



A uniform metrical theorem in multiplicative Diophantine approximation

Downloaded from: <https://research.chalmers.se>, 2026-04-25 09:27 UTC

Citation for the original published paper (version of record):

Björklund, M., Fregoli, R., Gorodnik, A. (2026). A uniform metrical theorem in multiplicative Diophantine approximation. *Forum of Mathematics, Sigma*, 14.
<http://dx.doi.org/10.1017/fms.2026.10172>

N.B. When citing this work, cite the original published paper.

RESEARCH ARTICLE

A uniform metrical theorem in multiplicative Diophantine approximation

Michael Björklund¹, Reynold Fregoli² and Alexander Gorodnik³

¹Department of Mathematics, Chalmers University of Technology, Gothenburg, Sweden; E-mail: michjo@chalmers.se.

²Department of Mathematics, University of Zürich, Switzerland and Department of Mathematics, University of Michigan, Ann Arbor, USA; E-mail: reynold.fregoli@math.uzh.ch.

³Department of Mathematics, University of Zürich, Switzerland; E-mail: alexander.gorodnik@math.uzh.ch (Corresponding author).

Received: 18 October 2022; **Revised:** 6 September 2023; **Accepted:** 15 October 2023

2020 Mathematics Subject Classification: 11K38, 11K60, 11J13, 11H46

Abstract

For Lebesgue generic $(x_1, x_2) \in \mathbb{R}^2$, we investigate the distribution of small values of products $q \cdot \|qx_1\| \cdot \|qx_2\|$ with $q \in \mathbb{N}$, where $\|\cdot\|$ denotes the distance to the closest integer. The main result gives an asymptotic formula for the number of $1 \leq q < T$ such that

$$a < q \cdot \|qx_1\| \cdot \|qx_2\| \leq b \quad \text{and} \quad \|qx_1\|, \|qx_2\| \leq c$$

for given parameters a, b, c satisfying certain growth conditions.

Contents

1	Introduction	2
2	Organisation of the paper	4
3	Tessellations	5
4	Group perturbations and controlled sets	7
5	Height function estimates	10
6	Correlations between the number of shifted lattice points in boxes	17
7	Mean counting within controlled sets	19
8	Smooth approximations	22
9	Proof of Theorem 1.3	31
10	Proof of Theorem 1.2	41
11	Proof of Theorem 1.1	48
A	Volume estimates	49
B	An auxiliary double sum	50
	References	55

MB was supported by Swedish VR-grant 11253320, RF and AG were supported by SNF grant 200021–182089.

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

1. Introduction

In this paper we will be interested in the basic problem of multiplicative Diophantine approximation regarding existence of small values for products

$$q \cdot \|qx_1\| \cdot \|qx_2\|, \quad \text{with } q \in \mathbb{N},$$

for Lebesgue generic $\underline{x} = (x_1, x_2) \in \mathbb{R}^2$, where $\|\cdot\|$ denotes the distance to the closest integer. The main result in this direction was established by Gallagher [8], who showed that for any nonincreasing function $\psi : \mathbb{N} \rightarrow \mathbb{R}_+$ satisfying $\sum_{q=1}^\infty \psi(q) \frac{\log q}{q} = \infty$, the inequality

$$q \cdot \|qx_1\| \cdot \|qx_2\| \leq \psi(q)$$

has infinitely many solutions $q \in \mathbb{N}$ for almost all $\underline{x} = (x_1, x_2) \in \mathbb{R}^2$. For instance, for almost all $\underline{x} = (x_1, x_2) \in \mathbb{R}^2$, there are infinitely many solutions $q \in \mathbb{N}$ to the inequality

$$q \cdot \|qx_1\| \cdot \|qx_2\| \leq (\log q)^{-2}.$$

The next natural question involves the distribution of the set of solutions

$$M(\underline{x}) := \{q \in \mathbb{N} : q \cdot \|qx_1\| \cdot \|qx_2\| \leq \psi(q)\}.$$

Wang and Yu [21] derived an explicit asymptotic formula for the cardinality $|M(\underline{x}) \cap [1, T]|$ as $T \rightarrow \infty$. Inhomogeneous versions of these problems were also studied by Chow and Technau [4, 5]. The above results are usually described as *asymptotic* Diophantine problems.

Here we explore the analogous *uniform* Diophantine problem that involves solutions to the inequalities

$$q \cdot \|qx_1\| \cdot \|qx_2\| \leq b, \quad 1 \leq q < T$$

for given parameters b and T . It should be noted that there is an essential difference between asymptotic problems and uniform problems. For instance, establishing a uniform version of the classical Khinchin theorem was addressed only recently in [15, 12, 13]. A somewhat different related uniform problem was also considered by Widmer [22]. His result concerns the sets

$$N(\underline{x}; b) := \{q \in \mathbb{N} : \|qx_1\| \cdot \|qx_2\| \leq b\},$$

and it follows from [22] that if $b \gg (\log T)^{4+\varepsilon}/T$ for some $\varepsilon > 0$, then for almost all $\underline{x} \in \mathbb{R}^2$,

$$|N(\underline{x}; b) \cap [1, T]| = 4 \cdot b(1 - \log(4b)) \cdot T + O_{\underline{x}, \varepsilon} \left(b^{2/3} (-\log b) \cdot T^{2/3} (\log T)^{4/3+\varepsilon/3} \right).$$

The condition on the parameter b is probably not optimal and is needed to ensure that the required Diophantine condition in [22] holds for almost all $\underline{x} \in \mathbb{R}^2$. However, it is not clear whether this method can give an asymptotic formula in a larger range.

Our first main result gives an asymptotic formula for the growth of the sets

$$L(\underline{x}; b) := \{q \in \mathbb{N} : q \cdot \|qx_1\| \cdot \|qx_2\| \leq b\}.$$

Theorem 1.1. *Let $\eta \in (1, 2)$. Then for every $\varepsilon > 0$ and for almost every $\underline{x} \in \mathbb{R}^2$,*

- when $(\ln T)^{-\eta} \leq b \leq (\ln T)^{-\eta/(3-\eta)}$,

$$|L(\underline{x}; b) \cap [1, T]| = 2 \cdot b \cdot (\ln T)^2 + O_{\underline{x}, \varepsilon} \left(b^{1-1/\eta} \cdot (\ln \ln T)^{6+\varepsilon} \ln T \right),$$

◦ when $b \geq (\ln T)^{-\eta/(3-\eta)}$,

$$|L(\underline{x}; b) \cap [1, T]| = 2 \cdot b \cdot (\ln T)^2 + O_{\underline{x}, \varepsilon} \left(b^{2/\eta} \cdot (\ln \ln T)^{6+\varepsilon} (\ln T)^2 \right).$$

Theorem 1.1 is proved in Section 11. It will be deduced from a more general result which we now describe.

Given parameters $T \geq 1$ and $a, b, c > 0$ satisfying

$$a < b < 1 \quad \text{and} \quad c < \frac{1}{2},$$

we consider the domains

$$\Omega = \Omega_{[1, T]}^{(a, b], c} := \left\{ (u, y) \in \mathbb{R}^2 \times [1, T) : \begin{array}{l} \max(|u_1|, |u_2|) \leq c \\ a < |u_1 u_2| \cdot y \leq b \end{array} \right\}, \tag{1.1}$$

and, for $\underline{x} = (x_1, x_2) \in \mathbb{R}^2$, the corresponding sets

$$Q_\Omega(\underline{x}) := \left\{ q \in [1, T) \cap \mathbb{N} : \begin{array}{l} \max(\|qx_1\|, \|qx_2\|) \leq c \\ a < q \cdot \|qx_1\| \cdot \|qx_2\| \leq b \end{array} \right\}.$$

It is natural to expect that the cardinality $|Q_\Omega(\underline{x})|$ grows as the volume of Ω . Our next theorem proves such an asymptotic formula:

Theorem 1.2. *Suppose that*

$$a \geq (\ln T)^{-\theta} \quad \text{and} \quad \zeta \cdot b \leq c^2 \quad \text{for some } \zeta, \theta > 0.$$

Let $\eta \in (1, 2)$. Then for every $\varepsilon > 0$ and for almost every $\underline{x} \in \mathbb{R}^2$,

◦ when $(c \cdot \ln T)^{-\eta} \leq b \leq (c \cdot \ln T)^{-\eta/(3-\eta)}$,

$$|Q_\Omega(\underline{x})| = \text{Vol}_3(\Omega) + O_{\underline{x}, \zeta, \theta, \varepsilon} \left(c^{-1} b^{1-1/\eta} \cdot (\ln \ln T)^{6+\varepsilon} \ln T + 1 \right),$$

◦ when $b \geq (c \cdot \ln T)^{-\eta/(3-\eta)}$,

$$|Q_\Omega(\underline{x})| = \text{Vol}_3(\Omega) + O_{\underline{x}, \zeta, \theta, \varepsilon} \left(b^{2/\eta} \cdot (\ln \ln T)^{6+\varepsilon} (\ln T)^2 + 1 \right).$$

To compare the above main term with the error term, we note that according to (A.4) in Appendix A,

$$\text{Vol}_3(\Omega) \sim 2 \cdot (\ln T)^2 \cdot (b - a) \quad \text{as } T \rightarrow \infty. \tag{1.2}$$

Due to the assumption on the parameter a , we cannot directly take $a = 0$ in the Theorem 1.2 to deduce almost sure asymptotics for the sets $L(\underline{x}; b)$ in Theorem 1.1. Instead we analyze the existence of lattice points in thin hyperbolic strips and show that when a decays sufficiently fast, the sets $L(\underline{x}, a)$ are empty for almost all $\underline{x} \in \mathbb{R}^2$ (see Section 11).

Our method also allows us to study the distribution of the scaled set $Q_\Omega(\underline{x})/T$ in the interval $[0, 1]$. For a compactly supported Lipschitz function $h : \mathbb{R} \rightarrow \mathbb{R}$ consider the weighted counting function

$$S_\Omega h(\underline{x}) := \sum_{q \in Q_\Omega(\underline{x})} h\left(\frac{q}{T}\right), \quad \underline{x} \in \mathbb{R}^2.$$

We shall show that $S_\Omega h(\underline{x})$ is asymptotic on average to the weighted mean

$$\mathcal{M}_\Omega(h) := \int_1^T h\left(\frac{y}{T}\right) \cdot \text{Vol}_2(\Omega(y)) \, dy, \tag{1.3}$$

where

$$\Omega(y) := \{x \in \mathbb{R}^2 : (x, y) \in \Omega\}.$$

We note that $\text{Vol}_2(\Omega(y))$ can be computed explicitly (see Lemma A.1). When $h = 1$ on the interval $[0, 1]$, the sum $S_\Omega h(\underline{x})$ is simply $|Q_\Omega(\underline{x})|$ and $\mathcal{M}_\Omega(h) = \text{Vol}_3(\Omega)$. For different oscillatory functions h , the asymptotics of $S_\Omega h(\underline{x})$ describe the distribution of $Q_\Omega(\underline{x})$ inside $[1, T)$.

We establish a bound for ‘subquadratic’ moments defined in terms of the function

$$\theta_\kappa(s) := s^2 / \ln(e + |s|)^{1+\kappa} \quad \text{with } \kappa > 0. \tag{1.4}$$

Theorem 1.3. *With the assumption as in Theorem 1.2, for every compactly supported Lipschitz function $h : \mathbb{R} \rightarrow \mathbb{R}$ and $\kappa > 0$,*

$$\int_{[0,1]^2} \theta_\kappa\left(S_\Omega h(\underline{x}) - \mathcal{M}_\Omega(h)\right) d\underline{x} \ll_{h,\zeta,\theta,\kappa} c^{-2} b \cdot (\ln \ln T)^{3+\kappa} (\ln T)^2.$$

In particular,

$$\int_{[0,1]^2} \theta_\kappa\left(|Q_\Omega(\underline{x})| - \text{Vol}_3(\Omega)\right) d\underline{x} \ll_{\zeta,\theta,\kappa} c^{-2} b \cdot (\ln \ln T)^{3+\kappa} (\ln T)^2.$$

Remark 1.4. In view of the volume asymptotics (1.2), for certain ranges of the parameters, the formula in Theorem 1.2 establishes an error term with essentially ‘square-root cancellation’.

Theorem 1.3 is proved in Section 9. Then Theorem 1.2 is deduced from it in Section 10 using a Borel-Cantelli argument. We outline the structure of the paper in Section 2 below.

2. Organisation of the paper

It will be convenient to consider a more geometric framework. For $\underline{x} \in \mathbb{R}^2$, we consider the unimodular lattice

$$\Lambda_{\underline{x}} := \{(\underline{p} + q\underline{x}, q) \in \mathbb{R}^2 \times \mathbb{R} : (\underline{p}, q) \in \mathbb{Z}^2 \times \mathbb{Z}\}. \tag{2.1}$$

Since $c < \frac{1}{2}$, the map

$$\Lambda_{\underline{x}} \cap \Omega \rightarrow Q_\Omega(\underline{x}) : (\underline{p} + q\underline{x}, q) \mapsto q$$

is a bijection. In particular, Theorem 1.2 can be viewed as an asymptotic lattice point counting problem for the cardinality $|\Lambda_{\underline{x}} \cap \Omega|$.

In order to produce an asymptotic formula for $|\Lambda_{\underline{x}} \cap \Omega|$, one usually aims to establish a bound on the second moment $\int_{[0,1]^2} \left| |\Lambda_{\underline{x}} \cap \Omega| - \text{Vol}_3(\Omega) \right|^2 d\underline{x}$. This idea goes back, for instance, to the work of W. Schmidt [17]. We also refer to the work Kleinbock, Shi, and Weiss [14], where this approach was developed for ergodic sums. However, our analysis in Section 5 suggests that a needed bound for the second moment is not attainable. As a substitute we establish a bound for ‘subquadratic’ moments $\int_{[0,1]^2} \theta_\kappa(|\Lambda_{\underline{x}} \cap \Omega| - \text{Vol}_3(\Omega)) d\underline{x}$, with θ_κ as in (1.4) and generalize the argument of [16, 17] to such subquadratic moments.

It will be convenient to view the domains $\Omega \subset \mathbb{R}^3$ as disjoint union of smaller domains. To construct such a decomposition, we use the action by the diagonal matrices

$$a(t) := \text{diag}(e^{t_1}, e^{t_2}, e^{-t_1-t_2}), \quad t = (t_1, t_2) \in \mathbb{R}_+^2.$$

In Section 3, we show that

$$\Omega = \bigsqcup_{n \in \mathcal{F}_\Omega} a(n)^{-1} \Delta_{\Omega,n}, \tag{2.2}$$

where \mathcal{F}_Ω is a finite subset of \mathbb{N}_o^2 and $\Delta_{\Omega,n}$ are finite subsets of \mathbb{R}^3 . We note that this tessellation procedure is different compared to the one used in our previous work [3]. While in [3] we used varying tessellations defined for each $a(t)$ -orbit, here we construct a tessellation directly in \mathbb{R}^3 . This has advantages and disadvantages. In particular, in our present construction the tiles $\Delta_{\Omega,n}$ have more complicated shape. This necessitates a new notion of regularity ((ε, γ, M) -controlled sets), which we develop in Section 4.

The decomposition (2.2) allows us to write

$$|\Lambda_{\underline{x}} \cap \Omega| = \sum_{n \in \mathcal{F}_\Omega} \widehat{\chi}_{\Delta_{\Omega,n}}(a(n)\Lambda_{\underline{x}}), \tag{2.3}$$

where $\widehat{\chi}_{\Delta_{\Omega,n}}$ denote the Siegel transforms of the characteristic function of the sets $\Delta_{\Omega,n}$. Hence, the original counting question is reformulated in terms of ergodic sums for the action of the group $\{a(t) : t \in \mathbb{R}^2\}$, albeit these sums are computed along a varying family of functions. The crucial ingredient of our proof is a quasi-independence estimate (Theorem 9.2) established in our previous work [2]. It gives a quantitative bound for the correlations of $\varphi \circ a(n)$, $n \in \mathbb{N}_o^2$, for smooth compactly supported functions φ on the space of unimodular lattices. It should be noted that this bound is only useful when $\min(n_1, n_2)$ is large, and one of the hardest parts of the present paper consists in treating the part of the sum (2.3) where $\min(n_1, n_2)$ is small.

In Section 5 we establish various nondivergence estimates for lattices $a(t)\Lambda_{\underline{x}}$ with $t \in \mathbb{R}_+^2$ and $\underline{x} \in [0, 1)^2$ (and somewhat more general lattices). Although this question fits in the general framework of non-divergence of unipotent flows (cf. [6, 19, 10]), we need much more precise information about the height functions along $a(t)\Lambda_{\underline{x}}$ (see, for instance, Corollary 5.5). The results of Section 5 will be used in the proof of the main theorem to estimate the part of the sum with small $\min(n_1, n_2)$ (this corresponds to the term $\mathcal{C}_\Omega^{(1)}$ in Section 9) and also in the construction of smooth approximations.

In Section 8 we build smooth compactly supported approximations for the functions $\widehat{\chi}_{\Delta_{\Omega,n}}$. Our main technical result here, used in the proof of the main theorem, is Lemma 8.5 (this corresponds to the term $\mathcal{C}_\Omega^{(2)}$ in Section 9). The proof of this lemma uses the notion of (ε, γ, M) -controlled sets and the nondivergence estimates from Section 5. Ultimately, its proof reduces to an arithmetic problem of estimating the number of points from $a(t)\Lambda_{\underline{x}}$ contained in an (ε, γ, M) -controlled set (see Lemma 7.1) and to an estimate for correlations of certain lattice counting functions (see Lemma 6.1). Those are handled in Sections 6–7.

The proof of Theorem 1.3 is completed in Section 9. In Section 10 we deduce Theorem 1.2 from Theorem 1.3 using a Borel-Cantelli argument. Finally, Theorem 1.1 is deduced from 1.2 in Section 11.

3. Tessellations

In this section we show how to decompose the ‘hyperboloid strips’

$$\Omega := \left\{ (\underline{x}, y) \in \mathbb{R}^2 \times \mathbb{R} : \begin{array}{l} a < |x_1 x_2| \cdot y \leq b, \\ \max(|x_1|, |x_2|) \leq c, \\ T_0 \leq y < T \end{array} \right\}. \tag{3.1}$$

defined for $1 \leq T_0 < T$, $0 < a < b$, and $c > 0$. We use the action of the semigroup of diagonal matrices of the form

$$a(t) := \begin{pmatrix} e^{t_1} & & \\ & e^{t_2} & \\ & & e^{-(t_1+t_2)} \end{pmatrix}, \quad \text{for } t = (t_1, t_2) \in \mathbb{R}_+^2. \tag{3.2}$$

Our aim below is to show that Ω can be written as a union of certain $a(t)$ -translates of smaller pieces $\Delta_{\Omega,n}$, where $n = (n_1, n_2) \in \mathbb{N}_o^2$, and whose dependence on the parameters is rather mild. More precisely, let

$$\Delta_{\Omega,n} := \left\{ (\underline{x}, y) : \begin{array}{l} a < |x_1 x_2| \cdot y \leq b \\ c e^{-1} < |x_i| \leq c, \text{ for } i = 1, 2 \\ T_0 e^{-(n_1+n_2)} \leq y < T e^{-(n_1+n_2)} \end{array} \right\} \tag{3.3}$$

and

$$\mathcal{F}_\Omega := \{n \in \mathbb{N}_o^2 : \alpha_\Omega \leq n_1 + n_2 < \beta_\Omega\}, \tag{3.4}$$

where

$$\alpha_\Omega := \ln\left(\frac{T_0 c^2}{b e^2}\right) \quad \text{and} \quad \beta_\Omega := \ln\left(\frac{T c^2}{a}\right). \tag{3.5}$$

Then, we have the following lemma:

Lemma 3.1. *For every $n \in \mathbb{N}_o^2$,*

$$\Delta_{\Omega,n} \neq \emptyset \implies n \in \mathcal{F}_\Omega$$

and

$$\Omega = \bigsqcup_{n \in \mathcal{F}_\Omega} a(n)^{-1} \Delta_{\Omega,n},$$

Furthermore, for all $n \in \mathcal{F}_\Omega$,

$$\Delta_{\Omega,n} \subset [-c, c]^2 \times \left(\frac{a}{c^2}, \frac{b e^2}{c^2}\right].$$

Proof. We first note that for every $n = (n_1, n_2) \in \mathbb{N}_o^2$,

$$a(n)^{-1} \Delta_{\Omega,n} = \left\{ (\underline{x}, y) : \begin{array}{l} a < |x_1 x_2| y \leq b \\ c e^{-(n_i+1)} < |x_i| \leq c e^{-n_i}, \text{ for } i = 1, 2 \\ T_0 \leq y < T \end{array} \right\},$$

so in particular, $a(n)^{-1} \Delta_{\Omega,n} \subset \Omega$ and the sets $\{a(n)^{-1} \Delta_{\Omega,n}\}_{n \in \mathbb{N}^d}$ are disjoint. Pick a point $(\underline{x}, y) \in \Omega$. Then there is clearly a unique $n \in \mathbb{N}_o^2$ such that

$$c \cdot e^{-(n_i+1)} < |x_i| \leq c \cdot e^{-n_i}, \quad \text{for both } i = 1, 2,$$

and thus $(\underline{x}, y) \in a(n)^{-1}\Delta_{\Omega,n}$. Since (\underline{x}, y) is arbitrary, this shows that

$$\Omega = \bigsqcup_{n \in \mathbb{N}_o^2} a(n)^{-1}\Delta_{\Omega,n}. \tag{3.6}$$

However, most of the sets $\Delta_{\Omega,n}$ in this union are empty. Indeed, suppose that $\Delta_{\Omega,n} \neq \emptyset$, and pick $(\underline{x}, y) \in \Delta_{\Omega,n}$. Then,

$$a < |x_1x_2| \cdot y < c^2 \cdot T \cdot e^{-(n_1+n_2)},$$

and thus $n_1 + n_2 < \ln\left(\frac{Tc^2}{a}\right)$. Similarly,

$$c^2 \cdot T_0 \cdot e^{-2} \cdot e^{-(n_1+n_2)} \leq |x_1x_2|y \leq b,$$

and thus $\ln\left(\frac{T_0c^2}{be^2}\right) \leq n_1 + n_2$. In other words, $\Delta_{\Omega,n}$ is nonempty only if $n \in \mathcal{F}_\Omega$, and thus the union in (3.6) can without loss of generality be restricted to \mathcal{F}_Ω .

Finally, suppose that $(\underline{x}, y) \in \Delta_{\Omega,n}$ for some n . Then $\underline{x} \in [-c, c]^2$ and

$$\frac{a}{c^2} < \frac{a}{|x_1x_2|} < y \leq \frac{b}{|x_1x_2|} \leq \frac{be^2}{c^2},$$

and thus

$$\Delta_{\Omega,n} \subset [-c, c]^2 \times \left(\frac{a}{c^2}, \frac{be^2}{c^2}\right]$$

independently of n . □

4. Group perturbations and controlled sets

The aim of this section is to establish regularity under small $SL_3(\mathbb{R})$ -perturbations for families of sets as in (3.3). Since the sets relevant to our problem degenerate in some directions, we will need to introduce a new notion of regularity.

We begin with some notation. If $E \subset \mathbb{R}^2 \times \mathbb{R}$, let

$$E^y = \{\underline{x} \in \mathbb{R}^2 : (\underline{x}, y) \in E\} \quad \text{and} \quad E_{\underline{x}} = \{y \in \mathbb{R} : (\underline{x}, y) \in E\},$$

for $y \in \mathbb{R}$ and $\underline{x} \in \mathbb{R}^2$. We refer to E^y as the *y-section* of E , and to $E_{\underline{x}}$ as the *x-section* of E .

The following definition captures a convenient form of regularity that we will use in the paper.

Definition 4.1 (Controlled set). Let $M > 1$ and $0 < \varepsilon < \gamma < M$. We say that a Borel set $E \subset \mathbb{R}^2 \times \mathbb{R}$ is (ε, γ, M) -controlled if either

$$E \subset [-M, M]^2 \times (\gamma, M] \quad \text{and} \quad \sup_{y \in [\gamma, 1]} \text{Vol}_2(E^y) \ll_M \max\left(\varepsilon, -\frac{\varepsilon}{\gamma} \ln\left(\frac{\varepsilon}{\gamma}\right)\right),$$

or if there is an interval $[\alpha, \beta]$ such that

$$E \subset [-M, M]^2 \times [\alpha, \beta], \quad \beta - \alpha \ll_M \varepsilon, \quad \text{and} \quad \alpha \geq \gamma/2$$

with implicit constants that are independent of ε and γ . If E satisfies the first set of conditions, we say that E is *type I* and if E satisfies the second set of conditions, we say that E is *type II*. We refer to the implicit constants above as the *bounds for E*.

Roughly speaking, a controlled set is type I if all of its y -sections have uniformly small volumes and it is type II if all of its \underline{x} -sections have uniformly small volumes.

Remark 4.2. Note that if $\varepsilon < \gamma' < \gamma$, $\varepsilon/\gamma' < 1/e$ and E is an (ε, γ, M) -controlled set, then E is also an $(\varepsilon, \gamma', M)$ -controlled set.

Let $M > 1$. For the rest of this section, let us fix the following real parameters:

$$0 < a < b \quad \text{and} \quad u_1^- < u_1^+ \leq 1/2, u_2^- < u_2^+ \leq 1/2 \quad \text{and} \quad 0 < \gamma < \delta \leq M. \tag{4.1}$$

We define the set

$$\Delta := \left\{ (\underline{x}, y) : \begin{array}{l} a < |x_1 x_2| \cdot y \leq b \\ u_i^- < |x_i| \leq u_i^+ \text{ for } i = 1, 2 \\ \gamma \leq y < \delta \end{array} \right\} \subset \mathbb{R}^2 \times \mathbb{R}. \tag{4.2}$$

We fix the max-norm on \mathbb{R}^3 and denote by $\|\cdot\|_{\text{op}}$ the operator norm on $\text{SL}_3(\mathbb{R})$ induced by the max-norm, in particular,

$$\|g \cdot \underline{x}\|_{\infty} \leq \|g\|_{\text{op}} \cdot \|\underline{x}\|_{\infty}, \quad \text{for } g \in \text{SL}_3(\mathbb{R}) \text{ and } \underline{x} \in \mathbb{R}^3$$

For $\varepsilon > 0$, we denote by V_{ε} the symmetric open neighborhoods around the identity in $\text{SL}_3(\mathbb{R})$:

$$V_{\varepsilon} := \{g \in \text{SL}_3(\mathbb{R}) : \|g - \text{id}\|_{\text{op}} < \varepsilon \quad \text{and} \quad \|g^{-1} - \text{id}\|_{\text{op}} < \varepsilon\}. \tag{4.3}$$

In other words,

$$\|g^{\pm 1} \cdot (\underline{x}, y) - (\underline{x}, y)\|_{\infty} < \varepsilon \cdot \|(\underline{x}, y)\|_{\infty}, \quad \text{for all } g \in V_{\varepsilon} \text{ and } (\underline{x}, y) \in \mathbb{R}^2 \times \mathbb{R}. \tag{4.4}$$

The following lemma is the main result of this section.

Lemma 4.3. *Let Δ be as in (4.2) with the assumptions (4.1). Then, for every*

$$0 < \varepsilon < \min\left(1/(2M), \gamma/(2M), a/(M^2 + 1)\right),$$

there are $(\varepsilon, \gamma/2, M + 1)$ -controlled sets $E_s \subset \mathbb{R}^2 \times \mathbb{R}$, $s = 1, \dots, 24$, such that

$$(g^{-1}\Delta \setminus \Delta) \sqcup (\Delta \setminus g^{-1}\Delta) \subset \bigcup_s E_s \quad \text{for all } g \in V_{\varepsilon}.$$

Furthermore, the bounds in Definition 4.1 for the sets E_s are independent of the parameters $a, b, u_1^{\pm}, u_2^{\pm}, \delta$.

Proof. Let us fix ε as above. For $g \in V_{\varepsilon}$ and $(\underline{x}, y) \in \mathbb{R}^2 \times \mathbb{R}$, we define $(\underline{x}(g), y(g))$ by

$$g \cdot (\underline{x}, y) := (\underline{x}(g), y(g)) = (x_1(g), x_2(g), y(g)). \tag{4.5}$$

Then

$$\max(|x_1(g) - x_1|, |x_2(g) - x_2|, |y(g) - y|) < \varepsilon \cdot \|(\underline{x}, y)\|_{\infty}, \tag{4.6}$$

for all $(\underline{x}, y) \in \mathbb{R}^2 \times \mathbb{R}$. In particular, for $(\underline{x}, y) \in \Delta$,

$$\max(|x_1(g) - x_1|, |x_2(g) - x_2|, |y(g) - y|) < M\varepsilon.$$

Hence, it follows that

$$u_i^- - M\varepsilon < |x_i(g)| \leq u_i^+ + M\varepsilon \text{ for } i = 1, 2 \quad \text{and} \quad \gamma - M\varepsilon \leq y(g) \leq \delta + M\varepsilon.$$

Also using that $\varepsilon < 1/(2M)$,

$$\begin{aligned} \left| |x_1 x_2| \cdot y - |x_1(g) x_2(g)| \cdot y(g) \right| &\leq \left| |x_1 x_2| \cdot y - |x_1(g) x_2| \cdot y \right| + \left| |x_1(g) x_2| \cdot y - |x_1(g) x_2(g)| \cdot y \right| \\ &\quad + \left| |x_1(g) x_2(g)| \cdot y - |x_1(g) x_2(g)| \cdot y(g) \right| \leq \varepsilon M^2. \end{aligned}$$

Therefore, we also have that

$$a - \varepsilon M^2 < |x_1(g) x_2(g)| \cdot y(g) \leq b + \varepsilon M^2.$$

We conclude that for $g \in V_\varepsilon$,

$$g^{-1} \Delta \subset \Delta_\varepsilon^+,$$

where

$$\Delta_\varepsilon^+ := \left\{ \begin{array}{l} a - \varepsilon M^2 < |x_1 x_2| \cdot y \leq b + \varepsilon M^2 \\ (\underline{x}, y) : u_i^-(\varepsilon) < |x_i| \leq u_i^+(\varepsilon) \text{ for } i = 1, 2 \\ \gamma - \varepsilon M \leq y < \delta + \varepsilon M \end{array} \right\},$$

where $u_i^-(\varepsilon) := u_i^- - \varepsilon M$ and $u_i^+(\varepsilon) := u_i^+ + \varepsilon M$. We note that

$$\Delta_\varepsilon^+ \subset (-1, 1)^2 \times [\gamma/2, M + 1], \tag{4.7}$$

because $\varepsilon < \gamma/(2M)$ and $\varepsilon < 1/(2M)$.

Let

$$E_1 := \{(\underline{x}, y) \in \Delta_\varepsilon^+ : a - \varepsilon M^2 < |x_1 x_2| \cdot y \leq a\}.$$

For all $y > 0$,

$$E_1^y \subset \left(\begin{array}{cc} u_1^+(\varepsilon) & 0 \\ 0 & u_2^+(\varepsilon) \end{array} \right) \cdot \left(\Xi \left(\frac{a}{y \cdot u_1^+(\varepsilon) u_2^+(\varepsilon)} \right) \setminus \Xi \left(\frac{a - \varepsilon M^2}{y \cdot u_1^+(\varepsilon) u_2^+(\varepsilon)} \right) \right),$$

where the set $\Xi(\cdot)$ is defined as in (A.1). It follows from (A.3) that

$$\begin{aligned} &\text{Vol}_2 \left(\Xi \left(\frac{a}{y \cdot u_1^+(\varepsilon) u_2^+(\varepsilon)} \right) \setminus \Xi \left(\frac{a - \varepsilon M^2}{y \cdot u_1^+(\varepsilon) u_2^+(\varepsilon)} \right) \right) \\ &\leq - \frac{4\varepsilon M^2}{y \cdot u_1^+(\varepsilon) u_2^+(\varepsilon)} \cdot \ln^- \left(\frac{a - \varepsilon M^2}{y \cdot u_1^+(\varepsilon) u_2^+(\varepsilon)} \right), \end{aligned}$$

where $\ln^-(z) := \ln \min(1, z)$. Since $\varepsilon < a/(M^2 + 1)$ and $u_1^+(\varepsilon), u_2^+(\varepsilon) < 1$, we have

$$\frac{a - \varepsilon M^2}{u_1^+(\varepsilon) u_2^+(\varepsilon)} > \varepsilon,$$

and thus

$$\text{Vol}_2((E_1)^y) \ll_M \frac{\varepsilon}{y} \cdot \ln^-\left(\frac{\varepsilon}{y}\right) \ll -\frac{2\varepsilon}{\gamma} \cdot \ln^-\left(\frac{2\varepsilon}{\gamma}\right) \tag{4.8}$$

for all $y \geq \gamma/2$. This verifies that the set E_1 is $(\varepsilon, \gamma/2, M + 1)$ -controlled.

Let

$$E_2 := \{(x, y) \in \Delta_\varepsilon^+ : u_1^-(\varepsilon) < x_1 \leq u_1^-\}.$$

Then for all y ,

$$\text{Vol}_2((E_1)^y) \leq M\varepsilon,$$

so that this set is also $(\varepsilon, \gamma/2, M + 1)$ -controlled. The set

$$E_3 := \{(x, y) \in \Delta_\varepsilon^+ : \gamma - M\varepsilon \leq y < \gamma\}.$$

is also obviously $(\varepsilon, \gamma/2, M + 1)$ -controlled of type II.

Finally, we observe that $\Delta_\varepsilon^+ \setminus \Delta = \Delta_\varepsilon^+ \cap \Delta^c$ is contained in a union of subsets of Δ_ε^+ , where each subset is defined by the negation of one of the inequalities that appears in the definition of Δ . Furthermore, in the case of $x_i, i = 1, 2$, we get two sets of inequalities $u_i^-(\varepsilon) < |x_i| \leq u_i^-$ and $u_i^+ < |x_i| \leq u_i^+(\varepsilon)$ which we view as four sets of inequalities in terms of x_i and consider the four corresponding subsets. This way we obtain twelve sets E_s . Since these sets are defined similarly to either E_1, E_2 , or E_3 , we can analyze them as above. Therefore, we conclude that all E_s 's are $(\varepsilon, \gamma/2, M + 1)$ -controlled. This verifies the claim of the lemma for the set $g^{-1}\Delta \setminus \Delta \subset \Delta_\varepsilon^+ \setminus \Delta$.

To handle the set $\Delta \setminus g^{-1}\Delta$, we observe that for $g \in V_\varepsilon$,

$$g^{-1}\Delta \supset \Delta_\varepsilon^-,$$

where

$$\Delta_\varepsilon^- := \left\{ (x, y) : \begin{array}{l} a + \varepsilon M^2 < |x_1 x_2| \cdot y \leq b - \varepsilon M^2 \\ u_i^- + \varepsilon M < |x_i| \leq u_i^+ - \varepsilon M \text{ for } i = 1, 2 \\ \gamma + \varepsilon M \leq y < \delta - \varepsilon M \end{array} \right\},$$

and $\Delta \setminus \Delta_\varepsilon^-$ is contained in the union of twelve $(\varepsilon, \gamma/2, M + 1)$ -controlled sets. This can be verified as above, so that we omit the details. □

5. Height function estimates

In this section we define a height function on the space $\widetilde{\mathcal{L}}_3$ consisting of all lattices in \mathbb{R}^3 and prove some technical level set estimates for this function that will be used later in our analysis of Siegel transforms.

Let $\widetilde{\mathcal{L}}_3$ denote the space of all lattices in \mathbb{R}^3 , endowed with the standard action of $\text{GL}_3(\mathbb{R})$. Given a bounded Borel function $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that the set $\{f \neq 0\}$ is bounded, we define the Siegel transform $\widehat{f} : \widetilde{\mathcal{L}}_3 \rightarrow \mathbb{R}$ by

$$\widehat{f}(\Lambda) := \sum_{\underline{\lambda} \in \Lambda \setminus \{0\}} f(\underline{\lambda}), \quad \text{for } \Lambda \in \widetilde{\mathcal{L}}_3.$$

Let $\mathcal{L}_3 \subset \widetilde{\mathcal{L}}_3$ denote the subspace of all unimodular lattices. This subspace is preserved by the restricted $\text{SL}_3(\mathbb{R})$ -action, and it is well-known that \mathcal{L}_3 carries a unique $\text{SL}_3(\mathbb{R})$ -invariant Borel probability mea-

sure, which we denote by μ . The following classical theorem of Siegel will play an important role in our analysis.

Theorem 5.1 (Siegel’s theorem). *Let f be a bounded Borel function with compact support. Then $\widehat{f} \in L^1(\mu)$ and*

$$\int_{\mathcal{L}_3} \widehat{f} d\mu = \int_{\mathbb{R}^3} f(\underline{z}) dz,$$

where the volume measure on \mathbb{R}^3 has been normalized so that $\text{Vol}_3([0, 1]^3) = 1$.

5.1. A height function on $\widetilde{\mathcal{L}}_3$

Let $\{e_1, e_2, e_3\}$ denote the standard (ordered) basis of \mathbb{R}^3 . We extend the max-norm with respect to this basis to the second exterior power $\mathbb{R}^3 \wedge \mathbb{R}^3$ as follows. If $\underline{u}, \underline{v} \in \mathbb{R}^3$ and

$$w = \underline{u} \wedge \underline{v} = w_{12} e_1 \wedge e_2 + w_{13} e_1 \wedge e_3 + w_{23} e_2 \wedge e_3,$$

then $\|w\|_\infty = \max(|w_{12}|, |w_{13}|, |w_{23}|)$.

Let Λ be a (not necessarily unimodular) lattice in \mathbb{R}^3 . We define the functions

$$\begin{aligned} s_1(\Lambda) &:= \inf\{\|\underline{\lambda}\|_\infty : \underline{\lambda} \in \Lambda \setminus \{0\}\}, \\ s_2(\Lambda) &:= \inf\{\|\underline{\lambda}_1 \wedge \underline{\lambda}_2\|_\infty : \underline{\lambda}_1 \wedge \underline{\lambda}_2 \neq 0, \underline{\lambda}_1, \underline{\lambda}_2 \in \Lambda \setminus \{0\}\}, \\ s_3(\Lambda) &:= \inf\{|\alpha| : \alpha e_1 \wedge e_2 \wedge e_3 = \underline{\lambda}_1 \wedge \underline{\lambda}_2 \wedge \underline{\lambda}_3 \neq 0, \underline{\lambda}_1, \underline{\lambda}_2, \underline{\lambda}_3 \in \Lambda \setminus \{0\}\}, \end{aligned}$$

and the height function

$$\text{ht}(\Lambda) := \min(s_1(\Lambda), s_2(\Lambda), s_3(\Lambda))^{-1}. \tag{5.1}$$

Since

$$\|g \cdot \underline{w}\|_\infty \leq \|g\|_{\text{op}} \cdot \|\underline{w}\|_\infty \quad \text{and} \quad \|g \cdot \underline{w}\|_\infty \geq \|g^{-1}\|_{\text{op}}^{-1} \cdot \|\underline{w}\|_\infty,$$

for all $g \in \text{GL}_3(\mathbb{R})$ and $\underline{w} \in \mathbb{R}^3$ (where $\|\cdot\|_{\text{op}}$ denotes the operator norm with respect to the max-norm on \mathbb{R}^3), we have

$$\|g\|_{\text{op}}^{-1} \cdot \text{ht}(\Lambda) \leq \text{ht}(g \cdot \Lambda) \leq \|g^{-1}\|_{\text{op}} \cdot \text{ht}(\Lambda), \tag{5.2}$$

for all $g \in \text{GL}_3(\mathbb{R})$ and $\Lambda \in \widetilde{\mathcal{L}}_3$. Before we proceed to the main topic of this section, we recall an important inequality due to Schmidt [18, Lemma 2].

Lemma 5.2. *For every bounded Borel function with compact support,*

$$|\widehat{f}(\Lambda)| \ll_{\text{supp}(f)} \|f\|_\infty \cdot \text{ht}(\Lambda), \quad \text{for all } \Lambda \in \widetilde{\mathcal{L}}_3,$$

where the implicit constants only depend on $\text{supp}(f)$.

5.2. Main results

We recall that

$$a(t) = \begin{pmatrix} e^{t_1} & & \\ & e^{t_2} & \\ & & e^{-(t_1+t_2)} \end{pmatrix}, \quad \text{for } t = (t_1, t_2) \in \mathbb{R}_+^2.$$

Given $\underline{x} = (x_1, x_2) \in \mathbb{R}^2$ and $r > 0$, we define the lattice

$$\Lambda_{\underline{x},r} := \left\{ \underline{p} + q(\iota(\underline{x}) + r\underline{e}_3) : \underline{p} \in \text{span}_{\mathbb{Z}}(\underline{e}_1, \underline{e}_2), q \in \mathbb{Z} \right\}, \tag{5.3}$$

where $\iota(\underline{x}) := (x_1, x_2, 0)$. We note that $\Lambda_{\underline{x},r}$ is a unimodular lattice if and only if $r = 1$.

Our first theorem in this section provides uniform upper bounds of the height function along $a(t)$ -orbits of lattices of the form $\Lambda_{\underline{x},r}$. Here and later in the paper, we use the notation

$$\lfloor t \rfloor := \min(t_1, t_2) \quad \text{for } t = (t_1, t_2) \in \mathbb{R}_+^2.$$

Theorem 5.3. *For all $r > 0$ and $t = (t_1, t_2) \in \mathbb{R}_+^2$,*

$$\sup \{ \text{ht}(a(t)\Lambda_{\underline{x},r}) : \underline{x} \in [0, 1)^2 \} \leq \max \left(e^{t_1+t_2}/r, e^{-\lfloor t \rfloor} \right).$$

Our second theorem roughly tells us that the map $t \mapsto \text{ht}(a(t)\Lambda_{\underline{x},r})$ is not large on a big volume set.

Theorem 5.4. *For every $r > 0$, $L > \max(1, 1/r)$ and $t \in \mathbb{R}_+^2$,*

$$\text{Vol}_2 \left(\{ \underline{x} \in [0, 1)^2 : \text{ht}(a(t)\Lambda_{\underline{x},r}) \geq L \} \right) \ll \max(r^{-1}, r^{-2})L^{-3} + r^{-1}L^{-2}e^{-\lfloor t \rfloor},$$

where the implicit constants are independent of L, r and t .

The theorems above provide useful upper bounds on certain integrals involving the height functions restricted to super-level sets. To state these bounds, we define the sets

$$M_{t,r}(\eta) := \{ \underline{x} \in [0, 1)^2 : \text{ht}(a(t)\Lambda_{\underline{x},r}) \geq \eta \} \tag{5.4}$$

for $\eta \geq 0$. We then have the following corollary, which we prove at the end of this section:

Corollary 5.5. *Let $\rho > 0$ and $\theta : [0, \infty) \rightarrow [0, \infty)$ be an increasing measurable function such that $u \mapsto \frac{\theta(u)}{u^2}$ is decreasing on $[1, \infty)$. Then, for every $t \in \mathbb{R}_+^2$, $r > \rho$ and $\eta \geq e^2\rho^{-1}$, we have*

$$\int_{M_{t,r}(\eta)} \theta(\text{ht}(a(t)\Lambda_{\underline{x},r})) \, d\underline{x} \ll \left(\max(r^{-1}, r^{-2})\eta^{-1} + r^{-1}e^{-\lfloor t \rfloor} \right) \cdot \int_{e^{-2}\eta}^{e^{t_1+t_2+1} \max(1, \rho^{-1})} \frac{\theta(u)}{u^3} \, du,$$

where the implicit constants are independent of ρ, η, r and t .

5.3. Lemmas about heights

Lemma 5.6 (Uniform Lower Bound for s_1). *For all $r > 0$ and $t = (t_1, t_2) \in \mathbb{R}_+^2$,*

$$\inf \{ s_1(a(t)\Lambda_{\underline{x},r}) : \underline{x} \in [0, 1)^2 \} \geq \min \left(e^{\lfloor t \rfloor}, r \cdot e^{-(t_1+t_2)} \right).$$

Proof. Let $\underline{x} = (x_1, x_2) \in [0, 1]^2$, $r > 0$, and $t = (t_1, t_2) \in \mathbb{R}_+^2$. We note that for every $\underline{p} = (p_1, p_2) \in \mathbb{Z}^2$ and $q \in \mathbb{Z}$, we have

$$\|a(t)(\underline{p} + q(\underline{u}(\underline{x}) + r\underline{e}_3))\|_\infty = \max\left(e^{t_1} \cdot |p_1 + qx_1|, e^{t_2} \cdot |p_2 + qx_2|, r \cdot |q| \cdot e^{-(t_1+t_2)}\right).$$

We assume that $\underline{p} + q(\underline{u}(\underline{x}) + r\underline{e}_3) \neq 0$. If $q = 0$, then $\underline{p} \neq 0$, and thus

$$\|a(t)(\underline{p} + q(\underline{u}(\underline{x}) + r\underline{e}_3))\|_\infty \geq e^{\lfloor t \rfloor}.$$

If $q \neq 0$, then

$$\|a(t)(\underline{p} + q(\underline{u}(\underline{x}) + r\underline{e}_3))\|_\infty \geq r \cdot e^{-(t_1+t_2)}.$$

We get a uniform lower bound by taking the minimum of these two bounds. □

Lemma 5.7 (Volume Bound for s_1). *For every $0 < \varepsilon < 1$, $r > 0$ and $t \in \mathbb{R}_+^2$, we have*

$$\text{Vol}_2\left(\{\underline{x} \in [0, 1]^2 : s_1(a(t)\Lambda_{\underline{x},r}) \leq \varepsilon\}\right) \ll \frac{\varepsilon^3}{r},$$

where the implicit constants are independent of ε, r and t .

Proof. Let $0 < \varepsilon < 1$ and $r > 0$ and $t = (t_1, t_2) \in \mathbb{R}_+^2$. Pick $\underline{x} = (x_1, x_2) \in [0, 1]^2$ such that

$$s_1(a(t)\Lambda_{\underline{x},r}) \leq \varepsilon.$$

This means that we can find a nonzero vector $\underline{p} + q(\underline{u}(\underline{x}) + r\underline{e}_3) \in \Lambda_{\underline{x},r}$ such that

$$\Lambda_{\underline{x}}(t) := e^{t_1}(p_1 + qx_1)\underline{e}_1 + e^{t_2}(p_2 + qx_2)\underline{e}_2 + re^{-(t_1+t_2)}q\underline{e}_3 \in a(t)\Lambda_{\underline{x},r}$$

satisfies $\|\Lambda_{\underline{x}}(t)\|_\infty \leq \varepsilon$. Writing this information coordinate-wise, we get the inequalities

$$|p_1 + qx_1| \leq \varepsilon \cdot e^{-t_1}, |p_2 + qx_2| \leq \varepsilon \cdot e^{-t_2}, |q| \leq \frac{\varepsilon \cdot e^{t_1+t_2}}{r}. \tag{5.5}$$

Since $\varepsilon < 1$ and $t_1, t_2 \geq 0$, there are no nonzero solutions for p when $q = 0$.

For $q \neq 0$, we note that since $\underline{x} = (x_1, x_2) \in [0, 1]^2$, there are at most $O(q^2)$ choices for $\underline{p} = (p_1, p_2) \in \mathbb{Z}^2$. Furthermore, for each such choice of \underline{p} ,

$$\text{Vol}_2\left(\{\underline{x} \in [0, 1]^2 : |p_1 + qx_1| \leq \varepsilon \cdot e^{-t_1}, |p_2 + qx_2| \leq \varepsilon \cdot e^{-t_2}\}\right) \ll \frac{\varepsilon^2 \cdot e^{-(t_1+t_2)}}{q^2},$$

where the implicit constants are independent of $\underline{p}, \varepsilon$ and t . In particular, upon summing over all possible \underline{p} and q , we see that the volume of the set of $\underline{x} \in [0, 1]^2$ for which the inequalities in (5.5) are satisfied is bounded from above by $O(\varepsilon^3/r)$, which finishes the proof. □

Let us introduce some notation which will be used in the proofs below. Given

$$r > 0, \underline{x} \in [0, 1]^2, \underline{p}^{(1)}, \underline{p}^{(2)} \in \text{span}_{\mathbb{Z}}(\underline{e}_1, \underline{e}_2), q_1, q_2 \in \mathbb{Z},$$

define $\underline{v}_1, \underline{v}_2 \in \Lambda_{\underline{x},r}$ by

$$\underline{v}_1 := \underline{p}^{(1)} + q_1(\underline{u}(\underline{x}) + r\underline{e}_3) \quad \text{and} \quad \underline{v}_2 := \underline{p}^{(2)} + q_2(\underline{u}(\underline{x}) + r\underline{e}_3). \tag{5.6}$$

Furthermore, given $t = (t_1, t_2) \in \mathbb{R}^2$, we define

$$\begin{aligned} \omega_{\underline{x}}(t) &:= a(t) \left(\left(\underline{p}^{(1)} + q_1(u(\underline{x}) + r\underline{e}_3) \right) \wedge \left(\underline{p}^{(2)} + q_2(u(\underline{x}) + r\underline{e}_3) \right) \right) \\ &= a(t) \left(\underline{p}^{(1)} \wedge \underline{p}^{(2)} + \left(q_2 \underline{p}^{(1)} - q_1 \underline{p}^{(2)} \right) \wedge \underline{x} + r \left(q_2 \underline{p}^{(1)} - q_1 \underline{p}^{(2)} \right) \wedge \underline{e}_3 \right). \end{aligned}$$

Note that

$$\omega_{\underline{x}}(t) = e^{t_1+t_2} \omega_{1,2}(\underline{x}) \underline{e}_1 \wedge \underline{e}_2 + r e^{-t_2} w_1 \underline{e}_1 \wedge \underline{e}_3 + r e^{-t_1} w_2 \underline{e}_2 \wedge \underline{e}_3,$$

where

$$\omega_{1,2}(\underline{x}) = m(\underline{p}) + (w_1 x_2 - w_2 x_1) \quad \text{and} \quad \underline{w} = w_1 \underline{e}_1 + w_2 \underline{e}_2 = q_2 \underline{p}^{(1)} - q_1 \underline{p}^{(2)},$$

and $m(\underline{p})$ is the unique integer such that $\underline{p}^{(1)} \wedge \underline{p}^{(2)} = m(\underline{p}) \underline{e}_1 \wedge \underline{e}_2$. In particular,

$$\|\omega_{\underline{x}}(t)\|_{\infty} = \max \left(e^{t_1+t_2} \cdot |m(\underline{p}) + (w_1 x_2 - w_2 x_1)|, e^{-t_2} \cdot r \cdot |w_1|, e^{-t_1} \cdot r \cdot |w_2| \right). \tag{5.7}$$

Note that if \underline{v}_1 and \underline{v}_2 are linearly independent, then $\omega_{\underline{x}}(t) \neq 0$, and thus either $\underline{w} \neq 0$ or $m(\underline{p}) \neq 0$.

Lemma 5.8 (Uniform Lower Bound for s_2). *For all $r > 0$ and $t = (t_1, t_2) \in \mathbb{R}_+^2$,*

$$\inf \{ s_2(a(t)\Lambda_{\underline{x},r}) : \underline{x} \in [0, 1]^2 \} \geq \min \left(r \cdot e^{-\lfloor t \rfloor}, e^{t_1+t_2} \right).$$

Proof. Note that $s_2(a(t)\Lambda_{\underline{x},r})$ is the minimum of $\|\omega_{\underline{x}}(t)\|_{\infty}$, when \underline{v}_1 and \underline{v}_2 , defined as in (5.6), vary over all linearly independent pairs of vectors in $\Lambda_{\underline{x},r}$. Hence, by (5.7), we need lower bounds on

$$\max \left(e^{t_1+t_2} \cdot |m(\underline{p}) + (w_1 x_2 - w_2 x_1)|, e^{-t_2} \cdot r \cdot |w_1|, e^{-t_1} \cdot r \cdot |w_2| \right),$$

when either $\underline{w} \neq 0$ or $m(\underline{p}) \neq 0$. If $\underline{w} \neq 0$, then since w_1 and w_2 are integers, we have

$$\max \left(e^{t_1+t_2} \cdot |m(\underline{p}) + (w_1 x_2 - w_2 x_1)|, e^{-t_2} \cdot r \cdot |w_1|, e^{-t_1} \cdot r \cdot |w_2| \right) \geq r \cdot e^{-\lfloor t \rfloor},$$

and if $m(\underline{p}) \neq 0$ and $\underline{w} = 0$, then

$$\max \left(e^{t_1+t_2} \cdot |m(\underline{p}) + (w_1 x_2 - w_2 x_1)|, e^{-t_2} \cdot r \cdot |w_1|, e^{-t_1} \cdot r \cdot |w_2| \right) \geq e^{t_1+t_2}.$$

We get a uniform lower bound by taking the minimum of these two bounds. □

Lemma 5.9 (Volume Bound for s_2). *For every $0 < \varepsilon < 1$, $r > 0$ and $t \in \mathbb{R}_+^2$, we have*

$$\text{Vol}_2 \left(\{ \underline{x} \in [0, 1]^2 : s_2(a(t)\Lambda_{\underline{x},r}) \leq \varepsilon \} \right) \ll \frac{\varepsilon^3}{r^2} + \frac{\varepsilon^2 \cdot e^{-\lfloor t \rfloor}}{r},$$

where the implicit constants are independent of ε, r and t .

Proof. Let $0 < \varepsilon < 1$, $r > 0$ and $t = (t_1, t_2) \in \mathbb{R}_+^2$. Pick $\underline{x} \in [0, 1]^2$ such that

$$s_2(a(t)\Lambda_{\underline{x},r}) \leq \varepsilon.$$

This means that we can find two linearly independent vectors \underline{v}_1 and \underline{v}_2 as in (5.6) such that $\|\omega_{\underline{x}}(t)\|_\infty \leq \varepsilon$. By (5.7), this results in the bounds

$$|m(\underline{p}) + (w_1x_2 - w_2x_1)| \leq \varepsilon \cdot e^{-(t_1+t_2)}, |w_1| \leq \frac{\varepsilon \cdot e^{t_2}}{r}, |w_2| \leq \frac{\varepsilon \cdot e^{t_1}}{r}. \tag{5.8}$$

We make two observations:

- If $q_1 = q_2 = 0$, then $\underline{w} = 0$ and $|m(\underline{p})| \leq \varepsilon \cdot e^{-(t_1+t_2)}$. Since $\varepsilon < 1$ and $t_1, t_2 \geq 0$ and $m(\underline{p})$ is an integer, we must have $m(\underline{p}) = 0$. This readily implies that $\underline{p}^{(1)} \wedge \underline{p}^{(2)} = 0$, so $\underline{p}^{(1)}$ and $\underline{p}^{(2)}$ are linearly dependent, and thus \underline{v}_1 and \underline{v}_2 are linearly dependent as well, contrary to our assumption.
- If $\underline{w} = 0$ and $(q_1, q_2) \neq (0, 0)$, then $q_2v_1 - q_1v_2 = 0$, which contradicts our assumption that \underline{v}_1 and \underline{v}_2 are linearly independent.

We can thus without loss of generality assume that $\underline{w} \neq 0$ and $(q_1, q_2) \neq (0, 0)$.

For a fixed $\underline{w} \neq 0$ and $m(\underline{p}) \in \mathbb{Z}$, we have

$$\text{Vol}_2\left(\left\{\underline{x} \in [0, 1]^2 : |m(\underline{p}) + (w_1x_2 - w_2x_1)| \leq \varepsilon \cdot e^{-(t_1+t_2)}\right\}\right) \ll \frac{\varepsilon \cdot e^{-(t_1+t_2)}}{\max(|w_1|, |w_2|)},$$

where the implicit constants are independent of \underline{p} and t . Furthermore, since $0 < \varepsilon < 1$ and $t_1, t_2 \geq 0$, it follows from (5.8) that there are at most $O(\max(|w_1|, |w_2|))$ choices for $m(\underline{p})$. We also note from (5.8) that there are

- $O\left(\frac{\varepsilon^2 \cdot e^{t_1+t_2}}{r^2}\right)$ choices for \underline{w} with $w_1, w_2 \neq 0$,
- $O\left(\frac{\varepsilon \cdot e^{t_2}}{r}\right)$ choices for \underline{w} with $w_2 = 0$,
- $O\left(\frac{\varepsilon \cdot e^{t_1}}{r}\right)$ choices for \underline{w} with $w_1 = 0$.

Summing over all of these choices, we get

$$\text{Vol}_2\left(\left\{\underline{x} \in [0, 1]^2 : s_2(a(t)\Lambda_{\underline{x},r}) \leq \varepsilon\right\}\right) \ll \frac{\varepsilon^3}{r^2} + \frac{\varepsilon^2 \cdot e^{-t_1}}{r} + \frac{\varepsilon^2 \cdot e^{-t_2}}{r},$$

which implies the claim. □

Finally, we note the following elementary result for s_3 , whose proof is left to the reader.

Lemma 5.10. *For all $t \in \mathbb{R}_+^2$ and $r > 0$,*

$$s_3(a(t)\Lambda_{\underline{x},r}) = r, \quad \text{for all } \underline{x} \in [0, 1]^2.$$

5.4. Proof of Theorem 5.3

By Lemma 5.6, Lemma 5.8 and Lemma 5.10, we have

$$\begin{aligned} \min(s_1(a(t)\Lambda_{\underline{x},r}), s_2(a(t)\Lambda_{\underline{x},r}), s_3(a(t)\Lambda_{\underline{x},r})) &\geq \min\left(e^{\lfloor t \rfloor}, r \cdot e^{-(t_1+t_2)}, r \cdot e^{-\lfloor t \rfloor}, e^{t_1+t_2}, r\right) \\ &\geq \min\left(r \cdot e^{-(t_1+t_2)}, e^{\lfloor t \rfloor}\right), \end{aligned}$$

for all $t = (t_1, t_2) \in \mathbb{R}_+^2$. Hence,

$$\text{ht}(a(t)\Lambda_{\underline{x},r}) \leq \max\left(e^{t_1+t_2}/r, e^{-\lfloor t \rfloor}\right),$$

for all $\underline{x} \in [0, 1]^2$.

5.5. Proof of Theorem 5.4

Let $L > 1$. We first note that

$$\begin{aligned} \text{Vol}_2\left(\{\underline{x} \in [0, 1]^2 : \text{ht}(a(t)\Lambda_{\underline{x},r}) \geq L\}\right) &= \text{Vol}_2\left(\bigcup_{i=1}^3 \left\{\underline{x} \in [0, 1]^2 : s_i(a(t)\Lambda_{\underline{x},r}) \leq \frac{1}{L}\right\}\right) \\ &\leq \sum_{i=1}^3 \text{Vol}_2\left(\left\{\underline{x} \in [0, 1]^2 : s_i(a(t)\Lambda_{\underline{x},r}) \leq \frac{1}{L}\right\}\right). \end{aligned}$$

By Lemma 5.7 and Lemma 5.9 (applied with $\varepsilon = \frac{1}{L}$), we have

$$\text{Vol}_2\left(\left\{\underline{x} \in [0, 1]^2 : s_1(a(t)\Lambda_{\underline{x},r}) \leq \frac{1}{L}\right\}\right) \ll \frac{1}{rL^3},$$

and

$$\text{Vol}_2\left(\left\{\underline{x} \in [0, 1]^2 : s_2(a(t)\Lambda_{\underline{x},r}) \leq \frac{1}{L}\right\}\right) \ll \frac{1}{r^2L^3} + \frac{e^{-\lfloor t \rfloor}}{rL^2}.$$

Furthermore, by Lemma 5.10, the last set in the sum is empty if $L > \frac{1}{r}$. Combining these estimates, we obtain the theorem.

5.6. Proof of Corollary 5.5

Let $r > \rho$ and $t \in \mathbb{R}_+^2$. We introduce the sets

$$B_{t,r}(i) = \{\underline{x} \in [0, 1]^2 : e^{i-1} < \text{ht}(a(t)\Lambda_{\underline{x},r}) \leq e^i\}, \quad \text{for } i \in \mathbb{Z}.$$

By Theorem 5.3,

$$\text{ht}(a(t)\Lambda_{\underline{x},r}) \leq \max\left(\frac{e^{t_1+t_2}}{r}, e^{-\lfloor t \rfloor}\right) \leq \max(1, \rho^{-1})e^{t_1+t_2}, \quad \text{for all } \underline{x} \in [0, 1]^2,$$

and thus $B_{t,r}(i)$ is empty for $i \geq t_1 + t_2 + \ln \max(1, \rho^{-1}) + 1$. Furthermore, by Theorem 5.4, applied with $L = e^{i-1}$ for $i \geq \ln(\rho^{-1}) + 1$, we have

$$\text{Vol}_2(B_{t,r}(i)) \ll \max(r^{-1}, r^{-2})e^{-3i} + r^{-1}e^{-\lfloor t \rfloor}e^{-2i}.$$

Hence, since θ is increasing, for $\eta \geq e^2\rho^{-1}$,

$$\begin{aligned} \int_{M_{t,r}(\eta)} \theta(\text{ht}(a(t)\Lambda_{\underline{x},r})) d\underline{x} &\leq \sum_{i=\lfloor \ln(\eta) \rfloor}^{\lceil t_1+t_2+\ln \max(1, \rho^{-1}) \rceil} \theta(e^i) \cdot \text{Vol}_2(B_{t,r}(i)) \\ &\ll \max(r^{-1}, r^{-2}) \left(\sum_{i=\lfloor \ln(\eta) \rfloor}^{\lceil t_1+t_2+\ln \max(1, \rho^{-1}) \rceil} \theta(e^i) \cdot e^{-3i} \right) \\ &\quad + r^{-1} \cdot \left(\sum_{i=\lfloor \ln(\eta) \rfloor}^{\lceil t_1+t_2+\ln \max(1, \rho^{-1}) \rceil} \theta(e^i) \cdot e^{-2i} \right) \cdot e^{-\lfloor t \rfloor}. \end{aligned}$$

By assumption, $u \mapsto \frac{\theta(u)}{u^2}$ is decreasing for $u \geq 1$, and thus

$$\sum_{i=\lfloor \ln(\eta) \rfloor}^{\lceil t_1+t_2+\ln \max(1,\rho^{-1}) \rceil} \theta(e^i) \cdot e^{-2i} \leq \int_{\lfloor \ln(\eta) \rfloor - 1}^{t_1+t_2+\ln \max(1,\rho^{-1})+1} \frac{\theta(e^u)}{e^{2u}} du \leq \int_{e^{-2\eta}}^{e^{t_1+t_2+1} \max(1,\rho^{-1})} \frac{\theta(u)}{u^3} du.$$

Also, in the summation range, we have

$$\sum_{i=\lfloor \ln(\eta) \rfloor}^{\lceil t_1+t_2+\ln \max(1,\rho^{-1}) \rceil} \theta(e^i) \cdot e^{-3i} \ll \eta^{-1} \cdot \int_{e^{-2\eta}}^{e^{t_1+t_2+1} \max(1,\rho^{-1})} \frac{\theta(u)}{u^3} du.$$

We conclude that

$$\int_{M_{t,r}(\eta)} \theta(\text{ht}(a(t)\Lambda_{\underline{x},r})) d\underline{x} \ll \left(\max(r^{-1}, r^{-2})\eta^{-1} + r^{-1}e^{-\lfloor t \rfloor} \right) \cdot \int_{e^{-2\eta}}^{e^{t_1+t_2+1} \max(1,\rho^{-1})} \frac{\theta(u)}{u^3} du,$$

which finishes the proof.

6. Correlations between the number of shifted lattice points in boxes

If $B \subset \mathbb{R}^2$ is a bounded Borel set, we define the counting function

$$N_B(\underline{x}) = \left| \left(\mathbb{Z}^2 + \underline{x} \right) \cap B \right|, \quad \text{for } \underline{x} \in \mathbb{R}^2. \tag{6.1}$$

Note that N_B is \mathbb{Z}^2 -periodic, and thus completely determined by its values on $[0, 1)^2$.

Our main result in this section reads as follows.

Lemma 6.1. *Let $M > 0$ and let $D^{(1)}$ and $D^{(2)}$ be Borel subsets of the square $[-M, M]^2$. Let $t = (t_1, t_2) \in \mathbb{R}_+^2$ with $t_1 \leq t_2$ and define*

$$D_t^{(i)} = \begin{pmatrix} e^{-t_1} & 0 \\ 0 & e^{-t_2} \end{pmatrix} D^{(i)} \quad \text{for } i = 1, 2.$$

Then, for all $(q_1, q_2) \in \mathbb{N}^2$,

$$\int_{[0,1]^2} N_{D_t^{(1)}}(q_1 \underline{x}) N_{D_t^{(2)}}(q_2 \underline{x}) d\underline{x} \leq F_t \left(\frac{\max(q_1, q_2)}{\gcd(q_1, q_2)} \right) \cdot \max \left(\text{Vol}_2(D^{(1)}), \text{Vol}_2(D^{(2)}) \right),$$

where F_t is defined as in (B.3).

We will derive Lemma 6.1 from the following general result.

Lemma 6.2. *Let B_1 and B_2 be bounded Borel sets in \mathbb{R}^2 and let q_1 and q_2 be relatively prime positive integers. Then,*

$$\int_{[0,1]^2} N_{B_1}(q_1 \underline{x}) N_{B_2}(q_2 \underline{x}) d\underline{x} \leq \left| \mathbb{Z}^2 \cap (q_2 B_1 - q_1 B_2) \right| \cdot \min \left(\frac{\text{Vol}_2(B_1)}{q_1^2}, \frac{\text{Vol}_2(B_2)}{q_2^2} \right).$$

Proof. Note that $\mathbb{Z}^2 \times \mathbb{Z}^2 = \bigsqcup_{\underline{k} \in \mathbb{Z}^2} E_{\underline{k}}(q_1, q_2)$, where

$$E_{\underline{k}}(q_1, q_2) := \{ (\underline{p}_1, \underline{p}_2) \in \mathbb{Z}^2 \times \mathbb{Z}^2 : q_2 \underline{p}_1 - q_1 \underline{p}_2 = \underline{k} \}, \quad \text{for } \underline{k} \in \mathbb{Z}^2.$$

Hence,

$$\begin{aligned} \int_{[0,1]^2} N_{B_1}(q_1 \underline{x}) N_{B_2}(q_2 \underline{x}) d\underline{x} &= \sum_{\underline{p}_1, \underline{p}_2} \int_{[0,1]^2} \chi_{B_1}(\underline{p}_1 + q_1 \underline{x}) \chi_{B_2}(\underline{p}_2 + q_2 \underline{x}) d\underline{x} \\ &= \sum_{\underline{k}} C_{\underline{k}}(q_1, q_2), \end{aligned}$$

where $C_{\underline{k}}(q_1, q_2) := \sum_{(\underline{p}_1, \underline{p}_2) \in E_{\underline{k}}(q_1, q_2)} \int_{[0,1]^2} \chi_{B_1}(\underline{p}_1 + q_1 \underline{x}) \chi_{B_2}(\underline{p}_2 + q_2 \underline{x}) d\underline{x}$.

Fix $\underline{k} \in \mathbb{Z}^2$ and $(\underline{p}'_1, \underline{p}'_2) \in E_{\underline{k}}(q_1, q_2)$. Then,

$$E_{\underline{k}}(q_1, q_2) = \left\{ (\underline{p}'_1 + q_1 \underline{l}, \underline{p}'_2 + q_2 \underline{l}) : \underline{l} \in \mathbb{Z}^2 \right\}. \tag{6.2}$$

Indeed, if $(\underline{p}_1, \underline{p}_2)$ is any point in $E_{\underline{k}}(q_1, q_2)$, then $q_2(\underline{p}_1 - \underline{p}'_1) = q_1(\underline{p}_2 - \underline{p}'_2)$. Since q_1 and q_2 are relatively prime integers, we must have $\underline{p}_1 - \underline{p}'_1 = q_1 \underline{l}$ and $\underline{p}_2 - \underline{p}'_2 = q_2 \underline{l}$ for some (unique) $\underline{l} \in \mathbb{Z}^2$, thus proving (6.2).

Hence, for a fixed choice of $(\underline{p}'_1, \underline{p}'_2) \in E_{\underline{k}}(q_1, q_2)$, the identity (6.2) allows us to rewrite the term $C_{\underline{k}}(q_1, q_2)$ as follows:

$$\begin{aligned} C_{\underline{k}}(q_1, q_2) &= \sum_{\underline{l}} \int_{[0,1]^2} \chi_{B_1}(\underline{p}'_1 + q_1(\underline{l} + \underline{x})) \chi_{B_2}(\underline{p}'_2 + q_2(\underline{l} + \underline{x})) d\underline{x} \\ &= \int_{\mathbb{R}^2} \chi_{B_1}(\underline{p}'_1 + q_1 \underline{x}) \chi_{B_2}(\underline{p}'_2 + q_2 \underline{x}) d\underline{x} \\ &= \int_{\mathbb{R}^2} \chi_{B_1}(\underline{p}'_1 - q_1 \underline{p}'_2 / q_2 + q_1 \underline{x}) \chi_{B_2}(q_2 \underline{x}) d\underline{x} \\ &= \int_{\mathbb{R}^2} \chi_{B_1}(\underline{k} / q_2 + q_1 \underline{x}) \chi_{B_2}(q_2 \underline{x}) d\underline{x}, \end{aligned}$$

where, in the last step, we have used the fact that $q_2 \underline{p}'_1 - q_1 \underline{p}'_2 = \underline{k}$. We conclude that

$$\begin{aligned} C_{\underline{k}}(q_1, q_2) &= \text{Vol}_2 \left(\left(\frac{1}{q_1} B_1 - \frac{\underline{k}}{q_1 q_2} \right) \cap \frac{1}{q_2} B_2 \right) \\ &= \left(\frac{1}{q_1 q_2} \right)^2 \cdot \text{Vol}_2((q_2 B_1 - \underline{k}) \cap q_1 B_2). \end{aligned}$$

In particular,

$$\begin{aligned} C_{\underline{k}}(q_1, q_2) &\leq \left(\frac{1}{q_1 q_2} \right)^2 \cdot \min(\text{Vol}_2(q_2 B_1), \text{Vol}_2(q_1 B_2)) \\ &= \min \left(\frac{\text{Vol}_2(B_1)}{q_1^2}, \frac{\text{Vol}_2(B_2)}{q_2^2} \right) \quad \text{for all } \underline{k} \in \mathbb{Z}^2, \end{aligned}$$

and

$$C_{\underline{k}}(q_1, q_2) = 0, \quad \text{for all } \underline{k} \notin q_2 B_1 - q_1 B_2.$$

Hence,

$$\sum_{k \in \mathbb{Z}^2} C_k(q_1, q_2) \leq |\mathbb{Z}^2 \cap (q_2 B_1 - q_1 B_2)| \cdot \min\left(\frac{\text{Vol}_2(B_1)}{q_1^2}, \frac{\text{Vol}_2(B_2)}{q_2^2}\right),$$

which finishes the proof. □

Proof of Lemma 6.1. Fix $(q_1, q_2) \in \mathbb{N}^2$ and write $q_1 = sq'_1$ and $q_2 = sq'_2$, where $s = \text{gcd}(q_1, q_2)$ and q'_1 and q'_2 are relatively prime. Since multiplication by positive integers on the torus $\mathbb{R}^2/\mathbb{Z}^2$ preserves the Lebesgue measure, we have

$$\int_{[0,1)^2} N_{D_t^{(1)}}(q_1 \underline{x}) N_{D_t^{(2)}}(q_2 \underline{x}) \, d\underline{x} = \int_{[0,1)^2} N_{D_t^{(1)}}(q'_1 \underline{x}) N_{D_t^{(2)}}(q'_2 \underline{x}) \, d\underline{x}.$$

If we apply Lemma 6.2 to the right-hand side with

$$B_1 = D_t^{(1)} \quad \text{and} \quad B_2 = D_t^{(2)},$$

and note that $q'_2 D_t^{(1)} - q'_1 D_t^{(2)} \subset [-M_1, M_1] \times [-M_2, M_2]$, where

$$M_i = 2 \max(q'_1, q'_2) \cdot M \cdot e^{-t_i} \quad \text{for } i = 1, 2.$$

we get

$$\begin{aligned} \int_{[0,1)^2} N_{D_t^{(1)}}(q'_1 \underline{x}) N_{D_t^{(2)}}(q'_2 \underline{x}) \, d\underline{x} &\leq |\mathbb{Z}^2 \cap (q'_2 D_t^{(1)} - q'_1 D_t^{(2)})| \cdot \min\left(\frac{\text{Vol}_2(D_t^{(1)})}{(q'_1)^2}, \frac{\text{Vol}_2(D_t^{(2)})}{(q'_2)^2}\right) \\ &\leq \frac{G(M_1, M_2)}{\max(q'_1, q'_2)^2} \cdot e^{-(t_1+t_2)} \cdot \max(\text{Vol}_2(D^{(1)}), \text{Vol}_2(D^{(2)})), \end{aligned}$$

where G is defined as in (B.2). Hence,

$$\int_{[0,1)^2} N_{D_t^{(1)}}(q'_1 \underline{x}) N_{D_t^{(2)}}(q'_2 \underline{x}) \, d\underline{x} \leq F_t(\max(q'_1, q'_2)) \cdot \max(\text{Vol}_2(D^{(1)}), \text{Vol}_2(D^{(2)})),$$

where F_t is defined as in (B.3). This finishes the proof. □

7. Mean counting within controlled sets

In this section we prove L^2 -bounds for Siegel transforms of indicator functions of controlled sets. These bounds will be useful later in Section 8 when we analyze smooth approximations of counting functions.

Lemma 7.1. *Let $M > 1$ and $0 < \varepsilon < 3\varepsilon < \gamma < 1$. We suppose that $E \subset \mathbb{R}^2 \times \mathbb{R}$ is an (ε, γ, M) -controlled set. Then, for all $t = (t_1, t_2) \in \mathbb{R}_+^2$ such that*

$$t_1 + t_2 > \max(1, -\ln(\gamma/2)),$$

we have

$$\int_{[0,1)^2} \widehat{\chi}_E(a(t)\Lambda_{\underline{x}})^2 \, d\underline{x} \ll_M e^{-(t_1+t_2)} + \max\left(\varepsilon, -\frac{\varepsilon}{\gamma} \ln\left(\frac{\varepsilon}{\gamma}\right)\right) \cdot \max(1, (t_1 + t_2))^2,$$

where the implicit constants depend only on M .

Proof. Let $E \subset \mathbb{R}^2 \times \mathbb{R}$ be a bounded Borel set, let $E(t) = a(t)^{-1}E$, and note that

$$\begin{aligned} \int_{[0,1]^2} \widehat{\chi}_E(a(t)\Lambda_{\underline{x}})^2 d\underline{x} &\leq \sum_{q_1, q_2 \in \mathbb{Z}} \sum_{p_1, p_2 \in \mathbb{Z}^2} \int_{[0,1]^2} \chi_{E(t)^{q_1}}(p_1 + q_1\underline{x}) \chi_{E(t)^{q_2}}(p_2 + q_2\underline{x}) d\underline{x} \\ &= \sum_{q_1, q_2 \in \mathbb{Z}} \int_{[0,1]^2} N_{D_{q_1}}(q_1\underline{x}) N_{D_{q_2}}(q_2\underline{x}) d\underline{x}, \end{aligned}$$

where N_{\bullet} is defined as in (6.1) and $D_q = E(t)^q$ for $q \in \mathbb{Z}$. We define

$$J := \{y \in \mathbb{R} : E^y \neq \emptyset\},$$

and note that for every $q \in \mathbb{Z}$,

$$E(t)^q = \begin{pmatrix} e^{-t_1} & \\ & e^{-t_2} \end{pmatrix} E^{q(t)},$$

where $q(t) = e^{-(t_1+t_2)}q$. In particular,

$$E(t)^q \neq \emptyset \iff q \in J_t := e^{t_1+t_2}J.$$

Hence,

$$\int_{[0,1]^2} \widehat{\chi}_E(a(t)\Lambda_{\underline{x}})^2 d\underline{x} \ll \sum_{q_1, q_2 \in J_t} \int_{[0,1]^2} N_{D_t^{(1)}}(q_1\underline{x}) N_{D_t^{(2)}}(q_2\underline{x}) d\underline{x},$$

where $D^{(i)} = E^{q_i(t)}$ for $i = 1, 2$, and $D_t^{(i)}$ is defined as in Lemma 6.1. The same lemma now tells us that

$$\begin{aligned} \int_{[0,1]^2} \widehat{\chi}_E(a(t)\Lambda_{\underline{x}})^2 d\underline{x} &\ll \sum_{q_1, q_2 \in J_t} F_t\left(\frac{\max(q_1, q_2)}{\gcd(q_1, q_2)}\right) \cdot \max(\text{Vol}_2(E^{q_1(t)}), \text{Vol}_2(E^{q_2(t)})) \\ &\ll \left(\sum_{q_1, q_2 \in J_t} F_t\left(\frac{\max(q_1, q_2)}{\gcd(q_1, q_2)}\right) \right) \cdot \sup_{y \in J} \text{Vol}_2(E^y), \end{aligned}$$

where F_t is defined as in (B.3).

The arguments up to this point have not made use of any special properties of E . In what follows, we will fix $0 < \varepsilon < 3\varepsilon < \gamma$ and assume that E is an (ε, γ, M) -controlled set (see Definition 4.1). The analysis will depend on whether E is type I or type II.

Let us first assume that E is type I. Then,

$$J \subset (\gamma, M] \quad \text{and} \quad \sup_{y \in J} \text{Vol}_2(E^y) \ll_M \max\left(\varepsilon, -\frac{\varepsilon}{\gamma} \ln\left(\frac{\varepsilon}{\gamma}\right)\right).$$

Furthermore, by Lemma B.1 (with $\alpha = \gamma$ and $\beta = M$),

$$\sum_{q_1, q_2 \in J_t} F_t\left(\frac{\max(q_1, q_2)}{\gcd(q_1, q_2)}\right) \ll_M e^{-(t_1+t_2)} + \max(1, -\ln(\gamma)) \cdot \max(1, t_1 + t_2),$$

provided that

$$t_1 + t_2 > \max(1, -\ln(\gamma)). \tag{7.1}$$

We conclude that if the conditions (7.1) hold, then

$$\int_{[0,1]^2} \widehat{\chi}_E(a(t)\Lambda_{\underline{x}})^2 d\underline{x} \ll (t_1 + t_2)^2 \cdot \max\left(\varepsilon, -\frac{\varepsilon}{\gamma} \ln\left(\frac{\varepsilon}{\gamma}\right)\right). \tag{7.2}$$

Let us now assume that E is type II. Then,

$$E \subset [-M, M]^2 \times [\alpha, \beta],$$

where $\frac{\gamma}{2} \leq \alpha$, and $\beta - \alpha \ll \varepsilon$. In particular,

$$J \subset [\alpha, \beta] \quad \text{and} \quad \sup_{y \in J} \text{Vol}_2(E^y) \ll M^2.$$

By Lemma B.1,

$$\begin{aligned} \sum_{q_1, q_2 \in J_t} F_t\left(\frac{\max(q_1, q_2)}{\gcd(q_1, q_2)}\right) &\ll_M e^{-(t_1+t_2)} + (\beta - \alpha) \cdot \max\left(1, \ln\left(\frac{\beta}{\alpha}\right)\right) \cdot \max(1, t_1 + t_2) \\ &\ll_M e^{-(t_1+t_2)} + \varepsilon \cdot \max\left(1, \ln\left(\frac{\beta}{\alpha}\right)\right) \cdot \max(1, (t_1 + t_2)), \end{aligned}$$

provided that

$$t_1 + t_2 > -\ln(\alpha). \tag{7.3}$$

Note that

$$\max\left(1, \ln\left(\frac{\beta}{\alpha}\right)\right) \leq \max\left(1, \frac{\beta - \alpha}{\alpha}\right).$$

If $\beta \leq 2\alpha$, the right-hand side is bounded from above by 1. Otherwise, the right-hand side is bounded from above by

$$\frac{\beta - \alpha}{\alpha} \ll \frac{\varepsilon}{\alpha} \ll \frac{\varepsilon}{\gamma},$$

since $\frac{\gamma}{2} \leq \alpha$. Hence,

$$\int_{[0,1]^2} \widehat{\chi}_E(a(t)\Lambda_{\underline{x}})^2 d\underline{x} \ll e^{-(t_1+t_2)} + \max\left(\varepsilon, \frac{\varepsilon^2}{\gamma}\right) \cdot \max(1, (t_1 + t_2)), \tag{7.4}$$

provided that (7.3) hold.

Now we combine (7.2) and (7.4). Since $3\varepsilon < \gamma$, we have

$$\frac{\varepsilon^2}{\gamma} \leq -\frac{\varepsilon}{\gamma} \ln\left(\frac{\varepsilon}{\gamma}\right).$$

Hence, if we combine (7.2) and (7.4), we get the uniform estimate (independent of whether E is type I or type II):

$$\int_{[0,1]^2} \widehat{\chi}_E(a(t)\Lambda_{\underline{x}})^2 d\underline{x} \ll e^{-(t_1+t_2)} + \max\left(\varepsilon, -\frac{\varepsilon}{\gamma} \ln\left(\frac{\varepsilon}{\gamma}\right)\right) \cdot \max(1, (t_1 + t_2))^2,$$

provided that (7.1) and (7.3) both hold. Note that since $\frac{\gamma}{2} \leq \alpha$, the second conditions in (7.1) and (7.3) are both satisfied if

$$t_1 + t_2 > \max(1, -\ln(\gamma/2)).$$

This finishes the proof. □

8. Smooth approximations

Let \mathcal{L}_3 denote the space of unimodular lattices in \mathbb{R}^3 . We can identify \mathcal{L}_3 with the homogeneous space $\mathrm{SL}_3(\mathbb{R})/\mathrm{SL}_3(\mathbb{Z})$ via the map $g \mathrm{SL}_3(\mathbb{Z}) \mapsto g\mathbb{Z}^3$. Fix a basis $\{Y_1, \dots, Y_8\}$ of the Lie algebra $\mathfrak{sl}_3(\mathbb{R})$. We adopt the following slight abuse of notation: for every $i = 1, \dots, 8$, let D_i both denote the differential operator

$$D_i \rho(g) = \frac{d}{dt} \rho(\exp(tY_i)g)|_{t=0}$$

on $C_b^\infty(\mathrm{SL}_3(\mathbb{R}))$ and the differential operator

$$D_i \varphi = \frac{d}{dt} \rho(\exp(tY_i)\Lambda)|_{t=0}$$

on $C_b^\infty(\mathcal{L}_3)$ (which we identify with the space of bounded and smooth right Γ -invariant functions on the group $\mathrm{SL}_3(\mathbb{R})$ via the map above). Differential operators can clearly be composed, and any composition of the D_1, \dots, D_8 can be rewritten as a linear combination of compositions of the form $D_m := D_1^{m_1} \circ \dots \circ D_8^{m_8}$ for some vector $m = (m_1, \dots, m_8) \in \mathbb{N}_0^8$ (with the convention that $D_0 = \mathrm{id}$). We define the norms

$$\|\rho\|_{C_b^s(\mathrm{SL}_3(\mathbb{R}))} := \max\{\|D_m \rho\|_\infty : m_1 + \dots + m_8 \leq s\}, \quad \text{for } \rho \in C_b^\infty(\mathrm{SL}_3(\mathbb{R}))$$

and

$$\|\varphi\|_{C^s(\mathcal{L}_3)} := \max\{\|D_m \varphi\|_\infty : m_1 + \dots + m_8 \leq s\}, \quad \text{for } \varphi \in C_b^\infty(\mathcal{L}_3).$$

Fix a right-invariant Riemannian metric on $\mathrm{SL}_3(\mathbb{R})$, and denote by Lip the corresponding Lipschitz semi-norm on \mathcal{L}_3 (viewed as the right quotient space $\mathrm{SL}_3(\mathbb{R})/\mathrm{SL}_3(\mathbb{Z})$). We define

$$\mathcal{N}_s(\varphi) := \max(\|\varphi\|_{C_b^s(\mathcal{L}_3)}, \mathrm{Lip}(\varphi)), \quad \text{for } \varphi \in C^\infty(\mathcal{L}_3). \tag{8.1}$$

For $\varepsilon > 0$, let V_ε denote the symmetric open neighborhoods around the identity in $\mathrm{SL}_3(\mathbb{R})$ defined in (4.3). For the rest of this paper, we fix a non-negative smooth function ρ_ε on $\mathrm{SL}_3(\mathbb{R})$ whose support is contained in V_ε and has integral one with respect to the Haar measure on $\mathrm{SL}_3(\mathbb{R})$. We leave it to the reader to verify that ρ_ε can be chosen so that for every integer $s \geq 1$, there is an integer $\sigma_s > 0$ such that

$$\|\rho_\varepsilon\|_{C_b^s(\mathrm{SL}_3(\mathbb{R}))} \ll \varepsilon^{-\sigma_s}, \quad \text{for all } \varepsilon \in (0, 1) \tag{8.2}$$

and where the implicit constants do not depend on ε .

By Lemma [1, Lemma 4.11], for every $L > 1$ there exists a smooth function $\eta_L : \mathcal{L}_3 \rightarrow [0, 1]$ such that

$$\{\mathrm{ht} \leq L/2\} \subset \{\eta_L = 1\} \subset \mathrm{supp}(\eta_L) \subset \{\mathrm{ht} \leq 2L\}, \tag{8.3}$$

where ht is the height function on \mathcal{L}_3 defined in (5.1), with the property that for every $s \geq 1$ and $m \in \mathbb{N}_o^8$ such that $m_1 + \dots + m_8 \leq s$, we have

$$D_m \eta_L \ll_s 1, \quad \text{for all } L \geq 1, \tag{8.4}$$

where the implicit constants only depend on s , but not on L .

If F is a locally bounded Borel function on \mathbb{R}^3 and $\rho \in C_c^\infty(G)$, we denote by $\rho * F$ the (action) convolution

$$(\rho * F)(u) := \int_G \rho(g) F(g^{-1}u) dm(g), \quad \text{for } u \in \mathbb{R}^3,$$

where m is a (fixed) Haar measure on $SL_3(\mathbb{R})$. We note that

$$D_m(\rho * F) := (D_m \rho) * F, \quad \text{for every } m \in \mathbb{N}_o^8. \tag{8.5}$$

Finally, if φ is a locally bounded function on \mathcal{L}_3 , we write

$$(\rho * \varphi)(u) = \int_G \rho(g) \varphi(g^{-1}\Lambda) dm(g), \quad \text{for } \Lambda \in \mathcal{L}_3.$$

Our first main result in this section now reads as follows.

Lemma 8.1. *Let B be a bounded subset of \mathbb{R}^3 and let $s \geq 1$ be an integer. Then, for every bounded Borel function $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that $\{f \neq 0\} \subset B$ and for every $L > 1$ and $\varepsilon \in (0, 1)$, the Siegel transform \hat{f} satisfies*

$$\mathcal{N}_s((\rho_\varepsilon * \hat{f}) \cdot \eta_L) \ll_{B,s} \varepsilon^{-\sigma_s} \cdot L \cdot \|f\|_\infty,$$

where the implicit constants only depend on B and s .

8.1. Proof of Lemma 8.1

We first establish pointwise upper bounds on $D_m(\rho_\varepsilon * \hat{f})$, for an arbitrary multi-index m .

Lemma 8.2. *Let B be a bounded subset of \mathbb{R}^3 and let $s \geq 1$ be an integer. Then, for every $0 < \varepsilon < 1$, $L > 1$ and for every bounded Borel function $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that $\{f \neq 0\} \subset B$ and for every multi-index $m = (m_1, \dots, m_8) \in \mathbb{N}_o^8$ such that $m_1 + \dots + m_8 \leq s$, we have*

$$|D_m((\rho_\varepsilon * \hat{f}) \cdot \eta_L)| \ll_{B,s} \varepsilon^{-\sigma_s} \cdot \|f\|_\infty \cdot L, \quad \text{for all } \Lambda \in \mathcal{L}_3,$$

where the implicit constants only depend on B and s .

Proof. Note that

$$D_m(\rho_\varepsilon * \hat{f}) = (D_m \rho_\varepsilon) * \hat{f} = \widehat{(D_m \rho_\varepsilon) * f}.$$

Hence, by Lemma 5.2,

$$|(D_m(\rho_\varepsilon * \hat{f}))(\Lambda)| \ll_B \|D_m(\rho_\varepsilon) * f\|_\infty \cdot ht(\Lambda) \ll \|\rho_\varepsilon\|_{C^s(SL_3(\mathbb{R}))} \cdot \|f\|_\infty \cdot ht(\Lambda),$$

for every $\Lambda \in \mathcal{L}_3$ and $m = (m_1, \dots, m_8)$ such that $m_1 + \dots + m_8 \leq s$, where the implicit constants only depend on the support of $D_m(\rho_\varepsilon) * f$, which is contained in $V_\varepsilon \cdot B$ (and thus in a ball of radius $2R$, where

R is the smallest radius of a ball enclosing B). By (8.2), $\|\rho_\varepsilon\|_{C_b^s(\mathrm{SL}_3(\mathbb{R}))} \ll \varepsilon^{-\sigma_s}$, and we obtain the bound

$$|D_m(\rho_\varepsilon * \widehat{f})(\Lambda)| \ll_{B,s} \varepsilon^{-\sigma_s} \cdot \|f\|_\infty \cdot \mathrm{ht}(\Lambda), \quad \text{for all } \Lambda \in \mathcal{L}_3,$$

Now the lemma follows from (8.4). □

Let us now discuss Lipschitz semi-norms. By definition, if $\Lambda \in \mathcal{L}_3$ and dist is a right-invariant distance on $\mathrm{SL}_3(\mathbb{R})$, we can induce a (noninvariant) distance on \mathcal{L}_3 by

$$\mathrm{dist}(\Lambda, h.\Lambda) = \inf\{\mathrm{dist}(\mathrm{id}, h\gamma) : \gamma \in \mathrm{Stab}_{\mathrm{SL}_3(\mathbb{R})}(\Lambda)\}, \quad \text{for } h \in \mathrm{SL}_3(\mathbb{R}).$$

The corresponding Lipschitz semi-norms on $C_b^\infty(\mathrm{SL}_3(\mathbb{R}))$ and $C_b^\infty(\mathcal{L}_3)$ are thus given by

$$\mathrm{Lip}_{\mathrm{SL}_3(\mathbb{R})}(\rho) = \sup\left\{\frac{|\rho(g) - \rho(hg)|}{\mathrm{dist}(\mathrm{id}, h)} : g, h \in \mathrm{SL}_3(\mathbb{R}), h \neq \mathrm{id}\right\},$$

for $\rho \in C_b^\infty(\mathrm{SL}_3(\mathbb{R}))$ and

$$\mathrm{Lip}_{\mathcal{L}_3}(\varphi) = \sup\left\{\frac{|\varphi(\Lambda) - \varphi(h.\Lambda)|}{\mathrm{dist}(\Lambda, h.\Lambda)} : \Lambda \in \mathcal{L}_3, h \notin \mathrm{Stab}_{\mathrm{SL}_3(\mathbb{R})}(\Lambda)\right\},$$

for $\varphi \in C_b^\infty(\mathcal{L}_3)$.

One checks that there is a constant $\theta > 0$ such that

$$\mathrm{Lip}_{\mathrm{SL}_3(\mathbb{R})}(\rho_\varepsilon) \ll \varepsilon^{-\theta}, \tag{8.6}$$

with implicit constants that are independent of ε . Upon possibly enlarging σ_s , we can (and will) always assume that $\theta < \sigma_s$. We also leave to the reader to show that the function η_L can be chosen so that

$$|\eta_L(\Lambda) - \eta_L(h.\Lambda)| \ll \mathrm{dist}(\Lambda, h.\Lambda), \tag{8.7}$$

with implicit constants that are independent of L .

Lemma 8.3. *Let B be a bounded subset of \mathbb{R}^3 . Then, for every $0 < \varepsilon < 1$ and for every bounded Borel function $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that $\{f \neq 0\} \subset B$ and $L \geq 1$,*

$$\mathrm{Lip}_{\mathcal{L}_3}((\rho_\varepsilon * \widehat{f}) \cdot \eta_L) \ll_B \varepsilon^{-\theta} \cdot L,$$

where the implicit constants only depend on B .

Proof. We use in the proof that

$$\|\rho_\varepsilon\|_{C_b^s(\mathrm{SL}_3(\mathbb{R}))} \ll \varepsilon^{-\theta} \quad \text{and} \quad \mathrm{Lip}_{\mathrm{SL}_3(\mathbb{R})}(\rho_\varepsilon) \ll \varepsilon^{-\theta}.$$

For every $h \in G$ and $\Lambda \in \mathcal{L}_3$, we have

$$\begin{aligned} \left|(\rho_\varepsilon * \widehat{f})(\Lambda) - (\rho_\varepsilon * \widehat{f})(h.\Lambda)\right| &\leq \int_{(V_\varepsilon \cup h^{-1}V_\varepsilon)} |\rho_\varepsilon(g) - \rho_\varepsilon(hg)| \cdot \widehat{f}(g^{-1}.\Lambda) \, dg \\ &\leq \mathrm{Lip}_{\mathrm{SL}_3(\mathbb{R})}(\rho_\varepsilon) \cdot \mathrm{dist}(\mathrm{id}, h) \cdot \int_{(V_\varepsilon \cup h^{-1}V_\varepsilon)} |\widehat{f}(g^{-1}.\Lambda)| \, dm(g) \\ &\ll \varepsilon^{-\theta} \cdot \mathrm{dist}(\mathrm{id}, h) \cdot \int_{(V_\varepsilon \cup h^{-1}V_\varepsilon)} |\widehat{f}(g^{-1}.\Lambda)| \, dm(g), \end{aligned}$$

where we in the last estimate have used (8.6). Note that the same computation goes through with $h_\gamma := h\gamma$ for $\gamma \in \text{Stab}_{\text{SL}_3(\mathbb{R})}(\Lambda)$, and thus

$$\left| (\rho_\varepsilon * \widehat{f})(\Lambda) - (\rho_\varepsilon * \widehat{f})(h.\Lambda) \right| \leq \varepsilon^{-\theta} \cdot \text{dist}(\Lambda, h.\Lambda) \cdot \int_{(V_\varepsilon \cup h\bar{\gamma}^{-1}V_\varepsilon)} |\widehat{f}(g^{-1}.\Lambda)| \, dm(g). \tag{8.8}$$

We also note that

$$\begin{aligned} ((\rho_\varepsilon * \widehat{f}) \cdot \eta_L)(\Lambda) - ((\rho_\varepsilon * \widehat{f}) \cdot \eta_L)(h.\Lambda) &= \left((\rho_\varepsilon * \widehat{f})(\Lambda) - (\rho_\varepsilon * \widehat{f})(h.\Lambda) \right) \cdot \eta_L(\Lambda) \\ &\quad + (\rho_\varepsilon * \widehat{f})(h.\Lambda) \cdot (\eta_L(\Lambda) - \eta_L(h.\Lambda)). \end{aligned} \tag{8.9}$$

We can without loss of generality assume that at least one of the points Λ and $h.\Lambda$ belong to $\text{supp}(\eta_L)$; otherwise the Lipschitz condition is trivially satisfied. The rest of our analysis is now divided into two cases. □

Case I: $\Lambda, h.\Lambda \in \text{supp}(\eta_L)$, so that $\text{ht}(\Lambda), \text{ht}(h.\Lambda) \leq 2L$

By Lemma 5.2 and (5.2), we have

$$|\widehat{f}(g^{-1}h.\Lambda)| \ll_B \|f\|_\infty \cdot \text{ht}(g^{-1}h.\Lambda) \leq \|f\|_\infty \cdot \|g\|_{\text{op}} \cdot \text{ht}(h.\Lambda), \quad \text{for all } g, h \in \text{SL}_3(\mathbb{R}),$$

and thus,

$$\begin{aligned} \int_{(V_\varepsilon \cup h\bar{\gamma}^{-1}V_\varepsilon)} |\widehat{f}(g^{-1}.\Lambda)| \, dg &\leq \int_{V_\varepsilon} |\widehat{f}(g^{-1}.\Lambda)| \, dg + \int_{V_\varepsilon} |\widehat{f}(g^{-1}h.\Lambda)| \, dg \\ &\ll_B (1 + \varepsilon) \cdot \|f\|_\infty \cdot \text{ht}(\Lambda) + (1 + \varepsilon) \cdot \text{ht}(h.\Lambda) \\ &\ll \|f\|_\infty \cdot \max(\text{ht}(\Lambda), \text{ht}(h.\Lambda)), \end{aligned} \tag{8.10}$$

for all $g \in V_\varepsilon$ and for all $h \in \text{SL}_3(\mathbb{R})$, where the implicit constants only depend on B (and not on $\varepsilon \in (0, 1)$). Hence, we deduce from (8.8) and (8.10) that the first term on the right-hand side in (8.9) is bounded above in absolute value by

$$\ll \|f\|_\infty \cdot \max(\text{ht}(\Lambda), \text{ht}(h.\Lambda)) \cdot \eta_L(\Lambda) \cdot \text{dist}(\Lambda, h.\Lambda) \cdot \varepsilon^{-\theta}, \tag{8.11}$$

while the second term is bounded above in absolute value by

$$\ll \|f\|_\infty \cdot \text{ht}(h.\Lambda) \cdot |\eta_L(\Lambda) - \eta_L(h.\Lambda)| \ll \|f\|_\infty \cdot \text{ht}(h.\Lambda) \cdot \varepsilon^{-\theta} \cdot \text{dist}(\Lambda, h.\Lambda), \tag{8.12}$$

where we in the last inequality have used (8.7). By our assumption, $\text{ht}(\Lambda) \leq 2L$ and $\text{ht}(h.\Lambda) \leq 2L$, so we conclude that

$$\left| (\rho_\varepsilon * \widehat{f})(\Lambda) - (\rho_\varepsilon * \widehat{f})(h.\Lambda) \right| \ll \|f\|_\infty \cdot \varepsilon^{-\theta} \cdot L \cdot \text{dist}(\Lambda, h.\Lambda).$$

Case II: $\Lambda \in \text{supp}(\eta_L)$ but $h.\Lambda \notin \text{supp}(\eta_L)$ or the opposite

We split this case into two subcases. Let us first assume that $\text{dist}(\Lambda, h.\Lambda) \geq 1$. Then, by the same analysis as above,

$$\begin{aligned} |(\rho_\varepsilon * \widehat{f})(\Lambda)\eta_L(\Lambda)| &\ll \varepsilon^{-\theta} \cdot \int_{V_\varepsilon} |\widehat{f}(g^{-1}.\Lambda)| \, dm(g) \\ &\ll_B \varepsilon^{-\theta} \cdot \|f\|_\infty \cdot \text{ht}(\Lambda) \leq \varepsilon^{-\theta} \cdot \|f\|_\infty \cdot L \\ &\ll \varepsilon^{-\theta} \cdot \|f\|_\infty \cdot L \cdot \text{dist}(\Lambda, h.\Lambda). \end{aligned}$$

If $\text{dist}(\Lambda, h.\Lambda) \leq 1$, we can choose h_γ such that $\text{dist}(\text{id}, h_\gamma) \leq 1$, and hence there is a compact set K , which is independent of ε (and h as long as $\text{dist}(\Lambda, h.\Lambda) \leq 1$), with the property that

$$V_\varepsilon \cup h_\gamma^{-1}V_\varepsilon \subset K.$$

Hence, in this case, by (8.6), Lemma 5.2 and (5.2),

$$\left| (\rho_\varepsilon * \widehat{f})(\Lambda) - (\rho_\varepsilon * \widehat{f})(h.\Lambda) \right| \ll_B \varepsilon^{-\theta} \cdot \text{dist}(\Lambda, h.\Lambda) \cdot \|f\|_\infty \cdot \left(\int_K \|g^{-1}\|_{\text{op}} \, dm(g) \right) \cdot \text{ht}(\Lambda).$$

Similarly, when $\text{dist}(\Lambda, h.\Lambda) \leq 1$,

$$|(\rho_\varepsilon * \widehat{f})(h.\Lambda)| \ll_K \varepsilon^{-\theta} \cdot \|f\|_\infty \cdot \text{ht}(\Lambda).$$

We conclude that in the difference (8.9) both terms are bounded above in absolute value by

$$\ll \|f\|_\infty \cdot \varepsilon^{-\theta} \cdot L \cdot \text{dist}(\Lambda, h.\Lambda),$$

where we for the second term have used (8.7). In both subcases, we see that

$$\left| (\rho_\varepsilon * \widehat{f})(\Lambda)\eta_L(\Lambda) - (\rho_\varepsilon * \widehat{f})(h.\Lambda)\eta_L(h.\Lambda) \right| \ll \|f\|_\infty \cdot \varepsilon^{-\theta} \cdot L \cdot \text{dist}(\Lambda, h.\Lambda).$$

The opposite case is handled in the same way by interchanging Λ and $h.\Lambda$, and we are done.

The proof of Lemma 8.1 follows upon combining Lemma 8.2 and Lemma 8.3.

8.2. Smooth approximations of counting functions

In this subsection we discuss smooth approximation of the counting functions that come out of our tessellation scheme in Section 3. We begin by recalling the notation. We have

$$\alpha_\Omega = \ln\left(\frac{T_0c^2}{be^2}\right), \quad \beta_\Omega = \ln\left(\frac{Tc^2}{a}\right), \tag{8.13}$$

and

$$\mathcal{F}_\Omega = \{n \in \mathbb{N}_o^2 : \alpha_\Omega \leq n_1 + n_2 < \beta_\Omega\}.$$

We also assume that

$$\zeta \cdot b \leq c^2 \quad \text{and} \quad c < 1/2 \tag{8.14}$$

for some $\zeta > 0$. Without loss of generality, $\zeta < 1$.

For $n \in \mathcal{F}_\Omega$ and a bounded measurable function $h : [0, \infty) \rightarrow \mathbb{R}$, we define

$$h_{\Omega,n}(\underline{u}, y) := h\left(\frac{e^{n_1+n_2} \cdot y}{T}\right) \chi_{\Delta_{\Omega,n}}(\underline{u}, y), \quad \text{for } (\underline{u}, y) \in \mathbb{R}^2 \times \mathbb{R},$$

where $\Delta_{\Omega,n}$ is defined in (3.3). By Lemma 3.1,

$$\Delta_{\Omega,n} \subset [-c, c]^2 \times \left[\frac{a}{c^2}, \frac{be^2}{c^2}\right] \subset [-1/2, 1/2]^2 \times (0, e^2 \zeta^{-1}]. \tag{8.15}$$

Let

$$\phi_{\Omega,n} := \widehat{h}_{\Omega,n} - \int_{\mathcal{L}_3} \widehat{h}_{\Omega,n} d\mu,$$

where μ is the unique $\text{SL}_3(\mathbb{R})$ -invariant probability measure on \mathcal{L}_3 .

Let ε_T be a decreasing function, which converges to zero as $T \rightarrow \infty$, and let ρ_{ε_T} be as in the previous subsection (with $\varepsilon = \varepsilon_T$). Let L_T be an increasing function, which tends to infinity with T , and define

$$f_{\Omega,n}(\underline{u}, y) := (\rho_{\varepsilon_T} * h_{\Omega,n})(\underline{u}, y), \quad \text{for } (\underline{u}, y) \in \mathbb{R}^2 \times \mathbb{R},$$

and

$$\varphi_{\Omega,n} := \widehat{f}_{\Omega,n} \cdot \eta_{L_T} - \int_{\mathcal{L}_3} \widehat{f}_{\Omega,n} \cdot \eta_{L_T} d\mu. \tag{8.16}$$

It follows from (8.15) that

$$\{f_{\Omega,n} \neq 0\} \subset [-1, 1]^2 \times [-1, e^2 \zeta^{-1} + 1] \tag{8.17}$$

for all n and for all sufficiently large T , so that by Theorem 5.1,

$$\int_{\mathcal{L}_3} \widehat{f}_{\Omega,n} \cdot \eta_L d\mu \leq \int_{\mathcal{L}_3} \widehat{f}_{\Omega,n} d\mu = \int_{\mathbb{R}^2 \times \mathbb{R}} f_{\Omega,n}(\underline{u}, y) d\underline{u} dy \ll_\zeta 1, \tag{8.18}$$

with implicit constants that are independent of T , ε_T and n . Then, (8.16), (8.17), (8.18), and Lemma 8.1 immediately imply the following result.

Lemma 8.4. *For every $s \geq 1$ and for all sufficiently large T ,*

$$\sup_{n \in \mathcal{F}_\Omega} \mathcal{N}_s(\varphi_{\Omega,n}) \ll_s \varepsilon_T^{-\sigma_s} \cdot L_T,$$

where the implicit constants only depend on $\|h\|_\infty$ and s , and not on T .

Let us now assume that h is a bounded Lipschitz continuous function on $[0, \infty)$. The rest of this section is devoted to the proof of the following lemma which roughly asserts that the smooth approximations of $\widehat{h}_{\Omega,n}$ above are good in the $L^2(\nu)$ -sense along a -orbits, where ν is the unique \mathbb{R}^2 -invariant measure on the torus $\mathcal{Y}_3 := \{\Lambda_{\underline{x}} : \underline{x} \in [0, 1)^2\} \subset \mathcal{L}_3$.

Lemma 8.5. *Suppose that*

$$\varepsilon_T < a\zeta^2/100 \quad \text{and} \quad L_T \geq 2e^2\zeta^{-1}.$$

Then, for all $n \in \mathbb{N}^2$ such that $n_1 + n_2 > \max(1, -\ln(a/2))$, we have

$$\begin{aligned} \|(\phi_{\Omega,n} - \varphi_{\Omega,n}) \circ a(n)\|_{L^2(\nu)} &\ll_{h,\zeta} \frac{\varepsilon_T}{a} \cdot \max(1, n_1 + n_2)^{1/2} + e^{-\frac{(n_1+n_2)}{2}} \\ &\quad + \max\left(\varepsilon_T, -\frac{\varepsilon_T}{a} \ln\left(\frac{\varepsilon_T}{a}\right)\right)^{1/2} \cdot \max(1, (n_1 + n_2)) \\ &\quad + \left(L_T^{-1/2} + e^{-\frac{|n|}{2}}\right) \cdot \max(1, n_1 + n_2)^{1/2}. \end{aligned}$$

Remark 8.6. The implicit function in the lemma depends only on $\|h\|_\infty$ and $\sup_{y \neq y'} \frac{|h(y) - h(y')|}{|y - y'|}$.

Proof. We observe that

$$\|(\phi_{\Omega,n} - \varphi_{\Omega,n}) \circ a(n)\|_{L^2(\nu)} \leq \|(\widehat{h}_{\Omega,n} - \widehat{f}_{\Omega,n} \cdot \eta_{L_T}) \circ a(n)\|_{L^2(\nu)} + \|\widehat{h}_{\Omega,n} - \widehat{f}_{\Omega,n} \cdot \eta_{L_T}\|_{L^1(\mu)}, \tag{8.19}$$

and

$$(\widehat{h}_{\Omega,n} - \widehat{f}_{\Omega,n} \cdot \eta_{L_T}) \circ a(n) = (\widehat{h}_{\Omega,n} - \widehat{f}_{\Omega,n}) \circ a(n) + (\widehat{f}_{\Omega,n} \cdot (1 - \eta_{L_T})) \circ a(n). \tag{8.20}$$

We estimate each of the the above terms separately.

First, we proceed with the estimate of $\|(\widehat{h}_{\Omega,n} - \widehat{f}_{\Omega,n}) \circ a(n)\|_{L^2(\nu)}$. We recall that

$$h_{\Omega,n}(\underline{x}, y) = h\left(\frac{e^{n_1+n_2} \cdot y}{T}\right) \cdot \chi_{\Delta_{\Omega,n}}(\underline{x}, y), \quad (\underline{x}, y) \in \mathbb{R}^2 \times \mathbb{R}.$$

For $g \in V_{\varepsilon_T}$, we write $g \cdot (\underline{x}, y) := (\underline{x}(g), y(g))$ as in (4.5). Then

$$h_{\Omega,n}(g \cdot (\underline{x}, y)) = h\left(\frac{e^{n_1+n_2} \cdot y(g)}{T}\right) \cdot \chi_{\Delta_{\Omega,n}}(\underline{x}(g), y(g)), \quad (\underline{x}, y) \in \mathbb{R}^2 \times \mathbb{R}.$$

It follows from (8.15) that

$$\max(\|\underline{x} - \underline{x}(g)\|_\infty, |y - y(g)|) \ll_\zeta \varepsilon_T, \quad \text{for all } g \in V_{\varepsilon_T}$$

provided that either $(\underline{x}, y) \in \Delta_{\Omega,n}$ or $(\underline{x}(g), y(g)) \in \Delta_{\Omega,n}$, which we assume from now on.

We observe that $|h_{\Omega,n} - h_{\Omega,n} \circ g|$ can be bounded by

$$\left| h\left(\frac{e^{n_1+n_2} \cdot y}{T}\right) - h\left(\frac{e^{n_1+n_2} \cdot y(g)}{T}\right) \right| \cdot \chi_{\Delta_{\Omega,n}}(\underline{x}, y) + \left| h\left(\frac{e^{n_1+n_2} y(g)}{T}\right) \right| \cdot \chi_{E_{\Omega,n}(g)},$$

where

$$E_{\Omega,n}(g) := \left(g^{-1}\Delta_{\Omega,n} \setminus \Delta_{\Omega,n}\right) \cup \left(\Delta_{\Omega,n} \setminus g^{-1}\Delta_{\Omega,n}\right).$$

Recall that $\Delta_{\Omega,n} = \emptyset$ unless $n \in \mathcal{F}_\Omega$ so that we may assume that $n \in \mathcal{F}_\Omega$, thus $n_1 + n_2 \leq \ln\left(\frac{Tc^2}{a}\right)$. Since h is Lipschitz continuous, for $n \in \mathcal{F}_\Omega$, we have

$$\begin{aligned} \left| h\left(\frac{e^{n_1+n_2} \cdot y}{T}\right) - h\left(\frac{e^{n_1+n_2} \cdot y(g)}{T}\right) \right| &\leq \frac{e^{n_1+n_2}}{T} \cdot |y - y(g)| \cdot \text{Lip}(h), \\ &\ll_\zeta \frac{c^2 \cdot \varepsilon_T}{a} \cdot \text{Lip}(h) \leq \frac{\varepsilon_T}{a} \cdot \text{Lip}(h). \end{aligned}$$

We note that the set $\Delta_{\Omega,n}$ is of the form (4.2), with

$$a = a, b = b, u_i^- = e^{-1}c, u_i^+ = c, \text{ for } i = 1, 2,$$

and with $\gamma = \gamma_{\Omega,n} = T_0 e^{-(n_1+n_2)}$ and $\delta = \delta_{\Omega,n} = T e^{-(n_1+n_2)}$, which, by (8.15) satisfy the bounds

$$2a < \frac{a}{c^2} < \gamma_{\Omega,n} \quad \text{and} \quad \delta_{\Omega,n} < \frac{be^2}{c^2} \leq e^2 \zeta^{-1}$$

Hence, by Lemma 4.3 (see also Remark 4.2) and the fact that $\varepsilon_T \leq a_T \zeta^2/100$, we can find $(\varepsilon_T, a, e^2 \zeta^{-1} + 1)$ -controlled sets $E_{\Omega,n}^{(s)}$, $s = 1, \dots, 24$, such that

$$E_{\Omega,n}(g) \subset \bigcup_s E_{\Omega,n}^{(s)}, \quad \text{for all } g \in V_{\varepsilon_T}.$$

Therefore, we conclude that for all $g \in V_\varepsilon$ and $n \in \mathcal{F}_\Omega$,

$$|h_{\Omega,n}(\underline{x}, y) - h_{\Omega,n}(g(\underline{x}, y))| \ll_\zeta \left(\frac{\varepsilon_T}{a} \cdot \chi_{\Delta_{\Omega,n}}(\underline{x}, y) + \sum_s \chi_{E_{\Omega,n}^{(s)}}(\underline{x}, y) \right) \cdot \text{Lip}(h),$$

provided that either $(\underline{x}, y) \in \Delta_{\Omega,n}$ or $(\underline{x}(g), y(g)) \in \Delta_{\Omega,n}$. In fact, this estimate holds for all $(\underline{x}, y) \in \mathbb{R}^2 \times \mathbb{R}$ since it holds trivially in the complementary set. Since $f_{\Omega,n} = \rho_{\varepsilon_T} * h_{\Omega,n}$, we obtain that

$$|h_{\Omega,n} - f_{\Omega,n}| \ll_\zeta \left(\frac{\varepsilon_T}{a} \cdot \chi_{\Delta_{\Omega,n}} + \sum_s \chi_{E_{\Omega,n}^{(s)}} \right) \cdot \text{Lip}(h).$$

for all $n \in \mathbb{N}_0^2$. We also have the corresponding bounds for the Siegel transforms:

$$|\widehat{h}_{\Omega,n} - \widehat{f}_{\Omega,n}| \ll_\zeta \left(\frac{\varepsilon_T}{a} \cdot \widehat{\chi}_{\Delta_{\Omega,n}} + \sum_s \widehat{\chi}_{E_{\Omega,n}^{(s)}} \right) \cdot \text{Lip}(h). \tag{8.21}$$

Therefore,

$$\|(\widehat{h}_{\Omega,n} - \widehat{f}_{\Omega,n}) \circ a(n)\|_{L^2(\nu)} \ll_\zeta \left(\frac{\varepsilon_T}{a} \cdot \|\widehat{\chi}_{\Delta_{\Omega,n}} \circ a(n)\|_{L^2(\nu)} + \sum_s \|\widehat{\chi}_{E_{\Omega,n}^{(s)}} \circ a(n)\|_{L^2(\nu)} \right) \cdot \text{Lip}(h). \tag{8.22}$$

It follows from (8.15) and Lemma 5.2 that

$$\|\widehat{\chi}_{\Delta_{\Omega,n}} \circ a(n)\|_{L^2(\nu)} \ll_\zeta \left(\int_{[0,1]^2} \text{ht}(a(n)\Lambda_{\underline{x}})^2 d\underline{x} \right)^{1/2},$$

where the implicit constants are independent of Ω and n . Furthermore, Corollary 5.5, applied with $\theta(u) = u^2$ and $\eta = e^2 \zeta^{-1}$, tells us that

$$\int_{[0,1]^2} \text{ht}(a(n)\Lambda_{\underline{x}})^2 d\underline{x} \ll_\zeta \max(1, n_1 + n_2),$$

We conclude that

$$\|\widehat{\chi}_{\Delta_{\Omega,n}} \circ a(n)\|_{L^2(\nu)} \ll \max(1, n_1 + n_2)^{1/2}, \tag{8.23}$$

for all $(n_1, n_2) \in \mathbb{N}^2$.

Next we estimate $\|\widehat{\chi}_{E_{\Omega,n}^{(s)}} \circ a(n)\|_{L^2(\nu)}$. Since the sets $E_{\Omega,n}^{(s)}$ are $(\varepsilon_T, a, e^2\zeta^{-1} + 1)$ -controlled, it follows from Lemma 7.1 that

$$\begin{aligned} \|\widehat{\chi}_{E_{\Omega,n}^{(s)}} \circ a(n)\|_{L^2(\nu)} &\ll_{\zeta} \left(e^{-(n_1+n_2)} + \max\left(\varepsilon_T, -\frac{\varepsilon_T}{a} \ln\left(\frac{\varepsilon_T}{a}\right)\right) \cdot \max(1, n_1 + n_2)^2 \right)^{1/2} \\ &\ll e^{-\frac{(n_1+n_2)}{2}} + \max\left(\varepsilon_T, -\frac{\varepsilon_T}{a} \ln\left(\frac{\varepsilon_T}{a}\right)\right)^{1/2} \max(1, n_1 + n_2), \end{aligned} \tag{8.24}$$

where the implicit constants are independent of n and T , provided that

$$n_1 + n_2 > \max(1, -\ln(a/2)). \tag{8.25}$$

If we now combine (8.23) and (8.24), we get from (8.22)

$$\begin{aligned} \|(\widehat{h}_{\Omega,n} - \widehat{f}_{\Omega,n}) \circ a(n)\|_{L^2(\nu)} &\ll_{\zeta,h} \frac{\varepsilon_T}{a} \cdot \max(1, n_1 + n_2)^{1/2} + e^{-\frac{(n_1+n_2)}{2}} \\ &\quad + \max\left(\varepsilon_T, -\frac{\varepsilon_T}{a} \ln\left(\frac{\varepsilon_T}{a}\right)\right)^{1/2} \cdot \max(1, (n_1 + n_2)), \end{aligned} \tag{8.26}$$

for all $n \in \mathbb{N}_o^2$ satisfying (8.25). This provides an estimate of the first term in (8.20).

Now we estimate the term $\|(\widehat{f}_{\Omega,n} \cdot (1 - \eta_{L_T})) \circ a(n)\|_{L^2(\nu)}$. It follows from (8.17) and Lemma 5.2 that

$$f_{\Omega,n} \ll_{h,\zeta} ht. \tag{8.27}$$

Furthermore,

$$\text{supp}(1 - \eta_{L_T}) \subset \left\{ ht \geq \frac{L_T}{2} \right\}. \tag{8.28}$$

Thus,

$$\|(\widehat{f}_{\Omega,n} \cdot (1 - \eta_{L_T})) \circ a(n)\|_{L^2(\nu)} \ll \left(\int_{M_{T,n}} (\text{ht}(a(n)\Lambda_{\underline{x}}))^2 d\underline{x} \right)^{1/2},$$

where

$$M_{T,n} := \left\{ \underline{x} \in [0, 1)^2 : \text{ht}(a(n)\Lambda_{\underline{x}}) \geq \frac{L_T}{2} \right\}.$$

Note that $M_{T,n} = M_{n,1}(L_T/2)$, where the latter set is defined as in (5.4). Hence, by Corollary 5.5, applied with $\theta(u) = u^2$, we get

$$\begin{aligned} \int_{M_{T,n}} (\text{ht}(a(n)\Lambda_{\underline{x}}))^2 d\underline{x} &\ll_{\zeta} \left(L_T^{-1} + e^{-\lfloor n \rfloor} \right) \cdot \int_{L_T e^{-2/2}}^{e^{n_1+n_2+1} \max(1, \zeta^{-1})} \frac{du}{u} \\ &\ll_{\zeta} \left(L_T^{-1} + e^{-\lfloor n \rfloor} \right) \cdot \max(1, n_1 + n_2), \end{aligned}$$

for all $n \in \mathbb{N}_o^2$, provided that $L_T \geq 2e^2\zeta^{-1}$. Hence, we conclude that

$$\|(\widehat{f}_{\Omega,n} \cdot (1 - \eta_{L_T})) \circ a(n)\|_{L^2(\nu)} \ll_{h,\zeta} \left(L_T^{-1/2} + e^{-\lfloor n \rfloor/2} \right) \cdot \max(1, n_1 + n_2)^{1/2}, \tag{8.29}$$

This proves (8.20) and hence provides an estimate of the first term in (8.19).

Finally, we estimate the second term in (8.19). We use that

$$\|\widehat{h}_{\Omega,n} - \widehat{f}_{\Omega,n} \cdot \eta_{L_T}\|_{L^1(\mu)} \leq \|\widehat{h}_{\Omega,n} - \widehat{f}_{\Omega,n}\|_{L^1(\mu)} + \|\widehat{f}_{\Omega,n} \cdot (1 - \eta_{L_T})\|_{L^1(\mu)}. \tag{8.30}$$

By (8.21) and Theorem 5.1,

$$\begin{aligned} \|\widehat{h}_{\Omega,n} - \widehat{f}_{\Omega,n}\|_{L^1(\mu)} &\ll_{\zeta} \left(\frac{\varepsilon_T}{a} \cdot \int_{\mathcal{L}_3} \widehat{\chi}_{\Delta_{\Omega,n}} d\mu + \sum_s \int_{\mathcal{L}_3} \widehat{\chi}_{E_{\Omega,n}^{(s)}} d\mu \right) \cdot \text{Lip}(h) \\ &\ll_h \frac{\varepsilon_T}{a} \cdot \text{vol}_3(\Delta_{\Omega,n}) + \sum_s \text{vol}_3(E_{\Omega,n}^{(s)}). \end{aligned}$$

It follows from the definition of $\Delta_{\Omega,n}$ that

$$\text{vol}_3(\Delta_{\Omega,n}) \ll b \ll_{\zeta} 1,$$

and since the sets $E_{\Omega,n}^{(s)}$ are $(\varepsilon_T, a, e^2\zeta^{-1} + 1)$ -controlled,

$$\text{vol}_3(E_{\Omega,n}^{(s)}) \ll_{\zeta} \max\left(\varepsilon_T, -\frac{\varepsilon_T}{a} \ln\left(\frac{\varepsilon_T}{a}\right)\right).$$

Hence,

$$\|\widehat{h}_{\Omega,n} - \widehat{f}_{\Omega,n}\|_{L^1(\mu)} \ll_{\zeta,h} \frac{\varepsilon_T}{a} + \max\left(\varepsilon_T, -\frac{\varepsilon_T}{a} \ln\left(\frac{\varepsilon_T}{a}\right)\right).$$

To estimate the second term in (8.30), we use (8.27) and (8.28) and obtain

$$\begin{aligned} \int_{\mathcal{L}_3} \widehat{f}_{\Omega,n} \cdot (1 - \eta_{L_T}) d\mu &\ll \int_{\{\text{ht} \geq \frac{L_T}{2}\}} \text{ht} d\mu \leq \|\text{ht}\|_{L^2(\mu)} \cdot \mu(\{\text{ht} \geq L_T/2\})^{1/2} \\ &\ll L_T^{-1/2} \cdot \|\text{ht}\|_{L^2(\mu)} \cdot \|\text{ht}\|_{L^1(\mu)}^{1/2}, \end{aligned}$$

where, in the second and third inequalities, we have used the Cauchy-Schwarz inequality and Markov’s inequality respectively. By [7, Lemma 3.10], $\text{ht} \in L^2(\mu)$ and thus

$$\int_{\mathcal{L}_3} \widehat{f}_{\Omega,n} \cdot (1 - \eta_{L_T}) d\mu \ll L_T^{-1/2}.$$

Therefore, we conclude that

$$\|\widehat{h}_{\Omega,n} - \widehat{f}_{\Omega,n} \cdot \eta_{L_T}\|_{L^1(\mu)} \ll_{\zeta,h} \frac{\varepsilon_T}{a} + \max\left(\varepsilon_T, -\frac{\varepsilon_T}{a} \ln\left(\frac{\varepsilon_T}{a}\right)\right) + L_T^{-1/2}. \tag{8.31}$$

This gives the estimate for the second term in (8.19). Finally, combining (8.26), (8.29), and (8.31), we deduce the lemma. \square

9. Proof of Theorem 1.3

9.1. Quantitative equidistribution of order two

We recall some notation from Section 8. For an integer $s \geq 1$, we defined the norms \mathcal{N}_s on the space $C_b^\infty(\mathcal{L}_3)$ of bounded smooth functions on \mathcal{L}_3 by

$$\mathcal{N}_s(\varphi) := \max(\|\varphi\|_{C^s(\mathcal{L}_3)}, \text{Lip}(\varphi)), \quad \text{for } \varphi \in C^\infty(\mathcal{L}_3).$$

We denote by $C_c^\infty(\mathcal{L}_3)$ the subspace of *compactly supported* smooth functions on \mathcal{L}_3 , and we recall that μ denotes the unique $\mathrm{SL}_3(\mathbb{R})$ -invariant probability measure on \mathcal{L}_3 . The following result is due to Kleinbock and Margulis [11] (see also [9] for the special case of one-parameter subgroups).

Theorem 9.1 (Kleinbock–Margulis) *There exist $\delta_o > 0$ and $s_o \geq 1$ such that for every $\varphi \in \mathbb{R} \cdot 1 + C_c^\infty(\mathcal{L}_3)$,*

$$\int_{[0,1]^2} \varphi(a(t)\Lambda_{\underline{x}}) d\underline{x} = \int_{\mathcal{L}_3} \varphi d\mu + O\left(e^{-\delta_o \lfloor t \rfloor} \mathcal{N}_{s_o}(\varphi)\right)$$

for all $t \in \mathbb{R}_+^2$, where the implicit constants are independent of t and φ .

In [2], the first and third author extended this result to products of two (and more) functions. For products of two functions, the exact statement is as follows:

Theorem 9.2. *There exist $\delta_o > 0$ and $s_o \geq 1$ such that for all $\varphi, \psi \in \mathbb{R} \cdot 1 + C_c^\infty(\mathcal{L}_3)$,*

$$\int_{[0,1]^2} \varphi(a(s)\Lambda_{\underline{x}})\psi(a(t)\Lambda_{\underline{x}}) d\underline{x} = \int_{\mathcal{L}_3} \varphi d\mu \int_{\mathcal{L}_3} \psi d\mu + O\left(e^{-\delta \min(\lfloor s \rfloor, \lfloor t \rfloor, \|s-t\|_\infty)} \mathcal{N}_{s_o}(\varphi)\mathcal{N}_{s_o}(\psi)\right),$$

for all $s, t \in \mathbb{R}_+^2$, where the implicit constants are independent of s, t .

An analogous formula for correlations but with different error term was also proven by Shi [20].

9.2. Notation

We recall that the parameters a, b and c are positive real numbers satisfying

$$a < b \quad \text{and} \quad c < \frac{1}{2} \quad \text{and} \quad c^2 > \zeta \cdot b \tag{9.1}$$

for some $\zeta > 0$. Without loss of generality, we may assume that $\zeta < 1$. Our goal is to analyze the sets

$$Q_\Omega(\underline{x}) = \left\{ q \in [T_0, T) \cap \mathbb{N} : \exists \underline{p} = (p_1, p_2) \in \mathbb{Z}^2 : \begin{array}{l} a < |p_1 + qx_1| |p_2 + qx_2| \cdot q \leq b \\ \max(|p_1 + qx_1|, |p_2 + qx_2|) \leq c \end{array} \right\}$$

defined for $\underline{x} = (x_1, x_2) \in \mathbb{R}^2$ and the sets

$$\Omega = \{(\underline{u}, y) \in \mathbb{R}^2 \times \mathbb{R} : a < |u_1 u_2| y \leq b, \|\underline{u}\|_\infty \leq c, T_0 \leq y < T\}.$$

We note that for the lattices $\Lambda_{\underline{x}} \subset \mathbb{R}^2 \times \mathbb{R}$ defined by

$$\Lambda_{\underline{x}} = \{(p + q\underline{x}, q) : \underline{p} \in \mathbb{Z}^2, q \in \mathbb{Z}\},$$

we have

$$q \in Q_\Omega(\underline{x}) \iff \exists \underline{p} \in \mathbb{Z}^2 \text{ such that } (p + q\underline{x}, q) \in \Lambda_{\underline{x}} \cap \Omega. \tag{9.2}$$

Furthermore, since $c < \frac{1}{2}$, there can be at most one $\underline{p} \in \mathbb{Z}^2$ for which the condition on the right-hand side holds.

According to Lemma 3.1, we have the decomposition

$$\Omega = \bigsqcup_{n \in \mathcal{F}_\Omega} a(n)^{-1} \Delta_{\Omega, n}, \tag{9.3}$$

where the sets

$$\Delta_{\Omega,n} \subset [-c, c]^2 \times \left(\frac{a}{c^2}, \frac{be^2}{c^2} \right] \tag{9.4}$$

are defined in (3.3), and \mathcal{F}_Ω is defined in (3.4)–(3.5).

Let us now fix a compactly supported Lipschitz function $h : \mathbb{R} \rightarrow \mathbb{R}$. By (9.2) and (9.3),

$$\begin{aligned} \sum_{q \in \mathcal{Q}_\Omega(\underline{x})} h\left(\frac{q}{T}\right) &= \sum_{(\underline{p}, q) \in \mathbb{Z}^2 \times \mathbb{Z}} h\left(\frac{q}{T}\right) \cdot \chi_\Omega(\underline{p} + q\underline{x}, q) = \sum_{(\underline{\lambda}_1, \underline{\lambda}_2) \in \Lambda_{\underline{x}}} h\left(\frac{\lambda_2}{T}\right) \cdot \chi_\Omega(\underline{\lambda}_1, \lambda_2) \\ &= \sum_{n \in \mathcal{F}_\Omega} \sum_{\underline{\lambda} \in \Lambda_{\underline{x}}} h\left(\frac{\lambda_2}{T}\right) \cdot \chi_{\Delta_{\Omega,n}}(a(n)\underline{\lambda}) = \sum_{n \in \mathcal{F}_\Omega} \widehat{h}_{\Omega,n}(a(n)\Lambda_{\underline{x}}), \end{aligned}$$

for every $\underline{x} \in \mathbb{R}^2$, where

$$h_{\Omega,n}(\underline{u}, y) := h\left(\frac{e^{n_1+n_2}y}{T}\right) \cdot \chi_{\Delta_{\Omega,n}}(\underline{u}, y), \quad (\underline{u}, y) \in \mathbb{R}^2 \times \mathbb{R}.$$

By Siegel’s Theorem (Theorem 5.1), we have

$$\begin{aligned} \int_{\mathcal{L}_3} \widehat{h}_{\Omega,n}(a(n)\Lambda) \, d\mu(\Lambda) &= \int_{\mathbb{R}} \int_{\mathbb{R}^2} h_{\Omega,n}(a(n)(\underline{u}, y)) \, d\underline{u} \, dy \\ &= \int_{\mathbb{R}} h\left(\frac{y}{T}\right) \cdot \left(\int_{\mathbb{R}^2} \chi_{a(n)^{-1}\Delta_{\Omega,n}}(\underline{u}, y) \, d\underline{u} \right) dy, \end{aligned}$$

and thus, by (9.3),

$$\begin{aligned} \sum_{n \in \mathcal{F}_\Omega} \int_{\mathcal{L}_3} \widehat{h}_{\Omega,n}(a(n)\Lambda) \, d\mu(\Lambda) &= \int_{\mathbb{R}} h\left(\frac{y}{T}\right) \cdot \left(\int_{\mathbb{R}^2} \sum_{n \in \mathcal{F}_\Omega} \chi_{a(n)^{-1}\Delta_{\Omega,n}}(\underline{u}, y) \, d\underline{u} \right) dy \\ &= \int_{\mathbb{R}} h\left(\frac{y}{T}\right) \cdot \left(\int_{\mathbb{R}^2} \chi_\Omega(\underline{u}, y) \, d\underline{u} \right) dy = \mathcal{M}_\Omega(h), \end{aligned}$$

where $\mathcal{M}_\Omega(h)$ is defined in (1.3). Therefore, we conclude that

$$\sum_{q \in \mathcal{Q}_\Omega(\underline{x})} h\left(\frac{q}{T}\right) - \mathcal{M}_\Omega(h) = \sum_{n \in \mathcal{F}_\Omega} \phi_{\Omega,n}(a(n)\Lambda_{\underline{x}}), \tag{9.5}$$

where

$$\phi_{\Omega,n}(\Lambda) := \widehat{h}_{\Omega,n} - \int_{\mathcal{L}_3} \widehat{h}_{\Omega,n} \, d\mu.$$

9.3. Partial moments

To simplify notation, we write

$$D_\Omega(\Lambda) := \sum_{n \in \mathcal{F}_\Omega} \phi_{\Omega,n}(a(n)\Lambda) \tag{9.6}$$

for a lattice Λ . Our goal is to analyze convex moments of the form

$$C_\Omega := \int_{\mathcal{L}_3} \theta_\kappa(D_\Omega) d\nu,$$

where the function θ_κ is defined in (1.4) and the measure ν on \mathcal{L}_3 is given by

$$\int_{\mathcal{L}_3} \varphi d\nu = \int_{[0,1]^2} \varphi(\Lambda_{\underline{x}}) d\underline{x}, \quad \text{for } \varphi \in C_c(\mathcal{L}_3).$$

The following properties of the function θ_κ can be readily verified:

- (P1) θ_κ is a convex function, increasing on $[0, \infty)$, and $\theta_\kappa(t) \leq t^2$ for all t .
- (P2) $\theta_\kappa(ct) \leq c^2 \cdot \theta_\kappa(t)$ for all $c \geq 1$ and t .
- (P3) $\theta_\kappa^{-1}(u) \ll_\kappa u^{1/2} (\ln^+ u)^{(1+\kappa)/2}$, for all $u \geq 0$, where $\ln^+ u := \ln \max(e, u)$.

It will be convenient to decompose the sum (9.6) further: for $\xi_T < \beta_\Omega$, we introduce the sets

$$\mathcal{G}_\Omega^+(\xi_T) := \{n \in \mathcal{F}_\Omega : \lfloor n \rfloor \geq \xi_T\} \quad \text{and} \quad \mathcal{G}_\Omega^-(\xi_T) := \{n \in \mathcal{F}_\Omega : \lfloor n \rfloor < \xi_T\}.$$

Then $\mathcal{F}_\Omega = \mathcal{G}_\Omega^+(\xi_T) \sqcup \mathcal{G}_\Omega^-(\xi_T)$ for all T . Let us fix $\varepsilon_T \in (0, 1)$ and $L_T > 1$, and let ρ_{ε_T} and η_{L_T} be defined as in Section 8. We introduce the functions

$$f_{\Omega,n} := \rho_{\varepsilon_T} * h_{\Omega,n}$$

on \mathbb{R}^3 , and consider

$$\varphi_{\Omega,n} := \widehat{f}_{\Omega,n} \cdot \eta_{L_T} - \int_{\mathcal{L}_3} \widehat{f}_{\Omega,n} \cdot \eta_{L_T} d\mu.$$

on \mathcal{L}_3 that provide a C_c^∞ -approximation for the functions $\phi_{\Omega,n}$. We now write:

$$\begin{aligned} D_\Omega &= \sum_{n \in \mathcal{G}_\Omega^-(\xi_T)} \phi_{\Omega,n} \circ a(n) + \sum_{n \in \mathcal{G}_\Omega^+(\xi_T)} (\phi_{\Omega,n} - \varphi_{\Omega,n}) \circ a(n) + \sum_{n \in \mathcal{G}_\Omega^+(\xi_T)} \varphi_{\Omega,n} \circ a(n) \\ &=: D_\Omega^{(1)} + D_\Omega^{(2)} + D_\Omega^{(3)}. \end{aligned} \tag{9.7}$$

Since θ_κ is convex and satisfies condition (P2), we have

$$C_\Omega \leq 3 \cdot \left(C_\Omega^{(1)} + C_\Omega^{(2)} + C_\Omega^{(3)} \right),$$

where

$$C_\Omega^{(k)} := \int_{\mathcal{L}_3} \theta_\kappa(D_\Omega^{(k)}) d\nu, \quad \text{for } k = 1, 2, 3.$$

In what follows, we will estimate these partial moments separately.

9.4. An upper bound on $C_\Omega^{(1)}$

Since θ_κ is convex and satisfies (P2), we have

$$\begin{aligned} C_\Omega^{(1)} &= \int_{\mathcal{L}_3} \theta_\kappa \left(\sum_{n \in \mathcal{G}_\Omega^-(\xi_T)} \phi_{\Omega,n} \circ a(n) \, dv \right) \\ &\leq |\mathcal{G}_\Omega^-(\xi_T)|^2 \cdot \int_{\mathcal{L}_3} \theta_\kappa \left(\frac{1}{|\mathcal{G}_\Omega^-(\xi_T)|} \sum_{n \in \mathcal{G}_\Omega^-(\xi_T)} \phi_{\Omega,n} \circ a(n) \, dv \right) \\ &\leq |\mathcal{G}_\Omega^-(\xi_T)| \cdot \sum_{n \in \mathcal{G}_\Omega^-(\xi_T)} \int_{\mathcal{L}_3} \theta_\kappa(\phi_{\Omega,n} \circ a(n)) \, dv. \end{aligned} \tag{9.8}$$

Furthermore,

$$\int_{\mathcal{L}_3} \theta_\kappa(\phi_{\Omega,n} \circ a(n)) \, dv \leq 2 \max \left(\int_{\mathcal{L}_3} \theta_\kappa(\widehat{h}_{\Omega,n} \circ a(n)) \, dv, \theta_\kappa \left(\int_{\mathcal{L}_3} \widehat{h}_{\Omega,n} \, d\mu \right) \right) \tag{9.9}$$

By Siegel’s Theorem (Theorem 5.1) and (9.4),

$$\int_{\mathcal{L}_3} \widehat{h}_{\Omega,n} \, d\mu = \int_{\mathbb{R}^2 \times \mathbb{R}} h_{\Omega,n}(\underline{u}, y) \, d\underline{u} dy \leq \|h\|_\infty \text{Vol}_3(\Delta_{\Omega,n}) \ll_h b \ll 1. \tag{9.10}$$

We stress that the implicit constants only depend on $\|h\|_\infty$ (which is assumed to be fixed throughout the proof).

Now, we observe that

$$|\widehat{h}_{\Omega,n}(a(n)\Lambda_{\underline{x}})| \leq \|h\|_\infty \cdot \widehat{\chi}_{\Delta_{\Omega,n}}(a(n)\Lambda_{\underline{x}}), \quad \text{for all } \underline{x} \in \mathbb{R}^2.$$

Furthermore, by (9.4),

$$\chi_{\Delta_{\Omega,n}}(\underline{u}, y) \leq \chi_{[-1,1]^2 \times [0,1]}(\underline{u}, r_\Omega), \quad \text{where } r_\Omega := \frac{c^2}{be^2},$$

so we conclude that

$$|\widehat{h}_{\Omega,n}(a(n)\Lambda_{\underline{x}})| \leq \|h\|_\infty \cdot \widehat{\chi}_{[-1,1]^2 \times [0,1]}(a(n)\Lambda_{\underline{x},r_\Omega}),$$

where the lattice $\Lambda_{\underline{x},r_\Omega}$ is defined as in (5.3). Thus, by Lemma 5.2,

$$|\widehat{h}_{\Omega,n}(a(n)\Lambda_{\underline{x}})| \ll_h \text{ht}(a(n)\Lambda_{\underline{x},r_\Omega}). \tag{9.11}$$

We note that $r_\Omega \geq e^{-2}\zeta$ by (9.1).

Let us now use these estimates to bound $C_\Omega^{(1)}$. We introduce the set

$$E_{T,n} := \{ \underline{x} \in [0, 1)^2 : \text{ht}(a(n)\Lambda_{\underline{x},r_\Omega}) \geq e^4 \zeta^{-1} \}.$$

By Corollary 5.5, applied with $\rho = e^{-2}\zeta$ and $\eta = e^4 \zeta^{-1}$, we have for all $n \in \mathbb{N}_>^2$,

$$\int_{E_{T,n}} \theta_\kappa(\text{ht}(a(n)\Lambda_{\underline{x},r_\Omega})) \, d\underline{x} \ll_\zeta r_\Omega^{-1} \cdot \int_{e^2 \zeta^{-1}}^\infty \frac{\theta_\kappa(u)}{u^3} \, du \ll_{\kappa,\zeta} r_\Omega^{-1},$$

where the implicit constants are independent of T and n . Here we used that the function $u \mapsto \frac{\theta_\kappa(u)}{u^3}$ is integrable on $[1, \infty)$. Hence, from (9.11), since θ_κ is increasing and satisfies (P2), we conclude that

$$\int_{E_{T,n}} \theta_\kappa(\widehat{h}_{\Omega,n}(a(n)\Lambda_{\underline{x}})) \, d\underline{x} \ll_h \int_{E_{T,n}} \theta_\kappa(\text{ht}(a(n)\Lambda_{\underline{x},r_\Omega})) \, d\underline{x} \ll_{\kappa,\zeta} r_\Omega^{-1}, \tag{9.12}$$

for all n .

It remains to estimate the integral over $E_{T,n}^c$. To do so, note that

$$\int_{E_{T,n}^c} \theta_\kappa(\widehat{h}_{\Omega,n}(a(n)\Lambda_{\underline{x}})) \, d\underline{x} \ll_h \int_{E_{T,n}^c \cap P_{\Omega,n}} \theta_\kappa(\text{ht}(a(n)\Lambda_{\underline{x},r_\Omega})) \, d\underline{x} \ll_{\kappa,\zeta} \text{Vol}_2(P_{\Omega,n}),$$

where

$$P_{\Omega,n} := \left\{ \underline{x} \in [0, 1]^2 : \widehat{h}_{\Omega,n}(a(n)\Lambda_{\underline{x}}) > 0 \right\}.$$

Using that $c < \frac{1}{2}$, and the upper bounds, which follow from (9.4), we obtain that

$$|\widehat{h}_{\Omega,n}(a(n)\Lambda_{\underline{x}})| \leq \|h\|_\infty \cdot \widehat{\chi}_{[-\frac{1}{2}, \frac{1}{2}]^2 \times [0, \frac{1}{2}]}(a(n)\Lambda_{\underline{x},r_\Omega/2}),$$

we see that

$$\begin{aligned} P_{\Omega,n} &\subseteq \left\{ \underline{x} \in [0, 1]^2 : a(n)\Lambda_{\underline{x},r_\Omega/2} \cap \left(\left[-\frac{1}{2}, \frac{1}{2} \right]^2 \times \left[0, \frac{1}{2} \right] \right) \neq \{0\} \right\} \\ &\subseteq \left\{ \underline{x} \in [0, 1]^2 : s_1(a(n)\Lambda_{\underline{x},r_\Omega/2}) \leq \frac{1}{2} \right\}, \end{aligned}$$

where s_1 is defined as in Subsection 5.1. Hence, by Lemma 5.7, applied with $\varepsilon = \frac{1}{2}$ and $r = r_\Omega/2$, we have

$$\text{Vol}_2(P_{\Omega,n}) \ll r_\Omega^{-1}, \quad \text{for all } n.$$

Therefore,

$$\int_{E_{T,n}^c} \theta_\kappa(\widehat{h}_{\Omega,n}(a(n)\Lambda_{\underline{x}})) \, d\underline{x} \ll_{h,\kappa,\zeta} r_\Omega^{-1}.$$

If we combine this estimate with (9.10) and (9.12), we deduce from (9.9) that

$$\int_{\mathcal{L}_3} \theta_\kappa(\varphi_{\Omega,n} \circ a(n)) \, dv \ll_{h,\kappa,\zeta} r_\Omega^{-1} \ll \frac{b}{c^2} \tag{9.13}$$

for all $n \in \mathbb{N}_o^2$. Hence, by (9.8),

$$\mathcal{C}_\Omega^{(1)} \leq |\mathcal{G}_\Omega^-(\xi_T)| \cdot \sum_{n \in \mathcal{G}_\Omega^-(\xi_T)} \int_{\mathcal{L}_3} \theta_\kappa(\varphi_{\Omega,n} \circ a(n)) \, dv \ll_{h,\kappa,\zeta} |\mathcal{G}_\Omega^-(\xi_T)|^2 \cdot \frac{b}{c^2}.$$

Since

$$|\mathcal{G}_\Omega^-(\xi_T)| \leq 2\xi_T \cdot (\beta_\Omega - \alpha_\Omega) \quad \text{and} \quad (\beta_\Omega - \alpha_\Omega)^2 \leq |\mathcal{F}_\Omega|,$$

we conclude that

$$|\mathcal{G}_\Omega^-(\xi_T)|^2 \ll \xi_T^2 \cdot |\mathcal{F}_\Omega|,$$

and

$$C_\Omega^{(1)} \ll_{h,\kappa,\zeta} \frac{b \cdot \xi_T^2}{c^2} \cdot |\mathcal{F}_\Omega|. \tag{9.14}$$

9.5. An upper bound on $C_\Omega^{(2)}$

Since $\theta_\kappa(u) \leq u^2$ for all u , we have

$$\begin{aligned} \int_{\mathcal{L}_3} \theta_\kappa(D_\Omega^{(2)}) \, d\nu &\leq \int_{\mathcal{L}_3} \left(\sum_{n \in \mathcal{G}_\Omega^+(\xi_T)} (\phi_{\Omega,n} - \varphi_{\Omega,n}) \circ a(n) \right)^2 \, d\nu \\ &\leq \left(\sum_{n \in \mathcal{G}_T^+} \|(\phi_{\Omega,n} - \varphi_{\Omega,n}) \circ a(n)\|_{L^2(\nu)} \right)^2. \end{aligned}$$

Let us assume that

$$\xi_T > \max(1, -\ln(a/2)). \tag{9.15}$$

Then for all $n \in \mathcal{G}_\Omega^+(\xi_T)$, we have $n_1 + n_2 > \max(1, -\ln(a/2))$. In particular, if we additionally assume that

$$\varepsilon_T < a\xi^2/100 \quad \text{and} \quad L_T \geq 2e^2\xi^{-1}, \tag{9.16}$$

then the conditions of Lemma 8.5 are satisfied, from which we conclude that

$$\begin{aligned} \|(\phi_{\Omega,n} - \varphi_{\Omega,n}) \circ a(n)\|_{L^2(\nu)} &\ll_{h,\zeta} \frac{\varepsilon_T}{a} \cdot \max(1, n_1 + n_2)^{1/2} \\ &\quad + e^{-(n_1+n_2)/2} + \max\left(\varepsilon_T, -\frac{\varepsilon_T}{a} \ln\left(\frac{\varepsilon_T}{a}\right)\right)^{1/2} \cdot \max(1, n_1 + n_2) \\ &\quad + \left(L_T^{-1/2} + e^{-\lfloor n \rfloor/2}\right) \cdot \max(1, n_1 + n_2)^{1/2} \end{aligned}$$

for all $n \in \mathcal{G}_\Omega^+(\xi_T)$. We note that

$$\begin{aligned} \sum_{n \in \mathcal{G}_\Omega^+(\xi_T)} \max(1, n_1 + n_2)^{1/2} &\leq \beta_\Omega^{1/2} \cdot |\mathcal{G}_\Omega^+(\xi_T)|, \\ \sum_{n \in \mathcal{G}_\Omega^+(\xi_T)} \max(1, n_1 + n_2) &\leq \beta_\Omega \cdot |\mathcal{G}_\Omega^+(\xi_T)|, \\ \sum_{n \in \mathcal{G}_\Omega^+(\xi_T)} e^{-(n_1+n_2)/2} &\leq 2e^{-\xi_T/2} \cdot |\mathcal{G}_\Omega^+(\xi_T)|, \\ \sum_{n \in \mathcal{G}_\Omega^+(\xi_T)} e^{-\lfloor n \rfloor/2} \max(1, n_1 + n_2)^{1/2} &\leq 2\beta_\Omega^{1/2} \cdot e^{-\xi_T/2} \cdot |\mathcal{G}_\Omega^+(\xi_T)|, \end{aligned}$$

and thus

$$\int_{\mathcal{L}_3} \theta_\kappa(D_\Omega^{(2)}) dv \leq \left(\sum_{n \in \mathcal{G}_\Omega^+(\xi_T)} \|(\phi_{\Omega,n} - \varphi_{\Omega,n}) \circ a(n)\|_{L^2(\nu)} \right)^2 \ll_{h,\zeta} \gamma \cdot |\mathcal{G}_\Omega^+(\xi_T)|^2 \leq \gamma \cdot \beta_\Omega^2 \cdot |\mathcal{F}_\Omega|, \tag{9.17}$$

where

$$\gamma := \beta_\Omega \cdot \max\left(\left(\frac{\varepsilon_T}{a}\right)^2, \beta_\Omega \cdot \max\left(\varepsilon_T, -\frac{\varepsilon_T}{a} \ln\left(\frac{\varepsilon_T}{a}\right)\right), L_T^{-1}, e^{-\xi_T}\right).$$

9.6. An upper bound on $\mathcal{C}_\Omega^{(3)}$

We have

$$\begin{aligned} \int_{\mathcal{L}_3} \left(D_\Omega^{(3)}\right)^2 dv &= \int_{\mathcal{L}_3} \sum_{m,n \in \mathcal{G}_\Omega^+(\xi_T)} \varphi_{\Omega,m} \circ a(m) \cdot \varphi_{\Omega,n} \circ a(n) dv \\ &= \sum_{\substack{m,n \in \mathcal{G}_\Omega^+(\xi_T) \\ \|m-n\| < \xi_T}} \int_{\mathcal{L}_3} \varphi_{\Omega,m} \circ a(m) \cdot \varphi_{\Omega,n} \circ a(n) dv \\ &\quad + \sum_{\substack{m,n \in \mathcal{G}_\Omega^+(\xi_T) \\ \|m-n\| \geq \xi_T}} \int_{\mathcal{L}_3} \varphi_{\Omega,m} \circ a(m) \cdot \varphi_{\Omega,n} \circ a(n) dv. \end{aligned} \tag{9.18}$$

We estimate the two sums on the right-hand side separately.

By Lemma 5.2 and (9.4), $|\varphi_{\Omega,n}| \ll_{h,\zeta} L_T$ for all n . Then, by the Cauchy-Schwarz inequality, we deduce

$$\begin{aligned} &\int_{\mathcal{L}_3} |\varphi_{\Omega,m} \circ a(m) \cdot \varphi_{\Omega,n} \circ a(n)| dv \\ &\ll_{h,\zeta} (\ln L_T)^{1+\kappa} \int_{\mathcal{L}_3} \ln(e + |\varphi_{\Omega,m} \circ a(m)|)^{-(1+\kappa)/2} \cdot |\varphi_{\Omega,m} \circ a(m)| \\ &\quad \cdot \ln(e + |\varphi_{\Omega,n} \circ a(n)|)^{-(1+\kappa)/2} \cdot |\varphi_{\Omega,n} \circ a(n)| dv \\ &\leq (\ln L_T)^{1+\kappa} \left(\int_{\mathcal{L}_3} \theta_\kappa(\varphi_{\Omega,m} \circ a(m)) dv \cdot \int_{\mathcal{L}_3} \theta_\kappa(\varphi_{\Omega,n} \circ a(n)) dv \right)^{1/2}. \end{aligned}$$

Hence, by (9.13), we have that

$$\int_{\mathcal{L}_3} |\varphi_{\Omega,m} \circ a(m) \cdot \varphi_{\Omega,n} \circ a(n)| dv \ll_{h,\zeta} \frac{b \cdot (\ln L_T)^{1+\kappa}}{c^2}.$$

so that

$$\sum_{\substack{m,n \in \mathcal{G}_\Omega^+(\xi_T) \\ \|m-n\| < \xi_T}} \int_{\mathcal{L}_3} |\varphi_{\Omega,m} \circ a(m) \cdot \varphi_{\Omega,n} \circ a(n)| dv \ll_{h,\zeta} \frac{b \cdot (\ln L_T)^{1+\kappa}}{c^2} \cdot |\mathcal{M}_\Omega(\xi_T)|, \tag{9.19}$$

where

$$\mathcal{M}_\Omega(\xi_T) := \{(m, n) \in \mathcal{G}_\Omega^+(\xi_T) : \|m - n\| < \xi_T\}.$$

For every $m \in \mathcal{G}_\Omega^+(\xi_T)$, there are $O(\xi_T^2)$ elements $n \in \mathcal{G}_\Omega^+(\xi_T)$ such that $\|m - n\| < \xi_T$, and thus $|\mathcal{M}_\Omega(\xi_T)| \ll \xi_T^2 \cdot |\mathcal{G}_\Omega^+(\xi_T)|$, where the implicit constants are independent of Ω . We thus see that

$$\sum_{\substack{m, n \in \mathcal{G}_\Omega^+(\xi_T) \\ \|m-n\| < \xi_T}} \int_{\mathcal{L}_3} \varphi_{\Omega, m} \circ a(m) \cdot \varphi_{\Omega, n} \circ a(n) \, dv \ll_{h, \zeta} \frac{b \cdot (\ln L_T)^{1+\kappa}}{c^2} \cdot \xi_T^2 \cdot |\mathcal{F}_\Omega|. \tag{9.20}$$

Let us now turn to the second sum in (9.18). By construction, the function $\varphi_{\Omega, n}$ belongs to $\mathbb{R} \cdot 1 + C_c^\infty(\mathcal{L}_3)$ and satisfies $\int_{\mathcal{L}_3} \varphi_{\Omega, n} \, d\mu = 0$. Hence, by Theorem 9.2, there exist $s \geq 1$ and $\delta > 0$ such that

$$\left| \int_{\mathcal{L}_3} \varphi_{\Omega, m} \circ a(m) \cdot \varphi_{\Omega, n} \circ a(n) \, dv \right| \ll e^{-\delta \min(\lfloor m \rfloor, \lfloor n \rfloor, \|m-n\|)} \mathcal{N}_s(\varphi_{\Omega, m}) \cdot \mathcal{N}_s(\varphi_{\Omega, n}),$$

where the implicit constants are independent of m, n, T . Furthermore, by Lemma 8.1, there exists a constant $\sigma_s > 0$ such that

$$\mathcal{N}_s(\varphi_{\Omega, n}) \ll_h \varepsilon_T^{-\sigma_s} \cdot L_T, \quad \text{for all } n,$$

where the implicit constants are independent of Ω and n . We conclude that for all $m, n \in \mathcal{G}_\Omega^+(\xi_T)$ such that $\|m - n\| \geq \xi_T$, we have

$$\left| \int_{\mathcal{L}_3} \varphi_{\Omega, m} \circ a(m) \cdot \varphi_{\Omega, n} \circ a(n) \, dv \right| \ll_h e^{-\delta \xi_T} \cdot \varepsilon_T^{-2\sigma_s} \cdot L_T^2,$$

and thus

$$\sum_{\substack{m, n \in \mathcal{G}_\Omega^+(\xi_T) \\ \|m-n\| \geq \xi_T}} \int_{\mathcal{L}_3} \varphi_{\Omega, m} \circ a(m) \cdot \varphi_{\Omega, n} \circ a(n) \, dv \ll_h e^{-\delta \xi_T} \cdot \varepsilon_T^{-2\sigma_s} \cdot L_T^2 \cdot |\mathcal{G}_\Omega^+(\xi_T)|^2. \tag{9.21}$$

If we now plug our estimates (9.20) and (9.21) into (9.18), we conclude that

$$\int_{\mathcal{L}_3} \theta_\kappa(D_\Omega^{(3)}) \, dv \ll_{h, \zeta} \frac{b \cdot (\ln L_T)^{1+\kappa} \cdot \xi_T^2}{c^2} \cdot |\mathcal{F}_\Omega| + e^{-\delta \xi_T} \cdot \varepsilon_T^{-2\sigma_s} \cdot L_T^2 \cdot |\mathcal{G}_\Omega^+(\xi_T)|^2.$$

Since $|\mathcal{G}_\Omega^+(\xi_T)| \ll \beta_\Omega^2$, where the implicit constants are independent of Ω , we have

$$\int_{\mathcal{L}_3} \theta_\kappa(D_\Omega^{(3)}) \, dv \ll_{h, \zeta} \left(\frac{b \cdot (\ln L_T)^{1+\kappa} \cdot \xi_T^2}{c^2} + e^{-\delta \xi_T} \cdot \varepsilon_T^{-2\sigma_s} \cdot L_T^2 \cdot \beta_\Omega^2 \right) \cdot |\mathcal{F}_\Omega|. \tag{9.22}$$

9.7. Putting it all together

Let us now summarize what we have so far:

$$\mathcal{C}_\Omega \leq 3 \cdot \left(\mathcal{C}_\Omega^{(1)} + \mathcal{C}_\Omega^{(2)} + \mathcal{C}_\Omega^{(3)} \right),$$

and by (9.14), (9.17), (9.22),

$$\begin{aligned}
 C_{\Omega}^{(1)} &\ll_{h,\kappa,\zeta} \frac{b \cdot \xi_T^2}{c^2} \cdot |\mathcal{F}_{\Omega}|, \\
 C_{\Omega}^{(2)} &\ll_{h,\zeta} \beta_{\Omega}^3 \cdot \max\left(\left(\frac{\varepsilon_T}{a}\right)^2, \beta_{\Omega} \cdot \max\left(\varepsilon_T, -\frac{\varepsilon_T}{a} \ln\left(\frac{\varepsilon_T}{a}\right)\right), L_T^{-1}, e^{-\xi_T}\right) \cdot |\mathcal{F}_{\Omega}|, \\
 C_{\Omega}^{(3)} &\ll_h \left(\frac{b \cdot (\ln L_T)^{1+\kappa} \cdot \xi_T^2}{c^2} + e^{-\delta \xi_T} \cdot \varepsilon_T^{-2\sigma_s} \cdot L_T^2 \cdot \beta_{\Omega}^2\right) \cdot |\mathcal{F}_{\Omega}|,
 \end{aligned}$$

provided that the parameters have been chosen so that

$$\xi_T \geq \max(1, -\ln(a/2)), \quad \varepsilon_T < a\zeta^2/100, \quad L_T \geq e^2\zeta^{-1}, \quad c^2 > \zeta \cdot b. \tag{9.23}$$

We further assume that

$$a \geq (\ln T)^{-\theta}, \quad \text{for some } \theta > 0.$$

Then, in particular, $\beta_{\Omega} = \ln\left(\frac{Tc^2}{a}\right) \ll_{\theta} \ln T$. In what follows, we will choose the parameters ξ_T , ε_T , and L_T , so that

$$C_{\Omega} \ll_{h,\kappa,\zeta} \left(\frac{b \cdot (\ln L_T)^{1+\kappa} \cdot \xi_T^2}{c^2} + O_{\theta,\rho}((\ln T)^{-\rho})\right) \cdot |\mathcal{F}_{\Omega}| \tag{9.24}$$

for all $\rho > 0$.

To prove (9.24), let

$$\xi_T = \xi_o \cdot \ln(\ln T), \quad L_T = (\ln T)^{3+\eta}, \quad \varepsilon_T = (\ln T)^{-\gamma}$$

for some *positive* constants ξ_o, η and γ that will be chosen later. With these choices, we see that

$$\begin{aligned}
 \beta_{\Omega}^3 \cdot \left(\frac{\varepsilon_T}{a}\right)^2 &\ll_{\theta} (\ln T)^{3-2(\gamma-\theta)}, \\
 \beta_{\Omega}^4 \cdot \max\left(\varepsilon_T, -\frac{\varepsilon_T}{a} \ln\left(\frac{\varepsilon_T}{a}\right)\right) &\ll_{\theta} (\ln T)^{4-(\gamma-\theta)} \cdot \ln(\ln T), \\
 \beta_{\Omega}^3 \cdot \max\left(L_T^{-1}, e^{-\xi_T}\right) &\ll_{\theta} \max\left((\ln T)^{-\eta}, (\ln T)^{3-\xi_o}\right), \\
 e^{-\delta \xi_T} \cdot \varepsilon_T^{-2\sigma_s} \cdot L_T^2 \cdot \beta_{\Omega}^2 &\ll_{\theta} (\ln T)^{-\delta \xi_o + 2\sigma_s \gamma + 2(3+\eta) + 2},
 \end{aligned}$$

for all sufficiently large T . We think of θ as fixed throughout the argument. Hence, if we pick η, γ sufficiently large and next ξ_o sufficiently large, then all of the right-hand sides above are $O_{\theta,\rho}((\ln T)^{-\rho})$ for any $\rho > 0$. Furthermore, one verifies that all the conditions in (9.23) are satisfied. Hence, (9.24) follows. Since

$$\frac{b}{c^2} \gg a \geq (\ln T)^{-\theta},$$

we obtain that

$$C_{\Omega} \ll_{h,\kappa,\zeta,\theta} \frac{b \cdot (\ln L_T)^{1+\kappa} \cdot \xi_T^2}{c^2} \cdot |\mathcal{F}_{\Omega}|. \tag{9.25}$$

Furthermore,

$$\xi_T \ll \ln \ln T \quad \text{and} \quad |\mathcal{F}_\Omega| \ll \beta_\Omega^2 - \alpha_\Omega^2,$$

where

$$\alpha_\Omega = \ln\left(\frac{T_0 c^2}{e^2 b}\right) \quad \text{and} \quad \beta_\Omega = \ln\left(\frac{T c^2}{a}\right).$$

Since, according to our assumptions,

$$\zeta \leq \frac{c^2}{b} < \frac{c^2}{a} \leq (\ln T)^\theta,$$

we conclude that

$$|\mathcal{F}_\Omega| \ll \left(\ln T + \ln\left(\frac{c^2}{a}\right)\right)^2 - \left(\ln T_0 + \ln\left(\frac{c^2}{e^2 b}\right)\right)^2 \ll_\zeta \mathcal{L}([T_0, T]) + \mathcal{M}([T_0, T]),$$

where

$$\mathcal{L}([T_0, T]) := (\ln T)^2 - (\ln T_0)^2 \quad \text{and} \quad \mathcal{M}([T_0, T]) := \ln T \ln \ln T.$$

Hence, it follows from (9.5) and (9.25) that

$$\int_{[0,1]^2} \theta_\kappa \left(S_\Omega h(\underline{x}) - \mathcal{M}_\Omega(h) \right) d\underline{x} = \int_{\mathcal{L}_3} \theta_\kappa(D_\Omega) d\nu \ll_{h,\kappa,\zeta,\theta} c^{-2} \cdot b \cdot (\ln \ln T)^{3+\kappa} \cdot \left(\mathcal{L}([T_0, T]) + \mathcal{M}([T_0, T]) \right). \quad (9.26)$$

In particular, in the case when $T_0 = 1$, we obtain Theorem 1.3.

10. Proof of Theorem 1.2

We work with the sets

$$\Omega = \Omega_{[T_0, T]}^{(a,b),c} := \left\{ (u, y) \in \mathbb{R}^2 \times [T_0, T] : \begin{array}{l} \max(|u_1|, |u_2|) \leq c \\ a < |u_1 u_2| \cdot y \leq b \end{array} \right\}$$

that depend on parameters $0 < a < b < 1, 0 < c < 1/2, 1 \leq T_0 < T$. For a lattice Λ , we consider the discrepancy function

$$D_{[T_0, T]}^{(a,b),c}(\Lambda) := \left| \Lambda \cap \Omega_{[T_0, T]}^{(a,b),c} \right| - \text{Vol}_3 \left(\Omega_{[T_0, T]}^{(a,b),c} \right).$$

Let us take in (9.26) a Lipschitz function h such that $h = 1$ on $[0, 1]$. Then

$$S_\Omega h(\underline{x}) = |\Lambda_{\underline{x}} \cap \Omega| \quad \text{and} \quad \mathcal{M}_\Omega(h) = \text{Vol}_3(\Omega),$$

so that it follows that

$$\int_{\mathcal{L}_3} \theta_\kappa \left(\left| D_{[T_0, T]}^{(a,b),c} \right| \right) d\nu \ll_{\kappa,\zeta,\theta} c^{-2} \cdot b \cdot (\ln \ln T)^{3+\kappa} \cdot \left(\mathcal{L}([T_0, T]) + \mathcal{M}([T_0, T]) \right). \quad (10.1)$$

We use this estimate for a Borel–Cantelli argument below.

For $s \in \mathbb{N}$, let us denote by \mathcal{I}_s the collection of intervals

$$I_{i,j} := \left[e^{2^i j}, e^{2^i(1+j)} \right). \tag{10.2}$$

with $(i, j) \in \mathbb{N}_0^2$ satisfying $2^i(1+j) < 2^s$. The following lemma is essentially [16, Lemma 1], although this lemma takes place in a slightly different setting. We provide a proof of our version for completeness.

Lemma 10.1. *For every $1 \leq N < 2^s$, there exists a subset $\mathcal{H}_N \subset \mathcal{I}_s$ such that*

$$|\mathcal{H}_N| \leq s \quad \text{and} \quad [1, e^N) = \bigsqcup_{I \in \mathcal{H}_N} I.$$

Proof. We write

$$N = \sum_{k=1}^p 2^{n_k},$$

where $0 \leq n_1 < n_2 < \dots < n_p < s$ are integers. In particular, $p \leq s$. Let

$$N_0 = 0 \quad \text{and} \quad N_m = \sum_{k=p-m+1}^p 2^{n_k}, \quad \text{for } m = 1, \dots, p,$$

so that $N_p = N$ and

$$[1, e^N) = \bigsqcup_{m=1}^p [e^{N_{m-1}}, e^{N_m}).$$

We claim that each interval $[e^{N_k}, e^{N_{k+1}})$ for $k = 0, \dots, p - 1$ is of the form $I_{i,j}$ for some index pair $(i, j) \in \mathcal{G}_s$. To see this, first note that $[1, e^{N_1}) = I_{i_1, j_1}$, where $i_1 = n_p$ and $j_1 = 0$. More generally, we see that

$$N_{m+1} = 2^{n_{p-m}} + \dots + 2^{n_p} = 2^{n_{p-m}}(1 + 2^{n_{p-m+1}-n_{p-m}} + \dots + 2^{n_p-n_{p-m}}).$$

Hence, if we define

$$i_m = n_{p-m} \quad \text{and} \quad j_m = 2^{n_{p-m+1}-n_{p-m}} + \dots + 2^{n_p-n_{p-m}}$$

for $1 \leq m \leq p - 1$, then $N_m = 2^{i_m} j_m$ and $N_{m+1} = 2^{i_m}(1 + j_m)$, and thus $[e^{N_m}, e^{N_{m+1}}) = I_{i_m, j_m}$. In particular,

$$[1, e^N) = \bigsqcup_{m=1}^p I_{i_m, j_m}.$$

so that we establish the lemma with $\mathcal{H}_N := \{I_{i_m, j_m} : 1 \leq m \leq p\}$. □

Lemma 10.2. *For every $s \in \mathbb{N}$,*

$$|\mathcal{I}_s| \ll 2^s, \quad \sum_{I \in \mathcal{I}_s} \mathcal{L}(I) \ll s \cdot 2^{2s}, \quad \sum_{I \in \mathcal{I}_s} \mathcal{M}(I) \ll s \cdot 2^{2s}.$$

Proof. We have

$$\mathcal{L}(I_{i,j}) = (2^i j + 2^i)^2 - (2^i j)^2 \ll 2^{2i} \cdot \max(1, j),$$

and

$$\mathcal{M}(I_{i,j}) \leq s \cdot 2^i(j + 1),$$

so that

$$\sum_{I \in \mathcal{I}_s} \mathcal{L}(I) \ll \sum_{i=0}^s 2^{2i} \left(\sum_{j=0}^{2^{s-i}-1} \max(1, j) \right) \ll \sum_{i=0}^s 2^{2i} (2^{s-i})^2 \ll s \cdot 2^{2s},$$

and

$$\sum_{I \in \mathcal{I}_s} \mathcal{M}(I) \ll s \cdot \sum_{i=0}^s 2^i \left(\sum_{j=0}^{2^{s-i}-1} (j + 1) \right) \ll s \cdot \sum_{i=0}^s 2^i (2^{s-i})^2 \ll s \cdot 2^{2s},$$

as required. □

We note that later in the argument below, we choose the parameters N and s so that $e^{N-1} \leq T < e^N$ and $2^{s-1} \leq N < 2^s$, so that $\ln T \ll 2^s \ll \ln T$.

We also use a similar argument to decompose the intervals $(a, b]$. For $M \in \mathbb{N}$ and $v \in \mathbb{N}$, we write

$$a_M := M^{-\sigma} \quad \text{and} \quad b_v := 2^{-v} \quad \text{with } \sigma > 0.$$

Let $t \in \mathbb{N}$. For $(i, j) \in \mathbb{N}_0 \times \mathbb{N}$ satisfying $2^i(1 + j) < 2^t$, we set

$$J_{i,j} := ((2^i(1 + j))^{-\sigma}, (2^i j)^{-\sigma}], \tag{10.3}$$

and for $i = 1, \dots, t - 1$,

$$J_{i,0} := (2^{-\sigma i}, 1]. \tag{10.4}$$

Additionally, we subdivide each of those intervals into smaller intervals using the points b_v with $v = 1, \dots, t - 1$. Let us denote the collection of all intervals that we obtain in this way by \mathcal{J}_t . Since each of the original intervals $J_{i,j}$ is subdivided into at most t subintervals, the following lemma follows immediately from Lemma 10.1.

Lemma 10.3. *For every $M \in \mathbb{N}$ with $2 \leq M < 2^t$ and $v \in \mathbb{N}$ with $M^{-\sigma} < 2^{-v}$, there exists a subset $\mathcal{G}_{M,v} \subset \mathcal{J}_t$ such that*

$$|\mathcal{G}_{M,v}| \leq t^2 \quad \text{and} \quad (a_M, b_v] = \bigsqcup_{J \in \mathcal{G}_{M,v}} J.$$

Additionally, we note that

Lemma 10.4. *Let $\rho \in (0, 1)$ and $\sigma \geq \rho^{-1}$. Then for every $t \in \mathbb{N}$,*

$$|\mathcal{J}_t| \ll t \cdot 2^t \quad \text{and} \quad \sum_{(a,b] \in \mathcal{J}_t} b^\rho \ll t.$$

Proof. We note that the number of the intervals $J_{i,j}$ is $O(2^t)$. Since each of them is subdivided into at most t subintervals, the first estimate follows. Regarding the second estimate, we observe that the sum

over intervals obtained from the subdivision of the interval $(u_1, u_2] = J_{i,j}$ is $O(u_2^{1/2})$, so that

$$\sum_{(a,b] \in \mathcal{J}_t} b^\rho \ll \sum_{i=0}^t \left(\sum_{j=1}^{2^{t-i}-1} (2^i j)^{-\sigma\rho} + 1 \right) \ll \sum_{i=0}^t 2^{-i\sigma\rho} \cdot 2^{(t-i)(1-\sigma\rho)} + t \ll t,$$

which proves the lemma. □

We recall our assumption that $a \geq (\ln T)^{-\theta}$. Later in the proof, we pick an integer M such that $a_M \leq a < a_{M-1}$ and $M < 2^t$, so that we may take the parameter t satisfying $(2^t - 1)^\sigma \leq (\ln T)^\theta$, in particular, $t \ll \ln \ln T \ll s$.

To interpolate the parameter c , we use

$$c_w := 2^{-w} \quad \text{with } w \in \mathbb{N}.$$

According to our assumptions, $c \geq \zeta^{1/2}(\ln T)^{-\theta/2}$, so that it is sufficient to consider $w \leq r$ with $r = O(\ln \ln T)$, in particular, $r \ll s$.

With these prerequisites, we are ready to set up a Borel–Cantelli argument. Let us fix $\kappa, \varepsilon > 0$, $\rho \in (0, 1)$, $\sigma \geq \rho^{-1}$, and an integer $s \geq 2$. Let $t(s), r(s) \in \mathbb{N}$ such that $t(s) \ll s$ and $r(s) \ll s$. We set

$$X_s := \left\{ \Lambda \in \mathcal{L}_3 : \sum_{w \leq r(s)} \sum_{I \in \mathcal{I}_s} \sum_{(a,b] \in \mathcal{J}_{t(s)}} c_w^2 \cdot b^{-(1-\rho)} \cdot \theta_\kappa \left(\left| D_I^{(a,b], c_w}(\Lambda) \right| \right) \geq s^{7+\kappa+\varepsilon} \cdot 2^{2s} \right\} \quad (10.5)$$

Then from Chebyshev’s inequality and (10.1),

$$\begin{aligned} \nu(X_s) &\leq \frac{1}{s^{7+\varepsilon} \cdot 2^{2s}} \sum_{w \leq r(s)} \sum_{I \in \mathcal{I}_s} \sum_{(a,b] \in \mathcal{J}_{t(s)}} c_w^2 \cdot b^{-(1-\rho)} \int_{\mathcal{L}_3} \theta_\kappa \left(\left| D_I^{(a,b], c_w} \right| \right) d\nu \\ &\ll_{\kappa, \zeta, \theta} \frac{s^{4+\kappa}}{s^{7+\kappa+\varepsilon} \cdot 2^{2s}} \left(\sum_{I \in \mathcal{I}_s} (\mathcal{L}(I) + \mathcal{M}(I)) \right) \left(\sum_{(a,b] \in \mathcal{J}_{t(s)}} b^\rho \right). \end{aligned}$$

Hence, it follows from Lemmas 10.2 and 10.4 that

$$\nu(X_s) \ll s^{-(1+\varepsilon)}. \quad (10.6)$$

Let us take an integer $1 \leq N < 2^s$, $1 \leq M < 2^t$, and $\nu \in \mathbb{N}$ such that $1/M < 2^{-\nu}$. We use the partitions of the intervals $[1, e^N)$ and $(a_M, b_\nu]$ provided by Lemmas 10.1 and 10.3. Then

$$\left| D_{[1, e^N)}^{(a_M, b_\nu], c_w} \right| = \left| \sum_{I \in \mathcal{H}_N} \sum_{J \in \mathcal{G}_{M, \nu}} D_I^{J, c_w} \right| \leq \frac{1}{|\mathcal{H}_N| |\mathcal{G}_{M, \nu}|} \sum_{I \in \mathcal{H}_N} \sum_{J \in \mathcal{G}_{M, \nu}} |\mathcal{H}_N| |\mathcal{G}_{M, \nu}| \cdot \left| D_I^{J, c_w} \right|.$$

Since θ_κ is increasing, convex and satisfies the property (P2), we have

$$\begin{aligned} \theta_\kappa\left(\left|D_{[1,e^N]}^{(a_M,b_v],c_w}\right|\right) &\leq \theta_\kappa\left(\frac{1}{|\mathcal{H}_N||\mathcal{G}_{M,v}|} \sum_{I \in \mathcal{H}_N} \sum_{J \in \mathcal{G}_{M,v}} |\mathcal{H}_N||\mathcal{G}_{M,v}| \cdot \left|D_I^{J,c_w}\right|\right) \\ &\leq \frac{1}{|\mathcal{H}_N||\mathcal{G}_{M,v}|} \sum_{I \in \mathcal{H}_N} \sum_{J \in \mathcal{G}_{M,v}} \theta_\kappa\left(|\mathcal{H}_N||\mathcal{G}_{M,v}| \cdot \left|D_I^{J,c_w}\right|\right) \\ &\leq |\mathcal{H}_N||\mathcal{G}_{M,v}| \cdot \sum_{I \in \mathcal{H}_N} \sum_{J \in \mathcal{G}_{M,v}} \theta_\kappa\left(\left|D_I^{J,c_w}\right|\right) \ll s^3 \cdot \sum_{I \in \mathcal{H}_N} \sum_{J \in \mathcal{G}_{M,v}} \theta_\kappa\left(\left|D_I^{J,c_w}\right|\right), \end{aligned}$$

where we have used the property (P2) with $c = |\mathcal{H}_N||\mathcal{G}_{M,v}| \leq s \cdot t(s)^2 \ll s^3$ in the last third inequality. This also implies that

$$b_v^{-(1-\rho)} \cdot \theta_\kappa\left(\left|D_{[1,e^N]}^{(a_M,b_v],c_w}\right|\right) \ll s^3 \cdot \sum_{I \in \mathcal{H}_N} \sum_{(a,b) \in \mathcal{G}_{M,v}} b^{-(1-\rho)} \cdot \theta_\kappa\left(\left|D_I^{J,c_w}\right|\right).$$

Therefore, for $\Lambda \notin X_s$,

$$\theta_\kappa\left(\left|D_{[1,e^N]}^{(a_M,b_v],c_w}(\Lambda)\right|\right) \ll c_w^{-2} \cdot b_v^{1-\rho} \cdot s^{10+\kappa+\varepsilon} \cdot 2^{2s}.$$

Then by the property (P3),

$$\begin{aligned} \left|D_{[1,e^N]}^{(a_M,b_v],c_w}(\Lambda)\right| &\ll_\kappa c_w^{-1} \cdot b_v^{(1-\rho)/2} \cdot s^{(10+\kappa+\varepsilon)/2} \cdot 2^s \cdot \left(\ln^+\left(c_w^{-2} \cdot b_v^{1-\rho} \cdot s^{10+\kappa+\varepsilon} \cdot 2^s\right)\right)^{(1+\kappa)/2} \\ &\ll c_w^{-1} \cdot b_v^{(1-\rho)/2} \cdot \left(\ln^+\left(c_w^{-2} \cdot b_v^{1-\rho}\right)\right)^{(1+\kappa)/2} s^{11/2+(2\kappa+\varepsilon)/2} \cdot (\ln s)^{(1+\kappa)/2} \cdot 2^s \\ &\ll_\zeta c_w^{-1} \cdot b_v^{(1-\rho)/2} \cdot s^{6+(3\kappa+\varepsilon)/2} \cdot (\ln s)^{1+\kappa} \cdot 2^s. \end{aligned} \tag{10.7}$$

for all $s \geq 2$ and $\Lambda \notin X_s$, where we have used that

$$\ln^+(x_1 \cdot x_2) \leq 2 \cdot \ln^+ x_1 \cdot \ln^+ x_2, \quad \text{for all } x_1, x_2 \geq 0,$$

and our assumption on parameters v and w . Since $\varepsilon, \kappa > 0$ are arbitrary, the last estimate can be restated as: for all $\varepsilon > 0, s \geq 2$, and $\Lambda \notin X_s$,

$$\left|D_{[1,e^N]}^{(a_M,b_v],c_w}(\Lambda)\right| \ll_{\zeta,\varepsilon} c_w^{-1} \cdot b_v^{(1-\rho)/2} \cdot s^{6+\varepsilon} \cdot 2^s. \tag{10.8}$$

Since by (10.6)

$$\sum_{s \geq 2} \nu(X_s) < \infty,$$

it follows from Borel-Cantelli's Lemma that there exists a conull Borel set $\Psi \subset [0, 1]^2$ and a measurable map $s_o : \Psi \rightarrow \mathbb{N}$ such that for all $\underline{x} \in \Psi$ and $s \geq s_o(\underline{x})$, we have $\Lambda_{\underline{x}} \notin X_s$ and the estimate (10.8) holds:

$$\left|D_{[1,e^N]}^{(a_M,b_v],c_w}(\Lambda_{\underline{x}})\right| \ll_{\zeta,\varepsilon} c_w^{-1} \cdot b_v^{1-\rho} \cdot s^{6+\varepsilon} \cdot 2^s, \tag{10.9}$$

for all integers $1 \leq N < 2^s, 1 \leq M < 2^{t(s)}, v$ such that $1/M < 2^{-v}$, and $w \leq r(s)$.

Now for general $T \geq 1$, we denote by N_T the positive integer such that

$$e^{N_T-1} \leq T < e^{N_T}.$$

We apply the above estimate with

$$s = \lfloor \log_2(N_T + 1) \rfloor + 1, \quad t = \lfloor \log_2((\ln T)^{\theta/\sigma} + 1) \rfloor + 1, \quad r = \lfloor \log_2(\zeta^{-1/2}(\ln T)^{\theta/2}) \rfloor + 1,$$

so that

$$2^s \ll N_T, \quad 2^t \ll N_T^{\theta/\sigma}, \quad r \ll s.$$

Let

$$T_o(\underline{x}) := \min \{T \geq 1 : \lfloor \log_2(N_T + 1) \rfloor \geq s_o(\underline{x})\}.$$

For $a \in ((\ln T)^{-\theta}, 1)$, we pick an integer $M < 2^t$ such that

$$a_M \leq a < a_{M-1}.$$

Note that then

$$a_{M-1} - a_M \ll M^{-1-\sigma} \ll a^{1+\sigma^{-1}}.$$

For the parameter $c \in (0, 1/2)$ satisfying $c \geq \zeta^{1/2}(\ln T)^{-\theta/2}$ we choose $w \leq r$ such that

$$c_w \leq c < c_{w-1}.$$

We observe that

$$\Omega_{[1, e^{N_T-1}]}^{(a_{M-1}, b_v], c_w} \subset \Omega_{[1, T]}^{(a, b_v], c} \subset \Omega_{[1, e^{N_T}]}^{(a_M, b_v], c_{w-1}}.$$

In particular,

$$\left| \Lambda \cap \Omega_{[1, T]}^{(a, b_v], c} \right| \leq \left| \Lambda \cap \Omega_{[1, e^{N_T}]}^{(a_M, b_v], c_{w-1}} \right|,$$

so that

$$D_{[1, T]}^{(a, b_v], c} \leq D_{[1, e^{N_T}]}^{(a_M, b_v], c_{w-1}} + \text{Vol}_3\left(\Omega_{[1, e^{N_T}]}^{(a_M, b_v], c_{w-1}}\right) - \text{Vol}_3\left(\Omega_{[1, e^{N_T-1}]}^{(a_{M-1}, b_v], c_w}\right).$$

According to the volume estimates from Lemma A.1,

$$\text{Vol}_3\left(\Omega_{[1, e^{N_T}]}^{(a_M, b_v], c_{w-1}}\right) = 2N_T^2(b_v - a_M) + O_{\zeta, \theta}\left(\ln T(\ln \ln T)(b_v - a_M) + 1\right),$$

and

$$\text{Vol}_3\left(\Omega_{[1, e^{N_T-1}]}^{(a_{M-1}, b_v], c_w}\right) = 2(N_T - 1)^2(b_v - a_{M-1}) + O_{\zeta, \theta}\left(\ln T(\ln \ln T)(b_v - a_{M-1}) + 1\right).$$

Then

$$\begin{aligned} \text{Vol}_3\left(\Omega_{[1, e^{N_T}]}^{(a_M, b_v], c_{w-1}}\right) - \text{Vol}_3\left(\Omega_{[1, e^{N_T-1}]}^{(a_{M-1}, b_v], c_w}\right) &\ll_{\zeta, \theta} N_T^2(a_{M-1} - a_M) + N_T b_v + \ln T(\ln \ln T)b_v + 1 \\ &\ll (\ln T)^2 a^{1+\sigma^{-1}} + \ln T(\ln \ln T)b_v + 1. \end{aligned}$$

Applying (10.9), we obtain that for $\underline{x} \in \Psi$ and $T \geq T_o(\underline{x})$,

$$\left| D_{[1, e^{N_T}]}^{(a_M, b_v], c_w}(\Lambda_{\underline{x}}) \right| \ll_{\zeta, \varepsilon} c_w^{-1} \cdot b_v^{(1-\rho)/2} \cdot (\ln N_T)^{6+\varepsilon} \cdot \ln T.$$

Combining those estimates, we conclude that

$$D_{[1, T]}^{(a, b_v], c}(\Lambda_{\underline{x}}) \ll_{\zeta, \theta, \varepsilon} (\ln T)^2 a^{1+\sigma^{-1}} + c^{-1} \cdot b_v^{(1-\rho)/2} \cdot (\ln \ln T)^{6+\varepsilon} \cdot \ln T + 1$$

for sufficiently large T . The lower bound on $D_{[1, T]}^{(a, b_v], c}$ is proved similarly. Ultimately, we conclude that

$$\begin{aligned} \left| \Lambda_{\underline{x}} \cap \Omega_{[1, T]}^{(a, b_v], c} \right| &= \text{Vol}_3 \left(\Omega_{[1, T]}^{(a, b_v], c} \right) \\ &\quad + O_{\zeta, \theta, \varepsilon} \left((\ln T)^2 a^{1+\sigma^{-1}} + c^{-1} \cdot b_v^{(1-\rho)/2} \cdot (\ln \ln T)^{6+\varepsilon} \cdot \ln T + 1 \right) \end{aligned}$$

for sufficiently large T .

Finally, for $b \in (a, 1)$, we pick $v \leq t$ such that $b_v \leq b < b_{v-1}$. Since

$$\left| \Lambda_{\underline{x}} \cap \Omega_{[1, T]}^{(a, b], c} \right| = \left| \Lambda_{\underline{x}} \cap \Omega_{[1, T]}^{(a, b_{v-1}], c} \right| - \left| \Lambda_{\underline{x}} \cap \Omega_{[1, T]}^{(b, b_{v-1}], c} \right|,$$

and

$$\text{Vol}_3 \left(\Omega_{[1, T]}^{(a, b], c} \right) = \text{Vol}_3 \left(\Omega_{[1, T]}^{(a, b_{v-1}], c} \right) - \text{Vol}_3 \left(\Omega_{[1, T]}^{(b, b_{v-1}], c} \right),$$

we deduce that

$$\begin{aligned} \left| \Lambda_{\underline{x}} \cap \Omega_{[1, T]}^{(a, b], c} \right| &= \text{Vol}_3 \left(\Omega_{[1, T]}^{(a, b], c} \right) \\ &\quad + O_{\zeta, \theta, \varepsilon} \left((\ln T)^2 \cdot b^{1+\sigma^{-1}} + c^{-1} \cdot b^{(1-\rho)/2} \cdot (\ln \ln T)^{6+\varepsilon} \cdot \ln T + 1 \right) \end{aligned}$$

for sufficiently large T . We recall that $\sigma \geq \rho^{-1}$, so that we get the best estimate when $\sigma = \rho^{-1}$:

$$\begin{aligned} \left| \Lambda_{\underline{x}} \cap \Omega_{[1, T]}^{(a, b], c} \right| &= \text{Vol}_3 \left(\Omega_{[1, T]}^{(a, b], c} \right) \\ &\quad + O_{\zeta, \theta, \varepsilon} \left((\ln T)^2 \cdot b^{1+\rho} + c^{-1} \cdot b^{(1-\rho)/2} \cdot (\ln \ln T)^{6+\varepsilon} \cdot \ln T + 1 \right). \end{aligned}$$

Let us suppose that $b \leq (c \cdot \ln T)^{-2/(3\rho+1)}$. Then one checks by a direct computation that the second summand in the error term dominates the first summand, so that we get

$$\left| \Lambda_{\underline{x}} \cap \Omega_{[1, T]}^{(a, b], c} \right| = \text{Vol}_3 \left(\Omega_{[1, T]}^{(a, b], c} \right) + O_{\zeta, \theta, \varepsilon} \left(c^{-1} \cdot b^{(1-\rho)/2} \cdot (\ln \ln T)^{6+\varepsilon} \cdot \ln T + 1 \right)$$

for $T \geq T_o(\underline{x})$. This provides a nontrivial estimate when $b \geq (c \cdot \ln T)^{-2/(\rho+1)}$. On the other hand, when $b \geq (c \cdot \ln T)^{-2/(3\rho+1)}$, we get the bound:

$$\left| \Lambda_{\underline{x}} \cap \Omega_{[1, T]}^{(a, b], c} \right| = \text{Vol}_3 \left(\Omega_{[1, T]}^{(a, b], c} \right) + O_{\zeta, \theta, \varepsilon} \left(b^{1+\rho} \cdot (\ln \ln T)^{6+\varepsilon} \cdot (\ln T)^2 + 1 \right)$$

for $T \geq T_o(\underline{x})$. These estimates hold for all T with explicit constant depending on \underline{x} . This gives Theorem 1.2 by choosing $\eta = 2/(\rho + 1)$.

11. Proof of Theorem 1.1

To prove the corollary, we investigate existence of points in lattices

$$\Lambda_{\underline{x}} = \{(\underline{p} + q\underline{x}, q) \in \mathbb{R}^2 \times \mathbb{R} : (\underline{p}, q) \in \mathbb{Z}^2 \times \mathbb{Z}\}$$

contained in very thin hyperbolic strips. For $T \geq 1$, let a_T be a positive function of T . We consider the domains

$$Y_T := \left\{(\underline{u}, y) \in \mathbb{R}^2 \times \mathbb{R} : |u_1 u_2| \cdot y \leq a_T, \max(|u_1|, |u_2|) \leq \frac{1}{2}, 1 \leq y < T\right\}. \tag{11.1}$$

Our main result in this section reads as follows.

Lemma 11.1. *Suppose that a_T is nonincreasing and $a_T = o((\ln T)^{-2})$ as $T \rightarrow \infty$. Then there is a conull set $Z \subset [0, 1]^2$ and a measurable function $T : Z \rightarrow [1, \infty)$ such that for every $\underline{x} \in Z$,*

$$\Lambda_{\underline{x}} \cap Y_T = \emptyset, \quad \text{for all } T \geq T(\underline{x}).$$

Proof. By our assumption on a , the family (Y_T) is decreasing. Hence, if we can show that for almost every $\underline{x} \in [0, 1]^2$, there is $T(\underline{x}) \geq 1$ such that $\Lambda_{\underline{x}} \cap Y_{T(\underline{x})} = \emptyset$, the lemma is established. Let us consider the counting function

$$N_T(\underline{x}) := |\Lambda_{\underline{x}} \cap Y_T|, \quad \text{for } \underline{x} \in [0, 1]^2,$$

and the sets

$$Y_T(q) := \left\{\underline{x} \in \mathbb{R}^2 : |x_1 x_2| \leq \frac{a}{q}, \max(|x_1|, |x_2|) \leq \frac{1}{2}\right\}.$$

Since the map $\underline{x} \mapsto q\underline{x}$ preserves the Haar measure on the $\mathbb{R}^2/\mathbb{Z}^2$, we note that upon unwrapping the the definition of the counting function N_T ,

$$\begin{aligned} \int_{[0,1]^2} N_T(\underline{x}) d\underline{x} &= \int_{[0,1]^2} \left(\sum_{q=1}^T |(\mathbb{Z}^2 + q\underline{x}) \cap Y_T(q)| \right) d\underline{x} \\ &= \int_{[0,1]^2} \left(\sum_{q=1}^T \sum_{p \in \mathbb{Z}^2} \chi_{Y_T(q)}(p + \underline{x}) \right) d\underline{x} = \sum_{q=1}^T \text{Vol}_2(Y_T(q)). \end{aligned}$$

Furthermore,

$$\Lambda_{\underline{x}} \cap Y_T \neq \emptyset \iff N_T(\underline{x}) \geq 1,$$

so that

$$\text{Vol}_2\left(\{\underline{x} \in [0, 1]^2 : \Lambda_{\underline{x}} \cap Y_T \neq \emptyset\}\right) \leq \int_{[0,1]^2} N_T(\underline{x}) d\underline{x} = \sum_{q=1}^T \text{Vol}_2(Y_T(q)).$$

It follows from (A.2) that for sufficiently large T ,

$$\begin{aligned} \text{Vol}_2(Y_T(q)) &= \frac{1}{4} \text{Vol}_2(\Xi(4a_T/q)) = 4a_T/q \cdot (1 - \ln(4a_T/q)) \\ &= \frac{4}{q} \cdot a_T(1 - \ln(4a_T)) + \frac{4 \ln(q)}{q} \cdot a_T. \end{aligned}$$

Hence,

$$\sum_{q=1}^T \text{Vol}_2(Y_T(q)) \ll (\ln T) \cdot a_T(1 - \ln(4a_T)) + (\ln T)^2 \cdot a_T.$$

By our assumption, the right-hand side tends to zero as $T \rightarrow \infty$, and thus we can find an increasing sequence (T_k) such that

$$\sum_{k=1}^{\infty} \text{Vol}_2\left(\{\underline{x} \in [0, 1]^2 : \Lambda_{\underline{x}} \cap Y_{T_k} \neq \emptyset\}\right) < \infty.$$

By Borel-Cantelli’s Lemma, there exists a conull subset $Z \subset [0, 1]^2$ such that for every $\underline{x} \in Z$, there is an index $k(\underline{x})$ such that $\Lambda_{\underline{x}} \cap Y_{T_{k(\underline{x})}} \neq \emptyset$. This finishes the proof. \square

Proof of Theorem 1.1. Take a nonincreasing function a_T such that $a_T = o((\ln T)^{-2})$ and consider the sets $\Omega = \Omega_{[1,T]}^{(a_T, b], 1/2}$. Then

$$L(\underline{x}; b) \cap [1, T) = L(\underline{x}; a_T) \cap [1, T) \bigsqcup Q_{\Omega}(\underline{x}).$$

By Lemma 11.1, $L(\underline{x}; a_T) = \emptyset$ for all almost all $\underline{x} \in \mathbb{R}^2$ and sufficiently large T (depending on \underline{x}), and thus $L(\underline{x}; b) \cap [1, T) = Q_{\Omega}(\underline{x})$. Hence, Theorem 1.1 follows from Theorem 1.2 and the volume formula (A.4). \square

A. Volume estimates

In this section we discuss some basic facts concerning the volumes of the sets

$$\Omega = \left\{ (\underline{x}, y) \in \mathbb{R}^2 \times [1, T) : \begin{array}{l} \max(|x_1|, |x_2|) \leq c \\ a < |x_1 x_2| \cdot y \leq b \end{array} \right\}$$

with $0 < a < b < 1$ and $c \leq 1/2$. We observe that these domains can be represented in terms of more basic sets

$$\Xi(\gamma) := \{\underline{x} \in [-1, 1]^2 : |x_1 x_2| \leq \gamma\}, \quad \gamma > 0. \tag{A.1}$$

Direct computation gives

$$\text{Vol}_2(\Xi(\gamma)) = 4 \max(1, \gamma \cdot (1 - \ln \gamma)) \quad \text{for } \gamma > 0. \tag{A.2}$$

In particular, it follows from the Mean Value Theorem that

$$\text{Vol}_2(\Xi(\gamma_2)) - \text{Vol}_2(\Xi(\gamma_1)) \leq 4 |\ln \min(1, \gamma_1)| \cdot (\gamma_2 - \gamma_1) \quad \text{for } \gamma_2 > \gamma_1 > 0. \tag{A.3}$$

We observe that the y -sections

$$\Omega(y) := \{\underline{x} \in \mathbb{R}^2 : (\underline{x}, y) \in \Omega\}, \quad \text{for } y \in [1, T).$$

can be written as

$$\Omega(y) = c \cdot \left(\Xi\left(\frac{b}{c^2 \cdot y}\right) \setminus \Xi\left(\frac{a}{c^2 \cdot y}\right) \right),$$

and thus

$$\text{Vol}_2(\Omega(y)) = c^2 \cdot \left(\text{Vol}_2\left(\Xi\left(\frac{b}{c^2 \cdot y}\right)\right) - \text{Vol}_2\left(\Xi\left(\frac{a}{c^2 \cdot y}\right)\right) \right), \quad \text{for all } y \in [1, T].$$

In view of this, the following lemma can be deduced from (A.2) by a direct computation:

Lemma A.1. *For $\max(1, \frac{b}{c^2}) \leq y < T$,*

$$\begin{aligned} \text{Vol}_2(\Omega(y)) &= \frac{4 \cdot \ln y}{y} \cdot (b - a) \\ &\quad + \frac{4}{y} \cdot \left((b - a) \cdot (1 + 2 \cdot \ln c) - b \ln b + a \ln a \right), \end{aligned}$$

so that when $b \ll c^2$,

$$\begin{aligned} \text{Vol}_3(\Omega) &= 2 \cdot (\ln T)^2 \cdot (b - a) \\ &\quad + 4 \cdot \ln T \cdot \left((b - a) \cdot (1 + 2 \cdot \ln c) - b \ln b + a \ln a \right) + O(1). \end{aligned}$$

In particular, if in addition $a \gg (\ln T)^{-\theta}$ for some $\theta > 0$, then

$$\text{Vol}_3(\Omega) = 2 \cdot (\ln T)^2 \cdot (b - a) + O\left(\ln T (\ln \ln T) \cdot (b - a) + 1\right). \tag{A.4}$$

B. An auxiliary double sum

This is a largely technical section where we collect some estimates on certain multiparameter sums that are used in Section 7. This part can be safely skipped on a first read.

We begin with a simple observation. For $\underline{u} = (u_1, u_2) \in \mathbb{R}_+^2$, let

$$N(\underline{u}) := |\mathbb{Z}^d \cap ([-u_1, u_1] \times [-u_2, u_2])|. \tag{B.1}$$

A simple counting argument that we leave to the reader shows that $N(\underline{u}) \ll G(\underline{u})$, where

$$G(\underline{u}) = \begin{cases} 1 & \text{if } \lfloor \underline{u} \rfloor < 1 \\ \lfloor \underline{u} \rfloor & \text{if } \lfloor \underline{u} \rfloor < 1 \leq \lceil \underline{u} \rceil \\ u_1 u_2 & \text{if } 1 \leq \lfloor \underline{u} \rfloor \end{cases}. \tag{B.2}$$

Let us fix a constant $M > 0$ for the rest of the section. For $t = (t_1, t_2) \in \mathbb{R}_+^2$, we define the function

$$F_t(q) = \frac{G(2Mqe^{-t_1}, 2Mqe^{-t_2})}{q^2} \cdot e^{-(t_1+t_2)}, \quad q \geq 1. \tag{B.3}$$

The explicit formula for G above tells us that

$$F_t(q) = \begin{cases} \frac{e^{-(t_1+t_2)}}{q^2} & \text{if } q < \frac{e^{\lfloor t \rfloor}}{2M} \\ 2M \frac{e^{-(2\lfloor t \rfloor + \lceil t \rceil)}}{q} & \text{if } \frac{e^{\lfloor t \rfloor}}{2M} \leq q < \frac{e^{\lceil t \rceil}}{2M} \\ 4M^2 e^{-2(t_1+t_2)} & \text{if } \frac{e^{\lceil t \rceil}}{2M} \leq q \end{cases}. \tag{B.4}$$

Our main goal in this section is to prove the following upper bound on an auxiliary double sum which involves the function F_t .

Lemma B.1. Let $0 < \alpha < \beta \leq M$ and let $t = (t_1, t_2) \in \mathbb{R}_+^2$. Define

$$\alpha_t = \alpha \cdot e^{t_1+t_2} \quad \text{and} \quad \beta_t = \beta \cdot e^{t_1+t_2}$$

and suppose that $\alpha < 1, \alpha_t \geq 1$. Then,

$$\sum_{q_1, q_2 = \alpha_t}^{\beta_t} F_t \left(\frac{\max(q_1, q_2)}{\gcd(q_1, q_2)} \right) \ll_M e^{-(t_1+t_2)} + (\beta - \alpha) \cdot \max \left(1, \ln \left(\frac{\beta}{\alpha} \right) \right) \cdot \max(1, t_1 + t_2),$$

where the implicit constants depend only on M .

Remark B.2. We adopt the following sum convention: If $1 \leq \gamma < \delta$, and m and n are integers such that

$$m < \gamma \leq m + 1 \quad \text{and} \quad n \leq \delta < n + 1,$$

then $\sum_{q=\gamma}^{\delta} := \sum_{q=m+1}^n$, where the right-hand side is defined to be zero if $m = n$.

B.1. Proof of Lemma B.1

The following standard estimates will be used in the proof. For $1 \leq \gamma < \gamma + 1 < \delta$ we have

$$\sum_{q=\gamma}^{\delta} \frac{1}{q} = \ln \left(\frac{\delta}{\gamma} \right) + \mathcal{O} \left(\frac{1}{\gamma} \right), \tag{B.5}$$

where the implicit constants are independent of γ and δ . In addition, the following elementary upper bound on the sum-of-divisors function holds

$$\frac{1}{n} \sum_{\substack{m=1 \\ m|n}}^n m \ll \ln(n). \tag{B.6}$$

Let us begin with the proof. We fix $t = (t_1, t_2) \in \mathbb{R}_+^2$ and $0 < \alpha < \beta \leq M$ such that

$$1 > \alpha \quad \text{and} \quad \alpha_t \geq 1 \quad \text{and} \quad e^{\lceil t \rceil} > 2M,$$

and we want to bound the double sum

$$S_t(\alpha, \beta) := \sum_{q_1, q_2 = \alpha_t}^{\beta_t} F_t \left(\frac{\max(q_1, q_2)}{\gcd(q_1, q_2)} \right)$$

from above. Note that $F_t(1) \ll_M e^{-(t_1+t_2)}$.

We first consider the case when $\beta_t - \alpha_t < 1$. Then, the double sum above contains at most one term (necessarily with $q_1 = q_2$), and thus

$$S_t(\alpha, \beta) \leq F_t(1) \ll_M e^{-(t_1+t_2)}. \tag{B.7}$$

Let us from now on assume that $\beta_t - \alpha_t \geq 1$, and define the sets

$$\begin{aligned} \mathcal{E}_0 &:= \left\{ (q_1, q_2) \in [\alpha_t, \beta_t]^2 \cap \mathbb{N}^2 : \frac{\max(q_1, q_2)}{\gcd(q_1, q_2)} < \frac{e^{\lfloor t \rfloor}}{2M} \right\}, \\ \mathcal{E}_1 &:= \left\{ (q_1, q_2) \in [\alpha_t, \beta_t]^2 \cap \mathbb{N}^2 : \frac{e^{\lfloor t \rfloor}}{2M} \leq \frac{\max(q_1, q_2)}{\gcd(q_1, q_2)} < \frac{e^{\lceil t \rceil}}{2M} \right\}, \\ \mathcal{E}_2 &:= \left\{ (q_1, q_2) \in [\alpha_t, \beta_t]^2 \cap \mathbb{N}^2 : \frac{e^{\lceil t \rceil}}{2M} \leq \frac{\max(q_1, q_2)}{\gcd(q_1, q_2)} \right\}, \end{aligned}$$

and the functions

$$S_t^{(k)}(\alpha, \beta) := \sum_{(q_1, q_2) \in \mathcal{E}_k} F_t \left(\frac{\max(q_1, q_2)}{\gcd(q_1, q_2)} \right), \quad \text{for } k = 0, 1, 2.$$

Note that

$$S_t(\alpha, \beta) = S_t^{(0)}(\alpha, \beta) + S_t^{(1)}(\alpha, \beta) + S_t^{(2)}(\alpha, \beta).$$

We will estimate these three partial sums separately below.

An upper bound for $S_t^{(0)}(\alpha, \beta)$

Note that if $\gcd(q_1, q_2) = d$, then

$$\left(\frac{\gcd(q_1, q_2)}{\max(q_1, q_2)} \right)^2 = \frac{d^2}{\max(q_1, q_2)^2} \leq \frac{d}{q_1 \cdot q_2'},$$

where $q_2 = d \cdot q_2'$. Hence,

$$\begin{aligned} S_t^{(0)}(\alpha, \beta) &= \sum_{(q_1, q_2) \in \mathcal{E}_0} \left(\frac{\gcd(q_1, q_2)}{\max(q_1, q_2)} \right)^2 \cdot e^{-(t_1+t_2)} \leq \sum_{q_1, q_2 = \alpha_t}^{\beta_t} \left(\frac{\gcd(q_1, q_2)}{\max(q_1, q_2)} \right)^2 \cdot e^{-(t_1+t_2)} \\ &\leq 2 \cdot \sum_{q_1 = \alpha_t}^{\beta_t} \left(\sum_{\substack{d=1 \\ d|q_1}}^{q_1} \sum_{\substack{q_2' = \alpha_t/d}}^{\beta_t/d} \frac{d}{q_1 \cdot q_2'} \right) \cdot e^{-(t_1+t_2)}. \end{aligned}$$

By (B.5) (with $\gamma = \frac{\alpha_t}{d}$ and $\delta = \frac{\beta_t}{d}$), we now see that

$$S_t^{(0)}(\alpha, \beta) \ll \sum_{q_1 = \alpha_t}^{\beta_t} \left(\frac{1}{q_1} \sum_{\substack{q_1=1 \\ d|q_1}}^{q_1} d \right) \cdot \left(1 + \ln \left(\frac{\beta}{\alpha} \right) \right) \cdot e^{-(t_1+t_2)},$$

and by (B.6),

$$\sum_{q=\alpha_t}^{\beta_t} \left(\frac{1}{q} \sum_{\substack{d=1 \\ d|q}} d \right) \cdot e^{-(t_1+t_2)} \ll \left(\sum_{q=\alpha_t}^{\beta_t} \ln(q) \right) \cdot e^{-(t_1+t_2)} \leq (\beta_t - \alpha_t) \cdot \ln(\beta_t) \cdot e^{-(t_1+t_2)} \\ \ll_M (\beta - \alpha) \cdot \max(1, t_1 + t_2),$$

since $\beta \leq M$. We conclude that

$$S_t^{(0)}(\alpha, \beta) \ll (\beta - \alpha) \cdot \max\left(1, \ln\left(\frac{\beta}{\alpha}\right)\right) \cdot \max(1, t_1 + t_2). \tag{B.8}$$

An upper bound for $S_t^{(1)}(\alpha, \beta)$

To simplify notation, let us assume that $t_1 \leq t_2$ so that $\lfloor t \rfloor = t_1$ and $\lceil t \rceil = t_2$. Then,

$$S_t^{(1)}(\alpha, \beta) = \left(\sum_{(q_1, q_2) \in \mathcal{E}_1} \frac{\gcd(q_1, q_2)}{\max(q_1, q_2)} \right) \cdot e^{-(2t_1+t_2)},$$

If $(q_1, q_2) \in \mathcal{E}_1$, then

$$\frac{e^{t_1}}{2M} \leq \frac{\max(q_1, q_2)}{\gcd(q_1, q_2)} \quad \text{and} \quad \max(q_1, q_2) \leq \beta_t = \beta \cdot e^{t_1+t_2},$$

and thus $\gcd(q_1, q_2) \leq 2M \cdot \beta \cdot e^{t_2}$. Let $\gamma_t = \min(\alpha_t, 2M \cdot \beta \cdot e^{t_2})$, and note that

$$S_t^{(1)}(\alpha, \beta) \leq \sum_{d=1}^{\gamma_t} \left(\sum_{\substack{q_1, q_2=\frac{\alpha_t}{d}}}^{\frac{\beta_t}{d}} \frac{1}{\max(q_1, q_2)} \right) \cdot e^{-(2t_1+t_2)} \\ + \sum_{d=\alpha_t}^{\beta_t} \left(\sum_{\substack{q_1, q_2=1}}^{\frac{\beta_t}{d}} \frac{1}{\max(q_1, q_2)} \right) \cdot e^{-(2t_1+t_2)} \\ \ll \left(\sum_{d=1}^{\gamma_t} \frac{\beta_t - \alpha_t}{d} \right) \cdot e^{-(2t_1+t_2)} + \left(\sum_{d=\alpha_t}^{\beta_t} \frac{\beta_t}{d} \right) \cdot e^{-(2t_1+t_2)} \\ \ll (\beta - \alpha) \cdot (1 + \ln(\gamma_t)) \cdot e^{-t_1} + \beta \cdot \left(\ln\left(\frac{\beta}{\alpha}\right) + \mathcal{O}\left(\frac{1}{\alpha_t}\right) \right) \cdot e^{-t_1},$$

where we in the last inequality have used (B.5) (with parameters $\gamma = 1, \delta = \gamma_t$ and $\gamma = \alpha_t, \delta = \beta_t$). Since

$$\ln(\gamma_t) \ll 1 + t_2,$$

we have

$$(\beta - \alpha) \cdot (1 + \ln(\gamma_t)) \cdot e^{-t_1} \ll_M (\beta - \alpha) \cdot \max(1, t_1 + t_2) \cdot e^{-\lfloor t \rfloor}.$$

For the term

$$\beta \cdot \left(\ln\left(\frac{\beta}{\alpha}\right) + \mathcal{O}\left(\frac{1}{\alpha_t}\right) \right) \cdot e^{-t_1},$$

we consider two separate cases. If $\beta/\alpha \geq e$, then $\ln(\beta/\alpha)$ dominates, and since $\beta - \alpha \gg \beta$, we find

$$\beta \cdot \left(\ln\left(\frac{\beta}{\alpha}\right) + \mathcal{O}\left(\frac{1}{\alpha_t}\right) \right) \cdot e^{-t_1} \ll_M (\beta - \alpha) \cdot \ln\left(\frac{\beta}{\alpha}\right) \cdot e^{-t_1}.$$

If $\beta/\alpha < e$, on the other hand, we deduce

$$\beta \cdot \left(\ln\left(\frac{\beta}{\alpha}\right) + \mathcal{O}\left(\frac{1}{\alpha_t}\right) \right) \cdot e^{-t_1} \ll_M \beta \cdot \ln\left(\frac{\beta}{\alpha}\right) \cdot e^{-t_1} + \frac{\beta}{\alpha} \cdot e^{-2t_1-t_2}.$$

On observing that for $1 < \beta/\alpha \leq e$

$$\beta \cdot \ln\left(\frac{\beta}{\alpha}\right) \ll (\beta - \alpha),$$

we now obtain

$$\beta \cdot \left(\ln\left(\frac{\beta}{\alpha}\right) + \mathcal{O}\left(\frac{1}{\alpha_t}\right) \right) \cdot e^{-t_1} \ll_M (\beta - \alpha) \cdot e^{-t_1} + e^{-2t_1-t_2}.$$

Then, by combining the previous estimates, we conclude that

$$S_t^{(1)}(\alpha, \beta) \ll_M (\beta - \alpha) \cdot \max\left(1, \ln\left(\frac{\beta}{\alpha}\right)\right) \cdot \max(1, t_1 + t_2) \cdot e^{-\lfloor t \rfloor} + e^{-(t_1+t_2)}. \tag{B.9}$$

An upper bound for $S_t^{(2)}(\alpha, \beta)$

The following crude estimate will suffice:

$$\begin{aligned} S_t^{(2)}(\alpha, \beta) &= \left(\sum_{(q_1, q_2) \in \mathcal{E}_2} 1 \right) \cdot e^{-2(t_1+t_2)} \leq \left(\sum_{q_1=\alpha_t}^{\beta_t} \sum_{q_2=\alpha_t}^{\beta_t} 1 \right) \cdot e^{-2(t_1+t_2)} \\ &\ll (\beta_t - \alpha_t)^2 \cdot e^{-2(t_1+t_2)} \ll_M (\beta - \alpha)^2 \ll \beta - \alpha, \end{aligned} \tag{B.10}$$

since $\beta \leq M$.

Putting it all together

If we now combine (B.7), (B.8), (B.9) and (B.10), we get

$$S_t(\alpha, \beta) \ll_M e^{-(t_1+t_2)} + (\beta - \alpha) \cdot \max\left(1, \ln\left(\frac{\beta}{\alpha}\right)\right) \cdot \max(1, t_1 + t_2),$$

where the implicit constants are independent of α and β , and hence the proof of Lemma B.1 is complete.

Acknowledgements. We would like to thank Shunsuke Usuki for pointing out issues in a previous version of the paper.

Competing interests. The authors have no competing interest to declare.

References

- [1] M. Björklund and A. Gorodnik, ‘Central limit theorems for Diophantine approximants’, *Math. Ann.* **374**(3–4) (2019), 1371–1437.
- [2] M. Björklund and A. Gorodnik, ‘Effective multiple equidistribution of translated measures’, *Int. Math. Res. Not. IMRN* **1** (2023), 210–242.
- [3] M. Björklund and A. Gorodnik, ‘Central limit theorems for generic lattice point counting’, *Selecta Math. (N.S.)* **29**(1) (2023), Paper No. 12, 44 pp.
- [4] A. Chow and N. Technau, *Littlewood and Duffin-Schaffer-type problems in Diophantine approximation*. To appear in Memoirs of American Mathematical Society, <https://arxiv.org/abs/2010.09069>.
- [5] S. Chow and N. Technau, ‘Counting multiplicative approximations’, *Ramanujan J.* **62**(1) (2023), 241–250.
- [6] S. G. Dani, ‘On orbits of unipotent flows on homogeneous spaces, II’, *Ergodic Theory Dynamical Systems* **6** (1986), 167–182.
- [7] A. Eskin, G. Margulis and S. Mozes, ‘Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture’, *Ann. Math.* **147**(1) (1998), 93–141.
- [8] P. Gallagher, ‘Metric simultaneous diophantine approximation’, *J. London Math. Soc.* **37** (1962), 387–390.
- [9] D. Y. Kleinbock and G. A. Margulis, *Bounded Orbits of Nonquasiunipotent Flows on Homogeneous Spaces*. Sinai’s Moscow Seminar on Dynamical Systems, 141–172, Amer. Math. Soc. Transl. Ser. 2, 171, Adv. Math. Sci., 28, Amer. Math. Soc., Providence, RI, 1996.
- [10] D. Y. Kleinbock and G. A. Margulis, ‘Flows on homogeneous spaces and Diophantine approximation on manifolds’, *Ann. Math.* **148**(1) (1998), 339–360.
- [11] D. Y. Kleinbock and G. A. Margulis, *On Effective Equidistribution of Expanding Translates of Certain Orbits in the Space of Lattices*. Number theory, analysis and geometry, 385–396, Springer, New York, 2012.
- [12] D. Kleinbock and A. Rao, ‘A zero-one law for uniform Diophantine approximation in Euclidean norm’, *Int. Math. Res. Not. IMRN* **8** (2022), 5617–5657.
- [13] D. Kleinbock and A. Rao, ‘Weighted uniform diophantine approximation of systems of linear forms’, *Pure Appl. Math. Q.* **18**(3) (2022), 1095–1112.
- [14] D. Kleinbock, R. Shi and B. Weiss, ‘Pointwise equidistribution with an error rate and with respect to unbounded functions’, *Math. Ann.* **367**(1–2) (2017), 857–879.
- [15] D. Kleinbock and N. Wadleigh, ‘A zero-one law for improvements to Dirichlet’s Theorem’, *Proc. Amer. Math. Soc.* **146**(5) (2018), 1833–1844.
- [16] W. Schmidt, ‘A metrical theorem in diophantine approximation’, *Canadian J. Math.* **12** (1960), 619–631. A metrical theorem in geometry of numbers.
- [17] W. Schmidt, ‘A metrical theorem in geometry of numbers’, *Trans. Amer. Math. Soc.* **95** (1960), 516–529.
- [18] W. Schmidt, ‘Asymptotic formulae for point lattices of bounded determinant and subspaces of bounded height’, *Duke Math. J.* **35** (1968), 327–339.
- [19] N. Shah, ‘Limit distributions of polynomial trajectories on homogeneous spaces’, *Duke Math. J.* **75** (1994), 711–732.
- [20] R. Shi, ‘Expanding cone and applications to homogeneous dynamics’, *Int. Math. Res. Not. IMRN* **9** (2021), 7060–7095.
- [21] Y. Wang and K. R. Yu, ‘A note on some metrical theorems in Diophantine approximation’, *Chinese Ann. Math.* **2**(1) (1981), 1–12, <https://www.overleaf.com/project/62e942de3cafd37a59f23419>.
- [22] M. Widmer, ‘Asymptotic Diophantine approximation: The multiplicative case’, *Ramanujan J.* **43** (2017), 83–93.