mathsf{a}$ is geodesic for every $n \in \mathbb{Z}$. An element $\mathsf{a} \in S$ is an extreme point of Q (or, equivalently, a vertex of L) if and only if $n\mathsf{a}$ is uniquely geodesic for every $n \in \mathbb{Z}$.*

Proof. We give the proof for $d = 2$ for clarity. Suppose $\mathsf{a} \in S$ is an interior point of Q, and suppose it lies in the sector determined by the extreme points $\mathsf{b}, \mathsf{c} \in S$. Then there is a unique positive multiple $\alpha\mathsf{a}$ that lies on the segment between b and c, so $\alpha\mathsf{a} = \beta\mathsf{b} + \gamma\mathsf{c}$, with $\alpha > 1$, $\beta + \gamma = 1$, with necessarily rational coefficients since $\mathsf{a}, \mathsf{b}, \mathsf{c}$ have integer coordinates. But then by clearing common denominators, we have an integer multiple of a expressed in terms of b and c with a strictly smaller wordlength. This shows that $n\mathsf{a}$ cannot be geodesic for all n.

Now suppose $n\mathsf{a}$ is not geodesic for some $n \geq 1$. Then we can write $n\mathsf{a} = \sum \alpha_i \mathsf{a}_i$ with $\sum \alpha_i < n$. Assume further that n is the smallest such value, so that a itself does not appear in this spelling. Then

$$\|\mathsf{a}\|_L \leq \sum \frac{\alpha_i}{n} \|\mathsf{a}_i\|_L < 1,$$

showing that a is interior to Q.

For the word $n\mathsf{a}$, any alternative spelling with the same length expresses a as a convex combination of other generators, and such an expression exists if and only if a is not extreme.

(The proof is the same in arbitrary dimension: replace b, c with the extreme points $\mathsf{b}_1, \ldots, \mathsf{b}_n$ in a cell of a triangulation.) □

Now we establish that the word metric limits to a norm, and that they differ by a bounded additive amount. As a consequence, the spheres in the word metric, once normalized, converge to a limit shape. This is a small special case of the theory for finitely-generated nilpotent groups and, more generally, lattices in Lie groups of polynomial growth (see Pansu and Breuillard [10, 3]). Here we give an elementary proof in terms of the combinatorial group theory and Euclidean geometry.

Lemma 3.2. *For $\mathsf{w} \in \mathbb{Z}^d$, $\|\mathsf{w}\|_L \leq |\mathsf{w}|$.*

Proof. We prove that $\max_{\mathsf{w} \in S_n} \|\mathsf{w}\|_L \leq n$ for all $n \geq 1$ by induction on n. If $n = 1$, this is immediate from the definition of L. When $n > 1$, we can always write $\mathsf{w} \in S_n$ as $\mathsf{w} = \mathsf{w}' + \mathsf{a}$ where $\mathsf{w}' \in S_{n-1}$ and $\mathsf{a} \in S$. But then $\|\mathsf{w}\|_L \leq \|\mathsf{w}'\|_L + \|\mathsf{a}\|_L \leq n$. □

We say that $\mathsf{v} \in \mathbb{Z}^d$ has a *simple spelling* in terms of the generating set if there is a geodesic spelling which uses only generators that are extreme points of some face of L. (This property is discussed further below in §5.3.) For instance in $\mathbb{Z}^2$, simple spellings are formed using only consecutive significant generators. We will denote by $\mathcal{P}_n$ the set of points in S_n which have a simple spelling.

Definition 3.3. Fix a triangulation of L with vertices at extreme points. For any $n \in \mathbb{N}$, the words represented by simple spellings of length n are

$$\mathcal{P}_n := \left\{ \sum a_i \mathsf{v}_i : \quad a_i \text{ non-negative integers, } \{\mathsf{v}_i\} \text{ define a simplex in } L, \ \sum a_i = n \right\}.$$

Note that if $\mathsf{p} \in \mathcal{P}_n$, then it lies in $nL \cap S_n$. This is because p belongs to the facet of nL that its extreme points v_i do, so $\|\mathsf{p}\|_L = n$. Its wordlength is at most n because we have a spelling of length n, so Lemma 3.2 ensures that the wordlength is equal to n. Note also that each $\mathsf{p} \in \mathcal{P}_n$ is at distance two in the word metric from the other elements that differ by $\mathsf{v}_i - \mathsf{v}_j$, and therefore $\mathcal{P}_n$ is 2-dense in nL with respect to the L metric as well.

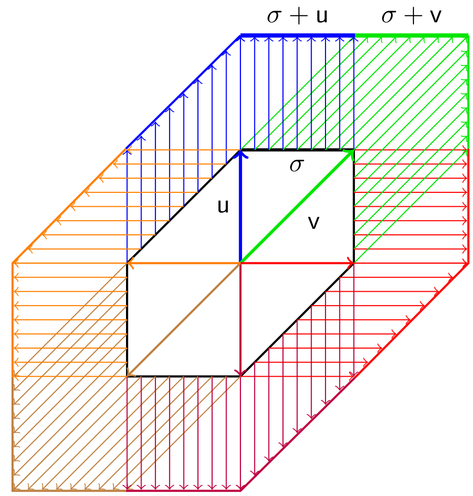


FIGURE 3. $Q + S = 2Q$.

Lemma 3.4 (Tiles)**.** *For any whole number $n \geq 2$,* $\quad (n-1)Q + S = nQ.$

Proof. We know that $(n-1)Q+Q = nQ$, by convexity of Q. Now consider $(n-1)Q+S$. This contains nS, and therefore the extreme points of nQ. Consider σ, a simplex in the (fixed) triangulation of L. For a particular i, the set $(n-1)\sigma + \mathsf{v}_i$ is a convex set (a copy of $(n-1)\sigma$). If the extreme points of σ are $\{\mathsf{v}_i\}$, then those translates overlap, covering $n\sigma$. Thus $(n-1)L + S$ includes nL. (See Figure 3.) But since each cone $\hat{\sigma}$ is contained in Q, one can similarly cone off to obtain the desired result. That is,

$$(n-1)Q + S = (n-1)\hat{L} + S = \left(\bigcup (n-1)\hat{\sigma}\right) + S = \bigcup \left((n-1)\hat{\sigma} + S\right) = nQ.$$

□

Lemma 3.5 (Bounded difference)**.** *There is a constant $K = K(S) \geq 1$ such that for all $\mathsf{w} \in \mathbb{Z}^d$,*

$$\|\mathsf{w}\|_L \leq |\mathsf{w}| < \|\mathsf{w}\|_L + K.$$

That is, S_n is contained in the annular region between $(n-K)L$ and nL.

Proof. Set $K = \max\{|x| : x \in Q \cap \mathbb{Z}^d\}$. This is the largest wordlength required to fill in the convex hull of the generators; for example, in Figure 4, we find $K = 3$. (Note that K can be arbitrarily large as the generating set S varies, but only depends on S.)

The ball of radius K then contains all of $Q \cap \mathbb{Z}^d$, and thus by the previous lemma we have that the ball of radius $K+1$ contains all of $2Q$, and so on until the ball of radius $n-1$ includes all lattice points in $(n-K)Q$. But since B_n is contained in nQ, this precisely means that the sphere is contained in $nQ \setminus (n-K)Q$, as required. □

Proposition 3.6 (Limit shape)**.** *As a Gromov-Hausdorff limit, we have*

$$\lim_{n\to\infty} \tfrac{1}{n} S_n = L.$$

Proof. For the forward inclusion, if $\mathsf{x}_n \in S_n$ is a sequence, then by the previous lemma,

$$\lim_{n\to\infty} \left\| \frac{\mathsf{x}_n}{n} \right\|_L = \lim_{n\to\infty} \frac{\|\mathsf{x}_n\|_L}{n} = \lim_{n\to\infty} \frac{|\mathsf{x}_n|}{n} = 1.$$

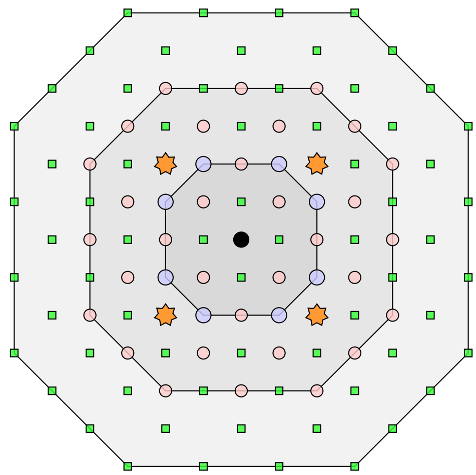


Figure 4. The *chess-knight metric*, with generators $\{(\pm 2, \pm 1), (\pm 1, \pm 2)\}$. The spheres of radius 1, 2, and 3 are shown entirely, as well as the first three dilates of Q. Four points from the sphere S_4 are shown, marked with stars, to illustrate the difference between the norm and the word metric: those points have $|w| = 4$ while $\|w\|_L = 4/3$. It takes three steps to fill in the lattice points in Q, so Lemma 3.5 shows that $|w|$ and $\|w\|$ never differ by more than 3.

The reverse inclusion follows from the fact that $\mathcal{P}_n \subseteq S_n$ is 2-dense in nL. $\square$

4. The limit measure

We will study the measure induced on L from the counting measure on S_n. First we note that the rank-one case is trivial: considering $(\mathbb{Z}, S)$ with $S = -S$, there are exactly as many negative integers in S_n as positive integers, by symmetry. Thus the counting measure on $\frac{1}{n} S_n$ limits to the uniform measure on the two-point set $L = \{\pm a\}$. We study $d \geq 2$ below.

Note that $\frac{1}{n} S_n \cap \Delta\sigma$ is near σ by Lemma 3.5. Recall that if μ_n, μ denote Borel probability measures on a space X, then the following are equivalent ([6, Thm 2.4]):

- For every bounded continuous $f : X \to \mathbb{R}$,
$$\lim_{n\to\infty} \int_X f\,d\mu_n = \int_X f\,d\mu.$$
- For every open $U \subseteq X$,
$$\liminf_{n\to\infty} \mu_n(U) \geq \mu(U).$$

In this setup, (μ_n) is said to *converge weakly* to μ. On the other hand, (μ_n) is said to *converge strongly* if
$$\lim_{n\to\infty} \sup_{A \subseteq X} \big|\mu_n(A) - \mu(A)\big| = 0.$$

Theorem 4.1 (Strong convergence on L)**.** *Define measures μ_n and μ on L by defining, for Lebesgue-measurable sets $\tau \subseteq L$,*

$$\mu_n(\tau) = \frac{\#(\frac{1}{n}S_n \cap \hat{\tau})}{\#S_n} = \frac{\#(S_n \cap \Delta\tau)}{\#S_n}; \qquad \mu(\tau) = \frac{\mathrm{Area}(\hat{\tau})}{\mathrm{Area}(Q)}.$$

Then $\mu_n \to \mu$ strongly.

From this we will immediately derive the weak convergence of measures on $\mathbb{R}^d$ needed to prove Theorem 1.1. Let $\mathcal{N}_r(A)$ denote the r-neighborhood of $A \subseteq \mathbb{R}^d$.

Corollary 4.2 (Weak convergence on $\mathbb{R}^d$)**.** *Define measures ν_n and ν on $\mathbb{R}^d$ by defining, for Lebesgue-measurable sets $A \subset \mathbb{R}^d$,*

$$\nu_n(A) = \frac{\#(\frac{1}{n}S_n \cap A)}{\#S_n}; \qquad \nu(A) = \mu(A \cap L).$$

Then $\nu_n \to \nu$ weakly.

Proof. Let $U \subseteq \mathbb{R}^d$ be open and set $\sigma = U \cap L$. Given $\epsilon > 0$, let $\sigma' \subseteq \sigma$ a closed subset such that $\mu(\sigma') > \mu(\sigma) - \epsilon$. Then for any metric inducing the standard topology, we can take large enough n so that $\mathcal{N}_{K/n}(\sigma') \subseteq U$. Then

$$\liminf_{n\to\infty} \nu_n(U) \geq \lim_{n\to\infty} \nu_n(\hat{\sigma}') = \lim_{n\to\infty} \mu_n(\sigma') = \mu(\sigma') \geq \mu(\sigma) - \epsilon = \nu(U) - \epsilon.$$

Let $\epsilon \to 0$ to get the desired inequality. □

Next we demonstrate that this suffices to prove the main theorem.

Proof of Theorem 1.1. Suppose that $f : \mathbb{Z}^d \to \mathbb{R}$ is asymptotic to a function $g : \mathbb{R}^d \to \mathbb{R}$ that is homogeneous of order k, meaning that $g(a\mathsf{x}) = a^k g(\mathsf{x})$, and $\lim_{\mathsf{x}\to\infty} \frac{f(\mathsf{x})}{g(\mathsf{x})} = 1$. This means that for all $\epsilon > 0$ there exists N such that

$$\mathsf{x} \notin B_N \implies (1-\epsilon)g(\mathsf{x}) \leq f(\mathsf{x}) \leq (1+\epsilon)g(\mathsf{x}).$$

But then we have

$$n > N \implies (1-\epsilon)\sum_{\mathsf{x}\in S_n} g(\mathsf{x}) \leq \sum_{\mathsf{x}\in S_n} f(\mathsf{x}) \leq (1+\epsilon)\sum_{\mathsf{x}\in S_n} g(\mathsf{x}),$$

which gives

$$(1-\epsilon)\frac{\sum_{\mathsf{x}\in S_n} g(\mathsf{x})}{n^k|S_n|} \leq \frac{\sum_{\mathsf{x}\in S_n} f(\mathsf{x})}{n^k|S_n|} \leq (1+\epsilon)\frac{\sum_{\mathsf{x}\in S_n} g(\mathsf{x})}{n^k|S_n|}.$$

Since ϵ was arbitrary, this means that

$$\lim_{n\to\infty} \frac{\sum_{\mathsf{x}\in S_n} f(\mathsf{x})}{n^k|S_n|} = \lim_{n\to\infty} \frac{\sum_{\mathsf{x}\in S_n} g(\mathsf{x})}{n^k|S_n|}.$$

But then $g(\mathsf{x})/n^k = g(\mathsf{x}/n)$, so

$$\frac{\sum_{\mathsf{x}\in S_n} g(\mathsf{x})}{n^k|S_n|} = \frac{\sum_{\mathsf{x}\in \frac{1}{n}S_n} g(\mathsf{x})}{|S_n|} = \int_{\mathbb{R}^d} g(\mathsf{x})\ d\nu_n(\mathsf{x}),$$

with respect to the measure defined above. Noting that ν_n and ν are supported on the compact set Q, weak convergence finishes the job:

$$\lim_{n\to\infty} \frac{\sum_{\mathsf{x}\in S_n} f(\mathsf{x})}{n^k|S_n|} = \lim_{n\to\infty} \int_{\mathbb{R}^d} g(\mathsf{x})\ d\nu_n(\mathsf{x}) = \int_{\mathbb{R}^d} g(\mathsf{x})\ d\nu(\mathsf{x}) = \int_L g(\mathsf{x})\ d\mu(\mathsf{x}),$$

as desired. □

4.1. **Rank two.** We will prove Theorem 4.1 in this section ($d = 2$) and the following section ($d > 2$). We begin by giving the necessary counting argument in $\mathbb{Z}^2$, because it has some features that are particular to that dimension.

In what follows, fix a side σ of L and let u and v be the names of its endpoints, which are necessarily integer points. The key step in proving Theorem 4.1 is to count the number of points of S_n in the sector $\Delta\sigma$ over an entire edge. First we get control on the geodesic spellings of large words in the sector.

Lemma 4.3 (Geodesic spellings)**.** *There is a uniform bound D_0 such that for sufficiently large words w in $\Delta\sigma$, there is a geodesic representative of the form $\mathsf{w} = a\mathsf{u} + b\mathsf{v} + \mathsf{w}'$, where $|\mathsf{w}'| < D_0$.*

Proof. Label the elements of S which lie on the line segment between u and v as $\mathsf{a}_1, \ldots, \mathsf{a}_r$. Then the first task is to show that all geodesic spellings of w are of the form

$$\mathsf{w} = a\mathsf{u} + b\mathsf{v} + \sum \alpha_i \mathsf{a}_i + \mathsf{w}'',$$

where $|\mathsf{w}''|$ is bounded. This is true because u, v, and the a_i are the only generators whose projection onto $\sigma^\perp$ (the line through the origin that is perpendicular to σ) is one. We know by Lemma 3.5 that the projection of w onto $\sigma^\perp$ is within K of $|\mathsf{w}|$, and thus there is a uniform bound on the number of other generators that can appear.

Next, write each of the a_i as $\frac{p_i}{q_i}\mathsf{u} + \frac{q_i - p_i}{q_i}\mathsf{v}$, and let $q = \operatorname{lcm}\{q_i\}$. Then without loss of generality, the coefficients α_i are at most q. (Otherwise, $q\mathsf{a}_i$ can be rewritten as an integer combination of u and v.) Finally, setting $\mathsf{w}' = \sum \alpha_i \mathsf{a}_i + \mathsf{w}''$ completes the proof. $\square$

Corollary 4.4 (Modifying geodesic spellings)**.** *For all but boundedly many words $\mathsf{w} \in \Delta_n\sigma$, there is a geodesic in the Cayley graph from the identity to w which passes through the points*

$$\mathsf{w}, \quad \mathsf{w} - \mathsf{v}, \quad \mathsf{w} - 2\mathsf{v}, \quad \ldots \quad \mathsf{w} - K\mathsf{v}.$$

Proof. First, we show that all but boundedly many lattice points in $\Delta_n\sigma$ have a geodesic spelling in which the coefficients of u and v are each at least K. We just consider the coefficient of v without loss of generality. Note that $\|\mathsf{v}\|_L = 1$ by definition of the L-norm. Thus words that are spelled $\mathsf{w} = a\mathsf{u} + b\mathsf{v} + \mathsf{w}'$ with $|\mathsf{w}'| < D_0$ are within $b + D_0$ of the line $\Delta\mathsf{u}$. Let $D = K + D_0$. Now note that $\mathcal{N}_D(\Delta_n\mathsf{u})$ has diameter $2D + 1$, so its number of lattice points is uniformly bounded. This shows that $\mathsf{w}, \mathsf{w} - \mathsf{v}, \ldots, \mathsf{w} - K\mathsf{v}$ are all metrically between w and e.

Finally, if a point is farther than K from a line, then moving it by a distance K will not cross over the line. Thus the modified spellings still represent points in the sector $\Delta\sigma$. $\square$

Then we get a very clean result: the integer points in $\Delta_n\sigma$ count, up to bounded additive error, the quantity we seek.

Lemma 4.5 (Sphere counting for $\mathbb{Z}^2$)**.**

$$\#(S_n \cap \Delta\sigma) \stackrel{+}{\asymp} \#(\mathbb{Z}^2 \cap \Delta_n\sigma).$$

Proof. Let $\Phi_n : \mathbb{Z}^2 \to \mathbb{Z}^2$ be given by $\Phi_n(\mathsf{w}) = \mathsf{w} - m\mathsf{v}$, where $m = |\mathsf{w}| - n$. That is, it modifies words by subtracting off copies of v when the wordlength differs from n.

Now consider applying Φ_n to words $\mathsf{w} \in \Delta_n\sigma$. For such words, as long as n is sufficiently large and w is D-far from $\Delta\mathsf{u}$, Corollary 4.4 guarantees that there is a geodesic representative using at least K copies of v. But we know that $0 \le |\mathsf{w}| - n \le K$, so this means that $\Phi_n(\mathsf{w}) = \mathsf{w} - m\mathsf{v}$ is a point on a geodesic path from e to w. Thus $|\Phi_n(\mathsf{w})| = |\mathsf{w}| - (|\mathsf{w}| - n) = n$, or in other words $\Phi_n(\mathsf{w}) \in S_n$.

This argument shows that, apart from a bounded number of points, Φ_n gives a bijection from $\mathbb{Z}^2 \cap \Delta_n\sigma$ to $S_n \cap \Delta\sigma$. Injectivity follows from Corollary 4.4; surjectivity is established by noting that if $|\mathsf{x}| = n$ and $\|\mathsf{x}\|_L = n - k$ for a point in the sector, then $\mathsf{x} + k\mathsf{v} \in \Delta_n\sigma$. □

Proof of Theorem 4.1 when $d = 2$. The region $\Delta_n\sigma$ is a quadrilateral with three of its four sides included, whose vertices have integer coordinates. Pick's Theorem says that for any polygonal region whose extreme points are integer points, the area is equal to the number of integer points in the interior plus half of the integer points on the boundary minus one ($A = i + \frac{b}{2} - 1$). Now $\Delta_n\sigma$ contains one of the two long boundary segments, and the number of integer points on the short boundary segments is uniformly bounded. Therefore, up to additive error, the number of integer points in $\Delta_n\sigma$ is equal to its area. But its area is exactly

$$\operatorname{Area}(\Delta_n\sigma) = n^2 \operatorname{Area}(\Delta_1\sigma) - (n-1)^2 \operatorname{Area}(\Delta_1\sigma) = (2n-1)\operatorname{Area}(\hat\sigma).$$

Thus we have $\#(S_n \cap \Delta\sigma) \overset{+}{\asymp} (2n-1)\operatorname{Area}(\hat\sigma)$, and summing over all sides gives $\#S_n = (2n-1)\operatorname{Area}(Q)$, which shows that

$$\frac{\#(S_n \cap \Delta\sigma)}{\#S_n} \to \frac{\operatorname{Area}(\hat\sigma)}{\operatorname{Area}(Q)}.$$

(Note that this also establishes the spherical growth asymptotics for $d = 2$, as in Theorem 1.4.)

To complete the proof it suffices to show that the estimate $\#(\mathbb{Z}^2 \cap \Delta_n\tau) \overset{+}{\asymp} \operatorname{Area}(\Delta_n\tau)$ is valid for small subarcs $\tau \subset \sigma$. Consider $\Delta_n\tau$, and approximate it by an integer trapezoid T_n in the following way: for the two vertices on nL, replace them with nearest-possible integer vertices on nL, and likewise for the two vertices on $(n-1)L$. (T_n is nondegenerate for sufficiently large n.) Then it is clear that both the area and the number of lattice points in T_n are boundedly close to those values for $\Delta_n\tau$, so we are done. □

This completes the proof that counting measure limits to cone measure on the polygon L.

4.2. **General rank.** For general rank d, we will get asymptotic comparisons rather than additive difference by carrying out the corresponding estimates. We obtain a limit shape $L = \lim_{n\to\infty} \frac{1}{n} S_n$ in $\mathbb{R}^d$ by taking the boundary polyhedron of the convex hull of the generators; we have a limiting distance on $\mathbb{R}^d$ via the norm induced by L; and finally, we obtain a measure on L as the limit of the counting measures on S_n, which is proportional to the Euclidean volume subtended by a facet. However, we no longer have Pick's Theorem to count the points in regions of the facets, so we must replace that part of the argument.

Recall that the generalization to higher dimensions of Pick's Theorem is by *Ehrhart polynomials*: the number of lattice points in the large dilates of a polytope with integer vertices is given by a polynomial formula in the dilation scalar. That is, there are coefficients $\{a_i\}$ depending on P such that

$$\#(\mathbb{Z}^d \cap nP) = a_d n^d + a_{n-1} n^{d-1} + \ldots + a_0$$

for any natural number n. Clearly $a_d = \mathrm{Vol}(P)$, and some of the other coefficients have known interpretations: a_{d-1} is given by the surface area of the boundary (relative to the fundamental volume of the lattice in each hyperplane slice through faces of P), and a_0 is the Euler characteristic. There remain quite a few open questions about the interpretations of the other coefficients.

It follows immediately that the number of lattice points in $\Delta_n\sigma$ is asymptotic to $d \cdot n^{d-1}\,\mathrm{Vol}(\Delta_1\sigma)$, by letting $P = \Delta_1\sigma$ and noting that $\Delta_n\sigma = nP \setminus (n-1)P$.

We would like to show that the limit measure on L is uniform on each face. First, by triangulating if necessary, we may assume that all the faces of L are simplices with integer vertices.

Lemma 4.6 (Lattice counting for $\mathbb{Z}^d$)**.** *If σ is a simplicial $(d-1)$-cell of L and τ is a simplex of the same dimension contained in σ, then*

$$\lim_{n\to\infty} \frac{\#(\mathbb{Z}^d \cap \Delta_n\tau)}{\#(\mathbb{Z}^d \cap \Delta_n\sigma)} = \frac{\mathrm{Vol}\,\Delta_1\tau}{\mathrm{Vol}\,\Delta_1\sigma}.$$

Proof. First, observe that all the lattice points in $\Delta_n\tau$ are contained in finitely many dilates of τ. That is,

$$\#(\mathbb{Z}^d \cap \Delta_n\tau) = \#\left(\mathbb{Z}^d \cap \bigcup_{j=1}^{q} k_j\tau\right).$$

Here, q is the number of hyperplanes parallel to σ between σ and the origin which contain lattice points. To see that this number is finite, consider the formula for the distance from a point to a plane, remembering that σ having integer vertices means that the hyperplane H containing σ has an equation with integer coefficients.

Let Λ_j be the subset $\mathbb{Z}^d \cap k_jH$. Note that all of the Λ_j are translates of some common lattice $\Lambda = \mathbb{Z}^d \cap H'$, where H' is the plane through the origin parallel to H. Let V be the covolume of Λ and let c be the minimal diameter of a fundamental domain for Λ, which exists because the set of possible diameters is discrete. We have

$$(k_j - c)^{d-1}\frac{\mathrm{Vol}\,\Delta_1\tau}{V} \le \#(\Lambda_j \cap k_j\tau) \le (k_j + c)^{d-1}\frac{\mathrm{Vol}\,\Delta_1\tau}{V},$$

and the same inequalities with the same constants holds for σ. But $n - 1 < k_j \le n$, so enlarging c by one, we can write

$$(n - c)^{d-1}\frac{\mathrm{Vol}\,\Delta_1\tau}{V} \le \#(\Lambda_j \cap k_j\tau) \le (n + c)^{d-1}\frac{\mathrm{Vol}\,\Delta_1\tau}{V}.$$

We can sum over j and get

$$q(n - c)^{d-1}\frac{\mathrm{Vol}\,\Delta_1\tau}{V} \le \#(\mathbb{Z}^d \cap \Delta_n\tau) \le q(n + c)^{d-1}\frac{\mathrm{Vol}\,\Delta_1\tau}{V}.$$

This means that

$$\#(\mathbb{Z}^d \cap \Delta_n\tau) \sim qn^{d-1}\frac{\mathrm{Vol}\,\Delta_1\tau}{V}.$$

Of course the same holds for $\tau = \sigma$.

Thus

$$\lim_{n\to\infty} \frac{\#(\mathbb{Z}^d \cap \Delta_n\tau)}{\#(\mathbb{Z}^d \cap \Delta_n\sigma)} = \frac{\mathrm{Vol}\,\Delta_1\tau}{\mathrm{Vol}\,\Delta_1\sigma},$$

as required. □

The next difference is that Lemma 4.5 no longer holds as stated, but is replaced by an asymptotic statement.

Lemma 4.7 (Sphere counting for $\mathbb{Z}^d$)**.**

$$\#(S_n \cap \Delta\sigma) = \#(\mathbb{Z}^d \cap \Delta_n\sigma) + O(n^{d-2}).$$

To prove this, run the same bijective argument as before on points that are outside of a D-neighborhood of the cone on the boundary of σ. The count of points close to the boundary is clearly lower-order, since they live in a region that measures length n in at most $d-2$ vector directions, and is bounded in the others.

This completes the proof of Theorem 4.1 for all d (since the $d=1$ case is elementary).

Example 4.8. We note that the first-order agreement of the algebraic and geometric count is best-possible. In the chess-knight generators for $\mathbb{Z}^2$, one can readily compute that for $n \geq 5$,

$$\#B_n = 14n^2 - 6n + 5 \ ; \qquad \#S_n = 28n - 20 \ ;$$

while the Ehrhart polynomials for the associated lattice point counts are

$$\#(\mathbb{Z}^2 \cap nQ) = 14n^2 + 6n + 1 \ ; \qquad \#(\mathbb{Z}^2 \cap \Delta_n L) = 28n - 8.$$

5. Applications

5.1. **Spheres versus balls.** From the sphere averages, we can quickly deduce the other averaging statement previewed in the introduction.

Proposition 1.3 (Spheres versus balls)**.** *For any function $f : \mathbb{Z}^d \to \mathbb{R}$ that is asymptotically homogeneous of order k,*

$$\lim_{n\to\infty} \frac{1}{|B_n|} \sum_{\mathsf{x}\in B_n} \frac{1}{n^k} f(\mathsf{x}) = \left(\frac{d}{d+k}\right) \lim_{n\to\infty} \frac{1}{|S_n|} \sum_{\mathsf{x}\in S_n} \frac{1}{n^k} f(\mathsf{x}).$$

That is, the coefficients of growth for sphere averages and ball averages are related by the simple expression $V_{g,S} = \left(\frac{d}{d+k}\right) v_{g,S}$.

Proof. We will repeatedly use the facts that $f \sim g$ and $g(a\mathsf{x}) = a^k g(\mathsf{x})$. Since $\frac{1}{n}B_n$ is uniformly filling in Q, the first equality is just a Riemann sum.

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{|B_n|} \sum_{\mathsf{x}\in B_n} \frac{1}{n^k} f(\mathsf{x}) &= \int_Q g(\mathsf{x}) \ d\,\mathrm{Vol}(\mathsf{x}) = \int_{I\times L} g(t\mathsf{x}) \ [d \cdot t^{d-1} dt] \ d\mu(\mathsf{x}) \\
&= d \cdot \left(\int_0^1 t^{d+k-1} \ dt \right) \int_L g(\mathsf{x}) \ d\mu(\mathsf{x}) \\
&= \left(\frac{d}{d+k}\right) \lim_{n\to\infty} \frac{1}{|S_n|} \sum_{\mathsf{x}\in S_n} \frac{1}{n^k} f(\mathsf{x}).
\end{aligned}$$

□

To put this result in context, we remark that there are three situations in which it is clear that the sphere average should equal the ball average. One case is that of any function averaged over a group with exponential growth, where almost all of the points on the ball will be concentrated on its boundary sphere. Alternately, for any function averaged over $\mathbb{Z}^d$, the points in the ball again become increasingly concentrated in the boundary as $d \to \infty$. Finally, sphere averages clearly equal ball averages for those size-like functions on $\mathbb{Z}^d$ with $k = 0$ (so f is close to a scale-invariant function). These last two cases provide a plausibility check on the $\frac{d}{d+k}$ factor in this statement.

5.2. **Higher moments.** Very natural examples of asymptotically homogeneous functions on $\mathbb{Z}^d$ are those built from powers of word metrics. Applying Proposition 1.3 to these, we get information on the expected word-length and expected location in $\mathbb{R}^d$ for elements in the ball B_n of radius n:

Corollary 5.1 (Expectations)**.** *For any finite generating set for $\mathbb{Z}^d$, the expected geodesic spelling length of words in the ball B_n of radius n is $\frac{d}{d+1}n$, and the expected location in $\mathbb{R}^n$ is on the polygon $\frac{d}{d+1}nL$ where L is the boundary of the convex hull of the chosen generating set.*

Thus, the expected position of a word in B_n is on $S_{\frac{dn}{d+1}}$, independent of the choice of generating set.

To see this in an example, consider $(\mathbb{Z}^2, \mathsf{std})$. As above, set up B_n as the union of S_j for $0 \le j \le n$, noting that $\#S_j = 4j$ for $j \ge 1$ and $\|\mathsf{x}\| = j$ for $\mathsf{x} \in S_j$. Thus the average wordlength over the ball is

$$\frac{4\sum_1^n j^2}{1+4\sum_1^n j} = \frac{4n^3+6n^2+2n}{6n^2+6n+3},$$

which grows like $\frac{2}{3}n$, as predicted. Though it was straightforward to calculate this directly for the simplest choice of generators, it is not apparent a priori how to proceed for an arbitrary generating set.

More generally we can compute higher moments by applying Theorem 1.3 to the functions $f(\mathsf{x}) = |\mathsf{x}|^p$ which are asymptotically homogeneous of order p:

Corollary 5.2 (Higher moments)**.** *The expected value of $|\mathsf{x}|^p$ over B_n is $\frac{d}{d+p}$.*

This tells us that the higher moments are independent of the choice of word metric as well.

5.3. **Asymptotic density.** The counting results can also be used to find the density of group elements with a particular property, (P). Recall that a group element $\mathsf{w} \in \mathbb{Z}^2$ is said to have a *simple spelling* if $\mathsf{w} = a\mathsf{v}_i + b\mathsf{v}_{i+1}$ for consecutive significant generators. We can verify that every word has a simple spelling with respect to the standard generators, whereas only 1 in 36 elements has a simple spelling in $S = \pm\{6\mathsf{e}_1, \mathsf{e}_1, 6\mathsf{e}_2, \mathsf{e}_2\}$ (compare Figure 2, where the words with simple spellings are those appearing on the bounding polygon). Let r be the number of sides of the polygon Q and let $A = \text{Area}(Q)$. There are rn simple spellings of length n, all representing different group elements. On the other hand, $\#S_n \stackrel{+}{\asymp} 2An$. Thus $\lim \frac{\#(S_n \cap (P))}{\#S_n} = \lim \frac{\#(B_n \cap (P))}{\#B_n} = \frac{r}{2A}$, or in other words:

Corollary 5.3. *For $(\mathbb{Z}^2, S)$, the density of words with simple spellings is $r/2A$.*

This does depend on the generating set—only on the convex hull, as usual, but not only on its area—and it holds uniformly at large word-lengths n, as well as when averaging over words of length $\le n$. As a check, recall that Pick's theorem says that $A = i + b/2 - 1$. We know that $r \le b$ and $i \ge 1$, which means $r/2A \le 1$, which is required for plausibility. Besides recovering the answers above, we also see for instance that with respect to the chess-knight generators (see Figure 4) the probability of simple spellings is $2/7$, which would have been extremely unpleasant to derive by hand.

We can likewise define simple spellings in higher dimensions (Def 3.3); the density of simple spellings could be expressed in terms of the asymptotics of number of ordered partitions of large integers n.

5.4. **Instability of geodesics.** For a geodesic space X with basepoint x_0, let $\mathcal{G}(x)$ be the set of all geodesics from x_0 to x. Then define the *instability function* $\mathbf{I}(x)$ to measure how far apart they can be, relative to the size of x:

$$\mathbf{I}(x) := 2\cdot\frac{\operatorname{diam}\mathcal{G}(x)}{d(x,x_0)} = \tfrac{2}{d(x,x_0)}\cdot \sup_{\alpha,\beta\in\mathcal{G}(x)} d_{\mathsf{Haus}}(\alpha,\beta),$$

where the metric on $\mathcal{G}(x)$ is given by Hausdorff distance. In general, $\operatorname{diam}\mathcal{G}(x)\le d(x,x_0)/2$, because any point on a geodesic from x_0 to x is within $d(x,x_0)/2$ of one of the endpoints, and therefore within that distance of any other geodesic. Thus the factor of 2 in the definition normalizes $\mathbf{I}$ so that $0\le \mathbf{I}(x)\le 1$ for any X, x_0, x.

Note that if X has unique geodesics, then $\mathbf{I}\equiv 0$. If X is δ-hyperbolic, then $\operatorname{diam}\mathcal{G}(x)\le\delta$ for all x, so $\mathbf{I}(x)\to 0$ as $x\to\infty$. The situation is different in $\mathbb{Z}^d$: only powers of single generators have unique spellings, so the directions in which $\mathbf{I}$ vanishes are precisely the vertex directions.

We can easily verify that the average instability for an arbitrary $(\mathbb{Z}^d,S)$ can be computed by $\int_L \mathbf{I}\;d\mu_L$. To see this, recall that any word x in a sector decomposes as $\mathsf{x}=\mathsf{x}_0+\mathsf{x}'$, where x_0 is spelled with generators in that sector and $|\mathsf{x}'|\le D_0$ has bounded size. But then any way of rearranging these letters is boundedly close to following a geodesic in the L-norm connecting the origin to x. Thus the fact that the word metric is asymptotic to the norm ensures that the instability function on $\mathbb{Z}^d$ is asymptotic to the instability function on $(\mathbb{R}^d,\|\cdot\|_L)$. The latter is clearly homogeneous of order zero, since both the numerator and denominator are homogeneous of order one.

For instance, the average instability of $\mathbb{Z}^2$ in its standard generators, with respect to either balls or to spheres, is $\frac{1}{2}$. We compute this by averaging $\mathbf{I}$ over the square L, where it takes the value 0 on the vertices, 1 on the midpoints of edges, and varies linearly in between.

Proposition 5.4 (Bounding instability)**.** *For any $(\mathbb{Z}^2,S)$, the average value of the instability is upper-bounded by half the average of the sidelengths of L, weighted by their cone measure:*

$$\lim_{n\to\infty}\frac{1}{|S_n|}\sum_{x\in S_n}\mathbf{I}(x)=\lim_{n\to\infty}\frac{1}{|B_n|}\sum_{x\in B_n}\mathbf{I}(x)\le\frac{1}{2}\sum_i \mu_L(\sigma_i)\cdot\|\sigma_i\|_L.$$

Proof. Consider a point $\mathsf{x}=a\mathsf{u}+b\mathsf{v}$ in a sector given by an edge σ of L with endpoints u and v. The farthest apart geodesics reaching x form a parallelogram with sides $a\mathsf{u}$ and $b\mathsf{v}$. Letting $c=\min(a,b)$, the vector $c(\mathsf{v}-\mathsf{u})$ reaches from the bottom of the parallelogram to the top, so the Hausdorff distance between the extreme geodesics is at most $c\|\mathsf{v}-\mathsf{u}\|=c\|\sigma\|$. The average value of c on σ is $1/4$, so the average value of $\mathbf{I}$ on σ is $\frac{1}{2}\|\sigma\|$. $\square$

In fact, one gets a similar bound for any $(\mathbb{Z}^d,S)$ in terms of the diameter of the faces in the norm.

Note that this does not give a tight bound in general; for instance this gives the trivial bound of 1 for $(\mathbb{Z}^2,\mathsf{std})$. However for the chess-knight generators we get the nontrivial bound $3/7$. Indeed we can force this upper bound to be small for any d: to build S, begin with nearest points to a Euclidean sphere of large radius and then add small vectors to form a generating set.

Proposition 5.5 (Stable presentations)**.** *For any d and any $\epsilon>0$, there exists a finite generating set S such that $\mathbf{I}(\mathsf{x})\le\epsilon$ for all sufficiently large $\mathsf{x}\in\mathbb{Z}^d$.*

5.5. **Sprawl and statistical hyperbolicity.** As is well known, the geometric condition called *hyperbolicity* (sometimes called word hyperbolicity, δ-hyperbolicity, or Gromov hyperbolicity) gives strong algebraic and geometric information about groups. Hyperbolicity is a large-scale invariant, so for finitely generated groups, whether or not the Cayley graph is hyperbolic does not depend on the choice of (finite) generating set. However, if we formulate a metric condition corresponding to hyperbolicity, then the measurements themselves depend on generators. We can quantify the degree to which certain triangle measurements resemble those in hyperbolic groups with a statistic we call the *sprawl* of a group (with respect to a generating set). We give a brief mention here, but develop some results and conjectures in [8].

The sprawl of a group measures the average distance between pairs of points on the spheres in the word metric, normalized by the radius, as the spheres get large. Sprawl thus gives a numerical measure of the asymptotic shape of spheres that can be studied for arbitrary finite presentations of groups. To be precise, let

$$E(G,S) := \lim_{n\to\infty} \frac{1}{|S_n|^2} \sum_{x,y\in S_n} \frac{1}{n} d(x,y),$$

provided this limit exists. One could make an analogous calculation in a symmetric space with respect to Haar measure. Note that since $0 \le d(x,y) \le 2n$, the value is always between 0 and 2. By way of interpretation, $E = 2$ means that generically one can pass through the origin when traveling between any two points on the sphere without taking a significant detour. This statistic is not quasi-isometry invariant but nonetheless captures interesting features of the large-scale geometry.

Hyperbolicity is often characterized with the slogan that "triangles are thin," meaning that the third side of a geodesic triangle must stay within bounded distance of the other two sides. In terms of $x, y \in S_n$, this says that the geodesic $\overline{xy}$ should be about as long as $d(x,0) + d(0,y)$, provided that $\overline{0x}$ and $\overline{0y}$ do not fellow-travel. Thus, if fellow-traveling is relatively rare in a Gromov hyperbolic space, then we will have $E = 2$. We show in [8] that for any non-elementary hyperbolic group with any generating set, $E(G,S) = 2$. (Recall that a hyperbolic group is called *elementary* if it is finite or has a finite-index cyclic subgroup.) Thus, $E < 2$ is an obstruction to hyperbolicity. We will say that a presentation (G,S) is *statistically hyperbolic* if $E(G,S) = 2$; this does not imply that G is a hyperbolic group, but only that this metric calculation works out *on average* as though it were. (For example, $F_2 \times \mathbb{Z}$ is statistically hyperbolic with respect to its standard generators.)

We can study statistical hyperbolicity for free abelian groups with the tools developed in this paper. For a function of several variables $f : (\mathbb{Z}^d)^m \to \mathbb{R}$ asymptotic to $g : (\mathbb{R}^d)^m \to \mathbb{R}$ with g homogeneous of order k, Theorem 1.1 tells us that

$$\lim_{n\to\infty} \frac{1}{|S_n|^m} \sum_{\mathsf{x}\in(S_n)^m} \frac{1}{n^k} f(\mathsf{x}) = \int_{L^m} g(\mathsf{x})\ d\mu^m(\mathsf{x}).$$

Thus it follows immediately from the main result of this paper that

$$E(\mathbb{Z}^d, S) = \int_{L^2} \|\mathsf{x} - \mathsf{y}\|_L \ d\mu^2$$

for all finite generating sets S. That means that we know exactly how sprawl depends on the generators. (There is an exact algorithm for computing this integral presented in [8].)

Corollary 5.6. *For any free abelian group $\mathbb{Z}^d$ with any finite generating set S, the sprawl statistic $E(\mathbb{Z}^d, S)$ is well defined and $E(\mathbb{Z}^d, S) < 2$.*

Proof. The sprawl is computed by integrating against a measure that is absolutely continuous with Lebesgue measure, and we are integrating a function whose maximum value is 2. But a small neighborhood A of any point on the polytope has positive measure, so $A \times A$ has positive measure, and on that set the integrand is strictly less than two. $\square$

With the usual tools for coarse geometry, $\mathbb{Z}^d$ would be indistinguishable from the Euclidean space $\mathbb{R}^d$, since they are quasi-isometric. Instead, let us compare the geometry of four groups: $\mathbb{R}^3$; $\mathbb{Z}^3$ with $\mathsf{std} = \{\pm \mathsf{e}_1, \pm \mathsf{e}_2, \pm \mathsf{e}_3\}$; $\mathbb{Z}^3$ with the "cube" generators $S = \{\pm \mathsf{e}_1 \pm \mathsf{e}_2 \pm \mathsf{e}_3\}$; and the free group F_3. These groups are arranged in order of sprawl, having E values $60/45 < 63/45 < 64/45 < 2$. We advocate the interpretation that they are also arranged from least to most hyperbolic, based on the prevalence of metric shortcuts.

References

[1] M. Beck and S. Robins, *Computing the continuous discretely. Integer-point enumeration in polyhedra.* Undergraduate Texts in Mathematics. Springer, New York, 2007.

[2] M. Benson, *Growth series of finite extensions of $\mathbf{Z}^n$ are rational.* Invent. Math. 73 (1983), no. 2, 251–269.

[3] E. Breuillard, *Geometry of groups of polynomial growth and shape of large balls.* arXiv:0704.0095

[4] D.Yu. Burago, *Periodic metrics.* Representation theory and dynamical systems, 205–210, *Adv. Soviet Math.*, 9, Amer. Math. Soc., Providence, RI, 1992.

[5] P. de la Harpe, *Topics in geometric group theory.* Chicago Lectures in Mathematics. University of Chicago Press, 2000.

[6] R. Durrett, *Probability: Theory and examples.* 2nd edition. Duxbury Press, 1996.

[7] G.H. Hardy and E.M. Wright. *An introduction to the theory of numbers.* Sixth edition. Oxford University Press, 2008.

[8] M. Duchin, C. Mooney, and S. Lelièvre, *Statistical hyperbolicity in groups.* arXiv:1104.4460.

[9] M. Duchin and C. Mooney, *Fine asymptotic geometry in the Heisenberg group.* arXiv:1106.5276.

[10] P. Pansu, *Croissance des boules et des géodésiques fermées dans les nilvariétés.* Ergodic Theory Dynam. Systems 3 (1983), no. 3, 415–445.