SPIRALING OF APPROXIMATIONS AND SPHERICAL AVERAGES OF SIEGEL TRANSFORMS

JAYADEV S. ATHREYA, ANISH GHOSH, AND JIMMY TSENG

ABSTRACT. We consider the question of how approximations satisfying Dirichlet's theorem spiral around vectors in $\mathbb{R}^d$. We give pointwise almost everywhere results (using only the Birkhoff ergodic theorem on the space of lattices). In addition, we show that for *every* unimodular lattice, on average, the directions of approximates spiral in a uniformly distributed fashion on the $d-1$ dimensional unit sphere. For this second result, we adapt a very recent proof of Marklof and Strömbergsson [19] to show a spherical average result for Siegel transforms on $\mathrm{SL}_{d+1}(\mathbb{R})/\mathrm{SL}_{d+1}(\mathbb{Z})$. Our techniques are elementary. Results like this date back to the work of Eskin-Margulis-Mozes [9] and Kleinbock-Margulis [14] and have wide-ranging applications. We also explicitly construct examples in which the directions are not uniformly distributed.

1. INTRODUCTION

It is a corollary of a classical theorem of Dirichlet [6], that, for every $\mathbf{x} \in \mathbb{R}^d$ ($d \geq 1$), there exist infinitely many $(\mathbf{p}, q) \in \mathbb{Z}^d \times \mathbb{N}$ such that

$$\|q\mathbf{x} - \mathbf{p}\| < C_d|q|^{-1/d}. \tag{1.1}$$

Here, $\| \ \|$ denotes the Euclidean norm on $\mathbb{R}^d$ and C_d is a constant depending only on d. If the L^∞-norm is used in (1.1), then C_d can be taken to be 1 for all d. In this paper, we are interested in the distribution of the *directions* of the approximates $(\mathbf{p}, q) \in \mathbb{Z}^d \times \mathbb{N}$ approaching $\mathbf{x}$, that is, the quantities

$$\theta(\mathbf{p}, q) := \frac{q\mathbf{x} - \mathbf{p}}{\|q\mathbf{x} - \mathbf{p}\|} \in \mathbb{S}^{d-1}.$$

Given $A \subset \mathbb{S}^{d-1}$, $T > 0$, we form the counting functions

$$N(\mathbf{x}, T) = \#\{(\mathbf{p}, q) \in \mathbb{Z}^d \times \mathbb{N}, 0 < q \leq T : \|q\mathbf{x} - \mathbf{p}\| < C_d|q|^{-1/d}\}$$

and

$$N(\mathbf{x}, T, A) = \#\{(\mathbf{p}, q) \in \mathbb{Z}^d \times \mathbb{N}, 0 < q \leq T : \|q\mathbf{x} - \mathbf{p}\| < C_d|q|^{-1/d}, \theta(\mathbf{p}, q) \in A\}.$$

Note that, while Dirichlet's theorem guarantees that $N(\mathbf{x}, T) \to \infty$ as $T \to \infty$, $N(\mathbf{x}, T, A)$ could, a priori, be 0 for all $T > 0$. Our first main theorem is

Theorem 1.1. *For $A \subset \mathbb{S}^{d-1}$, a measurable subset, and for almost every $\mathbf{x} \in \mathbb{R}^d$,*

$$\lim_{T \to \infty} \frac{N(\mathbf{x}, T, A)}{N(\mathbf{x}, T)} = \mathrm{vol}(A).$$

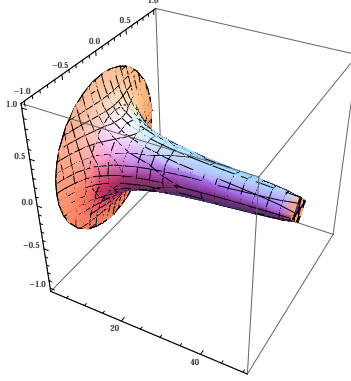*Here* $\mathrm{vol} := \mathrm{vol}_{\mathbb{S}^{d-1}}$ *is the Lebesgue probability measure on* $\mathbb{S}^{d-1}$.

FIGURE 1. The region P_{50} in $\mathbb{R}^{2+1}$

Corollary 1.2. *If* $\mathrm{vol}(A) > 0$, *the inequality*

$$\|q\mathbf{x} - \mathbf{p}\| < C_d|q|^{-1/d}, \theta(\mathbf{p}, q) \in A$$

has infinitely many solutions $(\mathbf{p}, q) \in \mathbb{Z}^d \times \mathbb{N}$ *for almost every* $\mathbf{x} \in \mathbb{R}^d$.

Remark. Logarithmic (in T) almost sure (in $\mathbf{x}$) asymptotics for $N(\mathbf{x}, T)$ follow from work of W. Schmidt [22]; see also [3] for a simple proof. We will show how the latter argument yields similar asymptotics for $N(\mathbf{x}, T, A)$ in §2.

1.1. **Lattices.** Theorem 1.1 is closely related to a general theorem about approximation of directions by lattice vectors. Fix $d \geq 1, c > 0$ and define the set

$$R := \left\{ \mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^d \times \mathbb{R} \ : \ \|\mathbf{v}_1\|^d |v_2| \leq c \right\},$$

which, for v_2 large enough, we may regard as a thinning region around the v_2-axis. And, for $T > 1$, identify pieces of R:

$$(1.2) \qquad P_T := \left\{ \mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^d \times \mathbb{R} \ : \ \|\mathbf{v}_1\|^d v_2 \leq c, 1 < v_2 \leq T \right\} \subset \mathbb{R}^{d+1}.$$

For a subset A of $\mathbb{S}^{d-1}$, we define the subset $P_{A,T}$ of P_T by

$$P_{A,T} := \left\{ \mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ v_2 \end{pmatrix} \in P_T : \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} \in A \right\}.$$

For $\Lambda \subset \mathbb{R}^{d+1}$, a unimodular lattice, define

$$N(\Lambda, T) := \#(\Lambda \cap P_T) \text{ and } N(\Lambda, T, A) := \#(\Lambda \cap P_{A,T}).$$

Recall that $X_{d+1} := \mathrm{SL}_{d+1}(\mathbb{R}) / \mathrm{SL}_{d+1}(\mathbb{Z})$ is the moduli space of unimodular lattices in $\mathbb{R}^{d+1}$ via the identification $g \mathrm{SL}_{d+1}(\mathbb{Z}) \mapsto g\mathbb{Z}^{d+1}$. With this identification, we endow X_{d+1} with the probability measure $\mu = \mu_{d+1}$ induced by the Haar measure on $\mathrm{SL}_{d+1}(\mathbb{R})$. Our second main theorem is

Theorem 1.3. *For* μ-*almost every* $\Lambda \in X_{d+1}$,

$$(1.3) \qquad \lim_{T \to \infty} \frac{N(\Lambda, T, A)}{N(\Lambda, T)} = \mathrm{vol}(A).$$

The relationship to Theorem 1.1 is given by the following standard construction: given $\mathbf{x} \in \mathbb{R}^d$, we form the matrix

$$h_{\mathbf{x}} = \begin{pmatrix} \mathrm{Id}_d & \mathbf{x} \\ 0 & 1 \end{pmatrix}$$

and the associated unimodular lattice in $\mathbb{R}^{d+1}$

$$\Lambda_{\mathbf{x}} := h_{\mathbf{x}}\mathbb{Z}^{d+1} = \left\{ \begin{pmatrix} q\mathbf{x} - \mathbf{p} \\ q \end{pmatrix} \ : \ \mathbf{p} \in \mathbb{Z}^d, q \in \mathbb{Z} \right\}.$$

Then we can view the approximates $(\mathbf{p}, q)$ of $\mathbf{x}$ satisfying Dirichlet's theorem with $1 \leq q \leq T$ as points of the lattice $\Lambda_{\mathbf{x}}$ in the regions P_T, and those with $\theta(\mathbf{p}, q) \in A$ as points of $\Lambda_{\mathbf{x}}$ in $P_{A,T}$ with $c = C_d$. That is, we have that $N(\mathbf{x}, T) = N(\Lambda_{\mathbf{x}}, T)$, and $N(\mathbf{x}, T, A) = N(\Lambda_{\mathbf{x}}, T, A)$, and, so, Theorem 1.1 can be reformulated as saying that (1.3) holds for $\Lambda_{\mathbf{x}}$ for almost every $\mathbf{x} \in \mathbb{R}^d$.

1.2. Average Spiraling.
We also have an L^1 (average) spiraling result on the space of lattices. Fix $0 \leq \epsilon < 1$ and $T > 0$, and define

$$(1.4) \qquad\qquad R_{\epsilon,T} := \{\mathbf{v} \in R \ : \ \epsilon T \leq v_2 \leq T\}$$

and, for a subset A of $\mathbb{S}^{d-1}$ with zero measure boundary,

$$(1.5) \qquad\qquad R_{A,\epsilon,T} := \left\{ \mathbf{v} \in R_{\epsilon,T} \ : \ \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} \in A \right\}.$$

For a unimodular lattice Λ, define

$$N(\Lambda, \epsilon, T) = \#\{\Lambda \cap R_{\epsilon,T}\}$$

and

$$N(\Lambda, A, \epsilon, T) = \#\{\Lambda \cap R_{A,\epsilon,T}\}.$$

Let dk denote the Haar measure on $K := K_{d+1} := \mathrm{SO}_{d+1}(\mathbb{R})$. Our third main theorem is

Theorem 1.4. *For every lattice $\Lambda \in X_{d+1}$, subset $A \subset \mathbb{S}^{d-1}$ with zero measure boundary, and $\epsilon > 0$, we have that*

$$(1.6) \qquad\qquad \lim_{T \to \infty} \frac{\int_K N(k^{-1}\Lambda, A, \epsilon, T) \ dk}{\int_K N(k^{-1}\Lambda, \epsilon, T) \ dk} = \mathrm{vol}(A).$$

Theorem 1.4 is derived from our result on *spherical averages of Siegel transforms*, Theorem 2.2, which we believe to be of independent interest.

1.3. Biased Spiraling.
On the other hand, we construct explicit examples of lattices Λ and directions $\mathbf{v}$ for which (non-averaged) equidistribution does not hold. Our fourth main theorem is

Theorem 1.5. *Let $d \geq 1$. There exists a lattice $\Lambda \in \mathrm{SL}_{d+1}(\mathbb{R})/\mathrm{SL}_{d+1}(\mathbb{Z})$, a set $A \subset \mathbb{S}^{d-1}$ with zero measure boundary, and a sequence $\{T_n\}$ for which*

$$\lim_{n \to \infty} \frac{N(\Lambda, A, \epsilon, T_n)}{N(\Lambda, \epsilon, T_n)} \neq \mathrm{vol}(A)$$

for every $1 > \epsilon \geq 0$.

For $d = 1$, note that $\mathbb{S}^0 := \{-1, 1\}$ and we define $\mathrm{vol}(\{-1\}) = \mathrm{vol}(\{1\}) = 1/2$.

Organization of the paper. In §2, we state Theorem 2.2, our result on spherical averages of Siegel transforms and use it to prove Theorem 1.4. We also prove Theorem 1.3. In §3, we prove Theorem 2.2. The construction of examples of nonuniform spiraling for certain lattice approximates (and the proof of Theorem 1.5) is carried out in §4. Despite the extensive literature on Dirichlet's theorem and its variants, as far as we are aware, our work is the first to study the problem of spiralling for lattice approximates. In a sequel [1], we prove multiparameter versions of the main theorems considered in the present work and related problems in Diophantine approximation and in [2] we establish versions of the main theorems in the wider generality of number fields.

2. Equidistribution on the space of lattices

In this section, we show how to reduce the proofs of Theorems 1.4 and 1.3 to equidistribution problems on the space of lattices. Theorem 1.3 is a consequence of this reduction and the Birkhoff ergodic theorem, which gives us almost everywhere equidistribution of trajectories for diagonal flows. For Theorem 1.4, our ergodic tool Theorem 2.2 will be of independent interest, as it gives an equidistribution theorem for spherical averages of Siegel transforms for quite general functions.

Recall the definition of the Siegel transform: given a lattice Λ in $\mathbb{R}^{d+1}$ and a bounded Riemann-integrable function f with compact support on $\mathbb{R}^{d+1}$, denote by $\widehat{f}$ its *Siegel transform*[1]:

$$\widehat{f}(\Lambda) := \sum_{\mathbf{v} \in \Lambda \setminus \{\mathbf{0}\}} f(\mathbf{v}).$$

Let $\mu = \mu_{d+1}$ be the probability measure on $X_{d+1} := \mathrm{SL}_{d+1}(\mathbb{R}) / \mathrm{SL}_{d+1}(\mathbb{Z})$ induced by the Haar measure on $\mathrm{SL}_{d+1}(\mathbb{R})$ and $d\mathbf{v}$ denote the usual volume measure on $\mathbb{R}^{d+1}$. (We also let $\mathrm{vol} := \mathrm{vol}_{\mathbb{R}^{d+1}}$ denote this volume measure and will make use of the subscript should the need to distinguish it from $\mathrm{vol}_{\mathbb{S}^{d-1}}$ arise.) We recall the classical Siegel Mean Value Theorem [23]:

Theorem 2.1. *Let f be as above.[2] Then $\widehat{f} \in L^1(X_{d+1}, \mu)$ and*

$$\int_{\mathbb{R}^{d+1}} f \ d\mathbf{v} = \int_{X_{d+1}} \widehat{f} \ d\mu.$$

Note that if f is the indicator function of a set $A \setminus \{\mathbf{0}\}$, then $\hat{f}(\Lambda)$ is simply the number of points in $\Lambda \cap (A \setminus \{\mathbf{0}\})$.

Let

$$g_t := \begin{pmatrix} e^t \, \mathrm{Id}_d & 0 \\ 0 & e^{-dt} \end{pmatrix} \in \mathrm{SL}_{d+1}(\mathbb{R})$$

and $e_1, \cdots, e_{d+1}$ be the standard basis of $\mathbb{R}^{d+1}$.

Note that if we set t so that $e^{dt} = T$, we have

$$g_t R_{\epsilon,T} = R_{\epsilon,1} =: R_\epsilon$$

and

$$g_t R_{A,\epsilon,T} = R_{A,\epsilon,1} =: R_{A,\epsilon}.$$

---

[1] One could define the Siegel transform only over primitive lattice points, in which case results analogous to Theorems 2.2 and 2.3 also hold (using, essentially, the same proof).

[2] This condition can be generalized to $f \in L^1(\mathbb{R}^{d+1})$.

We write $\mathbf{1}_{A,\epsilon,T}$ for the characteristic function of $R_{A,\epsilon,T}$ and $\mathbf{1}_{\epsilon,T}$ for the characteristic function of $R_{\epsilon,T}$, and drop the subscript when $T = 1$.

In view of the above discussions, we have

$$(2.1) \qquad N(\Lambda, A, \epsilon, e^t) = \widehat{\mathbf{1}}_{R_{A,\epsilon}}(g_{t/d}\Lambda)$$

$$(2.2) \qquad N(k^{-1}\Lambda, A, \epsilon, T) = \widehat{\mathbf{1}}_{R_{A,\epsilon}}(g_t k\Lambda).$$

Integrating these formulas with respect to t and k respectively, we obtain

$$(2.3) \qquad \int_0^S N(\Lambda, A, \epsilon, e^t)dt = \int_0^S \widehat{\mathbf{1}}_{R_{A,\epsilon}}(g_{t/d}\Lambda)dt$$

$$(2.4) \qquad \int_K N(k^{-1}\Lambda, A, \epsilon, T)dk = \int_K \widehat{\mathbf{1}}_{R_{A,\epsilon}}(g_t k\Lambda)dk.$$

2.1. **Proof of Theorem 1.3 and Theorem 1.1.** Let $s := \frac{\log 2}{d}$. (Note that s is a fixed constant). Moore's ergodicity theorem (see, for example, [4]) states that the action of g_t on X_{d+1} is ergodic, so, by the Birkhoff ergodic theorem, for any $h \in L^1(X_{d+1}, \mu)$, we have, for almost every $\Lambda \in X_{d+1}$,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} h(g_s^n \Lambda) = \int_{X_{d+1}} h d\mu.$$

By Siegel's mean value formula, the functions $\widehat{\mathbf{1}}_{P_{A,2}}$ and $\widehat{\mathbf{1}}_{P_2}$ are in $L^1(\mu)$. Write $Q_i = P_{2^i} \backslash P_{2^{i-1}}$, $Q_{A,i} = P_{A,2^i} \backslash P_{A,2^{i-1}}$. Then, since

$$g_{-s}Q_i = Q_{i+1} \text{ and } g_{-s}Q_{A,i} = Q_{A,i+1},$$

we have

$$\sum_{i=0}^{N-1} \widehat{\mathbf{1}}_{P_2}(g_s^i \Lambda) = \sum_{i=0}^{N-1} \#(\Lambda \cap Q_{i+1}) = \#(\Lambda \cap P_{2^N}) = N(\Lambda, 2^N)$$

$$\sum_{i=0}^{N-1} \widehat{\mathbf{1}}_{P_{A,2}}(g_s^i \Lambda) = \sum_{i=0}^{N-1} \#(\Lambda \cap Q_{A,i+1}) = \#(\Lambda \cap P_{A,2^N}) = N(\Lambda, 2^N, A).$$

By applying the Birkhoff ergodic theorem and the Siegel mean value theorem to these functions, we obtain, for almost every Λ,

$$\lim_{N \to \infty} \frac{1}{N} N(\Lambda, 2^N) = \text{vol}(P_2) \text{ and } \lim_{N \to \infty} \frac{1}{N} N(\Lambda, 2^N, A) = \text{vol}(P_{A,2}).$$

Note that if $F : [0, \infty) \to [0, \infty)$ is an increasing function, and $\frac{F(2^k)}{k} \to \log 2$, then

$$\lim_{T \to \infty} \frac{F(T)}{\log T} = 1.$$

Thus, we obtain (for almost every Λ)

$$\lim_{T \to \infty} \frac{1}{\log T} N(\Lambda, T, A) \bigg/ \frac{1}{\log T} N(\Lambda, T) = \frac{\text{vol}(P_{A,2})}{\text{vol}(P_2)}.$$

Since $\text{vol}(P_{A,2})/\text{vol}(P_2) = \text{vol}(A)$, we obtain Theorem 1.3. $\qquad\square$

This argument was used in [3] to obtain logarithmic (in T) asymptotics for $N(\Lambda, T)$ (and other related functions), and, as shown, also yields logarithmic asymptotics for $N(\Lambda, T, A)$. To prove Theorem 1.1, consider the collection of matrices $\{h_{\mathbf{x}} : \mathbf{x} \in \mathbb{R}^d\}$ forms the *horospherical* subgroup for $\{g_t\}$, and as such, the set of lattices $\{\Lambda_{\mathbf{x}} : \mathbf{x} \in \mathbb{R}^d\}$ is the *unstable* manifold for the action of $\{g_t\}_{t \geq 0}$

on X_{d+1}. In particular, for almost every $\mathbf{x} \in \mathbb{R}^d$, $\Lambda_{\mathbf{x}}$ is Birkhoff generic for the action of g_s, which following the above argument, yields Theorem 1.1 (see [3, Section 2.5.1]). $\qquad\square$

2.2. **Statement of results for Siegel transforms.** By (2.4), to prove Theorem 1.4, we need to show the equidistribution of the Siegel transforms of the sets $R_{A,\epsilon}$ and R_ϵ with respect to the integrals over g_t-translates of K. The main ergodic tool in this setting is our fifth main theorem, a result on the spherical averages of Siegel transforms:

Theorem 2.2. *Let f be a bounded Riemann-integrable function of compact support on $\mathbb{R}^{d+1}$. Then for any $\Lambda \in X_{d+1}$,*

$$\lim_{t \to \infty} \int_{K_{d+1}} \widehat{f}(g_t k \Lambda) \, \mathrm{d}k = \int_{X_{d+1}} \widehat{f} \, \mathrm{d}\mu.$$

We note that the above theorem is reminiscent of [9, Theorem 3.4] of Eskin-Margulis-Mozes, but the compact group and the one-parameter diagonal subgroup used there are different. In fact, a proof of this theorem in this spirit can be assembled from the work of Kleinbock-Margulis [14, Appendix], but we present an elementary proof relying on counting lattice points in balls. This proof is adapted from [19, Section 5.1], where the result is proved for balls around the origin (for the slightly different context of a cut-and-project quasicrystal).[3] For the proof of Theorem 2.2, we must adapt this proof for balls not containing the origin. This is done in Section 3 for the upper bound:

Theorem 2.3. *Let f be a bounded function of compact support in $\mathbb{R}^{d+1}$ whose set of discontinuities has zero Lebesgue measure. Then for any $\Lambda \in X_{d+1}$,*

$$\lim_{t \to \infty} \int_{K_{d+1}} \widehat{f}(g_t k \Lambda) \, \mathrm{d}k \leq \int_{X_{d+1}} \widehat{f} \, \mathrm{d}\mu.$$

REMARK 2.4. Since $\mathbb{R}^{d+1}$ is σ-compact, it follows immediately from the theorem that the assumption that f has compact support can be replaced with that of $f \in L^1(\mathbb{R}^{d+1})$—the other assumptions are still, however, necessary for the proof.

Corollary 2.5. *Let f be a bounded Riemann-integrable function of compact support in $\mathbb{R}^{d+1}$. Then for any $\Lambda \in X_{d+1}$,*

$$\lim_{t \to \infty} \int_{K_{d+1}} \widehat{f}(g_t k \Lambda) \, \mathrm{d}k \leq \int_{X_{d+1}} \widehat{f} \, \mathrm{d}\mu.$$

Proof. Immediate from the theorem and the Lebesgue criterion. $\qquad\square$

In this paper, we will focus on upper bounds, i.e. on the proof of Theorem 2.3. The lower bound, on the other hand, follows either from the methods in [15] or by applying the following equidistribution theorem (Theorem 2.6) of Duke, Rudnick and Sarnak (cf. [7]) (a simpler proof was given by Eskin and McMullen [10] using mixing and generalized by Shah [21]) and then approximating the Siegel transform $\widehat{f}$ from below by $h \in C_c(X_{d+1})$.

Theorem 2.6. *Let G be a non-compact semisimple Lie group and let K be a maximal compact subgroup of G. Let Γ be a lattice in G, let λ be the probabilty Haar measure on G/Γ, and let ν be any probability measure on K which is absolutely continuous with respect to a Haar measure on K. Let $\{a_n\}$ be a sequence of elements of G without accumulation points. Then for any $x \in G/\Gamma$ and any $h \in C_c(G/\Gamma)$,*

$$\lim_{n \to \infty} \int_K h(a_n k x) \, \mathrm{d}\nu(k) = \int_{G/\Gamma} h \, \mathrm{d}\lambda.$$

REMARK 2.7. One can replace $\mathrm{d}k$ by $\mathrm{d}\nu(k)$ in Theorems 2.2 and 2.3 without any changes to the proofs.

---

[3]For an introduction to [19], see [18]. Also see [17] and [20].

2.3. **Proof of Theorem 1.4.** We prove Theorem 1.4 using Theorem 2.2, while deferring the proof of the latter to Section 3. Thus, applying Theorem 2.2 to characteristic functions of $R_{A,\epsilon}$ and R_ϵ, we obtain

$$\lim_{t\to\infty} \int_K \widehat{\mathbf{1}}_{R_{A,\epsilon}}(g_t k\Lambda)\mathrm{d}\nu(k) = \int_{X_{d+1}} \widehat{\mathbf{1}}_{R_{A,\epsilon}}\mathrm{d}\mu = \mathrm{vol}(R_{A,\epsilon}),$$

where we have applied Siegel's mean value theorem in the last equality.[4] We apply this to numerator as well as denominator in (1.6) to get

$$\lim_{T\to\infty} \frac{\int_K N(k^{-1}\Lambda, A, \epsilon, T)\ \mathrm{d}k}{\int_K N(k^{-1}\Lambda, \epsilon, T)\ \mathrm{d}k} = \frac{\mathrm{vol}(R_{A,\epsilon})}{\mathrm{vol}(R_\epsilon)} = \mathrm{vol}(A),$$

which finishes the proof. $\qquad\square$

3. Proof of Theorem 2.3

As mentioned, to prove Theorem 2.2, we need only show the upper bound (Theorem 2.3):

$$(3.1) \qquad\qquad \lim_{t\to\infty} \int_{K_{d+1}} \widehat{f}(g_t k\Lambda)\ \mathrm{d}k \le \int_{X_{d+1}} \widehat{f}\ \mathrm{d}\mu.$$

We will approximate using step functions on balls (see Section 3.4), where we use the norm on $\mathbb{R}^{d+1} = \mathbb{R}^d \times \mathbb{R}$ given by the supremum of Euclidean norm in $\mathbb{R}^d = \mathrm{span}(e_1, \cdots, e_d)$ and by absolute value in $\mathbb{R} = \mathrm{span}(e_{d+1})$. Hence, balls will be open regions of $\mathbb{R}^{d+1}$ that are rods (i.e. solid cylinders). We need four cases: balls centered at $\mathbf{0} \in \mathbb{R}^{d+1}$, balls centered in $\mathrm{span}(e_{d+1})\backslash\{\mathbf{0}\}$, balls centered in $\mathrm{span}(e_1, \cdots, e_d)\backslash\{\mathbf{0}\}$, and all other balls. Since we will approximate using step functions, it suffices (as we show in Section 3.4) to assume that the balls in the second case do not meet $\mathbf{0}$ and in the last case do not meet $\mathrm{span}(e_{d+1}) \cup \mathrm{span}(e_1, \cdots, e_d)$.[5] Let $E := B(\boldsymbol{w}, r)$ be any such ball and χ_E be its characteristic function. By the monotone convergence theorem, we have

$$\int_{K_{d+1}} \widehat{\chi}_E(g_t k\Lambda)\ \mathrm{d}k = \sum_{\boldsymbol{v}\in\Lambda\backslash\{\mathbf{0}\}} \int_{K_{d+1}} \chi_{k^{-1}g_t^{-1}E}(\boldsymbol{v})\ \mathrm{d}k.$$

We show each case in turn. For the first case, we refer the reader to [19, Section 5.1], in which the desired result (with balls given by the Euclidean norm in $e_1, \cdots, e_{d+1}$) is shown for quasicrystals and is essentially the same for us. (Alternatively, a simplified version of the proof of the third case will also show the first case.) The proofs of the other three cases adapt this basic idea. For convenience of exposition, we show the fourth case before the third.

---

[4]There are two parameters $0 < r_1 < r_2$, easy to compute and depending on ε, such that, if we define the set W to be the union of all intervals with a terminal point on the sphere of radius r_1 in $\mathbb{R}^d \times \{0\}$, the other terminal point on the sphere of radius r_2 in $\mathbb{R}^d \times \{0\}$, lying in a ray emanating from the origin of $\mathbb{R}^d \times \{0\}$, and passing through a point of A, then we may regard the set $R_{A,\epsilon}$ as the region under the graph of the continuous function $v_2 = \frac{c}{\|\boldsymbol{v}_1\|^d}$ over W union a cylinder over the sphere of radius r_1. Since $\partial(A)$ has zero Lebesgue measure (on $\mathbb{S}^{d-1}$), so does $\partial(W)$ (on $\mathbb{R}^d$) using polar coordinates for Lebesgue measurable functions. As $R_{A,\epsilon}$ is bounded, it has finite Lebesgue measure. Applying Fubini's Theorem and noting that the cylinder over the sphere is a Jordan set, one sees that the compact set $\partial(R_{A,\epsilon})$ has zero Lebesgue measure (on $\mathbb{R}^{d+1}$). It is easy to see that a compact set of zero Lebesgue measure also has zero Jordan content. Therefore, the set $R_{A,\epsilon}$ is a Jordan set. Likewise, for R_ϵ. It is well-known that the characteristic functions of Jordan sets are Riemann-integrable and, thus, Theorem 2.2 applies to these functions.

[5]We note that the second and the fourth cases already suffice to show Theorem 1.4.

3.1. **The second case: balls centered in** $\mathrm{span}(e_{d+1})\backslash\{\mathbf{0}\}$. In this case, $\boldsymbol{w} = we_{d+1}$ for some $w \neq 0$. The proof is similar for whether w is positive or negative, so we may assume without loss of generality that $w > 0$. Let

$$\widetilde{B}^d := \widetilde{B}^d(r) := \{(x_1, \cdots, x_d)^t \mid R^2 x_1^2 + \cdots + R^2 x_d^2 < r^2\}$$

and

$$I(w) := I(w,r) := \{x_{d+1} \mid w - r < \frac{x_{d+1}}{R^d} < w + r\}$$

where $R := e^t$. Then $g_t^{-1}E$ is the rod given by

$$\widetilde{B}^d \times I(w).$$

Replacing the $< r^2$ with $= \rho^2$ for a $0 \leq \rho^2 < r^2$ and using the equation for the sphere $\tau\mathbb{S}^d$ where $\tau > 0$, we note that the intersection has at most two values for the x_{d+1}. Moreover, since we only need to consider balls not meeting $\mathbf{0}$, we may assume that the rod completely lies in the half-space determined by e_{d+1} and thus $x_{d+1} = \sqrt{\tau^2 - \rho^2/R^2} =: c(\rho)$ is our value of intersection provided that it lies in $I(w)$. Therefore, the intersection $\mathfrak{C}(\tau)$ of the rod with the sphere is a d-dimensional *cap*, through which each intersection (i.e. slice) with the affine hyperplane through and normal to $c(\rho)e_{d+1}$ is a $d-1$-dimensional sphere (with radius ρ/R). There are two types of caps. A *full cap* is a cap such that

(3.2) $c(\rho)|_{0 < \rho^2 < r^2} \subset I(w).$

We remark that the continuity of the function c means that (3.2) is interval inclusion. Otherwise, a cap is called an *end cap* because it is near one or the other end of the rod. Now let $\tau_- := R^d(w - r)$ and $\tilde{\tau}_+ := R^d(w + r)$. Then, for $0 < \tau < \tau_-$ and $\tau_+ := \sqrt{\tilde{\tau}_+^2 + r^2/R^2} < \tau$, the intersection between sphere and rod is empty. Since whether a cap is full or end depends only on (3.2), $\mathfrak{C}(\tau)$ are all full caps for $\tilde{\tau}_- := \sqrt{\tau_-^2 + r^2/R^2} < \tau < \tilde{\tau}_+$.

We are only interested in the case where R is large (and where r is small). Since

$$\tilde{\tau}_- - \tau_- = O(R^{-(d+2)})$$
$$\tau_+ - \tilde{\tau}_+ = O(R^{-(d+2)}),$$

the end caps are negligible (as we shall see below). We remark that these estimates also give the approximate "depth" of any full cap.

For R large, the d-dimensional volume of a full cap is nearly, but slightly larger than, that of the d-dimensional ball $\mathfrak{B}$ with boundary sphere exactly the slice through $c(r)$. More precisely,

(3.3) $\dfrac{\mathrm{vol}_{\tau\mathbb{S}^d}(\mathfrak{C}(\tau))}{\mathrm{vol}_d(\mathfrak{B})} \searrow \gamma(\tilde{\tau}_+) \geq 1$

as $\tau \nearrow \tilde{\tau}_+$ over the interval $(\tilde{\tau}_-, \tilde{\tau}_+)$. (Note that $\gamma(\tilde{\tau}_+) \searrow 1$ as $\tilde{\tau}_+ \nearrow \infty$.)

Let $\tilde{K}_d := \begin{pmatrix} \mathrm{SO}_d(\mathbb{R}) & 0 \\ 0 & 1 \end{pmatrix}$. Then $g_t^{-1}E$ is stabilized by every element of $\tilde{K}_d$. It is well-known that the unit sphere $\mathbb{S}^d$ can be realized as the homogeneous space $\tilde{K}_d\backslash K_{d+1}$. Consequently,

$$\int_{K_{d+1}} \chi_{k^{-1}g_t^{-1}E}(\boldsymbol{v}) \, \mathrm{d}k = \frac{1}{\mathrm{vol}_{\mathbb{S}^d}(\mathbb{S}^d)} \int_{\mathbb{S}^d} \chi_{s^{-1}g_t^{-1}E}(\boldsymbol{v}) \, \mathrm{dvol}_{\mathbb{S}^d}(s)$$

(3.4) $= \dfrac{\mathrm{vol}_{\mathbb{S}^d}(\mathbb{S}^d \cap \|\boldsymbol{v}\|^{-1}g_t^{-1}E)}{\mathrm{vol}_{\mathbb{S}^d}(\mathbb{S}^d)} =: A_R^E(\|\boldsymbol{v}\|)$

where the second equality follows from the correspondence $s\boldsymbol{v} \in g_t^{-1}E \iff s\boldsymbol{v}/\|\boldsymbol{v}\| \in \|\boldsymbol{v}\|^{-1}g_t^{-1}E \cap \mathbb{S}^d$. The invariance of the ratio of the volume measures on spheres of different radii implies that

$$A_R^E(\tau) = \frac{\mathrm{vol}_{\tau\mathbb{S}^d}(\tau\mathbb{S}^d \cap g_t^{-1}E)}{\mathrm{vol}_{\tau\mathbb{S}^d}(\tau\mathbb{S}^d)}$$

for $\tau > 0$. By (3.3) and the fact that the full caps are determined by polynomial equations, $A_R^E(\tau)$ is a strictly decreasing smooth function with respect to τ over the interval $(\tilde{\tau}_-, \tilde{\tau}_+)$.

To take care of the end caps, we shall replace our rod with a slightly larger one. More precisely, replace $I(w, r)$ with $I(w, r + \frac{1}{R^{2d+1}})$, which, for R large enough, lengthens the rod enough (see estimates above) so that all caps of the shorter rod are now full caps of the longer rod. In particular, $A_R^E(\tau)$ is strictly decreasing over the interval (τ_-, τ_+).

Let $B_{\text{Euc}}(\mathbf{0}, \tau)$ denote a ball of radius τ in $\mathbb{R}^{d+1}$ with respect to the Euclidean norm. Now it follows from the formula for A_R^E that

$$\sum_{\boldsymbol{v} \in \Lambda \backslash \{\mathbf{0}\}} A_R^E(\|\boldsymbol{v}\|) \le \int_{\tau_-}^{\tau_+} \#\big(B_{\text{Euc}}(\mathbf{0}, \tau) \cap \Lambda \backslash \{\mathbf{0}\}\big) \, (-\mathrm{d}A_R^E(\tau))$$

where the integral is the Riemann-Stieltjes integral. We remark that the integrability of the function $\#\big(B_{\text{Euc}}(\mathbf{0}, \tau) \cap \Lambda \backslash \{\mathbf{0}\}\big)$ follows from its monotonicity.

Now the well-known generalizations of the Gauss circle problem (or, alternatively, [19, Proposition 3.2]) show that

$$\#\big(B_{\text{Euc}}(\mathbf{0}, \tau) \cap \Lambda \backslash \{\mathbf{0}\}\big) \le (1 + \varepsilon) \operatorname{vol}(B_{\text{Euc}}(\mathbf{0}, \tau))$$

(3.5)
$$= (1 + \varepsilon) \operatorname{vol}(B_{\text{Euc}}(\mathbf{0}, 1)) \tau^{d+1}$$

for $\varepsilon \to 0$ as $R \to \infty$ (and hence as $\tau_- \to \infty$).

Also, by (3.3), we have that $C(\tau_-) := \operatorname{vol}_{\tau \mathbb{S}^d}(\mathfrak{C}(\tau_-)) \to \operatorname{vol}_d(\mathfrak{B})$ as $R \to \infty$ (and hence as $\tau_- \to \infty$). And we have

$$A_R^E(\tau) \le \frac{C(\tau_-)}{\tau^d \operatorname{vol}_{\mathbb{S}^d}(\mathbb{S}^d)}$$

over the interval (τ_-, τ_+). Consequently, it follows that

$$\sum_{\boldsymbol{v} \in \Lambda \backslash \{\mathbf{0}\}} A_R^E(\|\boldsymbol{v}\|) \le d \int_{\tau_-}^{\tau_+} (1 + \varepsilon) \operatorname{vol}(B_{\text{Euc}}(\mathbf{0}, 1)) \frac{C(\tau_-)}{\operatorname{vol}_{\mathbb{S}^d}(\mathbb{S}^d)} \, \mathrm{d}\tau$$

$$= d(1 + \varepsilon) \frac{\operatorname{vol}(B_{\text{Euc}}(\mathbf{0}, 1))}{\operatorname{vol}_{\mathbb{S}^d}(\mathbb{S}^d)} C(\tau_-)(\tau_+ - \tau_-).$$

Now $C(\tau_-)(\tau_+ - \tau_-)$ is the volume of a rod that has length within $O(\frac{1}{R^{d+1}})$ of the length of our original rod, but with cross-section volume $C(\tau_-)$. Let $R \to \infty$ and $\varepsilon \to 0$, we have

$$\sum_{\boldsymbol{v} \in \Lambda \backslash \{\mathbf{0}\}} A_R^E(\|\boldsymbol{v}\|) \le \operatorname{vol}(E),$$

as desired. (Note that $\frac{\operatorname{vol}(B_{\text{Euc}}(\mathbf{0},1))}{\operatorname{vol}_{\mathbb{S}^d}(\mathbb{S}^d)} = \frac{1}{d+1}$.)

3.2. The fourth case: all other balls.
The proof of this case is similar to the second case. The differences are as follows. In this case, $g_t^{-1}E$ is no longer invariant under every element of $\widetilde{K}_d$; however, we need this property only to show (3.4), which also follows because $\mathrm{d}vol_{\mathbb{S}^d}(s)$ is preserved under rotations. Caps are no longer such simple geometric objects (their boundaries are ellipsoids, not spheres); however, the proof is unaffected by this change.

3.3. The third case: balls centered in $\operatorname{span}(e_1, \cdots, e_d) \backslash \{\mathbf{0}\}$.
The difference between this case and the fourth case is that the ends of the rod do not both go to we_{d+1} for $w \to \infty$ or both for $w \to -\infty$, but one end goes to one direction and the other goes to the other. To take care of this issue, we must consider what happens near the origin. Fix the lattice Λ. Then, by the discreteness of the lattice, there is a ball $B_{\text{Euc}}(\mathbf{0}, \tau_0)$) that does not meet $\Lambda \backslash \{\mathbf{0}\}$ for some $\tau_0 > 0$ depending only on Λ. Since there are no relevant lattice points in any ball around the origin of radius smaller than τ_0, we need only consider $\tau \ge \tau_0$. Note that the rod and sphere meet in two connected components, each of which we will refer to as caps. This fact, however, does not affect the proof, except in minor ways as noted below.

Now recall the center of our ball E is $\boldsymbol{w}$, whose last coordinate is $w_{d+1} = 0$. Thus, $g_t^{-1}\boldsymbol{w} \to \boldsymbol{0}$. And since the rod $g_t^{-1}E$ is contracting in $x_1, \cdots, x_d$, our analysis of $A_R^E(\tau)$ for the fourth case also hold in this case for large enough R over the desired range $\tau \geq \tau_0$. In particular, for large R, we have that $A_R^E(\tau)$ is a smooth decreasing function of τ over (τ_0, τ_+) such that

$$A_R^E(\tau) \leq \frac{C(\tau_0)}{\tau^d \mathrm{vol}_{\mathbb{S}^d}(\mathbb{S}^d)}.$$

Here $C(\tau_0)$ is the volume of the two caps that is the relevant intersection. We may take care of end caps as in the previous cases.

Since, for R large, our analysis in the fourth case applies, we have from (3.3) that $C(\tau_0) = O(R^{-d})$ or, more precisely,

$$(3.6) \qquad\qquad \mathrm{vol}_d(\mathfrak{B}) \leq C(\tau_0) < \frac{d+1}{d}\, \mathrm{vol}_d(\mathfrak{B}).$$

Here, $\mathrm{vol}_d(\mathfrak{B})$ is twice the volume of the intersection of a d-hyperplane normal to e_{d+1} with the rod. Finally, as in the second case, for every $\varepsilon > 0$, there exists a (large) $\tau_1 > 0$ depending only on Λ such that for all $\tau \geq \tau_1$, we have that (3.5) holds. Hence, for large R, we have, as in the previous cases,

$$\sum_{\boldsymbol{v} \in \Lambda \backslash \{\boldsymbol{0}\}} A_R^E(\|\boldsymbol{v}\|) \leq \int_{\tau_0}^{\tau_+} \#\big(B_{\mathrm{Euc}}(\boldsymbol{0}, \tau) \cap \Lambda \backslash \{\boldsymbol{0}\}\big)\, (-\mathrm{d}A_R^E(\tau))$$

$$\leq const(\tau_1) \int_{\tau_0}^{\tau_1} (-\mathrm{d}A_R^E(\tau)) + (1+\varepsilon)\, \mathrm{vol}(B_{\mathrm{Euc}}(\boldsymbol{0}, 1)) \int_{\tau_1}^{\tau_+} \tau^{d+1}\, (-\mathrm{d}A_R^E(\tau))$$

$$\leq O(R^{-d}) + d(1+\varepsilon) \frac{\mathrm{vol}(B_{\mathrm{Euc}}(\boldsymbol{0}, 1))}{\mathrm{vol}_{\mathbb{S}^d}(\mathbb{S}^d)} C(\tau_0)(\tau_+ - \tau_1),$$

where the inequality for second integral follows as in previous cases. Since the function that counts lattice points in larger and larger balls is monotonically increasing, $const(\tau_1) := (1+\varepsilon)\, \mathrm{vol}(B_{\mathrm{Euc}}(\boldsymbol{0}, 1))\tau_1^{d+1}$, and thus the implicit constant depends only on r, ε, τ_0, and τ_1. Applying (3.6) and letting $R \to \infty$ and $\varepsilon \to 0$ yields the desired result. (Recall that $\mathrm{vol}_d(\mathfrak{B})$ is twice the volume of the intersection of a d-hyperplane normal to e_{d+1} with the rod.)

3.4. **Finishing the proof.** We may now approximate using step functions to obtain Theorem 2.3.[6] For the convenience of the reader, we give a proof. Let $F \subset \mathbb{R}^{d+1}$ be a compact subset with nonempty interior (i.e. $F^\circ \neq \emptyset$) and measure zero boundary (i.e. $\mathrm{vol}(\partial F) = 0$). Recall that our balls, which we have called rods, are given by the Euclidean norm in $e_1, \cdots, e_d$ and by absolute value in e_{d+1}—the unit ball defines a norm $\| \cdot \|_{\mathrm{rod}}$ on $\mathbb{R}^{d+1}$. Let us denote the rods defined in the beginning of this section (Section 3) as case-one, case-two, case-three, or case-four rods, respectively—recall these rods are open sets. We will refer to a sequence of sets as a *disjoint* sequence if the sets in the sequence are pairwise disjoint.

Lemma 3.1. *There exists a disjoint sequence $\{B_n\}_{n=1}^{\infty}$ of case-one, case-two, case-three, and case-four rods of $\mathbb{R}^{d+1}$ so that*

$$F^\circ \supset \coprod B_n$$
$$\mathrm{vol}(F) = \sum \mathrm{vol}(B_n).$$

Moreover, the radii of all the rods in the sequence may be chosen to be $\leq \eta$ for any choice of $\eta > 0$.

This lemma follows easily from a classical result: the Vitali covering theorem (see [12, Theorem 1.6] for example).

---

[6]The construction presented here is general and may be of independent interest.

Theorem 3.2 (Vitali Covering Theorem). *Let A be a subset of a doubling metric space (X, μ) and $\mathcal{F}$ be a collection of closed balls centered at A such that*

$$\inf\{r > 0 \mid \overline{B(a, r)} \in \mathcal{F}\} = 0$$

for each $a \in A$. Then there exists a countable subcollection $\{\overline{B_n}\}_{n=1}^\infty$ of pairwise disjoint closed balls such that

$$\mu\left(A \backslash \bigcup_{n=1}^\infty \overline{B_n}\right) = 0.$$

It will soon become apparent that we would like to mimic the construction of the Riemann integral, which, for the functions that we are interested in, is constructed over cubes. But, instead of cubes, we will use partitions (mod 0) consisting of pairwise disjoint open rods and will use the Vitali covering lemma (stated below). Let B_0 be a large enough open rod, not necessarily of any of the four cases, containing $\text{supp}(f)$ and let ρ_0 be its radius. Let $\mathcal{P}_0 = \{B_0\}$ be the initial partition.

3.4.1. *Refinements of partitions.* We recursively define refinements as follows: let $\mathcal{P}_j$ be a partition (mod 0) in which every element of $\mathcal{P}_{j-1}$ with radius $> 2^{-j}\rho_0$ is replaced with a sequence of pairwise disjoint open rods as constructed in Lemma 3.1 with $\eta \leq 2^{-j}\rho_0$.[7] We note that $\mathcal{P}_j$ is a countable union of case-one, case-two, case-three, and case-four rods only. Let us define the *size* of a partition $\mathcal{P}_j$ to be the supremum over all radii of elements in $\mathcal{P}_j$. Then the size of $\mathcal{P}_j \leq 2^{-j}\rho_0$. By construction, we have that

$$\bigcup_{B \in \mathcal{P}_j} B \subset \bigcup_{B \in \mathcal{P}_{j-1}} B$$

and, by Lemma 3.1, we have that both sets have Lebesgue measure equal to $\text{vol}(B_0)$ and thus the following set has Lebesgue measure zero

$$\mathcal{D}' = \overline{B_0} \Big\backslash \bigcap_{j=0}^\infty \bigcup_{B \in \mathcal{P}_j} B.$$

Now let $\mathcal{D}''$ denote the set of discontinuities of f, a set of Lebesgue measure zero by assumption. Let $\mathcal{D} := \mathcal{D}' \cup \mathcal{D}''$. Note that $\mathcal{D}$ is subset of $\overline{B_0}$, but it need not be compact.

3.4.2. *Approximating f by step functions.* We now wish to approximate f on the full measure set $\overline{B_0} \backslash \mathcal{D}$ by step functions over the rods of $\mathcal{P}_j$. Define the step functions on $\overline{B_0} \backslash \mathcal{D}$ as follows:

$$f_j := \sum_{B \in \mathcal{P}_j} \left(\sup f\big|_B\right) \chi_{B \backslash \mathcal{D}}.$$

For any point of $\overline{B_0} \backslash \mathcal{D}$, we note that the sum only has one term. Moreover, it is easy to see that these functions converge Lebesgue-a.e. to f:

Lemma 3.3. *The step functions $f_j \to f$ for every point of $\overline{B_0} \backslash \mathcal{D}$.*

3.4.3. *Proof for the case of step functions.* We now prove Theorem 2.3 for f_j. Take any total ordering of the set $\mathcal{P}_j = \{B_n\}_{n=1}^\infty$ and let

$$f_{j,n} := \sum_{i=1}^n \left(\sup f\big|_{B_i}\right) \chi_{B_i \backslash \mathcal{D}}.$$

Then we have that

$$\lim_{n \to \infty} f_{j,n} = f_j,$$

where the limit denotes pointwise convergence over the domain $\overline{B_0} \backslash \mathcal{D}$.

---

[7]Note that the many choices that we make in constructing the partition and its refinements do not affect the proof.

Let $M := \sup |f|_{\overline{B}_0}|$. Fix a lattice $g_t k\Lambda\backslash\{\mathbf{0}\}$. Since $M\chi_{\overline{B}_0\backslash\mathcal{D}} \leq M\chi_{\overline{B}_0}$, we have that

$$\sum_{\boldsymbol{v}\in g_t k\Lambda\backslash\{\mathbf{0}\}} M\chi_{\overline{B}_0\backslash\mathcal{D}}(\boldsymbol{v}) < \infty$$

because the same sum over $\chi_{\overline{B}_0}$ is the number of lattice points in this ball, which is finite. Now since $|f_{j,n}| \leq M\chi_{\overline{B}_0\backslash\mathcal{D}}$, dominated convergence implies that

$$(3.7) \qquad \sum_{\boldsymbol{v}\in g_t k\Lambda\backslash\{\mathbf{0}\}} f_j = \lim_{n\to\infty} \sum_{\boldsymbol{v}\in g_t k\Lambda\backslash\{\mathbf{0}\}} f_{j,n} < \infty.$$

Here, more precisely, $\boldsymbol{v} \in g_t k\Lambda\backslash\{\mathbf{0}\}\cap\overline{B}_0\backslash\mathcal{D}$. Furthermore, we note that (3.7) holds for every $k \in K_{d+1}$. Applying dominated convergence again, we conclude

$$(3.8) \qquad \int_{K_{d+1}} \widehat{f}_j(g_t k\Lambda) \, \mathrm{d}k = \lim_{n\to\infty} \int_{K_{d+1}} \sum_{\boldsymbol{v}\in g_t k\Lambda\backslash\{\mathbf{0}\}} f_{j,n} \, \mathrm{d}k.$$

Now define $\widetilde{f}_{j,n} := \sum_{i=1}^n \left(\sup f|_{B_i}\right)\chi_{B_i}$ and repeating the above with $\widetilde{f}_{j,n} + M\chi_{\mathcal{D}}$ (which is dominated by $2M\chi_{\overline{B}_0}$) in place of $f_{j,n}$ yields

$$(3.9) \qquad \int_{K_{d+1}} \widehat{f}_j(g_t k\Lambda) \, \mathrm{d}k \leq \lim_{n\to\infty} \int_{K_{d+1}} \sum_{\boldsymbol{v}\in g_t k\Lambda\backslash\{\mathbf{0}\}} \widetilde{f}_{j,n} \, \mathrm{d}k + \int_{K_{d+1}} \sum_{\boldsymbol{v}\in g_t k\Lambda\backslash\{\mathbf{0}\}} M\chi_{\mathcal{D}} \, \mathrm{d}k.$$

Note that

$$\lim_{n\to\infty} \widetilde{f}_{j,n} + M\chi_{\mathcal{D}}$$

is well defined and the limit denotes pointwise convergence on $\overline{B}_0$.

To handle the integral involving the zero Lebesgue measure set $\mathcal{D}$, we proceed as follows. Recall that Lebesgue measure is outer regular, which applied to $\mathcal{D}$ is the following:

$$\mathrm{vol}(\mathcal{D}) = \inf\{\mathrm{vol}(U) \mid U \supset \mathcal{D}, \ U \text{ open}\}.$$

It is easy to see that the subcollection of open rods

$$\mathfrak{T} := \left\{ B\left(x, \frac{1}{5}r\right) \mid B(x,r) \text{ is a case-one, case-two, case-three, or case-four rod} \right\}$$

is a basis of the usual topology of $\mathbb{R}^{d+1}$. Consequently, for every $\gamma' > 0$, there exists a family $\{B'_\alpha\} \subset \mathfrak{T}$ such that

$$\mathcal{D} \subset \cup B'_\alpha \quad \text{and} \quad \mathrm{vol}(\cup B'_\alpha) < \gamma'.$$

We now require a classical result: the Vitali covering lemma (see [11, Lemma 1.9] or [12, Theorem 1.16] for example).

Theorem 3.4 (Vitali covering lemma)**.** *Let $\mathfrak{C}$ be a collection of balls contained in a bounded subset of $\mathbb{R}^{d+1}$. Then there exists a finite or countably infinite subcollection of pairwise disjoint balls $\{B_m\}$ such that*

$$\bigcup_{B\in\mathfrak{C}} B \subset \bigcup_m 5B_m$$

where $5B_m$ is the ball concentric with B_m of 5 times the radius.

Moreover, the subcollection $\{B_m\}$ can be chosen so that

$$\sum_{m=1}^{\infty} \chi_{5B_m}(x) \leq C(d)$$

where C is a constant depending only on d.

The covering lemma allows us to cover $\mathcal{D}$ by a countable collection of our rods:

Lemma 3.5. *For every $\gamma > 0$, there exists (at most) a countable sequence $\{\widetilde{B}_m\}_{m=1}^{\infty}$ of case-one, case-two, case-three, and case-four rods such that*

$$\mathcal{D} \subset \cup_{m=1}^{\infty} \widetilde{B}_m \quad and \quad \sum_{m=1}^{\infty} \mathrm{vol}(\widetilde{B}_m) < \gamma.$$

Proof. Let $\{B_\alpha'\}$ (defined above) be $\mathfrak{C}$ and $\{B_m'\}_{m=1}^{\infty}$ be the subcollection in the Vitali covering lemma. Let $\widetilde{B}_m := 5B_m'$. Then $\{\widetilde{B}_m\}$ is a countable collection of case-one, case-two, case-three, and case-four rods. Consequently, we have

$$\mathrm{vol}(\mathcal{D}) \leq \sum_{m=1}^{\infty} \mathrm{vol}(\widetilde{B}_m) = \sum_{m=1}^{\infty} 5^{d+1}\, \mathrm{vol}(B_m') = 5^{d+1}\, \mathrm{vol}(\cup_{m=1}^{\infty} B_m') < 5^{d+1}\gamma'.$$

Our desired result is now immediate. □

Choose $\gamma > 0$ small. To bound the integral involving $\mathcal{D}$ in (3.9), we may approximate with the $\{\widetilde{B}_m\}_{m=1}^{\infty}$ from Lemma 3.5 as follows. We have

$$0 \leq \int_{K_{d+1}} \sum_{\boldsymbol{v} \in g_t k \Lambda \setminus \{\boldsymbol{0}\}} M\chi_{\mathcal{D}} \, \mathrm{d}k \leq \int_{K_{d+1}} \sum_{\boldsymbol{v} \in g_t k \Lambda \setminus \{\boldsymbol{0}\}} \sum_{m=1}^{\infty} M\chi_{\widetilde{B}_m} \, \mathrm{d}k$$

because

$$M\chi_{\mathcal{D}}(\boldsymbol{v}) \leq M\chi_{\cup_m \widetilde{B}_m}(\boldsymbol{v}) \leq \sum_{m=1}^{\infty} M\chi_{\widetilde{B}_m}(\boldsymbol{v})$$

for all $\boldsymbol{v}$. Now, replacing all instances of our use of the dominated convergence theorem by the monotone convergence theorem in the argument that we used to show (3.8), we have that

$$\int_{K_{d+1}} \sum_{\boldsymbol{v} \in g_t k \Lambda \setminus \{\boldsymbol{0}\}} \sum_{m=1}^{\infty} M\chi_{\widetilde{B}_m} \, \mathrm{d}k = \lim_{N \to \infty} \sum_{m=1}^{N} M \int_{K_{d+1}} \sum_{\boldsymbol{v} \in g_t k \Lambda \setminus \{\boldsymbol{0}\}} \chi_{\widetilde{B}_m} \, \mathrm{d}k,$$

where, a priori, the limit may be infinite—we will, however, show that the limit is finite and can be made arbitrarily small below.

Putting this together with (3.9) yields

$$(3.10) \qquad \int_{K_{d+1}} \widehat{f}_j(g_t k \Lambda) \, \mathrm{d}k \leq \lim_{n \to \infty} \sum_{i=1}^{n} \left(\sup f\big|_{B_i} \right) \int_{K_{d+1}} \sum_{\boldsymbol{v} \in g_t k \Lambda \setminus \{\boldsymbol{0}\}} \chi_{B_i} \, \mathrm{d}k$$

$$+ \lim_{N \to \infty} \sum_{m=1}^{N} M \int_{K_{d+1}} \sum_{\boldsymbol{v} \in g_t k \Lambda \setminus \{\boldsymbol{0}\}} \chi_{\widetilde{B}_m} \, \mathrm{d}k,$$

which is an inequality involving only characteristic functions on case-one, case-two, case-three, and case-four rods, and thus our results in the beginning of this section (Section 3) for these rods apply as follows.

The main term. We now estimate the first term of the right-hand side of (3.10), which we refer to as the *main term*. Consider the *ancillary step functions*

$$f_{j,n}' := \sum_{i=1}^{n} \left(\sup f\big|_{B_i} - M \right) \chi_{B_i}.$$

These $f_{j,n}'$ are dominated by $2M\chi_{\overline{B}_0}$. By the disjointness of the $\{B_i\}$, the function

$$f_j' := \lim_{n \to \infty} f_{j,n}'$$

is well-defined (and the limit denotes pointwise convergence). Now define

$$F_{j,n} := \sum_{i=1}^{n} M\chi_{B_i}$$

and
$$F_j := \lim_{n \to \infty} F_{j,n},$$
which is also well-defined function. Moreover, they satisfy

(3.11)
$$f_j' = \lim_{n \to \infty} \widetilde{f}_{j,n} - F_j.$$

We may apply the proof that we used to derive (3.8) with f_j' in place of f_j and $f_{j,n}'$ in place of $f_{j,n}$ to deduce
$$\int_{K_{d+1}} \widehat{f_j'}\,(g_t k \Lambda)\,\mathrm{d}k = \lim_{n \to \infty} \int_{K_{d+1}} \sum_{\boldsymbol{v} \in g_t k \Lambda \setminus \{\boldsymbol{0}\}} f_{j,n}'\,\mathrm{d}k = \lim_{n \to \infty} \int_{K_{d+1}} \widehat{f_{j,n}'}(g_t k \Lambda)\,\mathrm{d}k.$$

Moreover, these ancillary step functions satisfy
$$f_{j,n}' \geq f_{j,n+1}' \geq \cdots \geq -2M \chi_{\overline{B}_0}$$
for all n and, consequently, we have
$$\int_{K_{d+1}} \sum_{\boldsymbol{v} \in g_t k \Lambda \setminus \{\boldsymbol{0}\}} f_{j,n}'\,\mathrm{d}k \geq \int_{K_{d+1}} \sum_{\boldsymbol{v} \in g_t k \Lambda \setminus \{\boldsymbol{0}\}} f_{j,n+1}' \geq \mathrm{const},$$

which forms a monotonically decreasing sequence of real numbers (for every fixed t) and thus converges as $n \to \infty$ to a limit that is less than any element in the sequence. Namely, we have that
$$\lim_{t \to \infty} \lim_{n \to \infty} \int_{K_{d+1}} \sum_{\boldsymbol{v} \in g_t k \Lambda \setminus \{\boldsymbol{0}\}} f_{j,n}'\,\mathrm{d}k \leq \lim_{t \to \infty} \int_{K_{d+1}} \sum_{\boldsymbol{v} \in g_t k \Lambda \setminus \{\boldsymbol{0}\}} f_{j,n}'\,\mathrm{d}k$$

for every n. Applying our results for characteristic functions on case-one, case-two, case-three, and case-four balls, we have
$$\lim_{t \to \infty} \int_{K_{d+1}} \widehat{f_j'}\,(g_t k \Lambda)\,\mathrm{d}k \leq \sum_{i=1}^{n} \left(\sup f\big|_{B_i} - M \right) \mathrm{vol}(B_i)$$

for every n. Letting $n \to \infty$, we have

(3.12)
$$\lim_{t \to \infty} \int_{K_{d+1}} \widehat{f_j'}\,(g_t k \Lambda)\,\mathrm{d}k \leq \int_{\mathbb{R}^{d+1}} f_j\,\mathrm{d}\boldsymbol{v} - \sum_{i=1}^{\infty} M\,\mathrm{vol}(B_i)$$

Here we have used the properties of the Lebesgue integral, the fact that the $\{B_i\}$ are pairwise disjoint, and the fact that $\mathcal{D}$ has zero Lebesgue measure.

An easy modification of the above argument for $F_{j,n}$ in place of $f_{j,n}'$ yields

(3.13)
$$\lim_{t \to \infty} \int_{K_{d+1}} \widehat{F_j}(g_t k \Lambda)\,\mathrm{d}k \leq \sum_{i=1}^{\infty} M\,\mathrm{vol}(B_i).$$

Finally, applying (3.11), (3.12), (3.13), and the fact that the argument used to deduce (3.8) also works for $\widetilde{f}_{j,n}$ implies the following:

(3.14)
$$\lim_{t \to \infty} \lim_{n \to \infty} \sum_{i=1}^{n} \left(\sup f\big|_{B_i} \right) \int_{K_{d+1}} \sum_{\boldsymbol{v} \in g_t k \Lambda \setminus \{\boldsymbol{0}\}} \chi_{B_i}\,\mathrm{d}k \leq \int_{\mathbb{R}^{d+1}} f_j\,\mathrm{d}\boldsymbol{v}.$$

This handles the main term.

The null term. The second term of the right-hand side of (3.10), we refer to as the *null term* and handle exactly as $F_{j,n}$ to obtain
$$\lim_{t \to \infty} \lim_{N \to \infty} \sum_{m=1}^{N} M \int_{K_{d+1}} \sum_{\boldsymbol{v} \in g_t k \Lambda \setminus \{\boldsymbol{0}\}} \chi_{\widetilde{B}_m}\,\mathrm{d}k \leq \sum_{i=m}^{\infty} M\,\mathrm{vol}(\widetilde{B}_m),$$

where we note that it is immaterial that the $\{\widetilde{B}_m\}$ may not be pairwise disjoint by the second part of Theorem 3.4. Finally, letting $\gamma \to 0$ in the proceeding (which bounds the null term from above by 0) and applying (3.10) and (3.14) yields our desired result:

$$\lim_{t\to\infty} \int_{K_{d+1}} \widehat{\widetilde{f}_j}(g_t k \Lambda) \; \mathrm{d}k \le \int_{\mathbb{R}^{d+1}} f_j \; \mathrm{d}\mathbf{v}.$$

REMARK 3.6. Instead of introducing ancillary step functions, one could use the compactness of $\overline{B}_0$ to obtain a finite subcover of rods in $\mathcal{P}_j$ and $\{\widetilde{B}_m\}_{m=1}^{\infty}$ to approximate f_j.

3.4.4. *Proof for the general case.* The f_j's are dominated by $M\chi_{\overline{B}_0}$. Using an analogous argument to that for $f'_{j,n}$, we have that

$$\lim_{t\to\infty} \lim_{j\to\infty} \int_{K_{d+1}} \sum_{\mathbf{v} \in g_t k \Lambda \setminus \{\mathbf{0}\}} f_j \; \mathrm{d}k \le \lim_{t\to\infty} \int_{K_{d+1}} \sum_{\mathbf{v} \in g_t k \Lambda \setminus \{\mathbf{0}\}} f_j \; \mathrm{d}k$$

for every j. Now, applying our result for step functions, we have

$$\lim_{t\to\infty} \int_{K_{d+1}} \widehat{f}(g_t k \Lambda) \; \mathrm{d}k \le \int_{\mathbb{R}^{d+1}} f_j \; \mathrm{d}\mathbf{v}$$

for every j. Finally, we apply dominated convergence and Lemma 3.3 to obtain our desired result:

$$\lim_{t\to\infty} \int_{K_{d+1}} \widehat{f}(g_t k \Lambda) \; \mathrm{d}k \le \lim_{j\to\infty} \int_{\mathbb{R}^{d+1}} f_j \; \mathrm{d}\mathbf{v} = \int_{\mathbb{R}^{d+1}} f \; \mathrm{d}\mathbf{v}.$$

4. Nonuniform spiraling: proof of Theorem 1.5

In this section, we prove Theorem 1.5 by using continued fractions to construct a family of one-dimensional examples for which the directions are not uniformly distributed. For higher dimensions, we use non-minimal toral translations as examples.

We can strengthen the conclusion of Theorem 1.5:

Theorem 4.1. *Let $d \ge 1$. There exists a lattice $\Lambda \in \mathrm{SL}_{d+1}(\mathbb{R})/\mathrm{SL}_{d+1}(\mathbb{Z})$ and a set A of $\mathbb{S}^{d-1}$ for which*

$$\liminf_{T\to\infty} \frac{\#\{\Lambda \cap R_{A,\epsilon,T}\}}{\#\{\Lambda \cap R_{\epsilon,T}\}} > \mathrm{vol}(A)$$

and

$$\limsup_{T\to\infty} \frac{\#\{\Lambda \cap R_{-A,\epsilon,T}\}}{\#\{\Lambda \cap R_{\epsilon,T}\}} < \mathrm{vol}(A)$$

for every $1 > \epsilon \ge 0$.

4.1. **Proof of Theorems 1.5 and 4.1 in dimension one.** We prove Theorem 4.1, which also suffices to show Theorem 1.5. We must construct a lattice in $\mathbb{R}^2$ and pick a set A of $\mathbb{S}^0$ for that lattice. Let

$$A = \{-1\}.$$

To construct the lattice, we construct a number $x \in \mathbb{R}\backslash\mathbb{Q}$ using continued fractions (see [13] for an introduction) and form the associated unimodular lattice Λ_x. At the end, we will note that our method of construction provides a family of numbers, corresponding to a family of lattices, which satisfy the theorem. Using an analogous construction allows us to consider $A = \{1\}$ too.

Let x be the irrational number between zero and one for which

$$a_n := \begin{cases} 4 & \text{if } n \text{ is odd,} \\ n^n & \text{if } n \text{ is even} \end{cases}$$

is the nth continued fraction element (note $n \geq 1$). Since x is irrational, there is an unique $p \in \mathbb{Z}$ for which $|qx - p| < 1/2$, which, by forgetting p, we can regard as a rotation of the circle $\mathbb{R}/\mathbb{Z}$ by the unique representative of qx in the interval $(-1/2, 1/2)$. And therefore the only lattice points that matter for the region R from (1.2) are those $(p, q) \in \mathbb{Z} \times \mathbb{Z}$ coming from this rotation. Also, since the negation of a lattice point in R stays in R, we may, without loss of generality, consider lattice points with $q \in \mathbb{N}$. Finally, since x is positive, our lattice points will have $p \in \mathbb{N} \cup \{0\}$.

For some pairs of such (p, q), the ratio $\frac{qx-p}{\|qx-p\|}$ will be 1 and for others -1, which is equivalent to asking whether (p, q) is on one or the other side of the ray starting at the origin and going through $(x, 1)^T$, which is equivalent to asking whether $qx - p > 0$ or $qx - p < 0$, and which, if, for conciseness, we introduce the notation

$$q \cdot x$$

to denote the circle rotation above, is equivalent to asking whether $q \cdot x > 0$ or $q \cdot x < 0$.

Let p_n/q_n denote the nth convergent of x. We will use the following well-known facts about continued fractions and circle rotations:

(1) The rotations $q_{n-1} \cdot x$ and $q_n \cdot x$ alternate in sign.
(2)

$$p_n = a_n p_{n-1} + p_{n-2}$$
$$q_n = a_n q_{n-1} + q_{n-2}$$

(3)

$$q_n p_{n-1} - p_n q_{n-1} = -1^n$$

(4)

$$\frac{1}{q_n + q_{n+1}} < |q_n \cdot x| < \frac{1}{q_{n+1}}$$

(5) Convergents are best approximates (of the second kind):

$$|q_n \cdot x| < |q \cdot x|$$

for all $0 < q < q_{n+1}$.

The following is a general fact of the continued fraction of any irrational number:

Lemma 4.2. *We have*

$$\frac{a_{n+1}}{2} < \frac{|q_{n-1} \cdot x|}{|q_n \cdot x|} < a_{n+1} + 2.$$

Proof. Both inequalities follow from Facts (2) and (4).   □

For our particular number x, the lemma implies that

Corollary 4.3. *For n, an even number, we have*

$$2 < \frac{|q_{n-1} \cdot x|}{|q_n \cdot x|} < 6$$

and, for n, an odd number, we have

$$\frac{(n+1)^{n+1}}{2} < \frac{|q_{n-1} \cdot x|}{|q_n \cdot x|} < (n+1)^{n+1} + 2.$$

Since $x < 1/2$, we have that $1 \cdot x > 0$. Using facts about continued fractions, the usual conventions $q_{-1} = 0$, $p_{-1} = 1$, and that $p_0 = 0$ by construction, it follows that $q_0 = 1$. Consequently,

Lemma 4.4. *For n, an even integer, we have*

$$q_n \cdot x > 0$$

and, for n, an odd integer, we have

$$q_n \cdot x < 0.$$

Proof. As noted, $q_0 \cdot x > 0$. Fact (1) immediately implies the result. $\square$

Since the denominators of the convergents are strictly increasing, Fact (4) implies that (p_n, q_n) are in R and that, for n large enough, $q_n \cdot x$ is itself a rotation by a small angle (much smaller than angle x). However, there are other lattice points in R, which we now describe.

We will count relevant lattice points by induction; it is convenient to induct on n, the index of the convergents. Since we are considering a limit, we may start counting lattice points starting with some large q_n without affecting our result. Therefore, we may assume that $\frac{1}{q_n}$ is small.

We are interested in lattice points in R. Recall that these lattice points come from the above-mentioned rotations and hence lattice points in R are equivalent to rotations $q \cdot x$ for which $|q \cdot x| \leq \frac{1}{q}$. To help us count, let us enlarge the lattices points of interest to those corresponding to

$$(4.1) \qquad |q \cdot x| \leq \frac{1}{q_n}$$

for $q_n \leq q < q_{n+1}$ and exclude those not in R. It follows from Fact (4) that the only lattice points satisfying (4.1) from those corresponding to $0 \leq q < q_n$ are q_{n-1} and $2q_{n-1}$ on one side of 0, 0 itself, and one on the other side of 0, which we will say corresponds to $\tilde{q}$—since convergents are best approximates, we know that $|\tilde{q} \cdot x| > |q_{n-1} \cdot x|$.

For the initial step of the induction on n, we have chosen to ignore the lattice points corresponding to $q_{n-1}, 2q_{n-1}, 0$, and $\tilde{q}$ and, for an induction step, we have already counted the contribution from these points. It the in-between lattice points corresponding to $q_n \leq q < q_{n+1}$ that concern us. The division algorithm describes all such lattice points as follows: $q = mq_n + r$. The only remainders r of interest are the ones already chosen, namely $q_{n-1}, 2q_{n-1}, 0$, and $\tilde{q}$. By Fact (2) applied to x, there are always at least three in-between points—to be precise, these in-between points for a given remainder r correspond to $\{q_n + r, 2q_n + r, 3q_n + r, \cdots\}$. Moreover, the number of in-between points is either $a_{n+1} - 1$ or a_{n+1} depending on the remainder r. If $r_1 \cdot x$ and $r_2 \cdot x$ are adjacent on the circle for $0 \leq r_1 \neq r_2 < q_n$, then their in-between lattice points divides the interval formed by $r_1 \cdot x$ and $r_2 \cdot x$ up into equal length pieces with the sole exception of one piece which may be slightly longer—this observation follows from Fact (1) and the fact that convergents are best approximates and rotations are isometries.

Let us consider these in-between points. We claim that the only in-between point for the remainder $r = \tilde{q}$ that may be relevant corresponds to $q_n + \tilde{q}$. First note, by Fact (1), the fact that circle rotation is translation on the abelian group $\mathbb{R}/\mathbb{Z}$, and that this translation is an isometry, we have that $|(mq_n + \tilde{q}) \cdot x| > |(\ell q_n + \tilde{q}) \cdot x|$ for $a_{n+1} > m > \ell \geq 0$. For $m \geq 2$, we have

$$|(mq_n + \tilde{q}) \cdot x| > \frac{1}{2q_n}$$

by Fact (4), but we also have

$$\frac{1}{mq_n + \tilde{q}} < \frac{1}{2q_n},$$

which shows our claim. The lattice point corresponding to $q_n + \tilde{q}$ is only one point and may be ignored for the limit that we are computing.

We claim that the only in-between point for remainder $r = 2q_{n-1}$ that may be relevant corresponds to $q_n + 2q_{n-1}$. The proof is analogous to that for $\tilde{q}$. And the possible relevant lattice point can be ignored for the limit.

For the remaining two remainders, the behaviors differ greatly (by construction) for odd-indexed and even-indexed convergents; we consider these cases separately.

4.1.1. *Odd-indexed convergents.* Let n be odd. We will show that many of the in-between lattice points for the remainder 0 are in R, while very few of the in-between lattice points for remainder q_{n-1} are. Let us first consider the in-between points for q_{n-1}, which correspond to $\{mq_n + q_{n-1}\}$ for $0 < m < a_{n+1}$ by Fact (2). By Lemma 4.4, we have

$$0 < (mq_n + q_{n-1}) \cdot x < (\ell q_n + q_{n-1}) \cdot x < q_{n-1} \cdot x$$

for $0 < \ell < m < a_{n+1}$. Note that we have exactly $M := a_{n+1} - 1$ in-between points between $q_{n-1} \cdot x$ and 0 in the given range. Since circle rotation by x is an abelian group, these in-between points divide up the interval between 0 and $q_{n-1} \cdot x$ into equal length segments, except for the segment with 0 as an endpoint, which is slightly longer. Thus, we have

$$(4.2) \qquad\qquad q_{n-1} \cdot x - \frac{m}{M+1}(q_{n-1} \cdot x) < (mq_n + q_{n-1}) \cdot x$$

for $0 < m \le M$. Now to be excluded from R, a lattice point must satisfy the following condition

$$\frac{1}{mq_n + q_{n-1}} < (mq_n + q_{n-1}) \cdot x,$$

which is satisfied, as one can see by applying Fact (4) to (4.2), if the point satisfies

$$\frac{1}{mq_n} < \frac{1}{2q_n}\left(1 - \frac{m}{M+1}\right).$$

The latter condition, in turn, is equivalent to asking at which values of m is the parabola $-m^2 + (M+1)m - 2(M+1) > 0$. The answer is between the two roots, which, for M large enough, are as close as we like to 2 and $M-1$. Since n is odd, M can be chosen large. Thus, except, possibly, for four in-between points, the rest are excluded from R. We can ignore these four points for computing the limit.

Finally, to finish the odd-indexed case, we consider in-between points for 0. There are $N := a_{n+1}$ of such points in the given range (which divide up the segment between $\tilde{q} \cdot x$ and 0 into equal length pieces, except for a slightly longer piece with endpoint $\tilde{q} \cdot x$). From (4.2), we have that

$$(4.3) \qquad\qquad -\frac{m}{N}(q_{n-1} \cdot x) < (mq_n) \cdot x$$

for $0 < m < N$. Now for a lattice point to be in R, we need the following condition to hold:

$$-\frac{1}{mq_n} \le (mq_n) \cdot x,$$

which is satisfied, as one can see by applying Fact (4) to (4.3), if the point satisfies

$$-\frac{1}{mq_n} \le -\frac{m}{q_n N}.$$

Let L_n be the number of in-between lattice points for remainder 0 in R. Our calculation implies that $L_n \ge \lfloor (n+1)^{(n+1)/2} \rfloor$.

4.1.2. *Even-indexed convergents.* For any remainder, there are at most $a_{n+1} = 4$ in-between points. All of these can be ignored in the limit calculation.

4.1.3. *Finishing the proof of Theorem 4.1.* Thus, the lattice points that project onto A have count $\sum(L_{2k+1} + P_1)$ where P_1 is a natural number ≤ 5. While, all lattice points in R have count $\sum(L_{2k+1} + P_2)$ where P_2 is a natural number ≤ 10. It is clear that fixing an ϵ does not affect the proceeding. Therefore, we have shown that

$$\liminf_{T \to \infty} \frac{\#\{\Lambda \cap R_{A,\epsilon,T}\}}{\#\{\Lambda \cap R_{\epsilon,T}\}} = 1$$

and

$$\limsup_{T \to \infty} \frac{\#\{\Lambda \cap R_{-A,\epsilon,T}\}}{\#\{\Lambda \cap R_{\epsilon,T}\}} = 0.$$

The theorem is now immediate.

4.2. **Other numbers that satisfy Theorem 4.1 in dimension one.** It is clear that our construction of x via continued fractions in the proof of Theorem 4.1 is a general construction. Let $a_n(x_1)$ and $a_n(x_2)$ be the n-th elements of the continued fraction expansions of x_1 and x_2, respectively. Then we define $x_1 \# x_2$ to be the continued fraction whose elements are

$$a_{2n}(x_1 \# x_2) := a_n(x_1)$$
$$a_{2n+1}(x_1 \# x_2) := a_n(x_2)$$

for $n \in \mathbb{N} \cup \{0\}$. Since continued fractions are unique, the operation $\#$ is a well-defined (noncommutative) product of real numbers, which we refer to as the *continued fraction product*. It is clear from our construction above that, to satisfy the theorem, the number $x_1 \# x_2$ must have $a_n(x_2)$ grow faster than $a_n(x_1)$ as $n \to \infty$. Let us refer to such numbers as *unbalanced*.

Finally, reversing the order of the continued fraction product for our constructed number x will provide an example of a number satisfying the theorem for $A = \{1\}$.

4.3. **Proof of Theorems 1.5 and 4.1 in higher dimensions.** We use the well-known fact:

Lemma 4.5. *The toral translation by a vector $\boldsymbol{x} = (x_1, \cdots, x_d)^T$ is non-minimal if and only if there exist integers $k_1, \cdots, k_d$ not all zero such that $\sum k_i x_i \in \mathbb{Z}$.*

This is example is a simple observation. Let $d \geq 1$. Let $\boldsymbol{x}$ correspond to a non-minimal toral translation. Then, it follows that there is a primitive integer lattice vector $\boldsymbol{v}$ perpendicular (with respect to the usual dot product in $\mathbb{R}^{d+1}$) to the $d+1$-vector $(\boldsymbol{x}^T, 1)^T$. And $\boldsymbol{v} \neq (0, \cdots, 0, 1)^T$. Changing the basis of $\mathbb{Z}^{d+1}$ to $\{\boldsymbol{v}, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_{d+1}\}$ allows us to see that $(\boldsymbol{x}^T, 1)^T$ lies in the $X := \text{span}\{\boldsymbol{v}_2, \cdots, \boldsymbol{v}_{d+1}\}$ (thought of as a subspace of $\mathbb{R}^{d+1}$). Now it follows that the basis vectors $\{\boldsymbol{v}, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_{d+1}\}$ determine a parallelepiped of $d+1$-volume equal to 1. Hence it follows that the only lattice points of $\mathbb{Z}^{d+1}$ closer to $(\boldsymbol{x}^T, 1)^T$ than the Euclidean distance between v and X must lie on X.

Now let $Y := \text{span}\{\boldsymbol{e}_1, \cdots \boldsymbol{e}_d\}$ (thought of as a subspace of $\mathbb{R}^{d+1}$). The spaces X and Y do not coincide because their normal vectors are not in the same direction. Therefore $X \cap Y \cap \mathbb{Z}^{d+1}$ is a proper sublattice of $X \cap \mathbb{Z}^{d+1}$ and hence no lattice points in the thinning region R (after becoming thin enough) project onto $\mathbb{S}^{d-1}$ outside of this sublattice—the projection is onto a lower dimensional sphere $\mathbb{S}^{d-2}$. This proves the theorem for $d \geq 2$.

For $d = 1$, we note that the $(x, 1)^T$ is a rational vector and hence goes through an point of $\mathbb{Z}^2$. It easy to see that all integer lattice points close enough to $(x, 1)^T$ lie on the line through it—in this case, there is no projection at all. This proves the theorem for $d = 1$.

We conclude by remarking that an exaggerated version of the proceeding example is given by taking the unbalanced number x constructed in the proof of Theorem 4.1 in dimension one and forming $\boldsymbol{x} = (x, \cdots, x)^T$. This gives a higher dimensional example satisfying Theorem 4.1.

References

[1] Jayadev S. Athreya, Anish Ghosh, and Jimmy Tseng, *Spherical averages of Siegel transforms for higher rank diagonal actions and applications*, submitted.

[2] Jayadev S. Athreya, Anish Ghosh, and Jimmy Tseng, *Spiraling in number fields*, in preparation.

[3] Jayadev Athreya, Andrew Parrish, and Jimmy Tseng, *Ergodic Theory and Diophantine approximation for linear forms and translation surfaces*, preprint, 2014.

[4] B. Bekka and M. Mayer, *Ergodic Theory and Topological Dynamics of group actions on homogeneous spaces*, Cambridge University Press, 2000.

[5] S. G. Dani, *Divergent trajectories of flows on homogeneous spaces and Diophantine approximation.* J. Reine Angew. Math. 359 (1985), 55–89.

[6] L. G. P. Dirichlet, *Veral lgemeinerung eines Satzes aus der Lehre von den Kettenbrüchen nebst einige Anwendungen auf die Theorie der Zahlen*, S.-B. Preuss. Akad. Wiss. (1842), 93–95.

[7] W. Duke, Z. Rudnick, and P. Sarnak. Density of integer points on affine homogeneous varieties. Duke Math. J., 71(1):143–179, 1993.

[8] A. Eskin and G. Margulis, *Recurrence Properties of Random Walks on Finite Volume Homogeneous Manifolds*, Random walks and geometry, 431–444, Walter de Gruyter, Berlin, 2004.

[9] A. Eskin, G. Margulis and S. Mozes, *Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture*, Ann. of Math. 147 (1998), no. 1, 93–141.

[10] A. Eskin and C. McMullen, *Mixing, counting, and equidistribution in Lie groups*, Duke Math. J., 71(1):181–209, 1993.

[11] K. J. Falconer, "The geometry of fractal sets," Cambridge Tracts in Mathematics **85**, Cambridge University Press, Cambridge, 1986.

[12] J. Heinonen, "Lectures on analysis on metric spaces," Universitext, Springer-Verlag, New York, 2001.

[13] A. Khinchin, "Continued Fractions," The University of Chicago Press, Chicago, 1964.

[14] D. Kleinbock and G. Margulis, *Bounded orbits of nonquasiunipotent flows on homogeneous spaces*, Amer. Math. Soc. Transl. (1996), v. 171, 141-172.

[15] D. Kleinbock and G. Margulis, *Logarithm laws for flows on homogeneous spaces*, Invent. Math., 138 (1999), 451–494.

[16] D. Kleinbock, N. Shah and A. Starkov, *Dynamics of subgroup actions on homogeneous spaces of Lie groups and applications to number theory*, In Handbook of dynamical systems, Vol. 1A, pages 813–930. North-Holland, Amsterdam, 2002.

[17] J. Marklof and A. Strömbergsson, *The Boltzmann-Grad limit of the periodic Lorentz gas*, Annals of Mathematics 174 (2011) 225–298.

[18] J. Marklof and A. Strömbergsson, *The distribution of free path lengths in the periodic Lorentz gas and related lattice point problems*, Annals of Mathematics 172 (2010) 1949–2033.

[19] J. Marklof and A. Strömbergsson, *Free path lengths in quasicrystals*, Communications in Mathematical Physics 330 (2014) 723–755.

[20] J. Marklof and A. Strömbergsson, *The periodic Lorentz gas in the Boltzmann-Grad limit: asymptotic estimates*. Geom. Funct. Anal. 21 (2011), 560–647.

[21] Nimish Shah, *Limit distributions of expanding translates of certain orbits on homogeneous spaces*, Proc. Indian Acad. Sci., Math. Sci. 106 (1996) 105–125.

[22] W. Schmidt, *Asymptotic formulae for point lattices of bounded determinant and subspaces of bounded height*, Duke Math J. 35 (1968), 327–339.

[23] C. S. Siegel, *A mean value theorem in geometry of numbers*, Ann. Math. 46 (1945), 340–347.

J.S.A.: Department of Mathematics, University of Illinois Urbana-Champaign, 1409 W. Green Street, Urbana, IL 61801, USA
    E-mail address: `jathreya@illinois.edu`

A.G.: School of Mathematics, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400005 India
    E-mail address: `ghosh@math.tifr.res.in`

J.T.: School of Mathematics, University of Bristol, University Walk, Bristol, BS8 1TW UK
    E-mail address: `j.tseng@bristol.ac.uk`