Moscow Journal of Combinatorics and Number Theory
Vol. 9, No. 2, 2020
dx.doi.org/10.2140/moscow.2020.9.101

A dynamical Borel–Cantelli lemma via improvements to Dirichlet's theorem

Dmitry Kleinbock and Shucheng Yu

Let $X \cong \mathrm{SL}_2(\mathbb{R})/\mathrm{SL}_2(\mathbb{Z})$ be the space of unimodular lattices in $\mathbb{R}^2$, and for any $r \geq 0$ denote by $K_r \subset X$ the set of lattices such that all its nonzero vectors have supremum norm at least e^{-r}. These are compact nested subsets of X, with $K_0 = \bigcap_r K_r$ being the union of two closed horocycles. We use an explicit second moment formula for the Siegel transform of the indicator functions of squares in $\mathbb{R}^2$ centered at the origin to derive an asymptotic formula for the volume of sets K_r as $r \to 0$. Combined with a zero-one law for the set of the ψ-Dirichlet numbers established by Kleinbock and Wadleigh (*Proc. Amer. Math. Soc.* **146** (2018), 1833–1844), this gives a new dynamical Borel–Cantelli lemma for the geodesic flow on X with respect to the family of shrinking targets $\{K_r\}$.

1. Introduction

Let (X, μ) be a probability space, and let $\{a_s\}_{s\in\mathbb{R}}$ be a one-parameter measure-preserving flow on X. Given a family of measurable subsets $\{B_s\}_{s>0}$ of X with $\mu(B_s) \to 0$ as $s \to \infty$ (called *shrinking targets*), the *shrinking targets problem* asks for a dichotomy on whether generic orbits of $\{a_s\}_{s>0}$ would hit the shrinking targets indefinitely. That is, we are looking for a zero-one law for the measure of the limsup set

$$B_\infty := \limsup_{s\to\infty} a_{-s} B_s = \{x \in X \mid a_s x \in B_s \text{ for an unbounded set of } s > 0\}.$$

For any $n \in \mathbb{N}$ let

$$\widetilde{B}_n := \bigcup_{0\leq s<1} a_{-s} B_{n+s} \tag{1-1}$$

be the thickening of the shrinking targets $\{B_s\}_{n\leq s<n+1}$ along the flow $\{a_{-s}\}_{0\leq s<1}$. Note that $a_n x \in \widetilde{B}_n$ if and only if there exists some $s \in [n, n+1)$ such that $a_s x \in B_s$. We thus have

$$B_\infty = \limsup_{n\to\infty} a_{-n} \widetilde{B}_n = \{x \in X \mid a_n x \in \widetilde{B}_n \text{ infinitely often}\}, \tag{1-2}$$

and the classical Borel–Cantelli lemma implies

$$\sum_n \mu(\widetilde{B}_n) < \infty \quad \Longrightarrow \quad \mu(B_\infty) = 0. \tag{1-3}$$

Kleinbock was supported by NSF grants DMS-1600814 and DMS-1900560. Yu acknowledges that this project received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 754475).
MSC2010: primary 11J04, 37A17; secondary 11H60, 37D40.
Keywords: Siegel transform, dynamical Borel–Cantelli lemma.

On the other hand, following the terminology of [Chernov and Kleinbock 2001] we say the family of shrinking targets $\{B_s\}_{s>0}$ is *Borel–Cantelli* (*BC*) for the flow $\{a_s\}_{s>0}$ if $\mu(B_\infty) = 1$. Thus a necessary condition for $\{B_s\}_{s>0}$ to be BC for $\{a_s\}_{s>0}$ is that the sequence of its thickenings has divergent sum of measures, and we say $\{B_s\}_{s>0}$ satisfies *a dynamical Borel–Cantelli lemma for* $\{a_s\}_{s>0}$ if this is also a sufficient condition.

The shrinking targets problem for continuous time flow in the context of homogeneous spaces was first studied in [Sullivan 1982], where he established a logarithm law for the fastest rate of geodesic cusp excursions in finite-volume hyperbolic manifolds. Later using the exponential mixing rate and a smooth approximation argument, the first author and Margulis [Kleinbock and Margulis 1999] proved that the family of cusp neighborhoods $\{\Phi^{-1}(r(s), \infty)\}_{s>0}$ with divergent sum of measures is BC for any diagonalizable flow on $(G/\Gamma, \mu)$, where G is a connected semisimple Lie group without compact factors, $\Gamma < G$ is an irreducible lattice, and μ is the probability measure on $X = G/\Gamma$ coming from a Haar measure on G. Here Φ is a distance-like function on X [loc. cit., Definition 1.6] and $r(\cdot)$ is a quasi-increasing function [loc. cit., Section 2.4]. Later Maucourant [2006] obtained a similar dynamical Borel–Cantelli lemma for geodesic flows making excursions into shrinking hyperbolic balls (with a fixed center) on a finite-volume hyperbolic manifold. See [Athreya 2009] for a survey on shrinking targets problems in dynamical systems.

One main reason that such dynamical Borel–Cantelli lemmas have gained much attention is due to their connections to metric number theory, which were first explored in [Sullivan 1982]. Such connections were made more apparent later in [Kleinbock and Margulis 1999]. Let m, l be two positive integers and let $M_{m,l}(\mathbb{R})$ be the space of m by l real matrices. Given $\psi : [t_0, \infty) \to (0, \infty)$ a continuous nonincreasing function, let us define $W(\psi) \subset M_{m,l}(\mathbb{R})$, the set of ψ-approximable $m \times l$ real matrices such that $A \in W(\psi)$ if and only if there are infinitely many $\boldsymbol{q} \in \mathbb{Z}^l$ satisfying

$$\|A\boldsymbol{q} - \boldsymbol{p}\|^m < \psi(\|\boldsymbol{q}\|^l) \quad \text{for some } \boldsymbol{p} \in \mathbb{Z}^m,$$

where $\|\cdot\|$ is the supremum norm on respective Euclidean spaces. The classical Khinchin–Groshev theorem gives an exact criterion on when $W(\psi)$ has full or zero Lebesgue measure.

Theorem KG (Khinchin–Groshev). *Given a continuous nonincreasing* ψ, *the set* $W(\psi)$ *has full* (*resp.* *zero*) *Lebesgue measure if and only if the series* $\sum_k \psi(k)$ *diverges* (*resp. converges*).

See [Schmidt 1980] for more details. On the other hand, let $X = \mathrm{SL}_{m+l}(\mathbb{R})/\mathrm{SL}_{m+l}(\mathbb{Z})$ be the space of unimodular lattices in $\mathbb{R}^{m+l}$ and let $\Delta : X \to [0, \infty)$ be the function on X given by

$$\Delta(\Lambda) := \sup_{\boldsymbol{v} \in \Lambda \smallsetminus \{\boldsymbol{0}\}} \log\left(\frac{1}{\|\boldsymbol{v}\|}\right). \tag{1-4}$$

Note that $\Delta(\Lambda) \geq 0$ for any $\Lambda \in X$ due to Minkowski's convex body theorem, and for all $r \geq 0$ the sets

$$K_r := \Delta^{-1}([0, r]) \tag{1-5}$$

(of lattices such that all its nonzero vectors have supremum norm at least e^{-r}) are compact due to Mahler's compactness criterion; see, e.g., [Cassels 1997]. Following ideas of [Dani 1985], it was shown in [Kleinbock and Margulis 1999] that there exists a unique function $r = r_\psi : [s_0, \infty) \to \mathbb{R}$ depending on ψ (this was referred to as the Dani correspondence) such that $A \in M_{m,l}(\mathbb{R})$ is ψ-approximable if and only if the events $a_s\Lambda_A \in \Delta^{-1}(r(s), \infty)$ happen for an unbounded set of $s > s_0$, where

$$a_s = \operatorname{diag}(e^{s/m}, \dots, e^{s/m}, e^{-s/l}, \dots, e^{-s/l}),$$

with m copies of $e^{s/m}$ and l copies of $e^{-s/l}$, and

$$\Lambda_A = \begin{pmatrix} I_m & A \\ 0 & I_l \end{pmatrix} \mathbb{Z}^{m+l} \in X.$$

This way the first author and Margulis showed Theorem KG to be equivalent to a dynamical Borel–Cantelli lemma for the a_s-orbits making excursions into the cusp neighborhoods $\Delta^{-1}(r(s), \infty)_{s>s_0}$, and used this to give an alternative dynamical proof of Theorem KG based on mixing properties of the a_s-action on X; see [Kleinbock and Margulis 1999].

More recently, for a given ψ as above, the first author and Wadleigh [Kleinbock and Wadleigh 2018] studied the finer problem of improvements to Dirichlet's theorem. See [Davenport and Schmidt 1970a; 1970b] for the history of the problem of improving Dirichlet's theorem. Following the definition in [Kleinbock and Wadleigh 2018] an m by l real matrix A is called *ψ-Dirichlet* if the system of inequalities

$$\|A\boldsymbol{q} - \boldsymbol{p}\|^m < \psi(t) \quad \text{and} \quad \|\boldsymbol{q}\|^l < t$$

has solutions in $(\boldsymbol{p}, \boldsymbol{q}) \in \mathbb{Z}^m \times (\mathbb{Z}^l \smallsetminus \{\boldsymbol{0}\})$ for all sufficiently large t. Following the general scheme developed in [Kleinbock and Margulis 1999] they gave a dynamical interpretation of ψ-Dirichlet matrices. Namely, they showed that $A \in M_{m,l}(\mathbb{R})$ is not ψ-Dirichlet if and only if the events

$$a_s\Lambda_A \in K_{r(s)}$$

happen for an unbounded set of $s > s_0$, where a_s, Λ_A and $r = r_\psi$ are all as above. Hence in this case the family of shrinking targets is given by $\{K_{r(s)}\}_{s>s_0}$, and one is naturally interested in whether this family of shrinking targets is BC for the flow $\{a_s\}_{s>0}$.

However this dynamical interpretation is not helpful when it comes to determining necessary and sufficient conditions on ψ guaranteeing that almost every (almost no) A is ψ-Dirichlet. One of the main difficulties is that the shrinking targets $K_{r(s)}$ are far away from being $\mathrm{SO}_{m+l}(\mathbb{R})$-invariant, and thus applying the mixing properties of the a_s-action will involve certain Sobolev norms which are hard to control. Still, using a different method based on continued fractions the aforementioned conditions were found in [Kleinbock and Wadleigh 2018] for the case $m = l = 1$. Namely, the following was proved:

Theorem KW (Kleinbock–Wadleigh). *Let $\psi : [t_0, \infty) \to (0, \infty)$ be a continuous, nonincreasing function satisfying*

$$\text{the function } t \mapsto t\psi(t) \text{ is nondecreasing} \tag{1-6}$$

and

$$t\psi(t) < 1 \quad \text{for all } t \ge t_0. \tag{1-7}$$

Then if the series

$$\sum_n \frac{-(1 - n\psi(n))\log(1 - n\psi(n))}{n} \tag{1-8}$$

diverges (resp. converges), then Lebesgue-a.e. $x \in \mathbb{R}$ is not (resp. is) ψ-Dirichlet.

In this paper we use the above theorem to derive a dynamical Borel–Cantelli lemma for the diagonal flow $a_s := \operatorname{diag}(e^s, e^{-s})$ on $X := \mathrm{SL}_2(\mathbb{R})/\mathrm{SL}_2(\mathbb{Z})$. Let μ be the probability Haar measure on X, consider the function Δ on X as in (1-4), and define the sets K_r as in (1-5).

We now state our dynamical Borel–Cantelli lemma.

Theorem 1.1. *Let $r : [s_0, \infty) \to (0, \infty)$ be a continuous and nonincreasing function. Let $B_s = K_{r(s)}$ and let $B_\infty = \limsup_{t\to\infty} a_{-s} B_s$. Then we have*

$$\sum_n r(n) \log\left(\frac{1}{r(n)}\right) < \infty \quad \Longrightarrow \quad \mu(B_\infty) = 0.$$

If in addition we assume that the function $s \mapsto s + r(s)$ is nondecreasing, then we have

$$\sum_n r(n) \log\left(\frac{1}{r(n)}\right) = \infty \quad \Longrightarrow \quad \mu(B_\infty) = 1.$$

Comparing the statement of the above theorem with (1-3), one can guess that it can be approached by studying the thickenings

$$\widetilde{B}_n = \bigcup_{0\le s<1} a_{-s} B_{n+s} = \bigcup_{0\le s<1} a_{-s} K_{r(n+s)} \tag{1-9}$$

as in (1-1). We do it in several steps. In the beginning of Section 3 we prove an asymptotic measure formula for the sets K_r where r is small:

Theorem 1.2. *For any $0 < r < (\log 2)/2$ we have*

$$\mu(K_r) = \frac{4r^2 \log(1/r)}{\zeta(2)} + O(r^2),$$

where $\zeta(2) = \pi^2/6$ is the value of the Riemann zeta function at 2.

Here and hereafter for two positive quantities A and B, we will use the notation $A \ll B$ or $A = O(B)$ to mean that there is a constant $c > 0$ such that $A \le cB$, and we will use subscripts to indicate the dependence of the constant on parameters. We will write $A \asymp B$ for $A \ll B \ll A$.

The next step is to use Theorem 1.2 to estimate the measure of the thickening of K_r along the flow $\{a_{-s}\}_{0\le s<1}$ by bounding it from above and below by a finite union of a_s-translates of K_r. This is also done in Section 3 and yields the following result:

Theorem 1.3. *For any $0 < r < \log 1.01$ we have*

$$\mu\left(\bigcup_{0\le s<1} a_{-s} K_r\right) \asymp r \log\left(\frac{1}{r}\right).$$

The above asymptotic equality shows that the series appearing in Theorem 1.1 converges/diverges if and only if so does the series $\sum_n \mu(\widetilde{B}_n)$, where $\widetilde{B}_n$ is as in (1-9):

Corollary 1.4. *Let $r : [s_0, \infty) \to (0, \infty)$ be a nonincreasing function, and let $\widetilde{B}_n$ be as in (1-9). Then we have*

$$\sum_n \mu(\widetilde{B}_n) = \infty \quad \Longleftrightarrow \quad \sum_n r(n) \log\left(\frac{1}{r(n)}\right) = \infty.$$

Therefore, in view of (1-2) and (1-3), the convergence part of Theorem 1.1 is immediate from the Borel–Cantelli lemma. The divergence part however is trickier. Instead of using a dynamical approach as in [Kleinbock and Margulis 1999], our proof in Section 4 is non-dynamical and relies on Theorem KW and the Dani correspondence.

It remains to comment on our proof of Theorem 1.2. Instead of trying to describe the sets K_r explicitly in terms of coordinates and compute their measures directly, we adapt an indirect approach which relies on an explicit second moment formula of the Siegel transform of certain indicator functions. Recall that if f is a function on $\mathbb{R}^2$, its *primitive Siegel transform* is the function on X given by

$$\hat{f}(\Lambda) := \sum_{\boldsymbol{v}\in\Lambda_{\mathrm{pr}}} f(\boldsymbol{v}),$$

where Λ_{pr} is the set of primitive vectors of Λ. Clearly $\hat{f}(\Lambda) = \#(\Lambda_{\mathrm{pr}} \cap \mathcal{S})$ when f is the indicator function of a subset $\mathcal{S}$ of $\mathbb{R}^2$.

Let us briefly describe the history of the problem. The Siegel transform was originally defined by Siegel [1945] as the sum over all nonzero lattice points for unimodular lattices of any rank. In the same paper Siegel proved a mean value theorem for the Siegel transform, which in the primitive set-up amounts to

$$\int_X \hat{f}(\Lambda)\, d\mu(\Lambda) = \frac{1}{\zeta(2)} \int_{\mathbb{R}^2} f(\boldsymbol{x})\, d\boldsymbol{x} \tag{1-10}$$

for any bounded compactly supported f on $\mathbb{R}^2$. Since then there has been much work extending his result to higher moments. For example, Rogers [1955] proved a series of higher moment formulas, which in particular includes a second moment formula for the Siegel transform defined on the space of unimodular lattices of rank greater than 2. However, his result did not give a second moment formula on X as in our setting. For this setting, Schmidt [1960] proved an upper bound for the second moment of the primitive Siegel transform of indicator functions on $\mathbb{R}^2$. His bound was later logarithmically improved by Randol [1970] for discs centered at the origin and by Athreya and Margulis [2009] for general indicator functions building on Randol's bound. Athreya and Konstantoulas [2016] obtained similar bounds on the space of general symplectic lattices for a certain family of indicator functions. Continuing [Athreya and Konstantoulas 2016], Kelmer and the second author [Kelmer and Yu 2019] proved a second moment formula on the space of symplectic lattices $Y_n := \mathrm{Sp}(2n, \mathbb{R})/\mathrm{Sp}(2n, \mathbb{Z})$. In particular, when $n = 1$ we have $Y_1 = X$ and their formula also applies to our setting.[1] However, for our applications all these formulas are not explicit enough.

We now state an explicit second moment formula which we use to derive Theorem 1.2.

Theorem 1.5. *For any $r \geq 0$ let $\mathcal{S}_r$ be the open square with vertices given by $(\pm e^{-r}, \pm e^{-r})$, and let f_r be the indicator function of $\mathcal{S}_r$. Then we have*

$$\|\hat{f}_r\|_2^2 = \frac{8}{\zeta(2)}\left(e^{-2r} + \int_{\mathcal{D}_r}\left(\frac{e^{-r}}{x_1} + \frac{e^{-r}}{x_2} - \frac{1}{x_1x_2}\right) dx_1\, dx_2\right), \tag{1-11}$$

where

$$\mathcal{D}_r := \{\boldsymbol{x} = (x_1, x_2) \in \mathcal{S}_r \mid x_1 > 0,\ x_2 > 0,\ x_1 + x_2 > e^r\},$$

and $\|\cdot\|_2$ stands for the L^2-norm with respect to μ.

[1] See also [Fairchild 2019] for moment formulas of the Siegel–Veech transform recently obtained by Fairchild.

Remark 1.6. When $r \geq (\log 2)/2$ the region $\mathcal{D}_r$ is empty, and (1-11) simply reads as

$$\|\hat{f}_r\|_2^2 = \frac{8e^{-2r}}{\zeta(2)}.$$

We note that the latter equality in fact already follows from Siegel's mean value theorem, since in this case for any unimodular lattice there can only be at most one pair of primitive lattice points allowed in $\mathcal{S}_r$, which implies that $\hat{f}_r/2$ is an indicator function on X. When $0 \leq r < (\log 2)/2$, the region $\mathcal{D}_r$ is not empty, and it is not hard to compute the integral in (1-11) explicitly; see (3-5) below. In particular, plugging $r = 0$ into (1-11) we have $\|\hat{f}_0\|_2^2 = (12/\pi)^2 - 8 \approx 6.59$.

In Section 2 we prove a much more general second moment formula, see Theorem 2.1, with an arbitrary bounded measurable subset $\mathcal{S}$ of $\mathbb{R}^2$ in place of $\mathcal{S}_r$. Theorem 1.5 is derived from Theorem 2.1 by taking $\mathcal{S} = \mathcal{S}_r$.

2. The second moment formula

In this section, we prove Theorem 1.5 by establishing the following second moment formula for quite general subsets of $\mathbb{R}^2$.

Theorem 2.1. *Let $\mathcal{S}$ be a measurable bounded subset of $\mathbb{R}^2$, and let f be the indicator function of $\mathcal{S}$. Let $\widetilde{\mathcal{S}} = \{\boldsymbol{x} \in \mathbb{R}^2 \mid -\boldsymbol{x} \in \mathcal{S}\}$. Then we have*

$$\|\hat{f}\|_2^2 = \frac{1}{\zeta(2)}\left(\operatorname{area}(\mathcal{S}) + \operatorname{area}(\mathcal{S} \cap \widetilde{\mathcal{S}}) + \sum_{n \neq 0} \frac{\varphi(|n|)}{|n|} \int_{\mathcal{S}} |\mathcal{I}_{\boldsymbol{x}}^n| \, d\boldsymbol{x}\right),$$

where φ is the Euler's totient function, $\mathcal{I}_{\boldsymbol{x}}^n \subset \mathbb{R}$ is defined by

$$\mathcal{I}_{\boldsymbol{x}}^n := \left\{ t \in \mathbb{R} \;\middle|\; n\left(\frac{-x_2}{x_1^2 + x_2^2}, \frac{x_1}{x_1^2 + x_2^2}\right) + t(x_1, x_2) \in \mathcal{S} \right\},$$

and $|\mathcal{I}_{\boldsymbol{x}}^n|$ is the length of $\mathcal{I}_{\boldsymbol{x}}^n$ with respect to the Lebesgue measure on $\mathbb{R}$.

Before giving the proof let us make a few remarks about Theorem 2.1. First we note that for any bounded $\mathcal{S}$ there exists a sufficiently large $T > 0$ depending on $\mathcal{S}$ such that for any $|n| > T$ the set $\mathcal{I}_{\boldsymbol{x}}^n$ is empty for all $\boldsymbol{x} \in \mathcal{S}$. Thus the series on the right-hand side of (2-1) is a finite sum. Next we note that if we further assume $\mathcal{S}$ is symmetric with respect to the origin, then by symmetry we have $\mathcal{S} \cap \widetilde{\mathcal{S}} = \mathcal{S}$ and $|\mathcal{I}_{\boldsymbol{x}}^n| = |\mathcal{I}_{\boldsymbol{x}}^{-n}|$ for any $n \neq 0$. In particular, for such $\mathcal{S}$ we have the slightly simpler formula

$$\|\hat{f}\|_2^2 = \frac{2}{\zeta(2)}\left(\operatorname{area}(\mathcal{S}) + \sum_{n=1}^{\infty} \frac{\varphi(n)}{n} \int_{\mathcal{S}} |\mathcal{I}_{\boldsymbol{x}}^n| \, d\boldsymbol{x}\right). \tag{2-1}$$

Finally we note that for any $\Lambda \in X$ and f as in Theorem 2.1 we have

$$(\hat{f}(\Lambda))^2 = \hat{f}(\Lambda) + \hat{\chi}_{\mathcal{S} \cap \widetilde{\mathcal{S}}}(\Lambda) + \sum_{\substack{\boldsymbol{v}_1, \boldsymbol{v}_2 \in \Lambda_{\mathrm{pr}} \\ \text{lin. ind.}}} f(\boldsymbol{v}_1) f(\boldsymbol{v}_2).$$

Thus Theorem 2.1 together with (1-10) implies

$$\int_X \sum_{\substack{\boldsymbol{v}_1, \boldsymbol{v}_2 \in \Lambda_{\mathrm{pr}} \\ \text{lin. ind.}}} f(\boldsymbol{v}_1) f(\boldsymbol{v}_2) \, d\mu(\Lambda) = \frac{1}{\zeta(2)} \sum_{n \neq 0} \frac{\varphi(|n|)}{|n|} \int_{\mathcal{S}} |\mathcal{I}_{\boldsymbol{x}}^n| \, d\boldsymbol{x}. \tag{2-2}$$

It is worth pointing out that the above formula can be compared to its higher-dimensional analogue: when f is an indicator function of a bounded measurable subset $\mathcal{S}$ of $\mathbb{R}^k$ with $k \geq 3$, $X = \mathrm{SL}_k(\mathbb{R})/\mathrm{SL}_k(\mathbb{Z})$, and μ is the Haar probability measure on X, according to Rogers' second moment formula [1955] the left-hand side of (2-2) equals $(\mathrm{vol}(\mathcal{S})/\zeta(k))^2$. However, as we can see here the $k = 2$ case is much more complicated, with the answer depending on both the shape and the position of $\mathcal{S}$.

Coordinates and measures. We fix coordinates on $G = \mathrm{SL}_2(\mathbb{R})$ via the Iwasawa decomposition $G = KAN$ with

$$K = \{k_\theta \mid 0 \leq \theta < 2\pi\}, \quad A = \{a_s \mid s \in \mathbb{R}\}, \quad \text{and} \quad N = \{u_t \mid t \in \mathbb{R}\},$$

where

$$k_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad a_s = \begin{pmatrix} e^s & 0 \\ 0 & e^{-s} \end{pmatrix} \quad \text{and} \quad u_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

Explicitly, under coordinates $g = k_\theta a_s u_t$, μ is given by

$$d\mu(g) = \frac{1}{\zeta(2)} e^{2s}\, d\theta\, ds\, dt. \tag{2-3}$$

There is a natural identification between the homogeneous space G/N and $\mathbb{R}^2 \smallsetminus \{\mathbf{0}\}$ induced by the map $G \to \mathbb{R}^2 \smallsetminus \{\mathbf{0}\}$ sending $g = k_\theta a_s u_t \in G$ to

$$\boldsymbol{x}(g) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = g \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} e^s \cos\theta \\ e^s \sin\theta \end{pmatrix}, \tag{2-4}$$

the left column of g. The Lebesgue measure, $d\boldsymbol{x}$, on $\mathbb{R}^2 \smallsetminus \{\mathbf{0}\} \cong G/N$ can be expressed via the polar coordinates (s, θ) as

$$d\boldsymbol{x}(k_\theta a_s) = e^{2s}\, d\theta\, ds. \tag{2-5}$$

The second moment formula. In this subsection we prove Theorem 2.1, and with some more analysis we prove Theorem 1.5. As the first step of our computation we recall the following preliminary identity which relies on a standard unfolding argument. We note that one can find it in [Lang 1975, Chapter VIII, Section 1], and we include a short proof here to make the paper self-contained. See also [Kelmer and Yu 2019, Proposition 2.3] for a generalization to the space of symplectic lattices.

Lemma 2.2. *For any bounded and compactly supported function f on $\mathbb{R}^2$ and for any bounded $F \in L^2(X, \mu)$ we have*

$$\langle \hat{f}, F \rangle = \frac{1}{\zeta(2)} \int_{-\infty}^{\infty} \int_0^{2\pi} f(\boldsymbol{x}(k_\theta a_s)) \overline{\mathcal{P}_F(\boldsymbol{x}(k_\theta a_s))} e^{2s}\, d\theta\, ds,$$

where $\mathcal{P}_F$ is defined by

$$\mathcal{P}_F(\boldsymbol{x}(k_\theta a_s)) = \int_0^1 F(k_\theta a_s u_t \mathbb{Z}^2)\, dt$$

with k_θ, a_s and u_t as above, and $\langle \cdot, \cdot \rangle$ is the inner product on $L^2(X, \mu)$.

Proof. Let $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ and let $\Gamma_\infty = \Gamma \cap N$. Recall that there is an identification between Γ/Γ_∞ and $\mathbb{Z}^2_{\mathrm{pr}}$ sending $\gamma\Gamma_\infty$ to $\gamma\binom{1}{0}$. Using this identification, for any $\Lambda = g\mathbb{Z}^2$ with $g \in \mathrm{SL}_2(\mathbb{R})$ we can write

$$\hat{f}(\Lambda) = \sum_{\boldsymbol{v}\in\Lambda_{\mathrm{pr}}} f(\boldsymbol{v}) = \sum_{\boldsymbol{w}\in\mathbb{Z}^2_{\mathrm{pr}}} f(g\boldsymbol{w}) = \sum_{\gamma\in\Gamma/\Gamma_\infty} \tilde{f}(g\gamma), \tag{2-6}$$

where $\tilde{f}(g) := f\big(g\binom{1}{0}\big)$. We note that $\tilde{f}$ is a right N-invariant function on G. Let $\mathcal{F}_\Gamma$ be a fundamental domain for $X = G/\Gamma$, and let $\mathcal{F}_\infty$ be a fundamental domain for G/Γ_∞. Note that using the Iwasawa decomposition $G = KAN$ we can choose

$$\mathcal{F}_\infty = \{k_\theta a_s u_t \mid 0 < \theta < 2\pi,\ s \in \mathbb{R},\ 0 < t < 1\}. \tag{2-7}$$

Moreover, fix a set of coset representatives $\Sigma_\infty \subset \Gamma$ for Γ/Γ_∞, and note that $\bigcup_{\gamma\in\Sigma_\infty} \mathcal{F}_\Gamma\gamma$ is a disjoint union and forms a fundamental domain for G/Γ_∞. Now for any bounded $F \in L^2(X, \mu)$, using (2-3), (2-6), (2-7) and the facts that F is right Γ-invariant and $\tilde{f}$ is right N-invariant, we have

$$\begin{aligned}
\langle \hat{f}, F\rangle &:= \int_{\mathcal{F}_\Gamma} \hat{f}(g\mathbb{Z}^2)\overline{F(g\mathbb{Z}^2)}\, d\mu(g) = \sum_{\gamma\in\Gamma/\Gamma_\infty} \int_{\mathcal{F}_\Gamma} \tilde{f}(g\gamma)\overline{F(g\mathbb{Z}^2)}\, d\mu(g) \\
&= \sum_{\gamma\in\Sigma_\infty} \int_{\mathcal{F}_\Gamma\gamma} \tilde{f}(g)\overline{F(g\mathbb{Z}^2)}\, d\mu(g) = \int_{\bigsqcup_{\gamma\in\Sigma_\infty}\mathcal{F}_\Gamma\gamma} \tilde{f}(g)\overline{F(g\mathbb{Z}^2)}\, d\mu(g) \\
&= \int_{\mathcal{F}_\infty} \tilde{f}(g)\overline{F(g\mathbb{Z}^2)}\, d\mu(g) = \frac{1}{\zeta(2)} \int_{-\infty}^{\infty}\int_0^{2\pi}\int_0^1 \tilde{f}(k_\theta a_s u_t)\overline{F(k_\theta a_s u_t\mathbb{Z}^2)} e^{2s}\, dt\, d\theta\, ds \\
&= \frac{1}{\zeta(2)} \int_{-\infty}^{\infty}\int_0^{2\pi} f(\boldsymbol{x}(k_\theta a_s)) \int_0^1 \overline{F(k_\theta a_s u_t\mathbb{Z}^2)}\, dt\, e^{2s}\, d\theta\, ds.
\end{aligned}$$

Finally, we note that the above equalities can be justified since F is bounded and the defining series for $\hat{f}$ is absolutely convergent; see [Veech 1998, Lemma 16.10]. □

With this preliminary identity, we can now give:

Proof of Theorem 2.1. Using the relation (2-5) and Lemma 2.2 we have

$$\|\hat{f}\|_2^2 = \frac{1}{\zeta(2)} \int_{\mathbb{R}^2} f(\boldsymbol{x}(k_\theta a_s))\mathcal{P}_{\hat{f}}(\boldsymbol{x}(k_\theta a_s))\, d\boldsymbol{x} = \frac{1}{\zeta(2)} \int_{\mathcal{S}} \mathcal{P}_{\hat{f}}(\boldsymbol{x}(k_\theta a_s))\, d\boldsymbol{x}, \tag{2-8}$$

where

$$\mathcal{P}_{\hat{f}}(\boldsymbol{x}(k_\theta a_s)) = \int_0^1 \hat{f}(k_\theta a_s u_t\mathbb{Z}^2)\, dt,$$

with k_θ, a_s and u_t as before. First, by the definition of the primitive Siegel transform we have

$$\hat{f}(k_\theta a_s u_t\mathbb{Z}^2) = \#\left\{(m, n) \in \mathbb{Z}^2_{\mathrm{pr}} \,\middle|\, k_\theta a_s u_t \begin{pmatrix} m \\ n \end{pmatrix} \in \mathcal{S}\right\}.$$

Thus for $\boldsymbol{x}(k_\theta a_s) \in \mathcal{S}$ and $0 \le t < 1$ we have

$$\hat{f}(k_\theta a_s u_t\mathbb{Z}^2) = \sum_{(m,n)\in\mathbb{Z}^2_{\mathrm{pr}}} \chi_{I^{(m,n)}_{\boldsymbol{x}(k_\theta a_s)}}(t),$$

where

$$I^{(m,n)}_{\boldsymbol{x}(k_\theta a_s)} := \left\{ 0 \le t < 1 \;\middle|\; k_\theta a_s u_t \begin{pmatrix} m \\ n \end{pmatrix} \in \mathcal{S} \right\},$$

implying

$$\mathcal{P}_{\hat{f}}(\boldsymbol{x}(k_\theta a_s)) = \sum_{(m,n)\in\mathbb{Z}^2_{\mathrm{pr}}} |I^{(m,n)}_{\boldsymbol{x}(k_\theta a_s)}| = |I^{(1,0)}_{\boldsymbol{x}(k_\theta a_s)}| + |I^{(-1,0)}_{\boldsymbol{x}(k_\theta a_s)}| + \sum_{\substack{(m,n)\in\mathbb{Z}^2_{\mathrm{pr}} \\ n\neq 0}} |I^{(m,n)}_{\boldsymbol{x}(k_\theta a_s)}|.$$

Next, by direct computation we have for $\boldsymbol{x}(k_\theta a_s) = (x_1, x_2) = (e^s \cos\theta, e^s \sin\theta) \in \mathcal{S}$

$$k_\theta a_s u_t \begin{pmatrix} m \\ n \end{pmatrix} = n \begin{pmatrix} -e^{-s}\sin\theta \\ e^{-s}\cos\theta \end{pmatrix} + (m+nt) \begin{pmatrix} e^s\cos\theta \\ e^s \sin\theta \end{pmatrix} = n \begin{pmatrix} -x_2/(x_1^2+x_2^2) \\ x_1/(x_1^2+x_2^2) \end{pmatrix} + (m+nt) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \tag{2-9}$$

When $(m, n) = (1, 0)$ we have for $\boldsymbol{x}(k_\theta a_s) \in \mathcal{S}$

$$k_\theta a_s u_t \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

is contained in $\mathcal{S}$ for any $0 \le t < 1$. Thus $I^{(1,0)}_{\boldsymbol{x}(k_\theta a_s)} = [0, 1)$ and $|I^{(1,0)}_{\boldsymbol{x}(k_\theta a_s)}| = 1$ for any $\boldsymbol{x}(k_\theta a_s) \in \mathcal{S}$. Similarly, when $(m, n) = (-1, 0)$ we have for $\boldsymbol{x}(k_\theta a_s) \in \mathcal{S}$

$$k_\theta a_s u_t \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix}$$

is contained in $\mathcal{S}$ if and only if $\boldsymbol{x} \in \mathcal{S} \cap \widetilde{\mathcal{S}}$ with $\widetilde{\mathcal{S}}$ as in the theorem, implying $I^{(-1,0)}_{\boldsymbol{x}(k_\theta a_s)} = [0, 1)$ whenever $\boldsymbol{x} \in \mathcal{S} \cap \widetilde{\mathcal{S}}$.

When $n \neq 0$ by (2-9) we have for any integer m coprime to n

$$\begin{aligned} |I^{(m,n)}_{\boldsymbol{x}}| &= \left| \left\{ 0 \le t < 1 \;\middle|\; n\left(\frac{-x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2} \right) + (m+nt)(x_1, x_2) \in \mathcal{S} \right\} \right| \\ &= \left| \left\{ \frac{m}{n} \le t < 1 + \frac{m}{n} \;\middle|\; n\left(\frac{-x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2} \right) + nt(x_1, x_2) \in \mathcal{S} \right\} \right|. \end{aligned}$$

We note that as m runs through all the integers in each congruence class in $(\mathbb{Z}/|n|\mathbb{Z})^\times$, the intervals $[m/n, 1 + m/n)$ cover $\mathbb{R}$ exactly once. Thus for $n \neq 0$

$$\sum_{\substack{m\in\mathbb{Z} \\ (m,n)=1}} |I^{(m,n)}_{\boldsymbol{x}(k_\theta a_s)}| = \varphi(|n|) \left| \left\{ t \in \mathbb{R} \;\middle|\; n\left(\frac{-x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2} \right) + nt(x_1, x_2) \in \mathcal{S} \right\} \right| = \frac{\varphi(|n|)}{|n|} |\mathcal{I}^n_{\boldsymbol{x}}|,$$

where φ is the Euler's totient function and $\mathcal{I}^n_{\boldsymbol{x}}$ is as in Theorem 2.1. We thus have for $\boldsymbol{x} \in \mathcal{S}$

$$\mathcal{P}_{\hat{f}}(\boldsymbol{x}) = 1 + \chi_{\mathcal{S}\cap\widetilde{\mathcal{S}}}(\boldsymbol{x}) + \sum_{n\neq 0} \frac{\varphi(|n|)}{|n|} |\mathcal{I}^n_{\boldsymbol{x}}|.$$

We conclude the proof by plugging the above equation into (2-8). □

We can now give:

Proof of Theorem 1.5. To simplify notation for any $\boldsymbol{x} \in \mathbb{R}^2$, $t \in \mathbb{R}$, and $n \geq 1$ let

$$\boldsymbol{v}(\boldsymbol{x}, t, n) := n\left(\frac{-x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2}\right) + t(x_1, x_2).$$

First we note that

$$\|\boldsymbol{v}(\boldsymbol{x}, t, n)\|_2^2 = \frac{n^2}{x_1^2+x_2^2} + t^2(x_1^2+x_2^2) \geq \frac{n^2}{x_1^2+x_2^2},$$

where $\|\cdot\|_2$ stands for the standard Euclidean norm on $\mathbb{R}^2$. Thus for $\boldsymbol{x} \in \mathcal{S}_r$ and $n \geq 2$ we have

$$\|\boldsymbol{v}(\boldsymbol{x}, t, n)\| \geq \frac{\sqrt{2}}{2}\|\boldsymbol{v}(\boldsymbol{x}, t, n)\|_2 \geq \frac{\sqrt{2}}{\|\boldsymbol{x}\|_2} > e^r \geq e^{-r},$$

implying that $\mathcal{I}_{\boldsymbol{x}}^n$ is empty for any $\boldsymbol{x} \in \mathcal{S}_r$ and any $n \geq 2$. Here $\|\cdot\|$ stands for the supremum norm on $\mathbb{R}^2$, and for the third inequality we used the fact that $\|\boldsymbol{x}\|_2 < \sqrt{2}e^{-r}$, which follows from $\boldsymbol{x}$ being an element of $\mathcal{S}_r$. Since $\mathcal{S}_r$ is symmetric with respect to the origin, applying (2-1) to $f = f_r$ we get

$$\|\hat{f}_r\|_2^2 = \frac{8e^{-2r}}{\zeta(2)} + \frac{2}{\zeta(2)}\int_{\mathcal{S}_r} |\mathcal{I}_{\boldsymbol{x}}^1|\, d\boldsymbol{x} = \frac{8e^{-2r}}{\zeta(2)} + \frac{8}{\zeta(2)}\int_{\mathcal{S}_r^+} |\mathcal{I}_{\boldsymbol{x}}^1|\, d\boldsymbol{x}, \tag{2-10}$$

where $\mathcal{S}_r^+$ is the intersection of $\mathcal{S}_r$ with the first quadrant, and for the second equality we used the fact that $|\mathcal{I}_{(x_1,x_2)}^1| = |\mathcal{I}_{(\pm x_1,\pm x_2)}^1|$ which follows from the invariance of $\mathcal{S}_r$ under reflections around the coordinate axes. We note that for $\boldsymbol{x} \in \mathcal{S}_r^+$

$$\left(\frac{-x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2}\right) + t(x_1, x_2) \in \mathcal{S}_r$$

if and only if

$$-\frac{e^{-r}}{x_1} + \frac{x_2}{x_1(x_1^2+x_2^2)} < t < \frac{e^{-r}}{x_1} + \frac{x_2}{x_1(x_1^2+x_2^2)}$$

and

$$-\frac{e^{-r}}{x_2} - \frac{x_1}{x_2(x_1^2+x_2^2)} < t < \frac{e^{-r}}{x_2} - \frac{x_1}{x_2(x_1^2+x_2^2)}.$$

By direct computation if $r \geq (\log 2)/2$ then there is no $t \in \mathbb{R}$ satisfying above inequalities. Thus $\mathcal{I}_{\boldsymbol{x}}^1$ is empty, and the integral in the right-hand side of (2-10) is zero. If $0 \leq r < (\log 2)/2$, we define for any $\boldsymbol{x} \in \mathcal{S}_r^+$

$$L(\boldsymbol{x}) := \max\left\{-\frac{e^{-r}}{x_1} + \frac{x_2}{x_1(x_1^2+x_2^2)}, -\frac{e^{-r}}{x_2} - \frac{x_1}{x_2(x_1^2+x_2^2)}\right\},$$

$$U(\boldsymbol{x}) := \min\left\{\frac{e^{-r}}{x_1} + \frac{x_2}{x_1(x_1^2+x_2^2)}, \frac{e^{-r}}{x_2} - \frac{x_1}{x_2(x_1^2+x_2^2)}\right\}.$$

It is not hard to verify that as long as $0 \leq r < (\log 2)/2$, for $\boldsymbol{x} \in \mathcal{S}_r^+$ we have

$$L(\boldsymbol{x}) = -\frac{e^{-r}}{x_1} + \frac{x_2}{x_1(x_1^2+x_2^2)} \quad \text{and} \quad U(\boldsymbol{x}) = \frac{e^{-r}}{x_2} - \frac{x_1}{x_2(x_1^2+x_2^2)}.$$

Thus $\mathcal{I}_{\boldsymbol{x}}^1$ is nonempty if and only if $L(\boldsymbol{x}) < U(\boldsymbol{x})$ and whenever it is nonempty we have

$$\mathcal{I}_{\boldsymbol{x}}^1 = \left(-\frac{e^{-r}}{x_1} + \frac{x_2}{x_1(x_1^2+x_2^2)}, \frac{e^{-r}}{x_2} - \frac{x_1}{x_2(x_1^2+x_2^2)}\right).$$

By direct computation we have $L(\boldsymbol{x}) < U(\boldsymbol{x})$ if and only if $\boldsymbol{x} \in \mathcal{D}_r = \{(x_1, x_2) \in \mathcal{S}_r^+ \mid x_1 + x_2 > e^r\}$. Hence

$$\begin{aligned}\|\hat{f}_r\|_2^2 &= \frac{8e^{-2r}}{\zeta(2)} + \frac{8}{\zeta(2)} \int_{\mathcal{D}_r} \left(\left(\frac{e^{-r}}{x_2} - \frac{x_1}{x_2(x_1^2+x_2^2)}\right) - \left(-\frac{e^{-r}}{x_1} + \frac{x_2}{x_1(x_1^2+x_2^2)}\right)\right) dx_1\, dx_2 \\ &= \frac{8e^{-2r}}{\zeta(2)} + \frac{8}{\zeta(2)} \int_{\mathcal{D}_r} \left(\frac{e^{-r}}{x_1} + \frac{e^{-r}}{x_2} - \frac{1}{x_1 x_2}\right) dx_1\, dx_2. \end{aligned}$$ □

Besides the sets $\mathcal{S}_r$, another natural candidate to test formula (2-1) is the family of indicator functions of balls. For any $R > 0$ let $\mathcal{B}_R$ be the open ball of radius R centered at the origin, and let h_R be the indicator function of $\mathcal{B}_R$. We note that [Randol 1970] established an asymptotic formula for $\|\hat{h}_R\|_2^2$ for large R, and here we prove the following formula for $\|\hat{h}_R\|_2^2$:

Corollary 2.3. *For any $R > 0$ let h_R be as above. Then we have*

$$\|\hat{h}_R\|_2^2 = \frac{12R^2}{\pi} + \frac{48}{\pi} \sum_{n=1}^{\lfloor R^2 \rfloor} \varphi(n) \left(\frac{\sqrt{R^4 - n^2}}{n} + \arcsin\left(\frac{n}{R^2}\right) - \frac{\pi}{2}\right).$$

Proof. Since $\mathcal{B}_R$ is symmetric with respect to the origin, we can apply (2-1) to $\|\hat{h}_R\|_2^2$, and use $\zeta(2) = \pi^2/6$ to get

$$\|\hat{h}_R\|_2^2 = \frac{12R^2}{\pi} + \frac{12}{\pi^2} \sum_{n=1}^{\infty} \frac{\varphi(n)}{n} \int_{\mathcal{B}_R} |\mathcal{I}_{\boldsymbol{x}}^n|\, d\boldsymbol{x},$$

where

$$\mathcal{I}_{\boldsymbol{x}}^n := \left\{ t \in \mathbb{R} \;\middle|\; \left\| n\left(\frac{-x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2}\right) + t(x_1, x_2) \right\|_2 < R \right\}.$$

Using the polar coordinates, for any $(x_2, x_2) = (r\cos\theta, r\sin\theta) \in \mathcal{B}_R$ and $n \geq Rr$ we can write

$$\left\| n\left(\frac{-x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2}\right) + t(x_1, x_2) \right\|_2^2 = \frac{n^2}{r^2} + t^2 r^2 \geq R^2,$$

implying that $\mathcal{I}_{\boldsymbol{x}}^n$ is empty whenever $n \geq Rr = R\|\boldsymbol{x}\|_2$. In particular, $\mathcal{I}_{\boldsymbol{x}}^n$ is empty for any $\boldsymbol{x} \in \mathcal{B}_R$ if $n \geq R^2$. Similarly, for any $1 \leq n \leq \lfloor R^2 \rfloor$ the set $\mathcal{I}_{\boldsymbol{x}}^n$ is empty if $\|\boldsymbol{x}\|_2 \leq n/R$, and

$$\mathcal{I}_{\boldsymbol{x}}^n = \left(-\frac{\sqrt{R^2r^2 - n^2}}{r^2}, \frac{\sqrt{R^2r^2 - n^2}}{r^2}\right)$$

if $n/R < \|\boldsymbol{x}\|_2 < R$. Hence

$$\|\hat{h}_R\|_2^2 = \frac{12R^2}{\pi} + \frac{12}{\pi^2} \sum_{n=1}^{\lfloor R^2 \rfloor} \frac{\varphi(n)}{n} \int_0^{2\pi} \int_{n/R}^{R} \frac{2\sqrt{R^2r^2 - n^2}}{r^2} r\, dr\, d\theta$$

$$= \frac{12R^2}{\pi} + \frac{48}{\pi}\sum_{n=1}^{\lfloor R^2\rfloor}\varphi(n)\int_1^{R^2/n}\sqrt{1-r^{-2}}\,dr$$

$$= \frac{12R^2}{\pi} + \frac{48}{\pi}\sum_{n=1}^{\lfloor R^2\rfloor}\varphi(n)\left(\frac{\sqrt{R^4-n^2}}{n} + \arcsin\left(\frac{n}{R^2}\right) - \frac{\pi}{2}\right),$$

where for the second equality we applied a change of variable $(R/n)r \mapsto r$, and for the last equality we used the fact that $\int \sqrt{1-r^{-2}}\,dr = \sqrt{r^2-1} + \arcsin(1/r) + C$ for $r \geq 1$. □

3. Measure estimates of the shrinking targets

In this section, using the methods developed in the previous section, we prove Theorem 1.2 and then use it to derive Theorem 1.3 and Corollary 1.4.

Proof of Theorem 1.2. For any $r > 0$, let f_r be the indicator function of $\mathcal{S}_r$ as before. For any integer $k \geq 0$, let $B_r^k \subset X$ be the set of unimodular lattices having $2k$ nonzero primitive points in $\mathcal{S}_r$. First, we note that $K_r = B_r^0$ consists of lattices with no nonzero points in $\mathcal{S}_r$. Moreover, for any $\Lambda \in X$, there are at most two linearly independent primitive points of Λ inside $\mathcal{S}_r$. We thus have for any $r > 0$

$$\sum_{k=0}^{2}\mu(B_r^k) = 1, \tag{3-1}$$

and

$$\hat{f}_r = 2\chi_{B_r^1} + 4\chi_{B_r^2}.$$

Thus we can take the first moment and apply (1-10) to get

$$\mu(B_r^1) + 2\mu(B_r^2) = \frac{1}{2}\int_X \hat{f}_r(\Lambda)\,d\mu(\Lambda) = \frac{2e^{-2r}}{\zeta(2)}. \tag{3-2}$$

Taking the second moment of $\hat{f}_r$ we get

$$4\mu(B_r^1) + 16\mu(B_r^2) = \|\hat{f}_r\|_2^2. \tag{3-3}$$

Solving (3-1), (3-2) and (3-3) and applying Theorem 1.5 to (3-3), we get

$$\mu(K_r) = \mu(B_r^0) = 1 - \frac{2e^{-2r}}{\zeta(2)} + \frac{1}{\zeta(2)}\int_{\mathcal{D}_r}\left(\frac{e^{-r}}{x_1} + \frac{e^{-r}}{x_2} - \frac{1}{x_1x_2}\right)dx_1\,dx_2.$$

By direct computation we have for $0 < r < \frac{1}{2}\log 2$

$$\int_{\mathcal{D}_r}\left(\frac{e^{-r}}{x_1}+\frac{e^{-r}}{x_2}-\frac{1}{x_1x_2}\right)dx_1\,dx_2$$

$$= 2(1-r)(2e^{-2r}-1+r)+(2-2e^{-2r}-2r)\log(1-e^{-2r})-2r^2+\int_{1-e^{-2r}}^{e^{-2r}}\frac{\log t}{1-t}\,dt$$

$$= 2(1-r)(2e^{-2r}-1+r)+(2-2e^{-2r}-2r)\log(1-e^{-2r})-2r^2+\mathrm{Li}_2(1-e^{-2r})-\mathrm{Li}_2(e^{-2r}), \tag{3-4}$$

where $\mathrm{Li}_s(z) = \sum_{k=1}^{\infty} z^k/k^s$ is the polylogarithm function. Now for the term $\log(1-e^{-2r})$, using the Taylor expansion $e^{-2r} = 1-2r+2r^2+O(r^3)$, we get

$$\log(1-e^{-2r}) = \log(2r)+\log(1-r+O(r^2)) = \log(2r)-r+O(r^2).$$

Using the series representation $\mathrm{Li}_2(z) = \sum_{k=1}^{\infty} z^k/k^2$, we get $\mathrm{Li}_2(1-e^{-2r}) = 2r+O(r^2)$. Finally for the term $\mathrm{Li}_2(e^{-2r})$ we have the expansion, see [Wood 1992, Equation (9.7)],

$$\mathrm{Li}_2(e^{-2r}) = -2r(1-\log(2r))+\zeta(2)+O(r^2).$$

Plugging these into (3-4) and using the expansion $e^{-2r} = 1-2r+2r^2+O(r^3)$, we get

$$\int_{\mathcal{D}_r}\left(\frac{e^{-r}}{x_1}+\frac{e^{-r}}{x_2}-\frac{1}{x_1x_2}\right)dx_1\,dx_2 = 2-\zeta(2)-4r-4r^2\log r+O(r^2), \tag{3-5}$$

implying

$$\mu(K_r) = 1-\frac{2e^{-2r}}{\zeta(2)}+\frac{1}{\zeta(2)}(2-\zeta(2)-4r-4r^2\log r+O(r^2)) = -\frac{4r^2\log r}{\zeta(2)}+O(r^2),$$

finishing the proof. □

To estimate the measure of the thickening, we will need the following two preliminary lemmas. We note that by the Hajós–Minkowski theorem, see [Cassels 1997, IX.1.3], we have

$$K_0 = \Delta^{-1}\{0\} = \bigcup_{x\in[0,1)} \begin{pmatrix}1 & x\\ 0 & 1\end{pmatrix}\mathbb{Z}^2 \bigcup \begin{pmatrix}1 & 0\\ x & 1\end{pmatrix}\mathbb{Z}^2.$$

A simple observation is that any $\Lambda \in K_0$ contains either the point $(1,0)$ or the point $(0,1)$. Thus intuitively one shall expect that when r is small, lattices in K_r contain points close to either $(1,0)$ or $(0,1)$. For any $r>0$, let $\mathcal{A}_r \subset \mathbb{R}^2$ be the closed rectangle with vertices $(\pm\sqrt{e^{2r}-1}, e^r)$ and $(\pm\sqrt{e^{2r}-1}, e^{-r})$ and let $\mathcal{C}_r$ be the closed rectangle with vertices $(e^r, \pm\sqrt{e^{2r}-1})$ and $(e^{-r}, \pm\sqrt{e^{2r}-1})$; see Figure 1. The following lemma asserts that when r is small, then any $\Lambda \in K_r$ contains points either in $\mathcal{A}_r$ or in $\mathcal{C}_r$ (noting that $\mathcal{A}_r$ is a small rectangle containing $(0,1)$ and $\mathcal{C}_r$ is a small rectangle containing $(1,0)$).

Lemma 3.1. *Let $\mathcal{A}_r$ and $\mathcal{C}_r$ be as above. For any $0<r<\log 1.01$ and for any $\Lambda \in K_r$, we have $\Lambda_{\mathrm{pr}} \cap (\mathcal{A}_r \cup \mathcal{C}_r) \neq \varnothing$.*

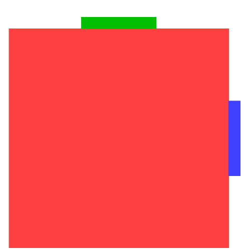

Figure 1. The square $\mathcal{S}_r$ (red), the rectangles $\mathcal{A}_r$ (green) and $\mathcal{C}_r$ (blue).

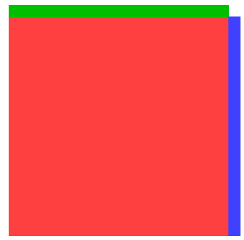

Figure 2. The square $\mathcal{S}_r$ (red), the rectangles $\mathcal{U}_r$ (green) and $\mathcal{R}_r$ (blue).

Proof. Let $\mathcal{U}_r$ be the closed rectangle with vertices $(\pm e^{-r}, e^{-r})$ and $(\pm e^{-r}, e^{r})$, and let $\mathcal{R}_r$ be the closed rectangle with vertices $(e^{-r}, \pm e^{-r})$ and $(e^{r}, \pm e^{-r})$; see Figure 2. Let

$$\widetilde{\mathcal{U}}_r := \{\boldsymbol{x} \in \mathbb{R}^2 \mid -\boldsymbol{x} \in \mathcal{U}_r\}.$$

Consider the rectangle $\mathcal{U}_r \sqcup \mathcal{S}_r \sqcup \widetilde{\mathcal{U}}_r$ and note that it has area 4. For any $\varepsilon > 0$ let $\mathcal{U}_{r,\varepsilon}$ be the open rectangle with vertices $(\pm e^{-r}, \pm(e^r + \varepsilon))$. Applying the Minkowski's convex body theorem to $\mathcal{U}_{r,\varepsilon}$ and letting ε approach zero, we see that for any $\Lambda \in X$, Λ_{pr} intersects $\mathcal{U}_r \sqcup \mathcal{S}_r \sqcup \widetilde{\mathcal{U}}_r$ nontrivially. Now let $\Lambda \in K_r$; since Λ has no nonzero point in $\mathcal{S}_r$ and Λ_{pr} is invariant under inversion, we have $\Lambda_{\mathrm{pr}} \cap \mathcal{U}_r \neq \varnothing$. Similarly we also have $\Lambda_{\mathrm{pr}} \cap \mathcal{R}_r \neq \varnothing$. Moreover, we note that for $0 < r < \log 1.01$, we have $\Lambda \cap \mathcal{U}_r = \Lambda_{\mathrm{pr}} \cap \mathcal{U}_r$ and $\Lambda \cap \mathcal{R}_r = \Lambda_{\mathrm{pr}} \cap \mathcal{R}_r$. This is because otherwise there would be some nonzero point $\boldsymbol{v} \in \Lambda \cap (\mathcal{U}_r \cup \mathcal{R}_r)$ and some integer $k \geq 2$ such that $\boldsymbol{v}/k \in \Lambda_{\mathrm{pr}}$, but $\boldsymbol{v} \in \mathcal{U}_r \cup \mathcal{R}_r$ and $k \geq 2$ imply that $\boldsymbol{v}/k \in \mathcal{S}_r$, contradicting the assumption that $\Lambda_{\mathrm{pr}} \cap \mathcal{S}_r = \varnothing$. Let $\boldsymbol{v}_1 = (t_1, 1 + v_1)$ be a point in $\Lambda_{\mathrm{pr}} \cap \mathcal{U}_r$ that is closest to the y-axis and let $\boldsymbol{v}_2 = (1 + v_2, t_2)$ be a point in $\Lambda_{\mathrm{pr}} \cap \mathcal{R}_r$ that is closest to the x-axis. We thus have $|t_i| \leq e^{-r}$ and $e^{-r} \leq 1 + v_i \leq e^r$ for $i = 1, 2$.

Let $\mathcal{P}_{\boldsymbol{v}_1,\boldsymbol{v}_2}$ be the parallelogram spanned by $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$. Then we have for $0 < r < \log 1.01$

$$|\mathcal{P}_{\boldsymbol{v}_1,\boldsymbol{v}_2}| = |(1 + v_1)(1 + v_2) - t_1 t_2| = (1 + v_1)(1 + v_2) - t_1 t_2 \leq e^{2r} + e^{-2r} < 3,$$

where $|\mathcal{P}_{\boldsymbol{v}_1,\boldsymbol{v}_2}|$ denotes the area of $\mathcal{P}_{\boldsymbol{v}_1,\boldsymbol{v}_2}$, and for the second equality we used that

$$(1 + v_1)(1 + v_2) \geq e^{-2r} \geq |t_1 t_2|.$$

Thus $|\mathcal{P}_{\boldsymbol{v}_1,\boldsymbol{v}_2}|$ equals 1 or 2. We claim that $|\mathcal{P}_{\boldsymbol{v}_1,\boldsymbol{v}_2}| = 1$. Suppose not; then $|\mathcal{P}_{\boldsymbol{v}_1,\boldsymbol{v}_2}| = 2$ and we have for $0 < r < \log 1.01$

$$t_1 t_2 = v_1 + v_2 + v_1 v_2 - 1 \leq 2(e^r - 1) + (e^r - 1)^2 - 1 < 0$$

and

$$|t_1 t_2| = 1 - v_1 - v_2 - v_1 v_2 \geq 1 - 2(e^r - 1) - (e^r - 1)^2 = 2 - e^{2r} > 0.9.$$

This implies $\min\{|t_1|, |t_2|\} > 0.9/e^{-r} > 0.9$. Since $t_1 t_2 < 0$, without loss of generality we may assume that $t_2 < 0$. Then we have $-e^{-r} \leq t_2 < -0.9$. On one hand, since $|\mathcal{P}_{\boldsymbol{v}_1,\boldsymbol{v}_2}| = 2$ and $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \Lambda_{\mathrm{pr}}$, we have

$$\boldsymbol{w} := \frac{\boldsymbol{v}_1 + \boldsymbol{v}_2}{2} = \left(\frac{t_1 + 1 + v_2}{2}, \frac{t_2 + 1 + v_1}{2}\right) \in \Lambda.$$

On the other hand, we have

$$0<\frac{t_1+1+v_2}{2}\le\frac{e^{-r}+e^{r}}{2}<e^{r},\quad 0<\frac{t_2+1+v_1}{2}<\frac{1+v_1}{2}\le\frac{e^{r}}{2}<e^{-r},$$

and $\boldsymbol{w}\notin\mathcal{S}_r$ implying $\boldsymbol{w}\in\mathcal{R}_r$. Thus $\boldsymbol{w}\in\Lambda\cap\mathcal{R}_r=\Lambda_{\mathrm{pr}}\cap\mathcal{R}_r$ is also a primitive vector of Λ. Moreover, since $-e^{-r}\le t_2<-0.9$, we have

$$0<\frac{t_2+1+v_1}{2}<\frac{e^{r}-0.9}{2}<\frac{1.01-0.9}{2}=0.055<|t_2|,$$

contradicting the assumption that $\boldsymbol{v}_2$ is the closest point in $\Lambda_{\mathrm{pr}}\cap\mathcal{R}_r$ to the x-axis. We thus have proved the claim, and it implies

$$|t_1t_2|=|v_1+v_2+v_1v_2|\le 2(e^{r}-1)+(e^{r}-1)^2=e^{2r}-1.$$

Hence we have

$$\min\{|t_1|,|t_2|\}\le\sqrt{|t_1t_2|}\le\sqrt{e^{2r}-1},$$

which implies $\Lambda_{\mathrm{pr}}\cap(\mathcal{A}_r\cup\mathcal{C}_r)\neq\varnothing$ finishing the proof. □

The following lemma states that for $r>0$ small, the orbits a_sK_r will completely leave the set K_r very shortly, and will remain separated for quite a long time.

Lemma 3.2. *For any* $0<r<\log 1.01$ *and any* $6r\le|s|\le\log 1.9$, *we have*

$$a_sK_r\cap K_r=\varnothing.$$

Proof. Suppose not, then there exists some $\Lambda\in a_sK_r\cap K_r$, and by definition the intersection of Λ_{pr} with $\mathcal{S}_r\cup a_s\mathcal{S}_r$ is empty. Without loss of generality we may assume that $s>0$. By Lemma 3.1 we have $\Lambda_{\mathrm{pr}}\cap(\mathcal{A}_r\cup\mathcal{C}_r)\neq\varnothing$ and similarly, $\Lambda_{\mathrm{pr}}\cap(a_s\mathcal{A}_r\cup a_s\mathcal{C}_r)\neq\varnothing$. We note that $a_s\mathcal{A}_r$ is the rectangle with vertices $(\pm e^{s}\sqrt{e^{2r}-1},e^{r-s})$ and $(\pm e^{s}\sqrt{e^{2r}-1},e^{-r-s})$. Since $e^{6r}\le e^{s}\le 1.9$ we have $a_s\mathcal{A}_r\subseteq\mathcal{S}_r$ implying $\Lambda_{\mathrm{pr}}\cap a_s\mathcal{C}_r\neq\varnothing$. Similarly, we have $\mathcal{C}_r\subseteq a_s\mathcal{S}_r$ and this implies $\Lambda_{\mathrm{pr}}\cap\mathcal{A}_r\neq\varnothing$ (see Figure 4). Let $\boldsymbol{v}_1\in\Lambda_{\mathrm{pr}}\cap\mathcal{A}_r$ and $\boldsymbol{v}_2\in\Lambda_{\mathrm{pr}}\cap a_s\mathcal{C}_r$, and let $\mathcal{P}_{\boldsymbol{v}_1,\boldsymbol{v}_2}$ be the parallelogram spanned by $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$. Then for $0<r<\log 1.01$ and $6r\le s\le\log 1.9$ we have

$$1<e^{s-2r}-(e^{2r}-1)e^{-s}\le|\mathcal{P}_{\boldsymbol{v}_1,\boldsymbol{v}_2}|\le e^{s+2r}+(e^{2r}-1)e^{-s}<2$$

contradicting the fact that $|\mathcal{P}_{\boldsymbol{v}_1,\boldsymbol{v}_2}|$ is a positive integer. □

We can now give:

Proof of Theorem 1.3. We prove the upper and lower bounds separately. For the upper bound, we first note that for any $\boldsymbol{v}\in\mathbb{R}^2$ we have $e^{-|s|}\|\boldsymbol{v}\|\le\|a_s\boldsymbol{v}\|\le e^{|s|}\|\boldsymbol{v}\|$. Hence for any $\Lambda\in X$ we have

$$|\Delta(a_s\Lambda)-\Delta(\Lambda)|\le|s|.$$

This implies that for any $s\in\mathbb{R}$ and any $r>0$

$$a_sK_r\subset K_{r+|s|}. \tag{3-6}$$

Figure 3. Figure 1 under the flow a_s: the rectangles $a_s\mathcal{S}_r$ (orange), $a_s\mathcal{A}_r$ (brown), and $a_s\mathcal{C}_r$ (purple).

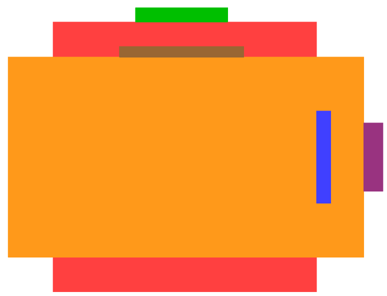

Figure 4. Figures 1 and 3 in one picture: the rectangle $a_s\mathcal{A}_r$ (brown) is contained in $\mathcal{S}_r$ (red), the rectangle $\mathcal{C}_r$ (blue) is contained in $a_s\mathcal{S}_r$ (orange).

Let $N = \lceil 1/r \rceil$. Using (3-6) and the fact that $1/N \le r$ we can estimate

$$\bigcup_{0\le s<1} a_{-s}K_r = \bigcup_{0\le i<N}\ \bigcup_{0\le t<1/N} a_{-i/N}a_{-t}K_r \subset \bigcup_{0\le i<N} a_{-i/N}K_{2r}.$$

Hence by Theorem 1.2 and since $N \asymp 1/r$ we have

$$\mu\biggl(\bigcup_{0\le s<1} a_{-s}K_r\biggr) \le \sum_{i=0}^{N-1}\mu(a_{-i/N}K_{2r}) \asymp r\log\Bigl(\frac{1}{r}\Bigr).$$

For the lower bound, for $0 < r < \log 1.01$ let $N = \lfloor 1/(6r)\rfloor$. First we have

$$\bigcup_{0\le i<\lfloor N\log 1.9\rfloor} a_{-i/N}K_r \subseteq \bigcup_{0\le s<1} a_{-s}K_r.$$

Moreover, for each $0 \le i < j < \lfloor N\log 1.9\rfloor$, we have $6r \le 1/N \le (j-i)/N < \log 1.9$; thus by Lemma 3.2 we have

$$a_{-i/N}K_r \cap a_{-j/N}K_r = a_{-j/N}(a_{(j-i)/N}K_r \cap K_r) = \varnothing.$$

Thus the union $\bigcup_{0\le i<\lfloor N\log 1.9\rfloor} a_{-i/N}K_r$ is disjoint and, again applying Theorem 1.2 and noting that $N \asymp 1/r$ we can estimate

$$\mu\biggl(\bigcup_{0\le s<1} a_{-s}K_r\biggr) \ge \sum_{i=0}^{\lfloor N\log 1.9\rfloor-1}\mu(a_{-i/N}K_r) \asymp r\log\Bigl(\frac{1}{r}\Bigr),$$

finishing the proof. □

Proof of Corollary 1.4. First we note that we can assume $\lim_{s\to\infty} r(s) = 0$ since otherwise both series would diverge. It follows that there exists $N > 0$ such that for any $n > N$, $0 < r(n) < \log 1.01$. Next, since $r(\cdot)$ is nonincreasing, for any $n > N$ we have

$$\bigcup_{0\le s<1} a_{-s} K_{r(n+1)} \subset \widetilde{B}_n \subset \bigcup_{0\le s<1} a_{-s} K_{r(n)}.$$

Moreover, since $n > N$ we have $0 < r(n+1) \le r(n) < \log 1.01$. Applying Theorem 1.3 to the left- and right-hand sides of the above inclusion relations we get

$$r(n+1)\log\left(\frac{1}{r(n+1)}\right) \ll \mu(\widetilde{B}_n) \ll r(n)\log\left(\frac{1}{r(n)}\right),$$

which finishes the proof. □

4. The dynamical Borel–Cantelli lemma

In this section we give the proof of Theorem 1.1 based on Theorem KW. Recall that for a given function $\psi : [t_0, \infty) \to (0, \infty)$ with $t_0 \ge 1$ fixed, we say a real number $x \in \mathbb{R}$ is *ψ-Dirichlet* if the system of inequalities

$$|qx - p| < \psi(t) \quad \text{and} \quad |q| < t$$

has a solution in $(p, q) \in \mathbb{Z} \times (\mathbb{Z} \smallsetminus \{0\})$ for all sufficiently large t. Let us denote by $D(\psi)$ the set of all ψ-Dirichlet numbers. Theorem KW gives a zero-one law for the Lebesgue measure of $D(\psi)$ as follows: if $\psi : [t_0, \infty) \to (0, \infty)$ is a continuous, nonincreasing function satisfying (1-6) and (1-7), then the series (1-8) diverges (resp. converges) if and only if the Lebesgue measure of $D(\psi)$ (resp. of $D(\psi)^c$) is zero.

For our purpose, we prove the following slightly modified version of *Dani correspondence*.

Lemma 4.1. *Let $\psi : [t_0, \infty) \to (0, \infty)$ be a continuous, nonincreasing function satisfying (1-6) and (1-7). Then there exists a unique continuous, nonincreasing function*

$$r = r_\psi : [s_0, \infty) \to (0, \infty), \quad \text{where } s_0 = \frac{\log t_0}{2} - \frac{\log \psi(t_0)}{2}$$

such that

$$\text{the function } s \mapsto s + r(s) \text{ is nondecreasing}, \tag{4-1}$$

and

$$\psi(e^{s-r(s)}) = e^{-s-r(s)} \quad \text{for all } s \ge s_0. \tag{4-2}$$

Conversely, given a continuous, nonincreasing function $r : [s_0, \infty) \to (0, \infty)$ satisfying (4-1), then there exists a unique continuous, nonincreasing function $\psi = \psi_r : [t_0, \infty) \to (0, \infty)$ with $t_0 = e^{s_0 - r(s_0)}$ satisfying (1-6), (1-7) and (4-2). Furthermore, if we assume $\lim_{t\to\infty} t\psi(t) = 1$ (or equivalently, $\lim_{s\to\infty} r(s) = 0$), then the series in (1-8) diverges if and only if the series

$$\sum_n r(n)\log\left(\frac{1}{r(n)}\right) \tag{4-3}$$

diverges.

Proof. The correspondence between $\psi = \psi_r$ and $r = r_\psi$ follows from the exact same construction as in [Kleinbock and Margulis 1999, Lemma 8.3], where $\psi(\cdot)$ and $r(\cdot)$ determine each other with the relations

$$e^s\psi(t) = e^{-r(s)} = e^{-s}t,$$

with s and t satisfying $s = (\log t)/2 - (\log\psi(t))/2$. The only difference is that here we require the two extra assumptions (1-6) and (1-7) on ψ which are respectively equivalent to the assumptions that $r(\cdot)$ is nonincreasing and $r(\cdot)$ is positive. We refer the reader to [Kleinbock and Margulis 1999, Lemma 8.3] for more details about this correspondence.

For the furthermore part, first we claim that the series in (1-8) diverges if and only if the integral

$$\int_{t_0}^{\infty} \frac{-(1-t\psi(t))\log(1-t\psi(t))}{t}\,dt \tag{4-4}$$

diverges. It suffices to show the function $G(t) := -\log(1-t\psi(t))(1-t\psi(t))$ is eventually nonincreasing in t. Note that the function $T \mapsto -T\log T$ is strictly increasing on the interval $(0, e^{-1})$. Since $\lim_{t\to\infty} t\psi(t) = 1$ and $t\psi(t) < 1$ for all $t \geq t_0$, there exists some $T_0 > t_0$ such that for all $t > T_0$, $0 < 1 - t\psi(t) < e^{-1}$. Moreover, together with the assumption (1-6) we get that $G(t)$ is nonincreasing in t for any $T > T_0$, finishing the proof the claim. Next, since $r(\cdot)$ is positive and nonincreasing, we have $0 < r(s) \leq r(s_0)$. Thus there exist constants $0 < c_1 < c_2$ such that for all $s \geq s_0$ and all $t \geq t_0$ with $s = (\log t)/2 - (\log\psi(t))/2$ we have

$$c_1 r(s) \leq 1 - t\psi(t) = 1 - e^{-2r(s)} \leq c_2 r(s).$$

This also implies

$$-\log(1-t\psi(t)) = -\log(r(s)) + O_{c_1,c_2}(1) \asymp_{c_1,c_2} -\log(r(s)),$$

where for the second estimate we used that $\lim_{s\to\infty} r(s) = 0$. Moreover, since $r(\cdot)$ is nonincreasing and continuous, it is differentiable at Lebesgue almost every $s \in \mathbb{R}$, and we denote by $r'(s)$ its derivative at $s \in \mathbb{R}$ whenever it exists. Using the relation $t = e^{s-r(s)}$ we get $dt/t = (1 - r'(s))\,ds$ for Lebesgue almost every $s \in \mathbb{R}$. We thus have

$$\int_{t_0}^{\infty} \frac{-(1-t\psi(t))\log(1-t\psi(t))}{t}\,dt \asymp_{c_1,c_2} \int_{s_0}^{\infty} -r(s)\log(r(s))(1-r'(s))\,ds \asymp \int_{s_0}^{\infty} -r(s)\log(r(s))\,ds,$$

where for the second estimate we used that $1 \leq 1 - r'(s) \leq 2$ for Lebesgue almost every $s \in \mathbb{R}$ which comes from the assumption (4-1) and that $r(\cdot)$ is nonincreasing. Finally, we conclude the proof by noting that the integral $\int_{s_0}^{\infty} -r(s)\log(r(s))\,ds$ diverges if and only if the series $\sum_n -r(n)\log(r(n))$ diverges since $\lim_{s\to\infty} r(s) = 0$ and $r(\cdot)$ is nonincreasing, which imply that the function $s \mapsto -r(s)\log(r(s))$ is eventually nonincreasing in s. □

As mentioned in the Introduction, we have the following dynamical interpretation of ψ-Dirichlet numbers.

Lemma 4.2 [Kleinbock and Wadleigh 2018, Proposition 4.5]. *Let $\psi : [t_0, \infty) \to (0, \infty)$ be a continuous and nonincreasing function satisfying (1-6) and (1-7). Let $r = r_\psi$ be as in Lemma 4.1. Then $x \in D(\psi)^c$ if and only if*

$$a_s\Lambda_x \in K_{r(s)} \quad \textit{for an unbounded set of } s, \tag{4-5}$$

where $a_s = \operatorname{diag}(e^s, e^{-s})$ *and*

$$\Lambda_x = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \mathbb{Z}^2 \in X$$

are as before.

Combining Theorem KW with Lemmas 4.1 and 4.2, we immediately have the following zero-one law.

Proposition 4.3. *Let* $r : [s_0, \infty) \to (0, \infty)$ *be continuous, nonincreasing, satisfying (4-1) and such that* $\lim_{s\to\infty} r(s) = 0$. *Then (4-5) holds for Lebesgue almost every (resp. almost no)* $x \in \mathbb{R}$ *provided that the series (4-3) diverges (resp. converges).*

To connect the above proposition with the corresponding property of almost every $\Lambda \in X$, we need an auxiliary lemma, which borrows some ideas from the work [Kleinbock and Rao 2019] of the first author with Anurag Rao.

Lemma 4.4. *Let* $r(\,\cdot\,)$ *be as in Proposition 4.3. For any* $c \in \mathbb{R}$ *and* $\lambda > 0$ *let*

$$r_{c,\lambda}(s) := r(s + c) - \lambda e^{-2(s+c)}$$

and define

$$D_{c,\lambda} := \{x \in \mathbb{R} \mid a_s \Lambda_x \in K_{r_{c,\lambda}(s)} \text{ for an unbounded set of } s\}.$$

If the series (4-3) diverges, then the set

$$D := \bigcap_{c\in\mathbb{R}} \bigcap_{\lambda>0} D_{c,\lambda}$$

has full Lebesgue measure.

Remark 4.5. We note that by our assumption $r_{c,\lambda}(\,\cdot\,)$ is not necessarily always positive, and the set $K_{r_{c,\lambda}(s)}$ is empty whenever $r_{c,\lambda}(s)$ is negative.

Proof of Lemma 4.4. For any function $f : [s_f, \infty) \to (0, \infty)$ with $s_f \geq 1$ we define

$$A_{\infty,f} := \{x \in \mathbb{R} \mid a_s \Lambda_x \in K_{f(s)} \text{ for an unbounded set of } s > s_f\}$$

and

$$N_f := \sum_{n\geq s_f} f(n) \log\left(\frac{1}{f(n)}\right).$$

First we note that the divergence of the series N_r is equivalent to the divergence of the series $N_{r_c/2}$ for any $c \in \mathbb{R}$, where $r_c(s) := r(s + c) = r_{c,0}(s)$. Moreover, it is clear that $(r_c/2)(\,\cdot\,)$ satisfies the assumptions in Proposition 4.3. Thus, by Proposition 4.3, if the series N_r diverges, then the set $A_{\infty,r_c/2}$ is of full Lebesgue measure for any $c \in \mathbb{R}$. On the other hand, for any $c \in \mathbb{R}$ and $\lambda > 0$ let $f_{c,\lambda}(s) = \lambda e^{-2(s+c)}$. It is easy to check that $f_{c,\lambda}|_{[s_{c,\lambda},\infty)}$ satisfies the assumptions in Proposition 4.3 with

$$s_{c,\lambda} := \max\left\{\frac{\log(2\lambda)}{2} - c,\, 1\right\},$$

and the series $N_{f_{c,\lambda}}$ converges for any $c \in \mathbb{R}$ and $\lambda > 0$. Thus by Proposition 4.3 the set $A_{\infty,f_{c,\lambda}}$ is of zero Lebesgue measure for any $c \in \mathbb{R}$ and $\lambda > 0$. Define

$$\bar{A} := \bigcap_{c\in\mathbb{R}} A_{\infty,r_c/2} \quad \text{and} \quad \underline{A} := \bigcup_{c\in\mathbb{R}} \bigcup_{\lambda>0} A_{\infty,f_{c,\lambda}}.$$

We note that since $r(\cdot)$ is nonincreasing, for any $c_1 < c_2$ we have $r_{c_1}/2 \geq r_{c_2}/2$ implying $A_{\infty,c_2/2} \subset A_{\infty,r_{c_1}/2}$. Hence the family of sets $\{A_{\infty,r_c/2}\}_{c\in\mathbb{R}}$ is nested and $\bar{A} = \lim_{c\to\infty} A_{\infty,r_c/2}$ is of full Lebesgue measure. Similarly, the family of sets $\{A_{\infty,f_{c,\lambda}}\}_{c\in\mathbb{R},\lambda>0}$ is also nested and the set

$$\underline{A} = \lim_{c\to-\infty} \lim_{\lambda\to\infty} A_{\infty,f_{c,\lambda}}$$

is of zero Lebesgue measure. Thus the set $\bar{A} \smallsetminus \underline{A}$ is of full Lebesgue measure and it suffices to show that $\bar{A} \smallsetminus \underline{A} \subset D$. That is, for any $x \in \bar{A} \smallsetminus \underline{A}$ we want to show that for any $c \in \mathbb{R}$ and any $\lambda > 0$ the events $a_s\Lambda_x \in K_{r_{c,\lambda}(s)}$ happen for an unbounded set of s. First we note that $x \in \bar{A}$ means that for any $c \in \mathbb{R}$ there exists an unbounded subset $S_c \subset \mathbb{R}$ such that $a_s\Lambda_x \in K_{r_c(s)/2}$ for any $s \in S_c$. Secondly, we note that $x \notin \underline{A}$ means that for any $c \in \mathbb{R}$ and $\lambda > 0$ there exists some constant $T_{c,\lambda} > 0$ such that for any $s \geq T_{c,\lambda}$ we have $a_s\Lambda_x \in \Delta^{-1}(f_{c,\lambda}(s), \infty)$. In particular, for any $s \in S_c \cap (T_{c,\lambda}, \infty)$ we have

$$f_{c,\lambda}(s) < \Delta(a_s\Lambda_x) \leq \frac{r_c(s)}{2}.$$

This implies

$$0 < \Delta(a_s\Lambda_x) \leq \frac{r_c(s)}{2} < \frac{r_c(s)}{2} + \frac{r_c(s)}{2} - f_{c,\lambda}(s) = r_{c,\lambda}(s)$$

for any $s \in S_c \cap (T_{c,\lambda}, \infty)$. Finally, we finish the proof by noting that since S_c is unbounded, the set $S_c \cap (T_{c,\lambda}, \infty)$ is also unbounded. □

We can now give:

Proof of Theorem 1.1. The convergent case follows directly from Corollary 1.4 and the classical Borel–Cantelli lemma, and we thus only need to prove the divergent case. Let $r : [s_0, \infty) \to (0, \infty)$ be continuous, nonincreasing, satisfying (4-1) and such that the series (4-3) diverges; we want to show that $\mu(B_\infty) = 1$. First we note that we can assume $\lim_{s\to\infty} r(s) = 0$, since otherwise the result would follow from the ergodicity of the flow $\{a_s\}_{s>0}$ on X. Let $D := \bigcap_{c\in\mathbb{R}} \bigcap_{\lambda>0} D_{c,\lambda}$ be as in Lemma 4.4 and define $B \subset X$ such that

$$B = \left\{ \begin{pmatrix} a & 0 \\ b & a^{-1} \end{pmatrix} \Lambda_x \in X \;\middle|\; b \in \mathbb{R},\ a > 0,\ x \in D \right\}.$$

We note that by Lemma 4.4 the set D has full Lebesgue measure. Thus the set $B \subset X$ is also of full measure (with respect to μ) and it suffices to show that $B \subset B_\infty$. First, by direct computation for

$$\Lambda = \begin{pmatrix} a & 0 \\ b & a^{-1} \end{pmatrix} \Lambda_x \in B$$

we have

$$a_s\Lambda = \begin{pmatrix} 1 & 0 \\ e^{-2s}a^{-1}b & 1 \end{pmatrix} a_{s+\log a}\Lambda_x. \tag{4-6}$$

Next, for any $y \in \mathbb{R}$ let

$$u_y^- = \begin{pmatrix} 1 & 0 \\ y & 1 \end{pmatrix}.$$

Note that for any $\boldsymbol{v} \in \mathbb{R}^2$, we have $\|u_y^- \boldsymbol{v}\| \le (|y|+1)\|\boldsymbol{v}\|$. This implies that for any $\Lambda \in X$

$$|\Delta(u_y^-\Lambda) - \Delta(\Lambda)| \le \log(1+|y|).$$

Using the above inequality, the relation (4-6), and the inequality $\log(1+x) < 2x$ for all $x > 0$, we get

$$|\Delta(a_s\Lambda) - \Delta(a_{s+\log a}\Lambda_x)| \le 2a^{-1}|b|e^{-2s}.$$

Since $x \in D$ for any $c \in \mathbb{R}$ and any $\lambda > 0$ we have $a_s\Lambda_x \in K_{r_{c,\lambda}(s)}$ for an unbounded set of s. In particular, taking $c = -\log a$, $\lambda = 2a^{-1}|b|$ we get

$$0 \le \Delta(a_s\Lambda) \le \Delta(a_{s-c}\Lambda_x) + \lambda e^{-2s} \le r_{c,\lambda}(s-c) + \lambda e^{-2s} = r(s)$$

for an unbounded set of s, finishing the proof. □

Acknowledgements

The authors would like to thank Anurag Rao, Nick Wadleigh and Cheng Zheng for many helpful conversations. Thanks are also due to the anonymous referee for a quick and careful report.

References

[Athreya 2009] J. S. Athreya, "Logarithm laws and shrinking target properties", *Proc. Indian Acad. Sci. Math. Sci.* **119**:4 (2009), 541–557. MR Zbl

[Athreya and Konstantoulas 2016] J. S. Athreya and I. Konstantoulas, "Discrepancy of general symplectic lattices", preprint, 2016. arXiv

[Athreya and Margulis 2009] J. S. Athreya and G. A. Margulis, "Logarithm laws for unipotent flows, I", *J. Mod. Dyn.* **3**:3 (2009), 359–378. MR Zbl

[Cassels 1997] J. W. S. Cassels, *An introduction to the geometry of numbers*, Springer, 1997. MR Zbl

[Chernov and Kleinbock 2001] N. Chernov and D. Kleinbock, "Dynamical Borel–Cantelli lemmas for Gibbs measures", *Israel J. Math.* **122** (2001), 1–27. MR Zbl

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

[Davenport and Schmidt 1970a] H. Davenport and W. M. Schmidt, "Dirichlet's theorem on diophantine approximation", pp. 113–132 in *Symposia Mathematica, IV* (Rome, 1968/69), Academic Press, London, 1970. MR Zbl

[Davenport and Schmidt 1970b] H. Davenport and W. M. Schmidt, "Dirichlet's theorem on diophantine approximation, II", *Acta Arith.* **16** (1970), 413–424. MR Zbl

[Fairchild 2019] S. K. Fairchild, "A higher moment formula for the Siegel–Veech transform over quotients by Hecke triangle groups", preprint, 2019. arXiv

[Kelmer and Yu 2019] D. Kelmer and S. Yu, "The second moment of the Siegel transform in the space of symplectic lattices", *Int. Math. Res. Not.* (online publication February 2019).

[Kleinbock and Margulis 1999] D. Y. Kleinbock and G. A. Margulis, "Logarithm laws for flows on homogeneous spaces", *Invent. Math.* **138**:3 (1999), 451–494. Correction in **211**:2 (2018), 855–862. MR Zbl

[Kleinbock and Rao 2019] D. Kleinbock and A. Rao, "A zero-one law for uniform Diophantine approximation in Euclidean norm", preprint, 2019. arXiv

[Kleinbock and Wadleigh 2018] D. Kleinbock and N. Wadleigh, "A zero-one law for improvements to Dirichlet's theorem", *Proc. Amer. Math. Soc.* **146**:5 (2018), 1833–1844. MR Zbl

[Lang 1975] S. Lang, $\mathrm{SL}_2(\mathbb{R})$, Addison-Wesley, Reading, MA, 1975. MR Zbl

[Maucourant 2006] F. Maucourant, "Dynamical Borel–Cantelli lemma for hyperbolic spaces", *Israel J. Math.* **152** (2006), 143–155. MR Zbl

[Randol 1970] B. Randol, "A group-theoretic lattice-point problem", pp. 291–295 in *Problems in analysis* (*papers dedicated to Salomon Bochner, 1969*), 1970. MR Zbl

[Rogers 1955] C. A. Rogers, "Mean values over the space of lattices", *Acta Math.* **94** (1955), 249–287. MR Zbl

[Schmidt 1960] W. M. Schmidt, "A metrical theorem in geometry of numbers", *Trans. Amer. Math. Soc.* **95** (1960), 516–529. MR Zbl

[Schmidt 1980] W. M. Schmidt, *Diophantine approximation*, Lecture Notes in Mathematics **785**, Springer, 1980. MR Zbl

[Siegel 1945] C. L. Siegel, "A mean value theorem in geometry of numbers", *Ann. of Math.* (2) **46** (1945), 340–347. MR Zbl

[Sullivan 1982] D. Sullivan, "Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics", *Acta Math.* **149**:3-4 (1982), 215–237. MR Zbl

[Veech 1998] W. A. Veech, "Siegel measures", *Ann. of Math.* (2) **148**:3 (1998), 895–944. MR Zbl

[Wood 1992] D. C. Wood, "The computation of polylogarithms", technical report 15-92*, University of Kent, Computing Laborabory, 1992, available at https://www.cs.kent.ac.uk/pubs/1992/110/content.pdf.

Received 2 Oct 2019. Revised 30 Dec 2019.

DMITRY KLEINBOCK:

kleinboc@brandeis.edu
Department of Mathematics, Brandeis University, Waltham, MA, United States

SHUCHENG YU:

yushucheng@campus.technion.ac.il
Department of Mathematics, Technion, Haifa, Israel