The n-point correlations between values of a linear form

Jens Marklof*†

Institut des Hautes Études Scientifique, 35, route de Chartres
F-91440 Bures-sur-Yvette, France

and

Laboratoire de Physique Théorique et Modèles Statistiques‡
Bat. 100, F-91405 Orsay Cedex, France

With an Appendix

The number of solutions of simultaneous quadratic equations

by Zeév Rudnick§

Raymond and Beverly Sackler School of Mathematical Sciences
Tel Aviv University, Tel Aviv 69978, Israel

July 1998, revised June 1999
Ergodic Theory and Dynamical Systems, to appear

Abstract. We show that the n-point correlation function for the fractional parts of a random linear form in m variables has a limit distribution with power-like tail. The existence of the limit distribution follows from the mixing property of flows on $\mathrm{SL}(m+1,\mathbb{R})/\mathrm{SL}(m+1,\mathbb{Z})$. Moreover, we prove similar limit theorems (i) for the probability to find the fractional part of a random linear form close to zero, and (ii) also for related trigonometric sums. For large m all of the above limit distributions approach the classical distributions for independent uniformly distributed random variables.

*Supported by the European Post-Doctoral Institute for the Mathematical Sciences and the European Commission (TMR Marie Curie Grant)

†Permanent address: School of Mathematics, University of Bristol, Bristol BS8 1TW, U.K. (Email: J.Marklof@Bristol.ac.uk)

‡UMR 8626, Université Paris XI – CNRS

§Supported in part by grants from the Israel Science Foundation and the Hermann Minkowski Center for Geometry at Tel Aviv University

Contents

1 Introduction

Consider an infinite sequence of numbers $\lambda_1, \lambda_2, \ldots$ which is uniformly distributed mod 1, i.e.

$$\lim_{N\to\infty} \frac{\#\{j \in [1,N] : \lambda_j \in I\}}{N} = |I| \tag{1.1}$$

for any subinterval I of the unit circle. Correlations on the scale of the average spacing $(1/N)$ between the λ_j can be measured by the *n-point correlation densities* [25, 30, 16],

$$R_n(\mathcal{S}, N) = \frac{1}{N} \sum_{\substack{j_1,\ldots,j_n=1 \\ j_\mu \neq j_\nu \forall \mu\neq\nu}}^{N} \prod_{a=1}^{n-1} \delta_N\big(s_a - N(\lambda_{j_a} - \lambda_{j_{a+1}})\big) \tag{1.2}$$

where $\mathcal{S} = (s_1, \ldots, s_{n-1}) \in \mathbb{R}^{n-1}$, and $\delta_N(x)$ is the Dirac mass on the circle of length N; it may be identified with a periodic superposition of Dirac masses on $\mathbb{R}$ via $\delta_N(x) = \sum_{\nu\in\mathbb{Z}} \delta(x + N\nu)$, where $\delta(x)$ denotes the standard Dirac delta distribution. The *two-point correlation density*

$$R_2(s, N) = \frac{1}{N} \sum_{\substack{j_1,j_2=1 \\ j_1\neq j_2}}^{N} \delta_N\big(s - N(\lambda_{j_1} - \lambda_{j_2})\big) \tag{1.3}$$

is for instance the density of all (not just nearest neighbours, that is) spacings. It should be pointed out that the $R_n(s, N)$ are obviously *not* probability densities, but contain the complete information to determine (through some combinatorial sieving) statistical measures such as the spacing distribution between nearest (or next-to-nearest etc.) neighbours, see [25, 16, 20] for details. Put furthermore

$$\overline{R}_n(\mathcal{B}, N) = \int_{\mathcal{B}} R_n(\mathcal{S}, N)\, d\mathcal{S} \tag{1.4}$$

where $\mathcal{B}$ is some bounded domain in $\mathbb{R}^{n-1}$; this represents the number of $(n-1)$-tuple spacings $N(\lambda_{j_1} - \lambda_{j_2}, \lambda_{j_2} - \lambda_{j_3}, \ldots, \lambda_{j_{n-1}} - \lambda_{j_n}) \bmod N\mathbb{Z}^{n-1}$ in $\mathcal{B}$. We will refer to $\overline{R}_n(\mathcal{B}, N)$ as the *n-point correlation function*. In the case when $\lambda_1, \lambda_2, \ldots, \lambda_N$ are independent uniformly distributed (on $[0,1]$) random variables (IUDRV), all n-point correlation functions converge to $|\mathcal{B}|$ almost surely, i.e.

$$\lim_{N\to\infty} \overline{R}_n(\mathcal{B}, N) = |\mathcal{B}| \quad \text{almost surely}, \tag{1.5}$$

where $|\mathcal{B}|$ denotes the Lebesgue measure of $\mathcal{B}$.

Our investigation here is concerned with the limiting behaviour ($N \to \infty$) of the correlation densities in the case when the sequence $\lambda_1, \ldots, \lambda_N$ is given by the fractional parts of the linear form in $d-1$ variables,

$$L_{\boldsymbol{m}} = \alpha_1 m_1 + \cdots + \alpha_{d-1} m_{d-1}, \tag{1.6}$$

at the integers $m_j = 1, \ldots, N_j$, where $j = 1, \ldots, d-1$, and $N_1 \cdots N_{d-1} = N$. The first immediate observation will be that $\overline{R}_n(\mathcal{B}, N)$ has in general no limit as $N \to \infty$, which is due to the strong correlations between the values at integers of $L_{\boldsymbol{m}}$. Limit distributions exist only, if the coefficients α_j or the cut-off parameter N are taken to be random. This fact was proven so far only in the case $d = 2$ for the consecutive spacing distribution (Bleher [2, 3], Greenman [12]) and the distribution of values in small random intervals (Mazel and Sinai [24]), where it follows from ergodic properties of the continued fraction map (or rather a natural extension of it). Our approach will avoid the use of a higher-dimensional analogue of continued fractions, and instead exploit the connection between values of linear forms and the dynamics of flows on the quotient space $\mathrm{SL}(d, \mathbb{R}) / \mathrm{SL}(d, \mathbb{Z})$, which is well known, see e.g. [6]. The central theorem that we shall use guarantees the equidistribution of measures concentrated along the unstable fibers of the flow, and may be regarded as a generalization of the equidistribution theorem for *closed* horocycles on $\mathrm{SL}(2, \mathbb{R}) / \mathrm{SL}(2, \mathbb{Z})$ (compare [32, 9, 15, 22, 23]). It will be derived as a corollary of the mixing property of flows on $\mathrm{SL}(d, \mathbb{R}) / \mathrm{SL}(d, \mathbb{Z})$, following ideas of Eskin and McMullen [9], Kleinbock and Margulis [19], and others (see references in [19]). Equidistribution theorems of this type are well known in the theory of unipotent flows and translates of measures on homogeneous spaces, which has been developed mainly by Margulis, Dani, Ratner, Shah, Mozes and Eskin (see Eskin's ICM review [8] for references and applications of the theory to other counting problems).

One physical motivation of our studies is the fact that the spacings between the fractional parts of a linear form mod 1 are in one-to-one correspondence with the spacings between eigenvalues of a d-dimensional quantum harmonic oscillator, where $\alpha_1, \ldots, \alpha_{d-1}$ are related to the normal modes of the classical oscillator [1, 27, 2, 3, 28, 12]. A further physical motivation is the interpretation of the fractional parts of a linear form as the quasi-energies of a certain model kicked quantum system and eigenvalues of related Schrödinger operators with quasi-periodic potentials, see [24, 10] for details and references.

For recent rigorous results on sequences, which behave more generic in the sense that their correlation functions converge, e.g. to those of IUDRVs (such as values at integers of polynomials of degree greater than one) or to those of random matrix ensembles (such as the zeros of zeta functions), see for instance [16, 20, 23, 30, 31, 33, 41, 42] and the surveys [4, 34, 35].

Acknowledgments. I thank E. Bogomolny, O. Bohigas, M. Kontsevitch, Z. Rudnick and Y.M. Suhov for helpful discussions.

2 Results

Basic definitions and notations. The expressions $x \ll_a y$ and $x = O_a(y)$ both mean *there exists a constant* C_a (which may depend on some additional parameter a) *such that* $|x| \leq C_a |y|$. The relation $x \asymp y$ means $x = O(y)$ *and* $y = O(x)$.

A function f on a measurable space M is called *piecewise continuous* if there exists an open (not necessarily connected) set $U \subset M$ such that (i) f is continuous on U and (ii) the complement of U has measure zero. f is called *piecewise constant* if there exist countably many open sets $U_i \subset M$ such that (i) f is constant on each U_i and (ii) the complement of $\bigcup_i U_i$ has measure zero. An example for a piecewise continuous/constant function is the characteristic function of an open set with boundary of measure zero.

⋆

We assume here and in the sequel that in the limit $N \to \infty$ the N_j are of the same order of magnitude, i.e. more precisely $N_j = d_j N^{1/(d-1)} + O(1)$ with positive constants $d_j > 0$ and $d_1 \cdots d_{d-1} = 1$. We denote by $\mathcal{B}$ a bounded open subset in $\mathbb{R}^{n-1}$ with boundary of measure zero.

Our main results are as follows (for details see Theorems 3.2 – 3.4, 3.11, 3.18 – 3.20):

□ The correlation function $\overline{R}_n(\mathcal{B}, N)$ has in general no limit for fixed $\alpha_1, \ldots, \alpha_{d-1}$ as $N \to \infty$.

□ If $\alpha_1, \ldots, \alpha_{d-1}$ are random with continuous joint probability density $h(\alpha_1, \ldots, \alpha_{d-1})$, then $\overline{R}_n(\mathcal{B}, N)$ has a limit distribution which is independent of h, i.e.

$$\lim_{N \to \infty} \mathrm{Prob}\{\overline{R}_n(\mathcal{B}, N) > X\} = \Psi_{n,d}(X)$$

where $\Psi_{n,d}(X)$ is continuous and has a power-like tail. More precisely,

$$\Psi_{n,d}(X) \asymp X^{-\frac{d}{n-1}} \quad \text{for } X \text{ large.}$$

□ For d large, the limit distribution localizes at the value $|\mathcal{B}|$, i.e.

$$d\Psi_{n,d}(X) \overset{d \to \infty}{\Longrightarrow} \delta(X - |\mathcal{B}|)\, dX.$$

This follows from the asymptotic behaviour of the moments of $\Psi_{n,d}(X)$ which are discussed in Section 3.8. Note that $\delta(X - |\mathcal{B}|)\,dX$ is the corresponding limit distribution for IUDRVs, because their n-point correlation functions converge almost surely to $|\mathcal{B}|$. This answers a question of Greenman on the level spacing distribution of high-dimensional harmonic oscillators [13].

□ Suppose now $\alpha_1, \ldots, \alpha_{d-1}$ are fixed and N is a random variable with a suitable probability distribution on $[1, M]$ (see Theorem 3.3 for details). Then, for almost all $\alpha_1, \ldots, \alpha_{d-1}$ (with respect to Lebesgue measure), we have

$$\lim_{M\to\infty} \mathrm{Prob}\{\overline{R}_n(\mathcal{B}, \,\cdot\,) > X\} = \Psi_{n,d}(X)$$

where $\Psi_{n,d}(X)$ is the same function as in the previous statements.

⋆

The above statements generalize analogous results of Bleher [2, 3] and Greenman [12] for the consecutive spacing distribution in the case $d = 2$. The two-dimensional case is easier to handle, because the spacings between nearest neighbours are determined by the continued fraction expansion of α_1, see [39] for a survey of results, some of which were later rediscovered in the Physics literature [27, 2, 3, 28, 12]. A particularly remarkable observation is that, for fixed α_1, the nearest-neighbour spacings can take at most three distinct values [40, 39]; this fact does not generalize to $d > 2$, where the number of distinct spacings is in general unbounded, compare [11] for details in the case $d = 3$.

We shall furthermore generalize a result of Mazel and Sinai [24] obtained in the case $d = 2$ for the distribution of fractional parts of linear forms in small subintervals $[\xi, \xi + \frac{\sigma}{N}]$ of the unit circle. Let us denote by $\mathcal{N}_\sigma(\xi, N)$ the number of points in such an interval, i.e.

$$\mathcal{N}_\sigma(\xi, N) = \sum_{j=1}^{N} \sum_{\nu\in\mathbb{Z}} \chi_\sigma\big(N(\lambda_j - \xi + \nu)\big), \tag{2.1}$$

where χ_σ is the characteristic function of $(0, \sigma]$.

Again assuming $\alpha_1, \ldots, \alpha_{d-1}$ are random with continuous joint probability density $h(\alpha_1, \ldots, \alpha_{d-1})$, we shall prove the following statements (compare Theorems 4.2 – 4.5):

□ If ξ is a random variable uniformly distributed in $[0, 1]$, $\mathcal{N}_\sigma(\xi, N)$ has a limit distribution which is independent of h, i.e.

$$\lim_{N\to\infty} \mathrm{Prob}\{\mathcal{N}_\sigma(\,\cdot\,, N) = K\} = P_K(\sigma)$$

for $K = 0, 1, 2, \dots$, where

$$\sum_{K=X}^{\infty} P_K(\sigma) \asymp X^{-(d+1)} \quad \text{for } X \text{ large.}$$

□ If $\xi = 0$, then $\mathcal{N}_\sigma(0, N)$ has a different limit distribution, i.e.

$$\lim_{N\to\infty} \mathrm{Prob}\{\mathcal{N}_\sigma(0, N) = K\} = P_{K,0}(\sigma)$$

for $K = 0, 1, 2, \dots$, where

$$\sum_{K=X}^{\infty} P_{K,0}(\sigma) \asymp X^{-d} \quad \text{for } X \text{ large.}$$

□ The expectation value of $P_{K,0}(\sigma)$ and expectation value and first moment of $P_K(\sigma)$ coincide with those for IUDRVs, but higher moments are different. For d large, however, both limit distributions are asymptotically Poisson, i.e.

$$P_K(\sigma) \overset{d\to\infty}{\Longrightarrow} \frac{\sigma^K}{K!}\, \mathrm{e}^{-\sigma}, \qquad P_{K,0}(\sigma) \overset{d\to\infty}{\Longrightarrow} \frac{\sigma^K}{K!}\, \mathrm{e}^{-\sigma}.$$

⋆

As was noted in [38, 24], the asymptotics of $\mathcal{N}_\sigma(\xi, N)$ is closely linked with the asymptotic behaviour of the trigonometric sums

$$W_N(\xi) = \frac{1}{N} \sum_{\nu=1}^{N} \sum_{j=1}^{N} \cos\big(2\pi\nu(\lambda_j + \xi)\big), \tag{2.2}$$

for which we obtain the following limit theorems, cf. [38] for the case $d = 2$ (assume $\alpha_1, \alpha_2, \dots$ are random as before):

□ If ξ is a random variable uniformly distributed in $[0, 1]$, $W_N(\xi)$ has a limit distribution, i.e. in particular

$$\lim_{N\to\infty} \mathrm{Prob}\{a < W_N < b\} = \Phi_d(a, b)$$

except possibly at the discontinuities of Φ_d (of which there are at most countably many), and we have the asymptotics

$$1 - \Phi_d(-X, X) \asymp X^{-(d+1)} \quad \text{for } X \text{ large.}$$

□ If $\xi = 0$, $d > 2$,

$$\lim_{N\to\infty} \mathrm{Prob}\{a < W_N(0) < b\} = \Phi_{d,0}(a, b)$$

except possibly at the discontinuities of $\Phi_{d,0}$ (of which there are at most countably many), and we have the asymptotics

$$1 - \Phi_{d,0}(-X, X) \asymp X^{-d} \quad \text{for } X \text{ large.}$$

In the case $d = 2$ we are only able to prove the above result, if the cut-off in the sum $W_N(0)$ is smoothed; see Section 4.1 for details.

The proof of these limit theorems (cf. Theorems 4.6 – 4.9) is essentially the same as for $\mathcal{N}_\sigma(\xi, N)$, combined with methods employed in [21].

⋆

The moments of the limit distributions discussed here are related to the number of solutions of simultaneous quadratic diophantine equations. These can be solved directly by elementary methods from number theory, as illustrated for the second moment of $\Phi_{d,0}$ in the Appendix by Zeév Rudnick; compare Katznelson [17, 18] for related results. Since, however, only finitely many of these moments do not diverge, this elementary approach gives no information on the existence of the full limit distributions. Here it is indeed necessary to apply the above-mentioned equidistribution theorems on $\mathrm{SL}(d, \mathbb{R}) / \mathrm{SL}(d, \mathbb{Z})$. From the latter approach, explicit formulas for the moments can be derived by means of classical reduction theory, see Sections 3.7, 3.8 and 4.4.

3 Correlation functions

3.1 Two-point correlations

The two-point correlation density has the representation (recall def. (1.3))

$$R_2^{\boldsymbol{\alpha}}(s, N) = \frac{1}{N} \sum_{m_1, n_1 = 1}^{N_1} \cdots \sum_{\substack{m_{d-1}, n_{d-1} = 1 \\ \boldsymbol{m} \neq \boldsymbol{n}}}^{N_{d-1}} \sum_{\nu \in \mathbb{Z}} \delta\big(s - N(L_{\boldsymbol{m}} - L_{\boldsymbol{n}} + \nu)\big). \tag{3.1}$$

Since the difference $k_j = m_j - n_j$ occurs exactly $N_j - |k_j|$ times, we can rewrite (3.1) as

$$R_2^{\boldsymbol{\alpha}}(s, N) = \sum_{\substack{-N_1 \leq k_1 \leq N_1 \\ \vdots \\ -N_{d-1} \leq k_{d-1} \leq N_{d-1} \\ (k_1, \ldots, k_{d-1}) \neq (0, \ldots, 0)}} \Big[\big(1 - \frac{|k_1|}{N_1}\big) \cdots \big(1 - \frac{|k_{d-1}|}{N_{d-1}}\big) \sum_{\nu \in \mathbb{Z}} \delta\big(s - N(L_{\boldsymbol{k}} + \nu)\big) \Big].$$

Let us now split up the above sums as

$$\sum_{\substack{(k_1,\dots,k_{d-1})\neq(0,\dots,0)\\ \nu}} = \sum_{(k_1,\dots,k_{d-1},\nu)\neq(0,\dots,0,0)} - \sum_{\substack{(k_1,\dots,k_{d-1})=(0,\dots,0)\\ \nu\neq 0}} ;$$

it is suggestive to use the notation $k_d = \nu$ so that, with the function

$$\tau_2(x) = \begin{cases} 1-|x| & x\in[-1,1] \\ 0 & x\notin[-1,1], \end{cases} \tag{3.2}$$

we eventually obtain

$$R_2^{\boldsymbol{\alpha}}(s,N) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d-\{\boldsymbol{0}\}} \tau_2\Big(\frac{k_1}{N_1}\Big)\cdots\tau_2\Big(\frac{k_{d-1}}{N_{d-1}}\Big)\,\delta(s-N\boldsymbol{\alpha k}) - \sum_{\nu\in\mathbb{Z}-\{0\}} \delta(s-N\nu), \tag{3.3}$$

with the abbreviation $\boldsymbol{\alpha k} = \alpha_1 k_1 + \cdots + \alpha_{d-1}k_{d-1} + k_d$. In the sequel we may ignore the second sum on the right-hand-side, since for any compact interval $\mathcal{B}$

$$\int_{\mathcal{B}} \sum_{\nu\in\mathbb{Z}-\{0\}} \delta(s-N\nu)\,ds = 0 \tag{3.4}$$

for N large enough.

The crucial observation is now that we have the representation

$$\begin{pmatrix} k_1 N_1^{-1} \\ \vdots \\ k_{d-1}N_{d-1}^{-1} \\ N\boldsymbol{\alpha k} \end{pmatrix} = \mathsf{M}\boldsymbol{k}, \tag{3.5}$$

where

$$\mathsf{M} = \begin{pmatrix} N_1^{-1} & & & \\ & \ddots & & \\ & & N_{d-1}^{-1} & \\ & & & N \end{pmatrix} \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ \alpha_1 & \cdots & \alpha_{d-1} & 1 \end{pmatrix} \tag{3.6}$$

is a matrix in $\mathrm{SL}(d,\mathbb{R})$. Let us put $\tilde{\boldsymbol{k}} = \mathsf{M}\boldsymbol{k}$ and define the distribution

$$D_2(s,\mathsf{M}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d-\{\boldsymbol{0}\}} \tau_2(\tilde{k}_1)\cdots\tau_2(\tilde{k}_{d-1})\,\delta(s-\tilde{k}_d). \tag{3.7}$$

We clearly get back to $R_2^{\boldsymbol{\alpha}}(s,N)$ if we choose M as above. In order to check that (3.7) is well-defined for arbitrary M, we integrate against the characteristic function $\chi_{\mathcal{B}}$ of some compact interval $\mathcal{B}$. It is then obvious that

$$\sum_{\boldsymbol{k}\in\mathbb{Z}^d-\{\boldsymbol{0}\}} \tau_2(\tilde{k}_1)\cdots\tau_2(\tilde{k}_{d-1})\,\chi_{\mathcal{B}}(\tilde{k}_d)$$

is a finite sum for any fixed $\mathsf{M}\in\mathrm{SL}(d,\mathbb{R})$. What is most important, $D_2(s,\mathsf{M})$ carries the following invariance property,

$$D_2(s,\mathsf{Mg}) = D_2(s,\mathsf{M}) \quad \text{for any } \mathsf{g}\in\Gamma, \tag{3.8}$$

with $\Gamma=\mathrm{SL}(d,\mathbb{Z})$, as can be readily verified. Hence $D_2(s,\mathsf{M})$ may be viewed as a distribution on the quotient space $\Sigma^d=\mathrm{SL}(d,\mathbb{R})/\Gamma$, which has finite volume with respect to Haar measure [37, 29], cf. also Section 3.7.

From relation (3.8) it is immediately clear that we cannot expect the two-point correlation density to converge. The exact asymptotic behaviour is determined by the geometry of Σ^d, which we shall discuss in more detail later on. Let us first see how the invariance properties of the two-point correlations carry over to higher correlation densities.

3.2 n-point correlations

Recall that the n-point correlation density is given by, compare (1.2),

$$R_n^{\boldsymbol{\alpha}}(\mathcal{S},N) = \frac{1}{N}\sum_{m_1^1,\ldots,m_1^n=1}^{N_1}\cdots\sum_{\substack{m_{d-1}^1,\ldots,m_{d-1}^n=1\\ \boldsymbol{m}^\mu\neq\boldsymbol{m}^\nu\forall\mu\neq\nu}}^{N_{d-1}}\;\sum_{\nu^1,\ldots,\nu^{n-1}\in\mathbb{Z}}\times$$
$$\times\prod_{a=1}^{n-1}\delta\big(s^a-N(L_{\boldsymbol{m}^a}-L_{\boldsymbol{m}^{a+1}}+\nu^a)\big), \tag{3.9}$$

where $\mathcal{S}=(s^1,\ldots,s^{n-1})$. In the sequel we shall always use bold letters and lower indices for vectors in configuration space (such as $\boldsymbol{m}=(m_1,m_2,\ldots)$) and script letters and upper indices for vectors in correlation space (such as $\mathcal{S}=(s^1,s^2,\ldots)$). Furthermore we write $\boldsymbol{x}^a=(x_1^a,x_2^a,\ldots)$ and $\mathcal{X}_j=(x_j^1,x_j^2,\ldots)$.

Let us now put

$$\boldsymbol{k}^1=\boldsymbol{m}^1-\boldsymbol{m}^2,\;\ldots,\;\boldsymbol{k}^{n-1}=\boldsymbol{m}^{n-1}-\boldsymbol{m}^n,$$

which in turn means that

$$\boldsymbol{m}^1=\boldsymbol{k}^1+\cdots+\boldsymbol{k}^{n-1}+\boldsymbol{m}^n,\;\ldots,\;\boldsymbol{m}^{n-1}=\boldsymbol{k}^{n-1}+\boldsymbol{m}^n,$$

or, relating all possible pairs $\boldsymbol{m}^a$, $\boldsymbol{m}^b$ $(a < b)$,

$$\boldsymbol{m}^a = \boldsymbol{k}^a + \cdots + \boldsymbol{k}^{b-1} + \boldsymbol{m}^b.$$

It is not hard to see that the multiplicity of $\mathcal{K}_j = (k_j^1, \ldots, k_j^{n-1})$ is

$$N_j - \max_{1\le a\le b\le n-1} \Big| \sum_{i=a}^{b} k_j^i \Big|,$$

and the n-point correlation density can thus be written as (compare (3.3) and put $k_d^a = \nu^a$)

$$R_n^{\boldsymbol{\alpha}}(\mathcal{S},N) = \sum_{\boldsymbol{k}^1,\ldots,\boldsymbol{k}^{n-1}\in\mathbb{Z}^d}^{(3.11)} \tau_n\Big(\frac{\mathcal{K}_1}{N_1}\Big)\cdots\tau_n\Big(\frac{\mathcal{K}_{d-1}}{N_{d-1}}\Big) \times \\ \times \delta(s^1 - N\boldsymbol{\alpha}\boldsymbol{k}^1)\cdots\delta(s^{n-1} - N\boldsymbol{\alpha}\boldsymbol{k}^{n-1}), \quad (3.10)$$

where the sum is restricted to

$$\sum_{j=a}^{b} \begin{pmatrix} k_1^j \\ \vdots \\ k_{d-1}^j \end{pmatrix} \neq \mathbf{0} \quad \text{for all } 1 \le a \le b \le n-1 \quad (3.11)$$

(no restrictions on k_d^j), and the function $\tau_n : \mathbb{R}^{n-1} \to \mathbb{R}^+$ is defined by

$$\tau_n(\mathcal{X}) = \begin{cases} 1 - \max\limits_{1\le a\le b\le n-1} \Big|\sum\limits_{i=a}^{b} x^i\Big| & \text{for } \max\limits_{1\le a\le b\le n-1} \Big|\sum\limits_{i=a}^{b} x^i\Big| \le 1 \\ 0 & \text{for } \max\limits_{1\le a\le b\le n-1} \Big|\sum\limits_{i=a}^{b} x^i\Big| > 1. \end{cases} \quad (3.12)$$

This function is clearly continuous and of compact support. Since we view the n-point density $R_n^{\boldsymbol{\alpha}}(\mathcal{S},N)$ as a distribution tested against compactly supported functions, it is not hard to see that — as before for $R_2(s,N)$ — we have

$$R_n^{\boldsymbol{\alpha}}(\mathcal{S},N) = \sum_{\boldsymbol{k}^1,\ldots,\boldsymbol{k}^{n-1}\in\mathbb{Z}^d}^{(3.14)} \tau_n\Big(\frac{\mathcal{K}_1}{N_1}\Big)\cdots\tau_n\Big(\frac{\mathcal{K}_{d-1}}{N_{d-1}}\Big) \times \\ \times \delta(s^1 - N\boldsymbol{\alpha}\boldsymbol{k}^1)\cdots\delta(s^{n-1} - N\boldsymbol{\alpha}\boldsymbol{k}^{n-1}), \quad (3.13)$$

for N large enough, where the sum is now restricted to

$$\sum_{j=a}^{b} \boldsymbol{k}^j \neq \mathbf{0} \quad \text{for all } 1 \le a \le b \le n-1. \quad (3.14)$$

It will turn out to be convenient if we understand the $N_j \in \mathbb{R}$ as real variables, taking relation (3.13) as a definition for $R_n^{\boldsymbol{\alpha}}(\mathcal{S}, N)$ at non-integer values.

Put $\tilde{\boldsymbol{k}}^a = \mathsf{M}\boldsymbol{k}^a$ with $\mathsf{M} \in \mathrm{SL}(d, \mathbb{R})$, and also set $\tilde{\mathcal{K}}_j = (\tilde{k}_j^1, \ldots, \tilde{k}_j^{n-1})$. The correct generalization of $D_2(s, \mathsf{M})$ in (3.7) is then

$$D_n(\mathcal{S}, \mathsf{M}) = \sum_{\mathcal{K}_1,\ldots,\mathcal{K}_d \in \mathbb{Z}^{n-1}}^{(3.14)} \tau_n(\tilde{\mathcal{K}}_1) \cdots \tau_n(\tilde{\mathcal{K}}_{d-1})\, \delta^{n-1}(\mathcal{S} - \tilde{\mathcal{K}}_d), \tag{3.15}$$

with

$$\delta^{n-1}(\mathcal{X}) = \delta(x^1) \cdots \delta(x^{n-1}).$$

We find that the invariance property

$$D_n(\mathcal{S}, \mathsf{Mg}) = D_n(\mathcal{S}, \mathsf{M}) \quad \text{for any } \mathsf{g} \in \Gamma, \tag{3.16}$$

holds for general $n > 2$ as well. Notice in particular that for any $\mathsf{g} \in \Gamma$ the set of conditions

$$\sum_{j=a}^{b} (\mathsf{g}\boldsymbol{k}^j) \neq \mathbf{0} \quad \text{for all } 1 \leq a \leq b \leq n-1 \tag{3.17}$$

is equivalent to the set (3.14).

We shall be particularly interested in the integrated densities

$$\overline{R}_n^{\boldsymbol{\alpha}}(\mathcal{B}, N) = \int_{\mathcal{B}} R_n^{\boldsymbol{\alpha}}(\mathcal{S}, N)\, d^{n-1}\mathcal{S},$$

$$\overline{D}_n(\mathcal{B}, \mathsf{M}) = \int_{\mathcal{B}} D_n(\mathcal{S}, \mathsf{M})\, d^{n-1}\mathcal{S},$$

for bounded open sets $\mathcal{B}$ with boundary of measure zero. Let us summarize the results of this section in the following proposition.

Proposition 3.1 *For N large enough, we have*

$$\overline{R}_n^{\boldsymbol{\alpha}}(\mathcal{B}, N) = \overline{D}_n(\mathcal{B}, \mathsf{M}),$$

with M given by (3.6).

3.3 Limit distributions

We will now state our main results; the proofs are given in the next two sections. In the sequel put $N_j = d_j N^{1/(d-1)}$, for arbitrary constants $d_j > 0$ ($d_1 \cdots d_{d-1} = 1$), so that N is the only free parameter. (Recall the N_j need not be restricted to integers.)

Theorem 3.2 *Let $\mathcal{B}$ be a bounded open subset in $\mathbb{R}^{n-1}$ with boundary of measure zero. Suppose $\alpha_1, \ldots, \alpha_{d-1} \in \mathbb{T}$ are random variables with continuous joint probability density $h(\alpha_1, \ldots, \alpha_{d-1})$. Then the limit*

$$\lim_{N\to\infty} \operatorname{Prob}\left\{ \overline{R}_n^{\boldsymbol{\alpha}}(\mathcal{B}, N) > X \right\}$$

exists for all $X > 0$ and is given by the continuous function (on $\mathbb{R}_+$)

$$\Psi_{n,d}(X) = \frac{\mu\left\{ \mathsf{M} \in \Sigma^d : \overline{D}_n(\mathcal{B}, \mathsf{M}) > X \right\}}{\mu(\Sigma^d)},$$

which is independent of h.

Tail estimates and moments of the limit distribution $\Psi_{n,d}(X)$ are given in Sections 3.6 and 3.8.

Theorem 3.3 *Let $\mathcal{B}$ be a bounded open subset in $\mathbb{R}^{n-1}$ with boundary of measure zero. Suppose now $(\alpha_1, \ldots, \alpha_{d-1}) \in \mathbb{T}^{d-1}$ are fixed and $N \in \mathbb{R}$ is a random variable distributed in $[1, M]$ with probability density $(N \log M)^{-1}$. Then, for almost all $\alpha_1, \ldots, \alpha_{d-1}$ (with respect to Lebesgue measure on $\mathbb{T}^{d-1}$), we have*

$$\lim_{M\to\infty} \operatorname{Prob}\left\{ \overline{R}_n^{\boldsymbol{\alpha}}(\mathcal{B}, N) > X \right\} = \Psi_{n,d}(X)$$

with the same limit $\Psi_{n,d}(X)$ as in Theorem 3.2. Alternatively, one may also take N to be a random integer distributed in $[1, M]$ with probability

$$\frac{\log(1 + N^{-1})}{\sum_{\nu=1}^{M} \log(1 + \nu^{-1})}.$$

We say $(\alpha_1, \ldots, \alpha_{d-1})$ is *badly approximable*, if there exists a constant $c > 0$ such that

$$\rho(\alpha_1 x_1 + \ldots + \alpha_{d-1} x_{d-1}) > c\big(\max\{|x_1|, \ldots, |x_{d-1}|\}\big)^{-(d-1)}$$

for all non-zero $(x_1, \ldots, x_{d-1}) \in \mathbb{Z}^{d-1}$, with $\rho(t) = \min_{n\in\mathbb{Z}}\{|t - n|\}$, see [6, 5, 36] for more details.

Theorem 3.4 *If $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{d-1})$ is badly approximable, then the set*

$$\{\overline{R}_n^{\boldsymbol{\alpha}}(\mathcal{B}, N) : N \geq 1\}$$

is bounded in $\mathbb{R}_+$. On the other hand, for almost all $\boldsymbol{\alpha}$ (with respect to Lebesgue measure) the function $\overline{R}_n^{\boldsymbol{\alpha}}(\mathcal{B}, N)$ is unbounded as $N \to \infty$.

The key to the proof of the above theorems are dynamical properties of flows on $\mathrm{SL}(d, \mathbb{R}) / \mathrm{SL}(d, \mathbb{Z})$.

3.4 Flows on $\mathrm{SL}(d,\mathbb{R})/\,\mathrm{SL}(d,\mathbb{Z})$

The following analysis holds as well for all other lattices Γ in $\mathrm{SL}(d,\mathbb{R})$, but we shall restrict our notation to the case $\Gamma = \mathrm{SL}(d,\mathbb{Z})$.

Define the flow Φ^t by the left action

$$\Phi^t : \quad \Sigma^d \to \Sigma^d,$$

$$\mathsf{M} \mapsto \begin{pmatrix} \mathrm{e}^{-t} & & & \\ & \ddots & & \\ & & \mathrm{e}^{-t} & \\ & & & \mathrm{e}^{(d-1)t} \end{pmatrix} \mathsf{M}.$$

As a result of Section 3.2, the behaviour of the n-point correlations is determined by the asymptotic distribution of the trajectory $\{\Phi^t(\mathsf{M}) : t \geq 0\}$. It is well know that this flow has the mixing property [26], which can be stated as follows. (Denote by μ the invariant measure on $\Sigma^d = \mathrm{SL}(d,\mathbb{R})/\Gamma$.)

Theorem 3.5 *Let $f, g \in L^2(\Sigma^d, \mu)$. Then*

$$\lim_{t\to\pm\infty} \int_{\Sigma^d} f(\Phi^t(\mathsf{M}))\, g(\mathsf{M})\, d\mu(\mathsf{M}) = \frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} f\, d\mu \int_{\Sigma^d} g\, d\mu.$$

This in particular implies ergodicity.

Theorem 3.6 *Let $f \in L^1(\Sigma^d, \mu)$. Then for μ-almost all M*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(\Phi^t(\mathsf{M}))\, dt = \frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} f\, d\mu.$$

Consider elements $\mathsf{M} \in \mathrm{SL}(d,\mathbb{R})$ which have a decomposition of the form

$$\mathsf{M} = \begin{pmatrix} \mathsf{A} & \boldsymbol{b} \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \mathrm{e}^{-s}\mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathrm{e}^{(d-1)s} \end{pmatrix} \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \boldsymbol{l} & 1 \end{pmatrix} \tag{3.18}$$

where $\mathsf{A} \in \mathrm{SL}(d-1,\mathbb{R})$, $\boldsymbol{b}$ a $(d-1)$-dimensional column vector, $s \in \mathbb{R}$, $\boldsymbol{l}$ a $(d-1)$-dimensional row vector, $\mathbf{1}$ a unit matrix and $\mathbf{0}$ zero vectors. This provides a local parametrization of $\mathrm{SL}(d,\mathbb{R})$. The Haar measure μ reads in these (local) coordinates

$$d\mu(\mathsf{M}) = \mathrm{e}^{d(d-1)s}\; ds\, dl_1 \cdots dl_{d-1}\, db_1 \cdots db_{d-1}\, d\mu_{d-1}(\mathsf{A}), \tag{3.19}$$

where μ_{d-1} is the Haar measure of $\mathrm{SL}(d-1,\mathbb{R})$. The action of the flow Φ^t on M reads

$$\Phi^t(\mathsf{M}) = \begin{pmatrix} \mathsf{A} & \boldsymbol{b}\,\mathrm{e}^{-dt} \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \mathrm{e}^{-(s+t)}\,1 & \mathbf{0} \\ \mathbf{0} & \mathrm{e}^{(d-1)(s+t)} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \boldsymbol{l} & 1 \end{pmatrix}. \tag{3.20}$$

With the abbreviation

$$\mathsf{M} = [s, \boldsymbol{l}, \boldsymbol{b}, \mathsf{A}]$$

the latter can be expressed as

$$\Phi^t([s, \boldsymbol{l}, \boldsymbol{b}, \mathsf{A}]) = [s+t, \boldsymbol{l}, \boldsymbol{b}\,\mathrm{e}^{-dt}, \mathsf{A}]. \tag{3.21}$$

Let us denote by $d(\mathsf{M}, \mathsf{M}')$ the distance between two points, which is induced by the right-invariant Riemannian metric on $\mathrm{SL}(d,\mathbb{R})$. Due to the right-invariance we see that the distance between the translates $\Phi^t(\mathsf{M})$, $\Phi^t(\mathsf{M}')$ (which we may view as two neighbouring choices of initial conditions) is given by

$$d(\Phi^t(\mathsf{M}), \Phi^t(\mathsf{M}')) = d([s, \boldsymbol{l}\,\mathrm{e}^{dt}, \boldsymbol{b}\,\mathrm{e}^{-dt}, \mathsf{A}], [s', \boldsymbol{l}'\,\mathrm{e}^{dt}, \boldsymbol{b}'\,\mathrm{e}^{-dt}, \mathsf{A}']).$$

Hence s and A characterize the neutral manifolds of the flow Φ^t, $\boldsymbol{l}$ the unstable and $\boldsymbol{b}$ the stable ones.

With the above parametrization, we have the following corollary of the ergodic theorem.

Corollary 3.7 *Let f be bounded and piecewise continuous on Σ^d. Then, for all $s \in \mathbb{R}$, $\mathsf{A} \in \mathrm{SL}(d-1,\mathbb{R})$ and almost all $\boldsymbol{l} \in \mathbb{T}^{d-1}$ (with respect to Lebesgue measure) we have*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f([s+t, \boldsymbol{l}, \mathbf{0}, \mathsf{A}])\, dt = \frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} f\, d\mu.$$

Proof. Since the function $f_{\mathsf{M}}(\mathsf{M}') = f(\mathsf{M}\mathsf{M}')$ is still right-Γ-invariant (as a function of M'), we may assume without loss of generality that $s = 0$, $\mathsf{A} = 1$.

Suppose first f is uniformly continuous. Every point of the form $[s, \boldsymbol{l}, \boldsymbol{b}, \mathsf{A}]$ which is in an ϵ-neighbourhood of the point $[0, \boldsymbol{l}, \mathbf{0}, 1]$, i.e.

$$d([0, \boldsymbol{l}, \mathbf{0}, 1], [s, \boldsymbol{l}, \boldsymbol{b}, \mathsf{A}]) < \epsilon,$$

stays in an ϵ-neighbourhood of the translated point $[t, \boldsymbol{l}, \mathbf{0}, 1]$, since

$$d(\Phi^t[0, \boldsymbol{l}, \mathbf{0}, 1], \Phi^t[s, \boldsymbol{l}, \boldsymbol{b}, \mathsf{A}]) = d([0, \mathbf{0}, \mathbf{0}, 1], [s, \mathbf{0}, \boldsymbol{b}\,\mathrm{e}^{-dt}, \mathsf{A}]) < \epsilon.$$

Hence, by uniform continuity, for every given $\delta > 0$ we find an $\epsilon > 0$ such that

$$|f(\Phi^t[0, \boldsymbol{l}, \mathbf{0}, 1]) - f(\Phi^t[s, \boldsymbol{l}, \boldsymbol{b}, \mathsf{A}])| < \delta$$

for all $t > 0$ and all $[s, \boldsymbol{l}, \boldsymbol{b}, \mathsf{A}]$ in the ϵ-neighbourhood. So for all T

$$\begin{aligned} \frac{1}{T}\int_0^T f(\Phi^t[s, \boldsymbol{l}, \boldsymbol{b}, \mathsf{A}])\, dt - \delta < \frac{1}{T}\int_0^T f(\Phi^t[0, \boldsymbol{l}, \mathbf{0}, 1])\, dt \\ < \frac{1}{T}\int_0^T f(\Phi^t[s, \boldsymbol{l}, \boldsymbol{b}, \mathsf{A}])\, dt + \delta. \quad (3.22) \end{aligned}$$

If therefore the limit

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T f([t, \boldsymbol{l}, \mathbf{0}, 1])\, dt$$

did not exist for almost all $\boldsymbol{l}$ and was not equal to $\mu(\Sigma^d)^{-1}\int_{\Sigma^d} f\, d\mu$, we would have a contradiction with Theorem 3.6, since δ can be made arbitrarily small. Thus the assertion holds for uniformly continuous functions. The extension to bounded piecewise continuous functions can be achieved by a standard measure-theoretic argument (approximation from above and below). □

Corollary 3.8 *Let f be bounded and piecewise continuous on Σ^d, and let $0 = t_0 < t_1 < t_2 \to \infty$ be a sequence with only accumulation point at infinity, such that $t_{j+1} - t_j \to 0$. Then, for all $s \in \mathbb{R}$, $\mathsf{A} \in \mathrm{SL}(d-1, \mathbb{R})$ and almost all $\boldsymbol{l} \in \mathbb{T}^{d-1}$ (with respect to Lebesgue measure) we have*

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t_j \le T} (t_{j+1} - t_j)\, f([s + t_j, \boldsymbol{l}, \mathbf{0}, \mathsf{A}]) = \frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} f\, d\mu.$$

Proof. Assume first f is continuously differentiable and of compact support. Therefore the function $F(t) = f([t, \boldsymbol{l}, \mathbf{0}, 1])$ is uniformly continuous, and we have in particular by the intermediate value theorem

$$F(t) = F(t_j) + F'(\xi)(t - t_j) \quad (3.23)$$

with some $\xi \in [t_j, t]$. Since F' is uniformly bounded we see that

$$\begin{aligned} \int_0^T f([t, \boldsymbol{l}, \mathbf{0}, 1])\, dt - \sum_{t_j \le T} (t_{j+1} - t_j)\, f([t_j, \boldsymbol{l}, \mathbf{0}, 1]) \\ \ll \sum_{t_j \le T} \int_{t_j}^{t_{j+1}} (t - t_j)\, dt = \frac{1}{2}\sum_{t_j \le T} (t_{j+1} - t_j)^2. \quad (3.24) \end{aligned}$$

Furthermore

$$\sum_{t_j \le T} (t_{j+1} - t_j)^2 \ll O(T^{\frac{1}{100}}) + \sum_{T^{\frac{1}{100}} \le t_j \le T} (t_{j+1} - t_j)^2$$
$$\ll O(T^{\frac{1}{100}}) + \sum_{T^{\frac{1}{100}} \le t_j \le T} o(1)(t_{j+1} - t_j) = o(T). \quad (3.25)$$

Hence the corollary holds for the above class of functions f. The extension to bounded piecewise continuous f is again possible by approximation from above/below. (Bounded functions can be approximated from above by linear combinations of compactly supported functions, which we have just discussed, and constant functions $f = \text{const}$ for which the theorem obviously holds.) $\square$

Following the lines of the proof of Corollary 5.2 in [23] (compare also [9, 19] and references therein), we exploit the mixing property to show the equidistribution of measures concentrated along the unstable manifold.

Corollary 3.9 *Let f be bounded and piecewise continuous on Σ^d, and h be continuous on the standard $(d-1)$-dimensional unit torus $\mathbb{T}^{d-1}$. Then, for every $s \in \mathbb{R}$, $\mathsf{A} \in \mathrm{SL}(d-1, \mathbb{R})$, we have*

$$\lim_{t\to\infty} \int_{\mathbb{T}^{d-1}} f([s+t, \boldsymbol{l}, \mathbf{0}, \mathsf{A}])\, h(\boldsymbol{l})\, d^{d-1}\boldsymbol{l} = \frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} f\, d\mu \int_{\mathbb{T}^{d-1}} h(\boldsymbol{l})\, d^{d-1}\boldsymbol{l},$$

where $d^{d-1}\boldsymbol{l} = dl_1 \cdots dl_{d-1}$.

Proof. As in the proof of Corollary 3.7 we may assume without loss of generality that $s = 0$, $\mathsf{A} = 1$.

Let f be a continuous function on $\mathrm{SL}(d, \mathbb{R})$, right-$\Gamma$-invariant and compactly supported when viewed as a function on Σ^d. Define furthermore the function H on $\mathrm{SL}(d, \mathbb{R})$ by

$$H(\mathsf{M}) = \tilde{h}(\boldsymbol{l})\, \frac{1}{\epsilon}\chi(\frac{s}{\epsilon})\, \frac{1}{\epsilon}\chi_\epsilon(\mathsf{A})\, \chi(b_1) \cdots \chi(b_{d-1}) \quad (3.26)$$

if M admits a representation of the form (3.18) and $H = 0$ if not, where χ is the characteristic function of the interval $[-\frac{1}{2}, \frac{1}{2}]$ and χ_ϵ the characteristic function of a ball in $\mathrm{SL}(d-1, \mathbb{R})$ centred at $\mathsf{A} = 1$ with volume ϵ. The function $\tilde{h}$ is continuous and compactly supported. (It will be later related to h by $h(\boldsymbol{l}) = \sum_{\boldsymbol{m} \in \mathbb{Z}^{d-1}} \tilde{h}(\boldsymbol{l} + \boldsymbol{m})$.) Thus H has compact support and the integral

$$\int_{\mathrm{SL}(d,\mathbb{R})} H(\mathsf{M})\, f(\Phi^t(\mathsf{M}))\, d\mu = \int_{\mathrm{SL}(d,\mathbb{R})} H([s, \boldsymbol{l}, \boldsymbol{b}, \mathsf{A}])\, f([s+t, \boldsymbol{l}, \boldsymbol{b}\, \mathrm{e}^{-dt}, \mathsf{A}])\, d\mu \quad (3.27)$$

is well defined.

Step A. f is uniformly continuous (since it is Γ-invariant and has compact support on Σ^d) and therefore we have

$$f([s+t,\boldsymbol{l},\boldsymbol{b}\,\mathrm{e}^{-dt},\mathsf{A}]) = f([s+t,\boldsymbol{l},\boldsymbol{0},\mathsf{A}]) + O(\mathrm{e}^{-dt}) \tag{3.28}$$

uniformly in $[s,\boldsymbol{l},\boldsymbol{b},\mathsf{A}]$ in the (compact) range of integration. So

$$\begin{aligned}\int_{\mathrm{SL}(d,\mathbb{R})} & H([s,\boldsymbol{l},\boldsymbol{b},\mathsf{A}])\, f([s+t,\boldsymbol{l},\boldsymbol{b}\,\mathrm{e}^{-dt},\mathsf{A}])\, d\mu \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}^{d-1}}\int_{\mathrm{SL}(d-1,\mathbb{R})} \tilde{h}(\boldsymbol{l})\, \frac{1}{\epsilon}\chi(\frac{s}{\epsilon})\, \frac{1}{\epsilon}\chi_\epsilon(\mathsf{A}) \times \\ &\quad \times f([s+t,\boldsymbol{l},\boldsymbol{0},\mathsf{A}])\, d\mu_{d-1}\, d^{d-1}\boldsymbol{l}\ \mathrm{e}^{d(d-1)s}\, ds + O(\mathrm{e}^{-dt}).\end{aligned} \tag{3.29}$$

Step B. We can rewrite (3.27) as

$$\int_{\mathrm{SL}(d,\mathbb{R})} H(\mathsf{M})\, f(\Phi^t(\mathsf{M}))\, d\mu = \int_{\Sigma^d} \Big(\sum_{\mathsf{g}\in\Gamma} H(\mathsf{Mg})\Big)\, f(\Phi^t(\mathsf{M}))\, d\mu \tag{3.30}$$

since $\Phi^t\circ f$ and μ are right-Γ-invariant. Because $G(\mathsf{M}) = \sum_{\mathsf{g}\in\Gamma} H(\mathsf{Mg})$ is also right-Γ-invariant and in $\mathrm{L}^2(\Sigma^d,\mu)$, the mixing property guarantees the existence of the limit $t\to\infty$ of (3.30), whose value is then given by

$$\begin{aligned}\frac{1}{\mu(\Sigma^d)}\int_{\Sigma^d}\Big(\sum_{\mathsf{g}\in\Gamma} H(\mathsf{Mg})\Big)\, d\mu \int_{\Sigma^d} f\, d\mu &= \frac{1}{\mu(\Sigma^d)}\int_{\mathrm{SL}(d,\mathbb{R})} H\, d\mu \int_{\Sigma^d} f\, d\mu \\ &= \frac{1}{\mu(\Sigma^d)}\int_{\Sigma^d} f\, d\mu \int_{\mathbb{R}^{d-1}} \tilde{h}(\boldsymbol{l})\, d^{d-1}\boldsymbol{l} \int_{\mathbb{R}} \frac{1}{\epsilon}\chi(\frac{s}{\epsilon})\, \mathrm{e}^{d(d-1)s}\, ds,\end{aligned} \tag{3.31}$$

where we have used the fact that

$$\int_{\mathrm{SL}(d-1,\mathbb{R})} \chi_\epsilon(\mathsf{A})\, d\mu_{d-1}(\mathsf{A}) = \epsilon.$$

Step C. From Steps A and B we conclude that

$$\begin{aligned}\lim_{t\to\infty} & \int_{\mathbb{R}}\int_{\mathbb{R}^{d-1}}\int_{\mathrm{SL}(d-1,\mathbb{R})} \tilde{h}(\boldsymbol{l})\, \frac{1}{\epsilon}\chi(\frac{s}{\epsilon})\, \frac{1}{\epsilon}\chi_\epsilon(\mathsf{A}) \times \\ &\quad \times f([s+t,\boldsymbol{l},\boldsymbol{0},\mathsf{A}])\, d\mu_{d-1}\, d^{d-1}\boldsymbol{l}\ \mathrm{e}^{d(d-1)s}\, ds \\ &= \frac{1}{\mu(\Sigma^d)}\int_{\Sigma^d} f\, d\mu \int_{\mathbb{R}^{d-1}} \tilde{h}(\boldsymbol{l})\, d^{d-1}\boldsymbol{l} \int_{\mathbb{R}} \chi(s)\, \mathrm{e}^{d(d-1)\epsilon s}\, ds.\end{aligned} \tag{3.32}$$

By the uniform continuity of f with respect to the right-invariant metric on $\mathrm{SL}(d,\mathbb{R})$, given any $\delta > 0$, we find an $\epsilon > 0$ such that

$$\big| f([s+t, \boldsymbol{l}, \mathbf{0}, \mathsf{A}]) - f([t, \boldsymbol{l}, \mathbf{0}, 1]) \big| < \delta$$

uniformly for all $s \in [-\epsilon, \epsilon]$, $\boldsymbol{l} \in \mathbb{R}^{d-1}$, A in a ball centred at $\mathsf{A} = 1$ of volume ϵ. We have the inclusion

$$\begin{aligned}
\int_{\mathbb{R}} \int_{\mathbb{R}^{d-1}} \int_{\mathrm{SL}(d-1,\mathbb{R})} & \tilde{h}(\boldsymbol{l}) \, \frac{1}{\epsilon} \chi(\frac{s}{\epsilon}) \, \frac{1}{\epsilon} \chi_\epsilon(\mathsf{A}) \, \times \\
& \times f([s+t, \boldsymbol{l}, \mathbf{0}, \mathsf{A}]) \, d\mu_{d-1} \, d^{d-1}\boldsymbol{l} \; \mathrm{e}^{d(d-1)s} \, ds - \delta \\
& < \int_{\mathbb{R}} \int_{\mathbb{R}^{d-1}} \int_{\mathrm{SL}(d-1,\mathbb{R})} \tilde{h}(\boldsymbol{l}) \, \frac{1}{\epsilon} \chi(\frac{s}{\epsilon}) \, \frac{1}{\epsilon} \chi_\epsilon(\mathsf{A}) \, \times \\
& \qquad \times f([t, \boldsymbol{l}, \mathbf{0}, 1]) \, d\mu_{d-1} \, d^{d-1}\boldsymbol{l} \, ds \\
& \qquad = \int_{\mathbb{R}^{d-1}} \tilde{h}(\boldsymbol{l}) \, f([t, \boldsymbol{l}, \mathbf{0}, 1]) \, d^{d-1}\boldsymbol{l} \\
& < \int_{\mathbb{R}} \int_{\mathbb{R}^{d-1}} \int_{\mathrm{SL}(d-1,\mathbb{R})} \tilde{h}(\boldsymbol{l}) \, \frac{1}{\epsilon} \chi(\frac{s}{\epsilon}) \, \frac{1}{\epsilon} \chi_\epsilon(\mathsf{A}) \, \times \\
& \qquad \times f([s+t, \boldsymbol{l}, \mathbf{0}, \mathsf{A}]) \, d\mu_{d-1} \, d^{d-1}\boldsymbol{l} \; \mathrm{e}^{d(d-1)s} \, ds + \delta. \qquad (3.33)
\end{aligned}$$

Because the limits $t \to \infty$ on the left- and the right-hand-side of the above inequality exist and differ only by 2δ, which can be made arbitrarily small, the limit of the inner term has to exist as well and is precisely given by

$$\lim_{t\to\infty} \int_{\mathbb{R}^{d-1}} \tilde{h}(\boldsymbol{l}) \, f([t, \boldsymbol{l}, \mathbf{0}, 1]) \, d^{d-1}\boldsymbol{l} = \frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} f \, d\mu \int_{\mathbb{R}^{d-1}} \tilde{h}(\boldsymbol{l}) \, d^{d-1}\boldsymbol{l}, \qquad (3.34)$$

or

$$\lim_{t\to\infty} \int_{\mathbb{T}^{d-1}} h(\boldsymbol{l}) \, f([t, \boldsymbol{l}, \mathbf{0}, 1]) \, d^{d-1}\boldsymbol{l} = \frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} f \, d\mu \int_{\mathbb{T}^{d-1}} h(\boldsymbol{l}) \, d^{d-1}\boldsymbol{l}, \qquad (3.35)$$

with $h(\boldsymbol{l}) = \sum_{\boldsymbol{m}\in\mathbb{Z}^{d-1}} \tilde{h}(\boldsymbol{l}+\boldsymbol{m})$. By standard measure-theoretic tricks (approximation from above and below), one can show that the above also holds for bounded piecewise continuous functions. □

3.5 Proof of Theorems 3.2 – 3.4

We begin with the following observation.

Lemma 3.10 *Fix $X > 0$. For $\mathcal{B}$ open with boundary of measure zero, the solutions M of the equation*

$$\overline{D}_n(\mathcal{B}, \mathsf{M}) = X \qquad (3.36)$$

form sets of (Haar) measure zero in Σ^d.

Proof. Cover Σ^d by a countable union of compact sets S. On each compact set S, $\overline{D}_n(\mathcal{B},\mathsf{M})$ has a representation as a *finite* sum

$$\overline{D}_n(\mathcal{B},\mathsf{M}) = \sum_{\mathcal{K}_1,\ldots,\mathcal{K}_d\in\mathbb{Z}^{n-1}}^{(3.14)} \tau_n(\tilde{\mathcal{K}}_1)\cdots\tau_n(\tilde{\mathcal{K}}_{d-1})\,\chi_{\mathcal{B}}(\tilde{\mathcal{K}}_d), \tag{3.37}$$

where $\chi_{\mathcal{B}}$ is the characteristic function of $\mathcal{B}$.

Suppose now the set of solutions $\mathsf{M}\in S$ of (3.36) has strictly positive measure. Then, since $\overline{D}_n(\mathcal{B},\mathsf{M})$ is piecewise continuous, we find a point M such that for all M' in a small open neighbourhood of M we have $\overline{D}_n(\mathcal{B},\mathsf{M}')=X$. Taking $\mathsf{M}'=\Phi^{-\epsilon}(\mathsf{M})$, this implies in particular $\overline{D}_n(\mathcal{B},\Phi^{-\epsilon}(\mathsf{M}))=X$, for all ϵ small enough. Now

$$\overline{D}_n(\mathcal{B},\Phi^{-\epsilon}(\mathsf{M})) = \sum_{\mathcal{K}_1,\ldots,\mathcal{K}_d\in\mathbb{Z}^{n-1}}^{(3.14)} \tau_n(\mathrm{e}^{\epsilon}\,\tilde{\mathcal{K}}_1)\cdots\tau_n(\mathrm{e}^{\epsilon}\,\tilde{\mathcal{K}}_{d-1})\,\chi_{\mathcal{B}}(\mathrm{e}^{-(d-1)\epsilon}\,\tilde{\mathcal{K}}_d),$$

and for ϵ small enough, we claim that

$$\overline{D}_n(\mathcal{B},\Phi^{-\epsilon}(\mathsf{M})) = \sum_{\mathcal{K}_1,\ldots,\mathcal{K}_d\in\mathbb{Z}^{n-1}}^{(3.14)} \tau_n(\mathrm{e}^{\epsilon}\,\tilde{\mathcal{K}}_1)\cdots\tau_n(\mathrm{e}^{\epsilon}\,\tilde{\mathcal{K}}_{d-1})\,\chi_{\mathcal{B}}(\tilde{\mathcal{K}}_d). \tag{3.38}$$

To see why (3.38) is indeed valid, suppose the contrary would be true. Then, since $\chi_{\mathcal{B}}$ is piecewise constant, we would find a constant $\delta>0$ such that we would have

$$\sum_{\mathcal{K}_1,\ldots,\mathcal{K}_d\in\mathbb{Z}^{n-1}}^{(3.14)} \tau_n(\mathrm{e}^{\epsilon}\,\tilde{\mathcal{K}}_1)\cdots\tau_n(\mathrm{e}^{\epsilon}\,\tilde{\mathcal{K}}_{d-1})\left[\chi_{\mathcal{B}}(\mathrm{e}^{-(d-1)\epsilon}\,\tilde{\mathcal{K}}_d)-\chi_{\mathcal{B}}(\tilde{\mathcal{K}}_d)\right]\geq\delta$$

for all $0<\epsilon<\epsilon_0$, ϵ_0 small. But this implies

$$\begin{aligned}&\lim_{\epsilon\to0}\left\{\overline{D}_n(\mathcal{B},\Phi^{-\epsilon}(\mathsf{M}))-\overline{D}_n(\mathcal{B},\mathsf{M})\right\}\\ &\quad=\lim_{\epsilon\to0}\sum_{\mathcal{K}_1,\ldots,\mathcal{K}_d\in\mathbb{Z}^{n-1}}^{(3.14)} \tau_n(\tilde{\mathcal{K}}_1)\cdots\tau_n(\tilde{\mathcal{K}}_{d-1})\left[\chi_{\mathcal{B}}(\mathrm{e}^{-(d-1)\epsilon}\,\tilde{\mathcal{K}}_d)-\chi_{\mathcal{B}}(\tilde{\mathcal{K}}_d)\right]\geq\delta,\end{aligned}$$

which means $\overline{D}_n(\mathcal{B},\mathsf{M}')$ is discontinuous in M, and thus contradict our assumption that $\overline{D}_n(\mathcal{B},\mathsf{M}')$ is constant in a small neighbourhood of M. Hence (3.38) holds.

Continuing with (3.38), we observe that for $\epsilon>0$

$$\tau_n(\mathrm{e}^{\epsilon}\,\mathcal{X})\leq\tau_n(\mathcal{X}),\quad\text{with strict inequality for all } \mathcal{X} \text{ with } \tau_n(\mathcal{X})\neq0.$$

Therefore $\overline{D}_n(\mathcal{B}, \Phi^{-\epsilon}(\mathsf{M})) < X$ for $\epsilon > 0$ arbitrarily small, a contradiction: The set of solutions of $\mathsf{M} \in S$ of (3.36) has to be of measure zero. Finally, a countable union of sets of measure zero has measure zero. The lemma follows. □

In the sequel we set $N_1 = c_1 \,\mathrm{e}^t$, $N_2 = c_2 \,\mathrm{e}^t$, ... , and thus N and t are related by

$$N = c_1 \cdots c_{d-1} \,\mathrm{e}^{(d-1)t},$$

with (arbitrary) positive constants c_j.

Proof of Theorem 3.2. Apply Corollary 3.9 with

$$s = \frac{1}{d-1} \log(c_1 \cdots c_{d-1}),$$

$$\mathsf{A} = (c_1 \cdots c_{d-1})^{1/(d-1)} \begin{pmatrix} c_1^{-1} & & \\ & \ddots & \\ & & c_{d-1}^{-1} \end{pmatrix},$$

and

$$f(\mathsf{M}) = \begin{cases} 1 & \text{if } \overline{D}_n(\mathcal{B}, \mathsf{M}) > X \\ 0 & \text{if } \overline{D}_n(\mathcal{B}, \mathsf{M}) \leq X \end{cases}$$

which, by virtue of Lemma 3.10 is piecewise continuous. Lemma 3.10 thus also implies the continuity of $\Psi_{n,d}(X)$. □

Proof of Theorem 3.3. As for Theorem 3.2, but use now Corollaries 3.7 and 3.8. □

Proof of Theorem 3.4. It follows from the results in [6] that the set $\{[t, \boldsymbol{\alpha}, \mathbf{0}, 1]\,\Gamma \,:\, t \geq 0\}$ is bounded in Σ^d for badly approximable $\boldsymbol{\alpha}$. Since $\overline{D}_n(\mathcal{B}, \mathsf{M})$ is bounded on compacta, the first assertion follows. The second statement follows from ergodicity and the fact that $\overline{D}_n(\mathcal{B}, \mathsf{M})$ is unbounded on Σ^d (see next section, Proposition 3.13 for an explicit lower bound). □

3.6 Tail estimates

Theorem 3.11 *For X_0 large enough, there exist constants $0 < C_1 \leq C_2 < \infty$ such that*

$$C_1 X^{-\frac{d}{n-1}} \leq \Psi_{n,d}(X) \leq C_2 X^{-\frac{d}{n-1}}$$

for all $X > X_0$, i.e. $\Psi_{n,d}(X)$ has a power-like tail.

Denote by A the subgroup of positive definite diagonal matrices

$$\mathsf{a}(a) = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_d \end{pmatrix} \in \mathrm{SL}(d,\mathbb{R}), \quad a_j > 0$$

and by N the subgroup of upper triangular matrices

$$\mathsf{n}(u) = \begin{pmatrix} 1 & u_{12} & \dots & u_{1d} \\ & \ddots & \ddots & \vdots \\ & & \ddots & u_{d-1,d} \\ & & & 1 \end{pmatrix} \in \mathrm{SL}(d,\mathbb{R}).$$

Every element $\mathsf{M} \in \mathrm{SL}(d,\mathbb{R})$ has the unique Iwasawa decomposition

$$\mathsf{M} = \mathsf{k}\,\mathsf{a}(a)\,\mathsf{n}(u), \tag{3.39}$$

with $\mathsf{k} \in \mathrm{SO}(d)$. The Haar measure in these coordinates reads [29]

$$d\mu = \rho(a)\; d\mathsf{k}\; d\mathsf{a}(a)\; d\mathsf{n}(u) \tag{3.40}$$

where $d\mathsf{k}$, $d\mathsf{a}$, $d\mathsf{n}$ are Haar measures of $\mathrm{SO}(d)$, A, N, respectively. For $\rho(a)$ one has [29]

$$\rho(a) = \prod_{1\leq i<j\leq d} \frac{a_i}{a_j} = \prod_{j=1}^{d} a_j^{d-2j+1}. \tag{3.41}$$

The following set is an example for a *Siegel set* [29],

$$\mathcal{S} = \{\mathsf{k}\,\mathsf{a}(a)\,\mathsf{n}(u) : \mathsf{k} \in \mathrm{SO}(d),\ 0 < a_j \leq \frac{2}{\sqrt{3}} a_{j+1}\ (j = 1, \dots, d-1), u \in \mathcal{F}_N\} \tag{3.42}$$

where $\mathcal{F}_N$ is a compact fundamental region of $N/(\Gamma \cap N)$. The above set has the property that it has finite Haar measure, contains one fundamental region of Γ and is itself contained in a finite union of fundamental regions. Therefore, we may obtain upper and lower bounds for $\Psi_{n,d}(X)$ by considering the quantity

$$\tilde{\Psi}_{n,d}(X) = \mu\Big\{\mathsf{M} \in \mathcal{S} : \overline{D}_n(\mathcal{B},\mathsf{M}) > X\Big\}$$

instead. What we will need next are bounds on $\overline{D}_n(\mathcal{B},\mathsf{M})$ in the asymptotic domains of the Siegel set $\mathcal{S}$. Fix a positive constant K. For $l \in \{1, \dots, d-1\}$, put

$$\mathcal{S}_{K,l} = \{\mathsf{M} = \mathsf{k}\,\mathsf{a}(a)\,\mathsf{n}(u) \in \mathcal{S} :\ a_l \leq K \leq \frac{2}{\sqrt{3}} a_{l+1}\}, \tag{3.43}$$

and for $l = 0, d$,

$$\mathcal{S}_{K,0} = \{\mathsf{M} = \mathsf{k}\,\mathsf{a}(a)\,\mathsf{n}(u) \in \mathcal{S} :\ K \leq \frac{2}{\sqrt{3}} a_1\}, \tag{3.44}$$

$$\mathcal{S}_{K,d} = \{\mathsf{M} = \mathsf{k}\,\mathsf{a}(a)\,\mathsf{n}(u) \in \mathcal{S} :\ a_d \leq K\}. \tag{3.45}$$

The sets $\mathcal{S}_{K,0}$ and $\mathcal{S}_{K,d}$ and clearly compact, for $a_1 \cdots a_d = 1$, and thus $\overline{D}_n(\mathcal{B}, \mathsf{M})$ is bounded on these two sets. An upper/lower bound for the limit distribution will follow from an upper/lower bound for $\overline{D}_n(\mathcal{B}, \mathsf{M})$ on the non-compact $\mathcal{S}_{K,l}$, $l = 1, \ldots, d-1$.

Lemma 3.12 *We have*

$$\mathcal{S} = \bigcup_{l=0}^{d} \mathcal{S}_{K,l}.$$

Proof. We use an inductive argument. Suppose $\mathsf{M} \notin \mathcal{S}_{K,0}$. Then $a_1 < \frac{\sqrt{3}}{2} K < K$. Hence either $\mathsf{M} \in \mathcal{S}_{K,1}$ or $a_2 < \frac{\sqrt{3}}{2} K < K$. In the latter case either $\mathsf{M} \in \mathcal{S}_{K,2}$ or $a_3 < \frac{\sqrt{3}}{2} K < K$, and so on. Finally, with $\mathsf{M} \notin \mathcal{S}_{K,0} \cup \cdots \cup \mathcal{S}_{K,d-2}$, we have either $\mathsf{M} \in \mathcal{S}_{K,d-1}$ or $a_d < \frac{\sqrt{3}}{2} K < K$. The latter condition implies $\mathsf{M} \in \mathcal{S}_{K,d}$. □

Proposition 3.13 *For K constant and small enough, we have (i) the upper bound*

$$\overline{D}_n(\mathcal{B}, \mathsf{M}) \ll_K \frac{1}{(a_1 \cdots a_l)^{n-1}}$$

uniformly for all $\mathsf{M} = \mathsf{k}\,\mathsf{a}(a)\,\mathsf{n}(u) \in \mathcal{S}_{K,l}$, and (ii) the lower bound

$$\frac{\gamma(\mathsf{k})}{(a_1 \cdots a_{d-1})^{n-1}} \ll_K \overline{D}_n(\mathcal{B}, \mathsf{M})$$

uniformly for all $\mathsf{M} = \mathsf{k}\,\mathsf{a}(a)\,\mathsf{n}(u) \in \mathcal{S}_{K,d-1}$. The function $\gamma(\mathsf{k})$ is positive and bounded for all $\mathsf{k} \in \mathrm{SO}(d)$, and non-zero on a set of strictly positive measure.

Proof. The function $\overline{D}_n(\mathcal{B}, \mathsf{M})$ is of the form

$$\overline{D}_n(\mathcal{B}, \mathsf{M}) = \sum_{\boldsymbol{k}^1, \ldots, \boldsymbol{k}^{n-1} \in \mathbb{Z}^d}^{(3.14)} F(\mathsf{M}\boldsymbol{k}^1, \ldots, \mathsf{M}\boldsymbol{k}^{n-1}) \tag{3.46}$$

with

$$F(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}) = \tau_n(\mathcal{X}_1) \cdots \tau_n(\mathcal{X}_{d-1})\, \chi_{\mathcal{B}}(\mathcal{X}_d).$$

In order to get rid of the restrictions (3.14) in the summations, we define sums

$$\overline{C}_\nu(\mathcal{B},\mathsf{M}) = \sum_{\boldsymbol{k}^1,\ldots,\boldsymbol{k}^{\nu-1}\in\mathbb{Z}^d} F_\nu(\mathsf{M}\boldsymbol{k}^1,\ldots,\mathsf{M}\boldsymbol{k}^{\nu-1}) \tag{3.47}$$

(without conditions in the summation) so that we can write

$$\overline{D}_n(\mathcal{B},\mathsf{M}) = \sum_{\nu=1}^{n} \overline{C}_\nu(\mathcal{B},\mathsf{M}) \tag{3.48}$$

with $F_n = F$ and suitable choices of F_ν for $\nu < n$. The F_ν can be chosen in such a way that they are piecewise continuous and of compact support. We will first calculate upper bounds for $\overline{C}_\nu(\mathcal{B},\mathsf{M})$.

Since

$$\mathsf{M}\boldsymbol{k}^j = \mathsf{k}\begin{pmatrix} a_1(k_1^j + u_{12}k_2^j + \ldots + u_{1d}k_d^j) \\ \vdots \\ a_d k_d^j \end{pmatrix}$$

with u_{ij} in compact sets, we have for $a_1,\ldots,a_l\to 0$, $a_{l+1},\ldots,a_d\to\infty$, the asymptotic behaviour (view the sum as a Riemann sum)

$$\sum_{\boldsymbol{k}^1,\ldots,\boldsymbol{k}^{\nu-1}\in\mathbb{Z}^d} F_\nu(\mathsf{M}\boldsymbol{k}^1,\ldots,\mathsf{M}\boldsymbol{k}^{\nu-1}) \sim \frac{\gamma_l(\mathsf{k})}{(a_1\cdots a_l)^{\nu-1}}, \tag{3.49}$$

with

$$\gamma_{l,\nu}(\mathsf{k}) = \int F(\mathsf{k}\begin{pmatrix} x_1^1 \\ \vdots \\ x_l^1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\ldots,\mathsf{k}\begin{pmatrix} x_1^{\nu-1} \\ \vdots \\ x_l^{\nu-1} \\ 0 \\ \vdots \\ 0 \end{pmatrix})dx. \tag{3.50}$$

We now find a constant $c_{\nu,K}$ such that for K small

$$|\overline{C}_\nu(\mathcal{B},\mathsf{M})| \le \frac{c_{\nu,K}}{(a_1\cdots a_l)^{\nu-1}}, \tag{3.51}$$

and in particular for $\nu = n$

$$\overline{D}_n(\mathcal{B},\mathsf{M}) \le \overline{C}_n(\mathcal{B},\mathsf{M}) \le \frac{c_{n,K}}{(a_1\cdots a_l)^{n-1}}. \tag{3.52}$$

This proves the upper bound. For a lower bound on $\mathcal{S}_{K,d-1}$ notice that $a_1,\ldots,a_{d-1}\to 0$ *implies* $a_d\to\infty$. Thus (K small)

$$\overline{C}_n(\mathcal{B},\mathsf{M}) \asymp_K \frac{\gamma_{d-1,n}(\mathsf{k})}{(a_1\cdots a_{d-1})^{n-1}}.$$

The lower bound given in the proposition follows from this relation and the estimate (3.51) for $\nu < n$, $l = d-1$. The function $\gamma(\mathsf{k})$ in the proposition is equal to $\gamma_{d-1,n}(\mathsf{k})$ in (3.50). $\square$

Before we proceed to prove Theorem 3.11, let us consider the following lemma.

Lemma 3.14 *Let $r_1 \geq r_2 \geq \cdots \geq r_{M-1} > r_M > -1$ be real numbers. Then*

$$\lim_{Y\to\infty} Y^{r_M+1} \int\cdots\int_{\substack{0<x_j<1\ (j=1,\dots,M)\\ \prod_{j=1}^M x_j<\frac{1}{Y}}} \prod_{j=1}^{M} x_j^{r_j}\,dx_j = \frac{1}{(r_M+1)\prod_{j=1}^{M-1}(r_j-r_M)}.$$

Remark. Replacing the conditions $0 < x_j < 1$ by the more general $0 < x_j < o_j$, for arbitrary constants $o_j > 0$, obviously changes the limit only by a positive constant.

Proof (by induction on M). We may assume with out loss of generality that we have strict inequalities $r_1 > r_2 > \cdots > r_{M-1} > r_M > -1$. The general result then follows from continuity and uniform convergence.

The case $M = 1$ is correct. Suppose now that the assertion of the lemma is true for all $M < N$. Integrating over x_N we see that

$$\begin{aligned}\int\cdots\int_{\substack{0<x_j<1\ (j=1,\dots,N)\\ \prod_{j=1}^N x_j<\frac{1}{Y}}} \prod_{j=1}^{N} x_j^{r_j}\,dx_j \\ = \frac{1}{r_N+1}\int\cdots\int_{0<x_j<1\ (j=1,\dots,N-1)} \min\{1, \frac{1}{Y\prod_{j=1}^{N-1}x_j}\}^{r_N+1} \prod_{j=1}^{N-1} x_j^{r_j}\,dx_j.\end{aligned} \tag{3.53}$$

The latter integral splits up into two terms,

$$\begin{aligned}\int\cdots\int_{\substack{0<x_j<1\ (j=1,\dots,N-1)\\ 1<\frac{1}{Y\prod_{j=1}^{N-1}x_j}}} \prod_{j=1}^{N-1} x_j^{r_j}\,dx_j \\ + Y^{-(r_N+1)}\int\cdots\int_{\substack{0<x_j<1\ (j=1,\dots,N-1)\\ 1>\frac{1}{Y\prod_{j=1}^{N-1}x_j}}} \prod_{j=1}^{N-1} x_j^{r_j-(r_N+1)}dx_j.\end{aligned} \tag{3.54}$$

By the induction hypothesis, the first contribution is of order $Y^{-(r_{N-1}+1)}$, which is non-leading since $r_N < r_{N-1}$. The second term can be written as

$$Y^{-(r_N+1)} \int \cdots \int_{0<x_j<1\ (j=1,\ldots,N-1)} \prod_{j=1}^{N-1} x_j^{r_j-(r_N+1)} dx_j$$
$$- Y^{-(r_N+1)} \int \cdots \int_{\substack{0<x_j<1\ (j=1,\ldots,N-1) \\ 1<\frac{1}{Y\prod_{j=1}^{N-1} x_j}}} \prod_{j=1}^{N-1} x_j^{r_j-(r_N+1)} dx_j, \quad (3.55)$$

where again by the induction hypothesis the second contribution is of order

$$Y^{-(r_N+1)} Y^{-[r_{N-1}-(r_N+1)+1]} = Y^{-(r_{N-1}+1)}$$

(use the new sequence $\tilde{r}_j = r_j - r_N - 1$, $j = 1, \ldots, N-1$; notice in particular $\tilde{r}_{N-1} > -1$). The first integral evaluates to

$$Y^{-(r_N+1)} \frac{1}{\prod_{j=1}^{N-1}(r_j - r_N)}.$$

□

Proof of Theorem 3.11. We begin with the *upper bound.* Proposition 3.13 yields the bound

$$\tilde{\Psi}_{n,d}(X) \leq \sum_l \mu\Big\{ \mathsf{M} \in \mathcal{S}_{K,l} : \frac{c}{(a_1 \cdots a_l)^{n-1}} > X \Big\}$$

for some constant $c > 0$ depending on K. An upper bound of the right-hand-side is thus

$$\sum_l \mu\Big\{ \mathsf{M} \in \mathcal{S} : \frac{c}{(a_1 \cdots a_l)^{n-1}} > X \Big\},$$

and we are left with the integral

$$I(X) = \mu\Big\{ \mathsf{M} \in \mathcal{S} : \frac{c}{(a_1 \cdots a_l)^{n-1}} > X \Big\}. \quad (3.56)$$

Now, (3.56) equals

$$I(X) = \int_{\substack{0<a_j\leq\frac{2}{\sqrt{3}}a_{j+1}(j=1,\ldots,d-1) \\ (a_1\cdots a_l)^{n-1}<\frac{c}{X} \\ \mathsf{k}\in \mathrm{SO}(d)}} \rho(a)\, d\mathsf{k}\, d\mathsf{a}(a). \quad (3.57)$$

Let us introduce new coordinates $b_j = a_j/a_{j+1}$ in terms of which the old ones read

$$a_j = (\prod_{\nu=1}^{d-1} b_\nu^\nu)^{-1/d} \prod_{\nu=j}^{d-1} b_\nu, \quad a_d = (\prod_{\nu=1}^{d-1} b_\nu^\nu)^{-1/d}. \quad (3.58)$$

In particular we have

$$a_1 \cdots a_l = \Big(\prod_{j=1}^{l} b_j^{j(d-l)}\Big)^{\frac{1}{d}} \Big(\prod_{j=l+1}^{d-1} b_j^{l(d-j)}\Big)^{\frac{1}{d}}. \tag{3.59}$$

The Haar measure on A reads in these coordinates $d\mathsf{a}(a) = \frac{db_j}{b_j}$ and so

$$\rho(a)\, d\mathsf{a}(a) = \prod_{j=1}^{d-1} b_j^{j(d-j)} \frac{db_j}{b_j}. \tag{3.60}$$

Thus (3.57) becomes now

$$I(X) = \int \prod_{j=1}^{d-1} b_j^{j(d-j)} \frac{db_j}{b_j}\, d\mathsf{k}. \tag{3.61}$$

with range of integration

$$0 < b_j < \frac{2}{\sqrt{3}}, \quad \Big(\prod_{j=1}^{l} b_j^{j(d-l)}\Big)^{\frac{n-1}{d}} \Big(\prod_{j=l+1}^{d-1} b_j^{l(d-j)}\Big)^{\frac{n-1}{d}} < \frac{c}{X}, \quad \mathsf{k} \in \mathrm{SO}(d).$$

A further change of variables

$$x_j = \begin{cases} b_j^{j(d-l)} & (j = 1, \dots, l) \\ b_j^{l(d-j)} & (j = l+1, \dots, d-1) \end{cases}$$

leads to

$$I(X) = \int_{\substack{0 < x_j < o_j (j=1,\dots,d-1) \\ (\prod_{\nu=1}^{d-1} x_\nu)^{\frac{n-1}{d}} < \frac{c}{X} \\ \mathsf{k} \in \mathrm{SO}(d)}} \prod_{j=1}^{l} x_j^{\frac{d-j}{d-l}} \frac{dx_j}{x_j} \prod_{j=l+1}^{d-1} x_j^{\frac{j}{l}} \frac{dx_j}{x_j}\, d\mathsf{k}, \tag{3.62}$$

with obvious constants $o_j > 0$. The application of Lemma 3.14 with $r_M = 0$ (all other powers are higher, $\frac{d-j}{d-l} > 1$ when $j = 1, \dots, l-1$, and $\frac{j}{l} > 1$ when $j = l+1, \dots, d-1$) and $Y = (\frac{X}{c})^{\frac{d}{n-1}}$ completes the proof of the upper bound.

A *lower bound* is, by virtue of Proposition 3.13,

$$J(X) = \mu\Big\{ \mathsf{M} \in \mathcal{S}_{K,d-1} : \frac{c'\gamma(\mathsf{k})}{(a_1 \cdots a_{d-1})^{n-1}} > X \Big\} \le \tilde{\Psi}_{n,d}(X)$$

for some constant $c' > 0$ depending on K. That is,

$$J(X) = \int_{\substack{0 < a_j \le \frac{2}{\sqrt{3}} a_{j+1} (j=1,\dots,d-1) \\ (a_1 \cdots a_{d-1})^{n-1} < \frac{c'\gamma(\mathsf{k})}{X} \\ a_{d-1} < K \\ \mathsf{k} \in \mathrm{SO}(d)}} \rho(a)\, d\mathsf{k}\, d\mathsf{a}(a). \tag{3.63}$$

Changing variables as before yields integral (3.61) with range of integration

$$0 < b_j < \frac{2}{\sqrt{3}}, \quad \Big(\prod_{j=1}^{d-1} b_j^j\Big)^{\frac{n-1}{d}} < \frac{c'\gamma(\mathsf{k})}{X}, \quad \Big(\prod_{j=1}^{d-1} b_j^j\Big)^{-1/d} b_{d-1} < K, \quad \mathsf{k} \in \mathrm{SO}(d).$$

Put $x_j = b_j^j$. This gives

$$J(X) = \int \prod_{j=1}^{d-1} x_j^{d-j} \frac{dx_j}{x_j}\, d\mathsf{k} \tag{3.64}$$

with range of integration

$$0 < x_j < \Big(\frac{2}{\sqrt{3}}\Big)^j \quad (j = 1, \ldots, d-1), \quad \Big(\prod_{\nu=1}^{d-1} x_\nu\Big)^{\frac{n-1}{d}} < \frac{c'\gamma(\mathsf{k})}{X},$$
$$x_{d-1}^{\frac{1}{d-1}} < K\Big(\prod_{\nu=1}^{d-1} x_\nu\Big)^{\frac{1}{d}}, \quad \mathsf{k} \in \mathrm{SO}(d).$$

Since $\gamma(\mathsf{k})$ non-zero on a set of strictly positive measure, we find a set in $\mathrm{SO}(d)$ of strictly positive measure for which $\gamma(\mathsf{k}) \geq \epsilon$, for some small constant $\epsilon > 0$. We also assume K is so small that $K^d \leq \frac{2}{\sqrt{3}}$. Notice also that $1 < (\frac{2}{\sqrt{3}})^j$. Altogether we have therefore

$$J(X) \gg \int \prod_{j=1}^{d-1} x_j^{d-j} \frac{dx_j}{x_j}, \tag{3.65}$$

with range of integration

$$0 < x_j < 1 \quad (j = 1, \ldots, d-2), \quad 0 < x_{d-1} < K^{d(d-1)},$$
$$\Big(\prod_{\nu=1}^{d-1} x_\nu\Big)^{\frac{n-1}{d}} < \frac{c'\epsilon}{X}, \quad x_{d-1}^{\frac{1}{d-1}} < K\Big(\prod_{\nu=1}^{d-1} x_\nu\Big)^{\frac{1}{d}}.$$

With the abbreviation $Y = (\frac{X}{c'\epsilon})^{\frac{d}{n-1}}$ these conditions can also be written as

$$0 < x_j < 1 \quad (j = 1, \ldots, d-2), \quad 0 < x_{d-1} < K^{d(d-1)},$$
$$x_{d-1} < \frac{1}{Y x_1 \cdots x_{d-2}}, \quad x_{d-1} < K^{d(d-1)} (x_1 \cdots x_{d-2})^{d-1},$$

which, since $(x_1 \cdots x_{d-2})^{d-1} < 1$ for $0 < x_j < 1$, are equivalent to

$$0 < x_j < 1 \quad (j = 1, \ldots, d-2),$$
$$0 < x_{d-1} < \frac{1}{Y x_1 \cdots x_{d-2}}, \quad x_{d-1} < K^{d(d-1)} (x_1 \cdots x_{d-2})^{d-1},$$

and thus

$$\int \prod_{j=1}^{d-1} x_j^{d-j} \frac{dx_j}{x_j}$$

$$= \int_{\substack{0<x_j<1 \\ (j=1,\dots,d-2)}} \min\left\{K^{d(d-1)}(x_1 \cdots x_{d-2})^{d-1}, \frac{1}{Y x_1 \cdots x_{d-2}}\right\} \prod_{j=1}^{d-2} x_j^{d-j} \frac{dx_j}{x_j}$$

$$= \int_{\substack{0<x_j<1 \\ (j=1,\dots,d-2)}} \min\left\{K^{d(d-1)}, \frac{1}{Y x_1^d \cdots x_{d-2}^d}\right\} \prod_{j=1}^{d-2} x_j^{2d-j-1} \frac{dx_j}{x_j}$$

$$= \int_{\substack{0<x_j<1 (j=1,\dots,d-2) \\ 0<x_{d-1}<K^{d(d-1)} \\ x_{d-1}\prod_{j=1}^{d-2} x_j^d < \frac{1}{Y}}} \prod_{j=1}^{d-2} x_j^{2d-j-1} \frac{dx_j}{x_j}\, dx_{d-1}.$$

A simple change of variables $\xi_j = x_j^d$ $(j = 1, \dots, d-2)$ permits the application of Lemma 3.14 with $r_M = 0$; this gives the correct asymptotics $\sim \frac{1}{Y}$. □

3.7 *Averages over* $\mathrm{SL}(d,\mathbb{R})/\mathrm{SL}(d,\mathbb{Z})$

Using classical reduction theory [37], we will now indicate how to calculate averages over $\mathrm{SL}(d,\mathbb{R})/\mathrm{SL}(d,\mathbb{Z})$ of sums of the type

$$G(\mathsf{M}) = \sum_{\boldsymbol{m}^1,\dots,\boldsymbol{m}^r \in \mathbb{Z}^d} F(\mathsf{M}\boldsymbol{m}^1, \dots, \mathsf{M}\boldsymbol{m}^r) \tag{3.66}$$

for $r < d$ and F compactly supported. The moments of the n-point correlation densities will turn out to be special cases of the above (see next section). Explicit formulas will be given only in the cases $r = 1, 2$. The general case is as elementary but more cumbersome to write down.

For a matrix $\mathsf{M} = (x_{ij})_{i,j=1,\dots,d}$ the Haar measure $d\mu$ on $\mathrm{SL}(d,\mathbb{R})$ can be written as

$$d\mu = \delta(1 - \det \mathsf{M}) \prod_{i,j=1}^{d} dx_{ij}. \tag{3.67}$$

With this choice, the volume of $\Sigma^d = \mathrm{SL}(d,\mathbb{R})/\mathrm{SL}(d,\mathbb{Z})$ is given by [37]

$$\mu(\Sigma^d) = \prod_{n=2}^{d} \zeta(n), \qquad \text{with} \quad \zeta(n) = \sum_{\nu=1}^{\infty} \frac{1}{\nu^n}. \tag{3.68}$$

Note that the Haar measure $d\mu$ defined in (3.67) differs from the one used in Section 3.5 by a positive constant, which is worked out in [7]. *Throughout* *this section, we shall mean by $d\mu$ always the Haar measure normalized such that (3.68) holds.* We furthermore set formally $\mu(\Sigma^1)=1$.

Consider now the decomposition

$$\begin{pmatrix} x_{11} & \cdots & x_{1d} \\ \vdots & \ddots & \vdots \\ x_{d1} & \cdots & x_{dd} \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{1r} & & & \\ \vdots & \ddots & \vdots & & \mathbf{0} & \\ x_{r1} & \cdots & x_{rr} & & & \\ x_{r+1,1} & \cdots & x_{r+1,r} & \det^{-1}(x_{ij})_{i,j\leq r} & & \\ \vdots & & \vdots & & 1 & \\ \vdots & & \vdots & \mathbf{0} & & \ddots \\ x_{d1} & \cdots & x_{dr} & & & & 1 \end{pmatrix} \times$$

$$\times \begin{pmatrix} 1 & & & a_{11} & \cdots & a_{1,d-r} \\ & \ddots & & \vdots & & \vdots \\ & & 1 & a_{r1} & \cdots & a_{r,d-r} \\ & & & \tilde{x}_{11} & \cdots & \tilde{x}_{1,d-r} \\ & & & \vdots & \ddots & \vdots \\ & & & \tilde{x}_{d-r,1} & \cdots & \tilde{x}_{d-r,d-r} \end{pmatrix} \tag{3.69}$$

where $\tilde{\mathsf{M}} = (\tilde{x}_{ij})_{i,j=1,\ldots,d-r}$ is in $\mathrm{SL}(d-r,\mathbb{R})$. This defines a set of new coordinates for $\mathrm{SL}(d,\mathbb{R})$, and, after a straightforward calculation, the Haar measure defined in (3.67) reads in these coordinates

$$d\mu = \Bigg(\prod_{\substack{i\leq d\\ j\leq r}} dx_{ij}\Bigg)\Bigg(\prod_{\substack{i\leq r\\ j\leq d-r}} da_{ij}\Bigg)\Bigg(\delta(1-\det\tilde{\mathsf{M}})\prod_{i,j=1}^{d-r} d\tilde{x}_{ij}\Bigg). \tag{3.70}$$

Let us start with the case $r=1$.

Theorem 3.15 *Let F be piecewise continuous and of compact support. Then*

$$\frac{1}{\mu(\Sigma^d)}\int_{\Sigma^d}\sum_{\boldsymbol{m}\in\mathbb{Z}^d} F(\mathsf{M}\boldsymbol{m})\, d\mu = F(\mathbf{0}) + \int_{\mathbb{R}^d} F(\boldsymbol{x})\, d\boldsymbol{x}.$$

Proof. Write $\boldsymbol{m}=t\boldsymbol{c}$ where $t=\gcd(m_1,\ldots,m_d)>0$. Then there is an element $\mathsf{g}\in\Gamma=\mathrm{SL}(d,\mathbb{Z})$ such that we have [37]

$$\boldsymbol{c} = \mathsf{g}\begin{pmatrix}1\\0\\ \vdots\\0\end{pmatrix}. \tag{3.71}$$

The stabilizer Γ_S of the unit vector on the right-hand side, i.e. the set of all $\mathsf{g} \in \Gamma$ with

$$\mathsf{g} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

is given by

$$\Gamma_S = \left\{ \begin{pmatrix} 1 & k_1 & \cdots & k_{d-1} \\ 0 & & & \\ \vdots & & \mathsf{g} & \\ 0 & & & \end{pmatrix} : k_i \in \mathbb{Z},\ \mathsf{g} \in \mathrm{SL}(d-1, \mathbb{Z}) \right\}. \tag{3.72}$$

Hence there is a one-to-one correspondence between the primitive vectors $\boldsymbol{c}$ and the coset Γ/Γ_S, and the integral under consideration can be rewritten as

$$\begin{aligned} &\frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \sum_{\boldsymbol{m} \in \mathbb{Z}^d} F(\mathsf{M}\boldsymbol{m})\, d\mu \\ &\qquad = F(\mathbf{0}) + \frac{1}{\mu(\Sigma^d)} \sum_{t=1}^{\infty} \int_{\Sigma^d} \sum_{\mathsf{g} \in \Gamma/\Gamma_S} F(\mathsf{M}\mathsf{g} \begin{pmatrix} t \\ 0 \\ \vdots \\ 0 \end{pmatrix})\, d\mu \\ &\qquad\qquad = F(\mathbf{0}) + \frac{1}{\mu(\Sigma^d)} \sum_{t=1}^{\infty} \int_{\mathrm{SL}(d,\mathbb{R})/\Gamma_S} F(\mathsf{M} \begin{pmatrix} t \\ 0 \\ \vdots \\ 0 \end{pmatrix})\, d\mu. \end{aligned} \tag{3.73}$$

In terms of the coordinates (3.69) for $r = 1$ a fundamental region for $\mathrm{SL}(d, \mathbb{R})/\Gamma_S$ is

$$\{x_{i1} \in \mathbb{R},\ a_{1j} \in [0, 1),\ \tilde{\mathsf{M}} \in \mathcal{F}_{\mathrm{SL}(d-1,\mathbb{R})/\,\mathrm{SL}(d-1,\mathbb{Z})}\},$$

where $\mathcal{F}_{\mathrm{SL}(d-1,\mathbb{R})/\,\mathrm{SL}(d-1,\mathbb{Z})}$ is some fundamental region for $\mathrm{SL}(d-1, \mathbb{R})/\,\mathrm{SL}(d-1, \mathbb{Z})$. Hence our integral yields

$$F(\mathbf{0}) + \frac{\mu(\Sigma^{d-1})}{\mu(\Sigma^d)} \sum_{t=1}^{\infty} \int_{\mathbb{R}^d} F(t\boldsymbol{x}) d\boldsymbol{x} = F(\mathbf{0}) + \frac{\mu(\Sigma^{d-1})}{\mu(\Sigma^d)} \sum_{t=1}^{\infty} \frac{1}{t^d} \int_{\mathbb{R}^d} F(\boldsymbol{x}) d\boldsymbol{x}, \tag{3.74}$$

which in turn equals

$$F(\mathbf{0}) + \frac{1}{\zeta(d)} \sum_{t=1}^{\infty} \frac{1}{t^d} \int_{\mathbb{R}^d} F(\boldsymbol{x}) d\boldsymbol{x} = F(\mathbf{0}) + \int_{\mathbb{R}^d} F(\boldsymbol{x}) d\boldsymbol{x}. \tag{3.75}$$

□

Remark on the proof. The above argumentation leading to (3.74) can in fact be used to calculate quickly volume formula (3.68), which was only used in the very last step to deduce (3.75): Relation (3.74) states that

$$\frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \sum_{\boldsymbol{m}\in\mathbb{Z}^d-\{\boldsymbol{0}\}} F(\mathsf{M}\boldsymbol{m})\, d\mu = C \int_{\mathbb{R}^d} F(\boldsymbol{x})\, d\boldsymbol{x} \tag{3.76}$$

with the (unknown, we pretend) constant

$$C = \frac{\mu(\Sigma^{d-1})}{\mu(\Sigma^d)}\, \zeta(d) < \infty,$$

which is independent of F. On the other hand, we also have

$$\frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \sum_{\boldsymbol{m}\in\mathbb{Z}^d-\{\boldsymbol{0}\}} \epsilon^d F(\epsilon\mathsf{M}\boldsymbol{m})\, d\mu = C \int_{\mathbb{R}^d} F(\boldsymbol{x})\, d\boldsymbol{x} \tag{3.77}$$

for every $\epsilon > 0$. Since C is independent of F (and thus also of ϵ), and since

$$\lim_{\epsilon\to 0} \epsilon^d \sum_{\boldsymbol{m}\in\mathbb{Z}^d-\{\boldsymbol{0}\}} F(\epsilon\mathsf{M}\boldsymbol{m}) = \lim_{\epsilon\to 0} \epsilon^d \sum_{\boldsymbol{m}\in\mathbb{Z}^d} F(\epsilon\mathsf{M}\boldsymbol{m}) = \int_{\mathbb{R}^d} F(\boldsymbol{x})\, d\boldsymbol{x}, \tag{3.78}$$

there is no other choice than $C = 1$, and so $\mu(\Sigma^d) = \zeta(d)\mu(\Sigma^{d-1}) = \zeta(d)\cdots\zeta(2)$. □

The last observation will simplify calculations for higher $r > 1$, e.g. for $r = 2$:

Theorem 3.16 *Assume $d > 2$, and let F be piecewise continuous and of compact support. Then*

$$\begin{aligned}
\frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \sum_{\boldsymbol{m}^1,\boldsymbol{m}^2\in\mathbb{Z}^d} & F(\mathsf{M}\boldsymbol{m}^1, \mathsf{M}\boldsymbol{m}^2)\, d\mu \\
&= F(\boldsymbol{0},\boldsymbol{0}) + \int_{\mathbb{R}^d} \big\{F(\boldsymbol{x},\boldsymbol{0}) + F(\boldsymbol{0},\boldsymbol{x})\big\}\, d\boldsymbol{x} \\
&\quad + \sum_{\substack{t^1,t^2=1\\ \gcd(t^1,t^2)=1}}^{\infty} \int_{\mathbb{R}^d} \big\{F(t^1\boldsymbol{x}, t^2\boldsymbol{x}) + F(t^1\boldsymbol{x}, -t^2\boldsymbol{x})\big\}\, d\boldsymbol{x} \\
&\quad + \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} F(\boldsymbol{x}^1,\boldsymbol{x}^2)\, d\boldsymbol{x}^1\, d\boldsymbol{x}^2.
\end{aligned}$$

In the case of $2 < r < d$ one has a similar sum over integrals of F over hyperplanes of dimension $0, d, 2d, \ldots, rd$.

Proof. Put $\boldsymbol{m}^j = t^j \boldsymbol{c}^j$ where $t^j = \gcd(m_1^j, \ldots, m_d^j) > 0$ for $j = 1, 2$. As above, we have for some $\mathsf{g} \in \Gamma = \mathrm{SL}(d, \mathbb{Z})$ the relation

$$(\boldsymbol{c}^1, \boldsymbol{c}^2) = \mathsf{g} \begin{pmatrix} 1 & p \\ 0 & q \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix}, \tag{3.79}$$

where either $(p,q) = (0,0)$, $(p,q) = (\pm 1, 0)$, $(p,q) = (0,1)$, or (p,q) with $p \neq 0$, $q > 1$, $\gcd(p,q) = \pm 1$. The first case corresponds to the contributions

$$\frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \sum_{\boldsymbol{m} \in \mathbb{Z}^d} \big\{ F(\boldsymbol{0}, \mathsf{M}\boldsymbol{m}) + F(\mathsf{M}\boldsymbol{m}, \boldsymbol{0}) \big\} \, d\mu,$$

and the second one to

$$\frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \sum_{\substack{t^1, t^2 = 1 \\ \gcd(t^1, t^2) = 1}}^{\infty} \sum_{\boldsymbol{m} \in \mathbb{Z}^d} \big\{ F(t^1 \mathsf{M}\boldsymbol{m}, t^2 \mathsf{M}\boldsymbol{m}) + F(t^1 \mathsf{M}\boldsymbol{m}, -t^2 \mathsf{M}\boldsymbol{m}) \big\} \, d\mu.$$

These cases can therefore be dealt with the identical argument of the previous proof; this explains all terms up to the last one. Assume therefore in the sequel $(p,q) = (0,1)$ or $(p \neq 0, q > 1)$ with $\gcd(p,q) = \pm 1$. The stabilizer Γ_S of the matrix on the right-hand-side of (3.79) (p, q fixed) reads now

$$\Gamma_S = \left\{ \begin{pmatrix} 1 & 0 & k_{11} & \cdots & k_{1,d-2} \\ 0 & 1 & k_{21} & \cdots & k_{2,d-2} \\ 0 & 0 & & & \\ \vdots & \vdots & & \mathsf{g} & \\ 0 & 0 & & & \end{pmatrix} : k_{ij} \in \mathbb{Z},\ \mathsf{g} \in \mathrm{SL}(d-2, \mathbb{Z}) \right\}, \tag{3.80}$$

and, by the same reasoning as in the previous proof, the integral

$$\frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \sum_{t^1, t^2} \sum_{\boldsymbol{c}^1, \boldsymbol{c}^2} F(t^1 \mathsf{M}\boldsymbol{c}^1, t^2 \mathsf{M}\boldsymbol{c}^2) \, d\mu,$$

where the second sum runs over $\boldsymbol{c}^1, \boldsymbol{c}^2$ corresponding to (p,q) of the above type, equals

$$\frac{1}{\mu(\Sigma^d)} \int_{\mathrm{SL}(d,\mathbb{R})/\Gamma_S} \sum_{t^1,t^2=1}^{\infty} \sum_{(p,q)} F(t^1 \mathsf{M} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, t^2 \mathsf{M} \begin{pmatrix} p \\ q \\ 0 \\ \vdots \\ 0 \end{pmatrix}) \, d\mu$$

$$= \frac{\mu(\Sigma^{d-2})}{\mu(\Sigma^d)} \sum_{t^1,t^2=1}^{\infty} \sum_{(p,q)} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} F(t^1 \boldsymbol{x}^1, t^2 p \boldsymbol{x}^1 + t^2 q \boldsymbol{x}^2) \, d\boldsymbol{x}^1 \, d\boldsymbol{x}^2$$

$$= C \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} F(\boldsymbol{x}^1, \boldsymbol{x}^2) \, d\boldsymbol{x}^1 \, d\boldsymbol{x}^2, \quad (3.81)$$

for some constant $C < \infty$. Hence we have

$$\frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \Bigg\{ \sum_{\boldsymbol{m}^1,\boldsymbol{m}^2 \in \mathbb{Z}^d} F(\mathsf{M}\boldsymbol{m}^1, \mathsf{M}\boldsymbol{m}^2) - F(\boldsymbol{0},\boldsymbol{0}) - \sum_{\boldsymbol{m} \in \mathbb{Z}^d} \big\{ F(\boldsymbol{0}, \mathsf{M}\boldsymbol{m}) + F(\mathsf{M}\boldsymbol{m}, \boldsymbol{0}) \big\}$$
$$- \sum_{\substack{t^1,t^2=1 \\ \gcd(t^1,t^2)=1}}^{\infty} \sum_{\boldsymbol{m} \in \mathbb{Z}^d} \big\{ F(t^1 \mathsf{M}\boldsymbol{m}, t^2 \mathsf{M}\boldsymbol{m}) + F(t^1 \mathsf{M}\boldsymbol{m}, -t^2 \mathsf{M}\boldsymbol{m}) \big\} \Bigg\} \, d\mu$$
$$= C \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} F(\boldsymbol{x}^1, \boldsymbol{x}^2) \, d\boldsymbol{x}^1 \, d\boldsymbol{x}^2. \quad (3.82)$$

We may replace on the left-hand-side $F(\boldsymbol{x}^1, \boldsymbol{x}^2)$ by $\epsilon^{2d} F(\epsilon \boldsymbol{x}^1, \epsilon \boldsymbol{x}^2)$ without changing the right-hand-side. Now observe that in the following only the first term is of leading order in ϵ, and in fact, converges to a Riemann integral:

$$\lim_{\epsilon \to 0} \epsilon^{2d} \Bigg\{ \sum_{\boldsymbol{m}^1,\boldsymbol{m}^2 \in \mathbb{Z}^d} F(\epsilon \mathsf{M}\boldsymbol{m}^1, \epsilon \mathsf{M}\boldsymbol{m}^2) - F(\boldsymbol{0},\boldsymbol{0}) - \sum_{\boldsymbol{m} \in \mathbb{Z}^d} \big\{ F(\boldsymbol{0}, \epsilon \mathsf{M}\boldsymbol{m}) + F(\epsilon \mathsf{M}\boldsymbol{m}, \boldsymbol{0}) \big\}$$
$$- \sum_{\substack{t^1,t^2=1 \\ \gcd(t^1,t^2)=1}}^{\infty} \sum_{\epsilon \boldsymbol{m} \in \mathbb{Z}^d} \big\{ F(\epsilon t^1 \mathsf{M}\boldsymbol{m}, \epsilon t^2 \mathsf{M}\boldsymbol{m}) + F(\epsilon t^1 \mathsf{M}\boldsymbol{m}, -\epsilon t^2 \mathsf{M}\boldsymbol{m}) \big\} \Bigg\} \, d\mu$$
$$= \lim_{\epsilon \to 0} \epsilon^{2d} \sum_{\boldsymbol{m}^1,\boldsymbol{m}^2 \in \mathbb{Z}^d} F(\epsilon \mathsf{M}\boldsymbol{m}^1, \epsilon \mathsf{M}\boldsymbol{m}^2) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} F(\boldsymbol{x}^1, \boldsymbol{x}^2) \, d\boldsymbol{x}^1 \, d\boldsymbol{x}^2; \quad (3.83)$$

thus $C = 1$. $\square$

3.8 Moments

We first need to work out the average of the function τ_n defined in (3.12).

Lemma 3.17

$$\int_{\mathbb{R}^{n-1}} \tau_n(\mathcal{X})\, d\mathcal{X} = 1.$$

Proof. Consider the identity

$$1 = \frac{1}{N^n} \sum_{m_1,\dots,m_n=1}^{N} 1. \tag{3.84}$$

We now use the argument of Section 3.2 by rewriting the sum over m_j as a sum over $k_1 = m_1 - m_2, \dots, k_{n-1} = m_{n-1} - m_n$ with multiplicities

$$N - \max_{1\le a\le b\le n-1} \Big|\sum_{i=a}^{b} k_i\Big|,$$

yielding

$$\frac{1}{N^n} \sum_{m_1,\dots,m_n=1}^{N} 1 = \frac{1}{N^{n-1}} \sum_{\mathcal{K}\in\mathbb{Z}^{n-1}} \tau_n\Big(\frac{\mathcal{K}}{N}\Big), \tag{3.85}$$

which in the limit $N \to \infty$ converges to the Riemann integral over τ_n. □

The first two moments of the limit distribution for the two-point correlations have the following explicit expression.

Theorem 3.18 *Assume for simplicity $\mathcal{B}$ is an interval symmetric about the origin. Then the expectation value is given by*

$$\int_0^\infty X \, d\Psi_{2,d}(X) = |\mathcal{B}|,$$

and the second moment reads for $d > 2$

$$\int_0^\infty X^2 \, d\Psi_{2,d}(X) = |\mathcal{B}|^2 + 2\Bigg\{ \Big(\frac{2}{3}\Big)^{d-1} + 2 \sum_{\substack{p,q=1\\ p<q\\ \gcd(p,q)=1}}^{\infty} \frac{1}{q^d}\Big(1 - \frac{1}{3}\frac{p}{q}\Big)^{d-1} \Bigg\} |\mathcal{B}|.$$

Proof. Notice that we have for the expectation value

$$\int_0^\infty X\, d\Psi_{2,d}(X) = \frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \overline{D}_2(\mathcal{B}, \mathsf{M})\, d\mu \tag{3.86}$$

which may be written as

$$\frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \left[\sum_{\boldsymbol{m}} F(\mathsf{M}\boldsymbol{m}) - F(\mathbf{0}) \right] d\mu \tag{3.87}$$

with

$$F(\boldsymbol{x}) = \tau_2(x_1) \cdots \tau_2(x_{d-1})\, \chi_{\mathcal{B}}(x_d). \tag{3.88}$$

The expectation value can now be readily calculated from Theorem 3.15. The second moment follows similarly from Theorem 3.16 applied to

$$\begin{aligned} \int_0^\infty X^2\, d\Psi_{2,d}(X) &= \frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \overline{D}_2(\mathcal{B}, \mathsf{M})^2\, d\mu \\ &= \frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \left[\sum_{\boldsymbol{m}} F(\mathsf{M}\boldsymbol{m}) - F(\mathbf{0}) \right]^2 d\mu \end{aligned} \tag{3.89}$$

with

$$F(\boldsymbol{x}^1, \boldsymbol{x}^2) = \tau_2(x_1^1)\, \tau_2(x_1^2) \cdots \tau_2(x_{d-1}^1)\, \tau_2(x_{d-1}^2)\, \chi_{\mathcal{B}}(x_d^1)\, \chi_{\mathcal{B}}(x_d^2). \tag{3.90}$$

□

The expectation value of the three-point correlations is as follows.

Theorem 3.19 *Let $d > 2$. Then*

$$\begin{aligned} \int_0^\infty X\, d\Psi_{3,d}(X) = |\mathcal{B}| &+ \sum_{\substack{p,q=1 \\ \gcd(p,q)=1}}^\infty \frac{1}{(p+q)^{d-1}} \int_{\mathbb{R}} \chi_{\mathcal{B}}(px, qx)\, dx \\ &+ 2 \sum_{\substack{p,q=1 \\ p<q \\ \gcd(p,q)=1}}^\infty \frac{1}{q^{d-1}} \int_{\mathbb{R}} \chi_{\mathcal{B}}(px, -qx)\, dx. \end{aligned}$$

Proof. We have

$$\int_0^\infty X\, d\Psi_{3,d}(X) = \frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \overline{D}_3(\mathcal{B}, \mathsf{M})\, d\mu \tag{3.91}$$

which we write as

$$\frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \Bigg[\sum_{\boldsymbol{m}^1,\boldsymbol{m}^2} F(\mathsf{M}\boldsymbol{m}^1, \mathsf{M}\boldsymbol{m}^2) \\ - \sum_{\boldsymbol{m}} \{F(\mathsf{M}\boldsymbol{m}, \mathbf{0}) + F(\mathbf{0}, \mathsf{M}\boldsymbol{m}) + F(\mathsf{M}\boldsymbol{m}, -\mathsf{M}\boldsymbol{m})\} \\ - F(\mathbf{0},\mathbf{0}) \Bigg] d\mu \quad (3.92)$$

with

$$F(\boldsymbol{x}^1, \boldsymbol{x}^2) = \tau_3(x_1^1, x_1^2) \cdots \tau_3(x_{d-1}^1, x_{d-1}^2)\, \chi_{\mathcal{B}}(x_d^1, x_d^2). \quad (3.93)$$

Apply again Theorem 3.16 and Lemma 3.17. □

We conclude the discussion of moments with a remark on the general case.

Theorem 3.20 *We have*

$$\int_0^\infty X^k \, d\Psi_{n,d}(X) \begin{cases} < \infty & \text{if } d > (n-1)k \\ = \infty & \text{if } d \leq (n-1)k \end{cases}$$

and, for n, k fixed,

$$\lim_{d\to\infty} \int_0^\infty X^k \, d\Psi_{n,d}(X) = |\mathcal{B}|^k,$$

which coincides with the kth moment of the Dirac distribution

$$d\Psi(X) = \delta(X - |\mathcal{B}|)\, dX.$$

Proof. The first assertion follows from the tail estimate given in Theorem 3.11. The asymptotic behaviour for large d is evident in the cases $n = 2, 3$, $k = 1, 2$ discussed above. The general case is similar, the only term, which does not depend on d, is $|\mathcal{B}|^k$. All other terms involve convergent series (for d large enough) whose limit vanishes as $d \to \infty$. □

Some further analysis shows that the limit value $|\mathcal{B}|^k$ is approached exponentially fast in d (k, n fixed); in the cases $n = 2, 3$, $k = 1, 2$ this is evident from the explicit formulas given above.

4 Values in small intervals

A statistic which is closely related to the n-point correlation density is the probability of finding K values of the linear form $L_{\boldsymbol{m}}$ in a random interval $[\xi, \xi+\frac{\sigma}{N}]$ of the unit circle, where ξ is a uniformly distributed random variable on $[0,1)$, and σ is a fixed constant measuring the size of the interval in units of the mean spacing. To be more precise, we are interested in the random variable

$$\mathcal{N}_\sigma^{\boldsymbol{\alpha}}(\xi, N) = \sum_{m_1=1}^{N_1} \cdots \sum_{m_{d-1}=1}^{N_{d-1}} \sum_{\nu\in\mathbb{Z}} \chi_\sigma\big(N(L_{\boldsymbol{m}} - \xi + \nu)\big), \tag{4.1}$$

χ_σ being the characteristic function of the interval $(0, \sigma]$, and in its probability distribution

$$\mathcal{P}_\sigma^{\boldsymbol{\alpha}}(N) = \operatorname{Prob}\Big\{\mathcal{N}_\sigma^{\boldsymbol{\alpha}}(\,\cdot\,, N) = K\Big\}. \tag{4.2}$$

The expectation value is clearly

$$\mathbb{E}\mathcal{N}_\sigma^{\boldsymbol{\alpha}}(\,\cdot\,, N) = \int_0^1 \mathcal{N}_\sigma^{\boldsymbol{\alpha}}(\xi, N)\, d\xi = \sigma, \tag{4.3}$$

and a short calculation shows that the nth moment can be written as

$$\begin{aligned}
\mathbb{E}\mathcal{N}_\sigma^{\boldsymbol{\alpha}}(\,\cdot\,, N)^n &= \int_0^1 \mathcal{N}_\sigma^{\boldsymbol{\alpha}}(\xi, N)^n\, d\xi \\
&= \frac{1}{N} \sum_{m_1^1,\ldots,m_1^n=1}^{N_1} \cdots \sum_{m_{d-1}^1,\ldots,m_{d-1}^n=1}^{N_{d-1}} \sum_{\nu^1,\ldots,\nu^{n-1}\in\mathbb{Z}} \times \\
&\qquad \times \int_0^\sigma \chi_\sigma\big(N(L_{\boldsymbol{m}^1} - L_{\boldsymbol{m}^n} + \nu^1) + \xi\big) \times \\
&\qquad\quad \times \chi_\sigma\big(N(L_{\boldsymbol{m}^2} - L_{\boldsymbol{m}^n} + \nu^2) + \xi\big) \cdots \\
&\qquad\qquad \cdots \chi_\sigma\big(N(L_{\boldsymbol{m}^{n-1}} - L_{\boldsymbol{m}^n} + \nu^{n-1}) + \xi\big)\, d\xi,
\end{aligned} \tag{4.4}$$

and is therefore related to the n-point correlations via

$$\mathbb{E}\mathcal{N}_\sigma^{\boldsymbol{\alpha}}(\,\cdot\,, N)^n = \int_{\mathbb{R}^{n-1}} C_n^{\boldsymbol{\alpha}}(\mathcal{S}, N)\, \rho_\sigma(\mathcal{S})\, d^{n-1}\mathcal{S} \tag{4.5}$$

with

$$\begin{aligned}
C_n^{\boldsymbol{\alpha}}(\mathcal{S}, N) = \sum_{\boldsymbol{k}^1,\ldots,\boldsymbol{k}^{n-1}\in\mathbb{Z}^d} &\tau_n\Big(\frac{\mathcal{K}_1}{N_1}\Big) \cdots \tau_n\Big(\frac{\mathcal{K}_{d-1}}{N_{d-1}}\Big) \times \\
&\times \delta(s^1 - N\boldsymbol{\alpha}\boldsymbol{k}^1) \cdots \delta(s^{n-1} - N\boldsymbol{\alpha}\boldsymbol{k}^{n-1}),
\end{aligned} \tag{4.6}$$

and

$$\rho_\sigma(\mathcal{S}) = \int_0^\sigma \chi_\sigma(s^1+\cdots+s^{n-1}+\xi)\times \\ \times\chi_\sigma(s^2+\cdots+s^{n-1}+\xi)\cdots\chi_\sigma(s^{n-2}+s^{n-1}+\xi)\,\chi_\sigma(s^{n-1}+\xi)\,d\xi,$$

which clearly has compact support. Its L^1-norm reads

$$\int_{\mathbb{R}^{n-1}} \rho_\sigma(\mathcal{S})\,d\mathcal{S} = \sigma^n. \tag{4.7}$$

The combinatorial argument which relates the n-point densities $C_n^{\boldsymbol{\alpha}}(\mathcal{S},N)$ and $R_n^{\boldsymbol{\alpha}}(\mathcal{S},N)$ can be found e.g. in [25, 30].

The invariance properties of $\mathcal{N}_\sigma^{\boldsymbol{\alpha}}(\,\cdot\,,N)$ itself can be expressed in terms of the group G defined as the semi-direct product $\mathrm{G}=\mathrm{SL}(d,\mathbb{R})\ltimes\mathbb{R}^d$ with multiplication law

$$\{\mathsf{M}_1,\boldsymbol{x}_1\}\{\mathsf{M}_2,\boldsymbol{x}_2\} = \{\mathsf{M}_1\mathsf{M}_2,\boldsymbol{x}_1+\mathsf{M}_1\boldsymbol{x}_2\}. \tag{4.8}$$

A natural action of this group on $\mathbb{R}^d$ is given by $\{\mathsf{M},\boldsymbol{x}\}\boldsymbol{y}=\boldsymbol{x}+\mathsf{M}\boldsymbol{y}$. We now define the function

$$\mathcal{V}_\sigma(\{\mathsf{M},\boldsymbol{x}\}) = \sum_{\boldsymbol{m}\in\mathbb{Z}^d} \chi_1(\tilde m_1)\cdots\chi_1(\tilde m_{d-1})\chi_\sigma(\tilde m_d), \tag{4.9}$$

where $\tilde{\boldsymbol{m}}=\{\mathsf{M},\boldsymbol{x}\}\boldsymbol{m}$. By a similar argument as in the last section, this function can be shown to be right-invariant under the discrete subgroup $\Lambda=\mathrm{SL}(d,\mathbb{Z})\ltimes\mathbb{Z}^d$ and may thus be viewed as a function on the manifold $\Omega=\mathrm{G}/\Lambda$. We clearly have

$$\mathcal{N}_\sigma^{\boldsymbol{\alpha}}(\xi,N) = \mathcal{V}_\sigma(\{\mathsf{M},\boldsymbol{x}\}) \tag{4.10}$$

for

$$\{\mathsf{M},\boldsymbol{x}\} = \left\{ \begin{pmatrix} N_1^{-1} & & & \\ & \ddots & & \\ & & N_{d-1}^{-1} & \\ & & & N \end{pmatrix} \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ \alpha_1 & \cdots & \alpha_{d-1} & 1 \end{pmatrix}, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ N\xi \end{pmatrix} \right\}.$$

As for the n-point correlations, the representation (4.10) implies that we cannot expect $\mathcal{P}_\sigma^{\boldsymbol{\alpha}}(N)$ to converge, since it behaves similar as a function on Σ^d along some (in general infinite) trajectory.

The following observation will be useful later on (Theorem 4.2).

Lemma 4.1 *Fix some non-negative integer K. The solutions M of the equation*

$$\mathcal{V}_\sigma(\{\mathsf{M},\boldsymbol{x}\}) = K \tag{4.11}$$

form a set with boundary of measure zero in Ω.

Proof. As in the proof of Lemma 3.10, cover Ω with countably many compact sets and consider the above equation on each set separately. On each compact set the sum representing $\mathcal{V}_\sigma(\{\mathsf{M},\boldsymbol{x}\})$ is a finite superposition of piecewise constant functions which take values in the non-negative integers. The boundary of the set of solutions of (4.11) is therefore contained in the set of discontinuities, and thus of measure zero. □

4.1 Limit distributions. Tail estimates

The following results are again a consequence of equidistribution (this time on $\mathrm{SL}(d,\mathbb{R}) \ltimes \mathbb{R}^d$), as we shall see in the next section. In the sequel put again $N_j = d_j N^{1/(d-1)}$, for arbitrary constants $d_j > 0$ ($d_1 \cdots d_{d-1} = 1$). We define $\mathcal{N}_\sigma(\xi, N)$ at non-integer N_j by simply replacing N_j by its integral part in the definition (4.1).

Theorem 4.2 *Suppose $\alpha_1, \ldots, \alpha_{d-1} \in \mathbb{T}$ are random variables with continuous joint probability density $h(\alpha_1, \ldots, \alpha_{d-1})$, and ξ is a uniformly distributed random variable on $[0,1)$. Then the limit*

$$\lim_{N\to\infty} \operatorname{Prob}\Big\{\mathcal{N}_\sigma(\,\cdot\,,N) = K\Big\}$$

exists for all $\sigma \in \mathbb{R}_+$, $K = 0,1,2,\ldots,$ and is given by

$$P_K(\sigma) = \frac{\mu_{\mathrm{G}}\Big\{\{\mathsf{M},\boldsymbol{x}\} \in \Omega : \mathcal{V}_\sigma(\{\mathsf{M},\boldsymbol{x}\}) = K\Big\}}{\mu_{\mathrm{G}}(\Omega)},$$

which is independent of h.

Note that $P_K(\sigma)$ is well defined, due to Lemma 4.1.

Theorem 4.3 *For K_0 large enough, there exist constants $0 < C_1(\sigma) \leq C_2(\sigma) < \infty$ such that*

$$C_1(\sigma) X^{-(d+1)} \leq \sum_{K=X}^{\infty} P_K(\sigma) \leq C_2(\sigma) X^{-(d+1)}$$

for all $X > X_0$, i.e. $P_K(\sigma)$ has a power-like tail.

The above theorems generalize a result of Mazel and Sinai [24] valid for $d = 2$, constant h, $\sigma \leq 1$, which was obtained by different methods. They are able to give an explicit formula for the limit distribution since in the range $\sigma \leq 1$ the sum

$$\sum_{\boldsymbol{m} \in \mathbb{Z}^2} \chi_1(\tilde{m}_1)\chi_\sigma(\tilde{m}_2),$$

defining $\mathcal{V}_\sigma(\{\mathsf{M}, \boldsymbol{x}\})$ has a particularly tractable form. In principle, it should also be possible to calculate explicitly the limit distribution in the above range for higher $d > 2$, but this requires a more detailed study of the geometry of the fundamental domain than we want to carry out here.

The probability of finding values close to zero has the following limiting behaviour.

Theorem 4.4 *Suppose $\alpha_1, \ldots, \alpha_{d-1} \in \mathbb{T}$ are random variables with continuous joint probability density $h(\alpha_1, \ldots, \alpha_{d-1})$. Then the limit*

$$\lim_{N\to\infty} \operatorname{Prob}\Big\{ \mathcal{N}_\sigma(0, N) = K \Big\}$$

exists for all $\sigma \in \mathbb{R}_+$, $K = 0, 1, 2, \ldots$, and is given by

$$P_{K,0}(\sigma) = \frac{\mu\Big\{ \mathsf{M} \in \Sigma^d : \mathcal{V}_\sigma(\{\mathsf{M}, \mathbf{0}\}) = K \Big\}}{\mu(\Sigma^d)},$$

which is independent of h.

Theorem 4.5 *For K_0 large enough, there exist constants $0 < C_1(\sigma) \leq C_2(\sigma) < \infty$ such that*

$$C_1(\sigma) X^{-d} \leq \sum_{K=X}^{\infty} P_{K,0}(\sigma) \leq C_2(\sigma) X^{-d}$$

for all $X > X_0$, i.e. $P_{K,0}(\sigma)$ has a power-like tail.

The proofs of the last two theorems are almost identical with those for the two-point correlations (and will therefore be omitted), since $\mathcal{V}_\sigma(\{\mathsf{M}, \mathbf{0}\})$ can be identified with a function on Σ^d.

We conclude this section with the limit theorems of the trigonometric sums $W_N(\xi)$,

$$W_N(\xi) = \frac{1}{N} \sum_{\nu=1}^{N} \sum_{m_1=1}^{N_1} \cdots \sum_{m_{d-1}=1}^{N_{d-1}} \cos\big(2\pi\nu(L_{\boldsymbol{m}} + \xi)\big), \tag{4.12}$$

compare (2.2); recall that $N = N_1 \cdots N_{d-1}$, hence $\frac{1}{N}$ normalizes exactly by the square-root of number of summands involved.

Theorem 4.6 *Suppose $\alpha_1, \ldots, \alpha_{d-1} \in \mathbb{T}$ are random variables with continuous joint probability density $h(\alpha_1, \ldots, \alpha_{d-1})$, and ξ is a uniformly distributed random variable on $[0,1)$. Then there is a function $\Phi_d(a,b)$, decreasing in a, increasing in b and continuous except for at most countably many $a, b \in \mathbb{R}$, such that*

$$\lim_{N\to\infty} \operatorname{Prob}\left\{ a < W_N < b \right\} = \Phi_d(a,b)$$

except possibly at the discontinuities of $\Phi_d(a,b)$.

Theorem 4.7 *For X_0 large enough, there exist constants $0 < C_1 \leq C_2 < \infty$ such that*

$$C_1 X^{-(d+1)} \leq 1 - \Phi_d(-X, X) \leq C_2 X^{-(d+1)}$$

for all $X > X_0$.

Theorem 4.8 *Let $d > 2$. Suppose $\alpha_1, \ldots, \alpha_{d-1} \in \mathbb{T}$ are random variables with continuous joint probability density $h(\alpha_1, \ldots, \alpha_{d-1})$. Then there is a function $\Phi_{d,0}(a,b)$, decreasing in a, increasing in b and continuous except for at most countably many $a, b \in \mathbb{R}$, such that*

$$\lim_{N\to\infty} \operatorname{Prob}\left\{ a < W_N(0) < b \right\} = \Phi_{d,0}(a,b)$$

except possibly at the discontinuities of $\Phi_{d,0}(a,b)$.

We are only able to prove the above theorem in the case $d = 2$, when W_N is replaced by the smoothed sum

$$\tilde{W}_N(\xi) = \frac{1}{N} \sum_{\nu=1}^{\infty} \sum_{m_1=1}^{N_1} \cdots \sum_{m_{d-1}=1}^{N_{d-1}} f(\frac{\nu}{N}) \cos\big(2\pi\nu(\xi + L_{\boldsymbol{m}})\big), \tag{4.13}$$

where the cut-off function f is of Schwartz class, e.g. smooth and compactly supported. The limit distribution will then depend on f.

Theorem 4.9 *For X_0 large enough, there exist constants $0 < C_1 \leq C_2 < \infty$ such that*

$$C_1 X^{-d} \leq 1 - \Phi_{d,0}(-X, X) \leq C_2 X^{-d}$$

for all $X > X_0$.

The proof of these limit theorems combines arguments from the proofs of Theorems 4.2 – 4.3 with arguments from [21] and will only be sketched in Section 4.3.

4.2 *Flows on* $\mathrm{SL}(d,\mathbb{R}) \ltimes \mathbb{R}^2 / \mathrm{SL}(d,\mathbb{Z}) \ltimes \mathbb{Z}^2$

In analogy with the flow Φ^t discussed in the previous section we shall now consider the flow

$$\Theta^t : \quad \Omega \to \Omega,$$

$$\{\mathsf{M}, \boldsymbol{x}\} \mapsto \left\{ \begin{pmatrix} \mathrm{e}^{-t} & & & \\ & \ddots & & \\ & & \mathrm{e}^{-t} & \\ & & & \mathrm{e}^{(d-1)t} \end{pmatrix}, \mathbf{0} \right\} \{\mathsf{M}, \boldsymbol{x}\}.$$

In term of the coordinates $\{\mathsf{M}, \boldsymbol{x}\} = \{[s, \boldsymbol{l}, \boldsymbol{b}, \mathsf{A}], \boldsymbol{x}\}$ (recall the parametrization (3.18)) the action of the flow reads

$$\Theta^t(\{[s, \boldsymbol{l}, \boldsymbol{b}, \mathsf{A}], \boldsymbol{x}\}) = \left\{ [s+t, \boldsymbol{l}, \boldsymbol{b}\,\mathrm{e}^{-dt}, \mathsf{A}], \begin{pmatrix} x_1\,\mathrm{e}^{-t} \\ \vdots \\ x_{d-1}\,\mathrm{e}^{-t} \\ x_d\,\mathrm{e}^{(d-1)t} \end{pmatrix} \right\}$$

Let us now understand the following generalization of Corollary 3.9.

Theorem 4.10 *Let f be bounded and piecewise continuous on Ω, and h be continuous on the standard $(d-1)$-dimensional unit torus $\mathbb{T}^{d-1}$. Then, for every $s \in \mathbb{R}$, $\mathsf{A} \in \mathrm{SL}(d-1,\mathbb{R})$, we have*

$$\lim_{t\to\infty} \int_0^1 \int_{\mathbb{T}^{d-1}} f(\left\{ [s+t, \boldsymbol{l}, \mathbf{0}, \mathsf{A}], \begin{pmatrix} \mathbf{0} \\ \xi\,\mathrm{e}^{(d-1)t} \end{pmatrix} \right\})\, h(\boldsymbol{l})\, d^{d-1}\boldsymbol{l}\, d\xi$$
$$= \frac{1}{\mu_{\mathrm{G}}(\Omega)} \int_\Omega f\, d\mu_{\mathrm{G}} \int_{\mathbb{T}^{d-1}} h(\boldsymbol{l})\, d^{d-1}\boldsymbol{l},$$

where μ_{G} denotes the Haar measure of G.

Proof. As in the proof of Corollary 3.7 we may assume without loss of generality that $s = 0$, $\mathsf{A} = 1$ (same argument).

Since the function

$$F(\mathsf{M}) = \int_{\mathbb{T}^d} f(\{\mathsf{M}, \mathbf{0}\}\{1, \boldsymbol{x}\})\, d^d\boldsymbol{x}$$

satisfies the conditions of Corollary 3.9, we have

$$\begin{aligned}\lim_{t\to\infty}\int_{\mathbb{T}^d}\int_{\mathbb{T}^{d-1}} & f(\{[t,\boldsymbol{l},\mathbf{0},1],\mathbf{0}\}\{1,\boldsymbol{x}\})\,h(\boldsymbol{l})\,d^{d-1}\boldsymbol{l}\,d^d\boldsymbol{x}\\ &= \frac{1}{\mu(\Sigma^d)}\int_{\Sigma^d} F\,d\mu\int_{\mathbb{T}^{d-1}} h(\boldsymbol{l})\,d^{d-1}\boldsymbol{l}\\ &= \frac{1}{\mu_{\mathrm{G}}(\Omega)}\int_{\Omega} f\,d\mu_{\mathrm{G}}\int_{\mathbb{T}^{d-1}} h(\boldsymbol{l})\,d^{d-1}\boldsymbol{l}. \qquad (4.14)\end{aligned}$$

(Notice that we may think of Ω as a product Σ^d times the unit torus $\mathbb{T}^d$, hence we have in particular $d\mu_{\mathrm{G}} = d\mu\, d^d\boldsymbol{x}$.)

Let us denote by $d_{\mathrm{G}}(\,\cdot\,,\,\cdot\,)$ the distance induced by the right-invariant metric on G. Then we have (with the abbreviation $\boldsymbol{lx} = l_1x_1+\cdots+l_{d-1}x_{d-1}$)

$$\begin{aligned} & d_{\mathrm{G}}\left(\{[t,\boldsymbol{l},\mathbf{0},1],\mathbf{0}\}\left\{1,\begin{pmatrix}x_1\\ \vdots\\ x_{d-1}\\ x_d\end{pmatrix}\right\},\left\{[t,\boldsymbol{l},\mathbf{0},1],\begin{pmatrix}0\\ \vdots\\ 0\\ (\boldsymbol{lx}+x_d)\,\mathrm{e}^{(d-1)t}\end{pmatrix}\right\}\right)\\ &= d_{\mathrm{G}}\left(\left\{[0,\mathbf{0},\mathbf{0},1],\begin{pmatrix}x_1\,\mathrm{e}^{-t}\\ \vdots\\ x_{d-1}\,\mathrm{e}^{-t}\\ 0\end{pmatrix}\right\},\left\{[0,\mathbf{0},\mathbf{0},1],\begin{pmatrix}0\\ \vdots\\ 0\\ 0\end{pmatrix}\right\}\right) \ll \mathrm{e}^{-t}, \qquad (4.15)\end{aligned}$$

uniformly for $x_j \in [0,1)$. Therefore, if f is continuous and has compact support, we obtain by uniform continuity

$$\begin{aligned}&\lim_{t\to\infty}\int_{\mathbb{T}^d}\int_{\mathbb{T}^{d-1}} f(\{[t,\boldsymbol{l},\mathbf{0},1],\mathbf{0}\}\{1,\boldsymbol{x}\})\,h(\boldsymbol{l})\,d^{d-1}\boldsymbol{l}\,d^d\boldsymbol{x}\\ &= \lim_{t\to\infty}\int_{\mathbb{T}^d}\int_{\mathbb{T}^{d-1}} f\Big(\Big\{[t,\boldsymbol{l},\mathbf{0},1],\begin{pmatrix}\mathbf{0}\\ (\boldsymbol{lx}+x_d)\,\mathrm{e}^{(d-1)t}\end{pmatrix}\Big\}\Big)\,h(\boldsymbol{l})\,d^{d-1}\boldsymbol{l}\,d^d\boldsymbol{x}. \qquad (4.16)\end{aligned}$$

Substituting in the last integral x_d by $\xi = x_d + \boldsymbol{lx}$ yields

$$\begin{aligned}&\int_{\mathbb{T}^d}\int_{\mathbb{T}^{d-1}} f\Big(\Big\{[t,\boldsymbol{l},\mathbf{0},1],\begin{pmatrix}\mathbf{0}\\ (\boldsymbol{lx}+x_d)\,\mathrm{e}^{(d-1)t}\end{pmatrix}\Big\}\Big)\,h(\boldsymbol{l})\,d^{d-1}\boldsymbol{l}\,d^d\boldsymbol{x}\\ &\qquad = \int_0^1\int_{\mathbb{T}^{d-1}} f\Big(\Big\{[t,\boldsymbol{l},\mathbf{0},1],\begin{pmatrix}\mathbf{0}\\ \xi\,\mathrm{e}^{(d-1)t}\end{pmatrix}\Big\}\Big)\,h(\boldsymbol{l})\,d^{d-1}\boldsymbol{l}\,d\xi, \qquad (4.17)\end{aligned}$$

which gives the theorem for compactly supported continuous functions. The result for bounded piecewise continuous f follows from approximation from above/below as before in the proof of Corollary 3.9. $\square$

4.3 Proof of Theorems 4.2 – 4.9

Proof of Theorem 4.2. Proceed similarly as for the Theorem 3.2, but use Lemma 4.1 instead of Lemma 3.10. □

For the proof of Theorem 4.3 we need the following bounds on $\mathcal{V}_\sigma(\{\mathsf{M}, \boldsymbol{x}\})$.

Proposition 4.11 *For K constant and small enough, we have (i) the upper bound*

$$\mathcal{V}_\sigma(\{\mathsf{M}, \boldsymbol{x}\}) \ll_K \frac{\gamma_{l,\max}(x_{l+1}a_{l+1}, \ldots, x_d a_d)}{a_1 \cdots a_l}$$

uniformly for all $\mathsf{M} = \mathsf{k}\,\mathsf{a}(a)\,\mathsf{n}(u) \in \mathcal{S}_{K,l}$, $\boldsymbol{x} \in [-\frac{1}{2}, \frac{1}{2}]^d$, and (ii) the lower bound

$$\frac{\gamma_{d-1}(\mathsf{k}, x_d a_d)}{a_1 \cdots a_{d-1}} \ll_K \mathcal{V}_\sigma(\{\mathsf{M}, \boldsymbol{x}\})$$

uniformly for all $\mathsf{M} = \mathsf{k}\,\mathsf{a}(a)\,\mathsf{n}(u) \in \mathcal{S}_{K,d-1}$, $\boldsymbol{x} \in [-\frac{1}{2}, \frac{1}{2}]^d$. The function $\gamma_{l,\max}(t_{l+1}, \ldots, t_d) \geq 0$ is compactly supported and bounded on $\mathbb{R}^{d-l}$, and non-zero on a set of strictly positive measure; the function $\gamma_{d-1}(\mathsf{k}, t_d) \geq 0$ is compactly supported and bounded on $\mathrm{SO}(d) \times \mathbb{R}$, and non-zero on a set of strictly positive measure.

Proof. The function $\mathcal{V}_\sigma(\{\mathsf{M}, \boldsymbol{x}\})$ has a representation of the form

$$\mathcal{V}_\sigma(\{\mathsf{M}, 0\}\{1, \boldsymbol{x}\}) = \sum_{\boldsymbol{m} \in \mathbb{Z}^d} F(\mathsf{M}(\boldsymbol{m} + \boldsymbol{x}))$$

with F compactly supported, non-negative and piecewise constant. We have for $a_1, \ldots, a_l \to 0$, $a_{l+1}, \ldots, a_d \to \infty$, $\boldsymbol{x} \in [-\frac{1}{2}, \frac{1}{2}]^d$, the asymptotic relation (recall the discussion in the proof of Proposition 3.13)

$$\mathcal{V}_\sigma(\{\mathsf{M}, 0\}\{1, \boldsymbol{x}\}) \sim \frac{\gamma_l(\mathsf{k}, a_{l+1}x_{l+1}, \ldots, a_d x_d)}{a_1 \cdots a_l}, \tag{4.18}$$

with

$$\gamma_l(\mathsf{k}, t_{l+1}, \ldots, t_d) = \int F(\mathsf{k} \begin{pmatrix} x_1 \\ \vdots \\ x_l \\ t_{l+1} \\ \vdots \\ t_d \end{pmatrix}) dx. \tag{4.19}$$

The *upper bound* follows now if we put

$$\gamma_{\max,l}(t_{l+1}, \ldots, t_d) = \max_{\mathsf{k} \in \mathrm{SO}(d)} \gamma_l(\mathsf{k}, t_{l+1}, \ldots, t_d).$$

The *lower bound* is obtained, for K constant and small enough, since $a_1, \ldots, a_{d-1} \to 0$ implies $a_d \to \infty$. $\square$

Proof of Theorem 4.3. By virtue of Proposition 4.11 we can give *upper bounds* on $\sum_{K\geq X} P_K(\sigma)$ (apply the same argument as in the proof of Theorem 3.11) by considering the integrals ($l = 1, \ldots, d-1$)

$$I(X) = \int_{\substack{0<a_j \leq \frac{2}{\sqrt{3}} a_{j+1} (j=1,\ldots,d-1) \\ \frac{\gamma_{\max,l}(a_{l+1}x_{l+1},\ldots,a_d x_d)}{a_1\cdots a_l} > X \\ \mathsf{k}\in \mathrm{SO}(d) \\ \boldsymbol{x}\in[-\frac12,\frac12]^d}} \rho(a)\, d\mathsf{a}(a)\, d\mathsf{k}\, d\boldsymbol{x}, \tag{4.20}$$

where $d\boldsymbol{x}$ is the standard Lebesgue measure. We relax the condition $\boldsymbol{x} \in [-\frac12, \frac12]^d$ immediately to $\boldsymbol{x} \in \mathbb{R}^d$.

Integration over the variables $\mathsf{k}, x_1, \ldots, x_l$ is trivial, and substituting $x_j = t_j a_j^{-1}$ for $j = l+1, \ldots, d$ yields

$$I(X) \ll \int_{\substack{0<a_j \leq \frac{2}{\sqrt{3}} a_{j+1} (j=1,\ldots,d-1) \\ \frac{\gamma_{\max,l}(t_{l+1},\ldots,t_d)}{a_1\cdots a_l} > X \\ (t_{l+1},\ldots,t_d)\in\mathbb{R}^{d-l}}} \frac{\rho(a)}{a_{l+1}\cdots a_d}\, d\mathsf{a}(a) dt_{l+1}\cdots dt_d. \tag{4.21}$$

Since $\gamma_{\max,l}$ is compactly supported, we have for some constant $c > 0$

$$I(X) \ll \int_{\substack{0<a_j \leq \frac{2}{\sqrt{3}} a_{j+1} (j=1,\ldots,d-1) \\ \frac{c}{a_1\cdots a_l} > X}} \frac{\rho(a)}{a_{l+1}\cdots a_d}\, d\mathsf{a}(a)$$

$$= \int_{\substack{0<a_j \leq \frac{2}{\sqrt{3}} a_{j+1} (j=1,\ldots,d-1) \\ \frac{c}{a_1\cdots a_l} > X}} a_1\cdots a_l\, \rho(a)\, d\mathsf{a}(a).$$

Changing variables as before (proof of Theorem 3.11) yields the integral

$$\int_{\substack{0<x_j<o_j (j=1,\ldots,d-1) \\ (\prod_{\nu=1}^{d-1} x_\nu)^{\frac1d} < \frac{c}{X}}} \prod_{j=1}^{l} x_j^{\frac{d-j}{d-l}+\frac1d} \frac{dx_j}{x_j} \prod_{j=l+1}^{d-1} x_j^{\frac{j}{l}+\frac1d} \frac{dx_j}{x_j}, \tag{4.22}$$

with suitable constants o_j, and the expected upper bound follows from Lemma 3.14 with $r_M = \frac1d$.

By proposition 4.11, the *lower bound* to be proved is obtained from the integral

$$J(X) = \int_{\substack{0<a_j \leq \frac{2}{\sqrt{3}} a_{j+1} (j=1,\ldots,d-1) \\ a_{d-1} < K < \frac{2}{\sqrt{3}} a_d \\ \frac{\gamma_{d-1}(\mathsf{k}, a_d x_d)}{a_1\cdots a_{d-1}} > X \\ \mathsf{k}\in \mathrm{SO}(d) \\ \boldsymbol{x}\in[-\frac12,\frac12]^d}} \rho(a)\, d\mathsf{a}(a)\, d\mathsf{k}\, d\boldsymbol{x}. \tag{4.23}$$

The parameter K is here related to the asymptotic regime $\mathcal{S}_{K,l}$ of the Siegel domain, and is of course not related to the K used in $P_K(\sigma)$.

As above we integrate over $x_1, \ldots, x_{d-1}$ and substitute $x_d = t a_d^{-1}$. This gives

$$J(X) = \int_{\substack{0<a_j\leq \frac{2}{\sqrt{3}}a_{j+1}\,(j=1,\ldots,d-1) \\ a_{d-1}<K<\frac{2}{\sqrt{3}}a_d \\ \frac{\gamma_{d-1}(\mathsf{k},t)}{a_1\cdots a_{d-1}}>X \\ \mathsf{k}\in\mathrm{SO}(d) \\ |t|<\frac{a_d}{2}}} \frac{\rho(a)}{a_d}\, d\mathsf{a}(a)\, d\mathsf{k}\, dt. \tag{4.24}$$

We have $K < \frac{2}{\sqrt{3}} a_d$ and so

$$J(X) \geq \int_{\substack{0<a_j\leq \frac{2}{\sqrt{3}}a_{j+1}\,(j=1,\ldots,d-1) \\ a_{d-1}<K \\ \frac{\gamma_{d-1}(\mathsf{k},t)}{a_1\cdots a_{d-1}}>X \\ \mathsf{k}\in\mathrm{SO}(d) \\ |t|<\frac{\sqrt{3}}{4}K}} \frac{\rho(a)}{a_d}\, d\mathsf{a}(a)\, d\mathsf{k}\, dt. \tag{4.25}$$

After the usual change of variables we find that

$$J(X) \gg \int \prod_{j=1}^{d-1} x_j^{d-j+\frac{1}{d}} \frac{dx_j}{x_j}\, d\mathsf{k}\, dt \tag{4.26}$$

with range of integration

$$0 < x_j < (\frac{2}{\sqrt{3}})^j \quad (j = 1, \ldots, d-1), \quad (\prod_{\nu=1}^{d-1} x_\nu)^{\frac{1}{d}} < \frac{\gamma_{d-1}(\mathsf{k},t)}{X},$$
$$x_{d-1}^{\frac{1}{d-1}} < K(\prod_{\nu=1}^{d-1} x_\nu)^{\frac{1}{d}}, \quad \mathsf{k} \in \mathrm{SO}(d), \quad |t| < \frac{\sqrt{3}}{4}K.$$

The asymptotics of this last integral can be worked out with the same method used in the end of the proof of Theorem 3.11. The only crucial difference is that in the application of Lemma 3.14 we have $r_M = \frac{1}{d}$ instead of 0. $\square$

Proof of Theorems 4.4 – 4.5. Almost identical with the proofs for the two-point correlations, since $\mathcal{V}_\sigma(\{\mathsf{M}, \mathbf{0}\})$ can be identified with a function on Σ^d. $\square$

Sketch of the proof of Theorems 4.6 – 4.9. We can write $W_N(\xi)$ as

$$W_N(\xi) = \frac{1}{2N} \sum_{\nu\in\mathbb{Z}} \sum_{j=1}^{N} \chi(\frac{\nu}{N}) \exp\big(2\pi \mathrm{i}\, \nu(\lambda_j + \xi)\big) - \frac{1}{2}, \tag{4.27}$$

where χ is the characteristic function of the interval $[-1,1]$. Let us now replace χ with a smooth, compactly supported function f and study

$$\tilde{W}_N(\xi) = \frac{1}{2N} \sum_{\nu\in\mathbb{Z}} \sum_{j=1}^{N} f(\frac{\nu}{N}) \exp\big(2\pi\mathrm{i}\,\nu(\lambda_j+\xi)\big) - \frac{1}{2} f(0) \tag{4.28}$$

instead. Using the Poisson summation formula for the sum over ν we obtain

$$\tilde{W}_N(\xi) = \frac{1}{2} \sum_{\nu\in\mathbb{Z}} \sum_{j=1}^{N} \hat{f}\big(N(\lambda_j+\xi+\nu)\big) - \frac{1}{2} f(0), \tag{4.29}$$

$\hat{f}$ being the Fourier transform of f. Reviewing the proofs of Theorems 4.2 – 4.5 one can conclude that Theorems 4.6 – 4.9 indeed hold for $\tilde{W}_N(\xi)$. Finally the density argument in [21] (Section 7) shows that the smoothness condition can be dropped. The main condition for this argument to work is that the expectation and variance satisfy (note that $\int h(\boldsymbol{\alpha})\, d^{d-1}\boldsymbol{\alpha} = 1$ by definition)

$$\lim_{N\to\infty} \mathbb{E}\tilde{W}_N = \lim_{N\to\infty} \int_0^1 \int_{\mathbb{T}^{d-1}} h(\boldsymbol{\alpha})\, \tilde{W}_N(\xi)\, d\xi\, d^{d-1}\boldsymbol{\alpha} = 0, \tag{4.30}$$

$$\lim_{N\to\infty} \mathbb{E}\tilde{W}_N^2 = \lim_{N\to\infty} \int_0^1 \int_{\mathbb{T}^{d-1}} h(\boldsymbol{\alpha})\, |\tilde{W}_N(\xi)|^2\, d\xi\, d^{d-1}\boldsymbol{\alpha} = \frac{1}{4} \int f(t)^2\, dt, \tag{4.31}$$

independent of the smoothness of f, which can be readily verified. In the case $\xi = 0$, the expectation value still vanishes,

$$\lim_{N\to\infty} \mathbb{E}\tilde{W}_N = \frac{1}{2} \lim_{N\to\infty} \int_{\mathbb{T}^{d-1}} h(\boldsymbol{\alpha})\, \tilde{W}_N(0)\, d^{d-1}\boldsymbol{\alpha} = 0. \tag{4.32}$$

but it is more complicated to obtain analogous relations for the variance. By a standard density argument, we may assume without loss of generality h has a finite Fourier expansion and consider each term of the expansion separately. For N large, the zeroth Fourier coefficient corresponds to

$$\begin{aligned} &\int_0^1 \int_{\mathbb{T}^{d-1}} |\tilde{W}_N(\xi)|^2\, d\xi\, d^{d-1}\boldsymbol{\alpha} \\ &\quad \sim \frac{1}{(2N)^2} \sum_{\substack{0<m_1\leq d_1 N^{1/(d-1)} \\ \vdots \\ 0<m_{d-1}\leq d_{d-1} N^{1/(d-1)}}} \sum_{\substack{0<n_1\leq d_1 N^{1/(d-1)} \\ \vdots \\ 0<n_{d-1}\leq d_{d-1} N^{1/(d-1)}}} \sum_{\substack{\nu,\mu\in\mathbb{Z} \\ \nu m_j = \mu n_j}} f(\frac{\nu}{N})\, f(\frac{\mu}{N}) \\ &\qquad - \frac{1}{4} f(0)^2. \end{aligned} \tag{4.33}$$

The latter equals

$$\frac{1}{(2N)^2} \sum_{\substack{0<m_1\leq d_1 N^{1/(d-1)} \\ \vdots \\ 0<m_{d-1}\leq d_{d-1} N^{1/(d-1)}}} \sum_{\substack{0<n_1\leq d_1 N^{1/(d-1)} \\ \vdots \\ 0<n_{d-1}\leq d_{d-1} N^{1/(d-1)}}} \sum_{\substack{\nu,\mu\neq 0 \\ \nu m_j=\mu n_j}} f(\frac{\nu}{N})\, f(\frac{\mu}{N}). \tag{4.34}$$

Generalizing the derivation given in the Appendix to general piecewise continuous and compactly supported cut-off functions f, we find for $d>2$

$$\lim_{N\to\infty} \int_0^1 \int_{\mathbb{T}^{d-1}} |\tilde{W}_N(0)|^2 \, d\xi \, d^{d-1}\boldsymbol{\alpha}$$
$$= \left\{ \frac{1}{2} \sum_{\substack{\nu,\mu=1 \\ \gcd(\nu,\mu)=1 \\ \mu\leq\nu}}^{\infty} \frac{1}{\nu^d} \int f(t) f(\frac{\mu}{\nu} t) \, dt \right\} - \frac{1}{4} \int f(t)^2 \, dt, \tag{4.35}$$

while in the case $d=2$ we observe a logarithmic divergence which is why our argument fails in this case. For $f\geq 0$ and $f(t)\geq f(t')$ for $|t|<|t'|$ we have

$$\int f(t) f(\frac{\mu}{\nu} t) \, dt \leq \int f(t)^2 \, dt$$

Using the relation (see Appendix)

$$\sum_{\substack{\nu,\mu=1 \\ \gcd(\nu,\mu)=1 \\ \mu\leq\nu}}^{\infty} \frac{1}{\nu^d} = \frac{\zeta(d-1)}{\zeta(d)}$$

we obtain the upper bound (cf. Appendix)

$$\lim_{N\to\infty} \int_0^1 \int_{\mathbb{T}^{d-1}} |\tilde{W}_N(0)|^2 \, d\xi \, d^{d-1}\boldsymbol{\alpha} \leq \frac{1}{4} \left(2\frac{\zeta(d-1)}{\zeta(d)} - 1 \right) \int f(t)^2 \, dt \tag{4.36}$$

which in turn gives also an upper bound for

$$\lim_{N\to\infty} \int_0^1 \int_{\mathbb{T}^{d-1}} h(\boldsymbol{\alpha}) \, |\tilde{W}_N(0)|^2 \, d\xi \, d^{d-1}\boldsymbol{\alpha},$$

proportional to $\int f(t)^2 \, dt$, which is all we need for our density argument to work. But in fact terms of the form

$$\int_0^1 \int_{\mathbb{T}^{d-1}} \mathrm{e}^{2\pi\mathrm{i}(k_1\alpha_1+\cdots+k_{d-1}\alpha_{d-1})} \, |\tilde{W}_N(0)|^2 \, d\xi \, d^{d-1}\boldsymbol{\alpha}$$

actually vanish when at least one of the k_j is not zero (we will not prove this here), so

$$\lim_{N\to\infty} \int_0^1 \int_{\mathbb{T}^{d-1}} h(\boldsymbol{\alpha}) \, |\tilde{W}_N(0)|^2 \, d\xi \, d^{d-1}\boldsymbol{\alpha} \tag{4.37}$$

coincides with the limit (4.35). $\square$

4.4 Moments

The first two moments of Φ_d, $\Phi_{d,0}$ and their relatives $P_K(\sigma)$, $P_{K,0}(\sigma)$ follow from the calculations carried out in the previous Proof of Theorems 4.4 – 4.5. Equivalently, they are as well an immediate consequence of Theorems 3.15 and 3.16; Theorem 3.15 — for the first moment of $\Phi_{d,0}(a,b)$ and $P_{K,0}(\sigma)$, and for the second moment of $\Phi_d(a,b)$ and $P_K(\sigma)$; and Theorem 3.15 — for the second moment of $\Phi_{d,0}(a,b)$ and $P_{K,0}(\sigma)$. Note that for instance

$$\sum_{K=0}^{\infty} K^n \, P_{K,0}(\sigma) = \frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \mathcal{V}_\sigma(\{\mathsf{M}, \mathbf{0}\})^n \, d\mu. \tag{4.38}$$

Recall also that $\sum_{K=0}^{\infty} K^n \, P_K(\sigma)$ is related to the expectation value of the n-point correlations, rel. (4.5).

The results are as follows.

Theorem 4.12

$$\sum_{K=0}^{\infty} K \, P_K(\sigma) = \sigma.$$

$$\sum_{K=0}^{\infty} K^2 \, P_K(\sigma) = \sigma^2 + \sigma.$$

These moments coincide with those of a Poisson distribution. The third moment will deviate similarly as the expectation value three-point correlations, compare (4.5) and Theorem 3.19.

Theorem 4.13

$$\sum_{K=0}^{\infty} K \, P_{K,0}(\sigma) = \sigma.$$

$$\sum_{K=0}^{\infty} K^2 \, P_{K,0}(\sigma) = \sigma^2 + \left(2 \, \frac{\zeta(d-1)}{\zeta(d)} - 1\right)\sigma \qquad (d > 2).$$

In the limit $d \to \infty$ all moments converge to the moments of a Poisson distribution:

Theorem 4.14 *Fix $n \in \mathbb{N}$. Then*

$$\lim_{d\to\infty} \sum_{K=0}^{\infty} K^n \, P_K(\sigma) = \sum_{K=0}^{\infty} K^n \, \frac{\sigma^K}{K!} \, \mathrm{e}^{-\sigma}$$

and

$$\lim_{d\to\infty} \sum_{K=0}^{\infty} K^n \, P_{K,0}(\sigma) = \sum_{K=0}^{\infty} K^n \, \frac{\sigma^K}{K!} \, \mathrm{e}^{-\sigma} \, .$$

Proof. The nth moment of $P_K(\sigma)$ is related via (4.5) to the expectation value of the n-point correlations, which converge to the limit given by IUDRVs (Theorem 3.20). The analogue of rel. (4.5) of course exists also for IUDRVs; this proves the first statement.

Let us turn to the second assertion. Recall that due to (4.38) we have to calculate ($n < d$)

$$\frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \sum_{\boldsymbol{m}^1,\ldots,\boldsymbol{m}^n \in \mathbb{Z}^d} F(\mathsf{M}\boldsymbol{m}^1, \ldots, \mathsf{M}\boldsymbol{m}^n) \, d\mu \tag{4.39}$$

with

$$F(\boldsymbol{x}^1, \ldots, \boldsymbol{x}^n) = \prod_{j=1}^{n} \chi_1(x_1^j) \cdots \chi_1(x_{d-1}^j) \, \chi_\sigma(x_d^j). \tag{4.40}$$

As in the proof Theorems 3.15 and 3.16, we put $\boldsymbol{m}^j = t^j \boldsymbol{c}^j$ where $t^j = \gcd(m_1^j, \ldots, m_d^j) > 0$ for $j = 1, \ldots, n < d$. All terms with $t^j > 1$ will later lead to contributions which vanish as $d \to \infty$.

For any given $(\boldsymbol{c}^1, \ldots, \boldsymbol{c}^n)$ we find some $\mathsf{g} \in \Gamma = \mathrm{SL}(d, \mathbb{Z})$ such that

$$(\boldsymbol{c}^1, \ldots, \boldsymbol{c}^n) = \mathsf{g} \begin{pmatrix} 1 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & a_{nn} \\ 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}, \tag{4.41}$$

with certain restrictions on the a_{ij}. It is for example possible that the jth column consists only of zeros, i.e. $a_{1j} = 0, \ldots, a_{jj} = 0$. This would eventually lead to some integral over $F(\boldsymbol{x}^1, \ldots, \boldsymbol{x}^n)$ with $\boldsymbol{x}^j = \mathbf{0}$. But for our special choice of F in (4.40) we find easily that $F(\boldsymbol{x}^1, \ldots, \boldsymbol{x}^n) = 0$ whenever $\boldsymbol{x}^j = \mathbf{0}$, and thus the resulting integral vanishes.

Suppose now that $|a_{ij}| > 1$. This will lead to contributions which will vanish when $d \to \infty$, and we will not deal with that case any further as well.

We are thus left with $a_{ij} = \pm 1$. In this case it is always possible to find an element $\mathsf{g} \in \Gamma = \mathrm{SL}(d, \mathbb{Z})$ such that the a_{ij} are of the following form:

(a) In every column there is only one non-zero $a_{ij} = \pm 1$.

(b) In every row there is at least one $a_{ij} \geq 0$.

Terms with negative $a_{ij} = -1$ would, however, lead to contributions containing terms of the form

$$F(\ldots, \boldsymbol{x}, \ldots, -\boldsymbol{x}, \ldots),$$

which are zero due to our special choice (4.40) of F.

Let us therefore consider the only remaining case when all a_{ij} are zero or $+1$. The set of upper triangular $(n \times n)$-matrices with only one $+1$ in every column is in one-to-one correspondence with the set partitions of $\{1, \ldots, n\}$ into ν disjoint non-empty subsets $[\mathcal{F}_1, \ldots, \mathcal{F}_\nu]$, where $\nu = 1, \ldots, n$: To a given partition $[\mathcal{F}_1, \ldots, \mathcal{F}_\nu]$ we can associate the matrix with coefficients (denote by i_s is the smallest element of $\mathcal{F}_s$)

$$a_{ij} = \begin{cases} 1 & \text{if } i = i_s \text{ and } j \in \mathcal{F}_s, \text{ for some } s = 1, \ldots, \nu \\ 0 & \text{otherwise,} \end{cases}$$

From this and a generalization of the arguments for Theorem 3.16 we obtain (denote by $\mathcal{P}_\nu$ the collection of all partitions into ν sets)

$$\begin{aligned} \frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \sum_{\boldsymbol{m}^1, \ldots, \boldsymbol{m}^n \in \mathbb{Z}^d} & F(\mathsf{M}\boldsymbol{m}^1, \ldots, \mathsf{M}\boldsymbol{m}^n)\, d\mu \\ = \sum_{\nu=1}^{n} \sum_{[\mathcal{F}_1, \ldots, \mathcal{F}_\nu] \in \mathcal{P}_\nu} & \int_{\mathbb{R}^\nu} F(\boldsymbol{x}^1, \ldots, \boldsymbol{x}^n) \Big|_{\substack{\boldsymbol{x}^i = \boldsymbol{x}^j \text{ if } i,j \text{ are} \\ \text{in the same set } \mathcal{F}_s}} d\boldsymbol{x}^{i_1} \cdots d\boldsymbol{x}^{i_\nu} \\ & + \Big[\text{terms of lower order in } d\Big] \qquad (4.42) \end{aligned}$$

The reason why all coefficients in front of the integrals are $C = 1$ follows from a consequent application of the argument given at the very end of Section 3.7.

The integration can be carried out term by term, and our final result is

$$\frac{1}{\mu(\Sigma^d)} \int_{\Sigma^d} \sum_{\boldsymbol{m}^1, \ldots, \boldsymbol{m}^n \in \mathbb{Z}^d} F(\mathsf{M}\boldsymbol{m}^1, \ldots, \mathsf{M}\boldsymbol{m}^n)\, d\mu = \sum_{\nu=1}^{n} |\mathcal{P}_\nu|\, \sigma^\nu, \qquad (4.43)$$

where $|\mathcal{P}_\nu|$ is the number of partitions of $\{1, \ldots, n\}$ into ν sets. The right-hand-side of (4.43) represents exactly the nth moment of a Poisson distribution. $\square$

Finally, we have for the moments of $\Phi_d(X, \infty)$:

Theorem 4.15

$$\int_{-\infty}^{\infty} X d\Phi_d(X, \infty) = 0.$$

$$\int_{-\infty}^{\infty} X^2 d\Phi_d(X, \infty) = \frac{1}{4}.$$

Theorem 4.16

$$\int_{-\infty}^{\infty} X d\Phi_{d,0}(X, \infty) = 0.$$

$$\int_{-\infty}^{\infty} X^2 d\Phi_{d,0}(X, \infty) = \frac{1}{4}\left(2\,\frac{\zeta(d-1)}{\zeta(d)} - 1\right) \qquad (d > 2).$$

Again, higher moments converge for $d \to \infty$ to the moments of the limit distribution corresponding to IUDRVs. Note that this distribution is not Gaussian; some of its properties are discussed in [38].

A Appendix: The number of solutions of simultaneous quadratic equations. By Zeév Rudnick

We consider the number of integer solutions of the system of quadratic equations $ax_i = by_i$, $i = 1, \dots, f$ where the variables lie in a large box. By using elementary methods from number theory, we can show

Theorem A.1 *For $f \geq 1$ let $N_f(T)$ be the number of positive integer solutions of the system*

$$ax_i = by_i, \quad i = 1, \dots, f, \qquad a, b < T, \quad x_i, y_i < T^{1/f}. \tag{A.1}$$

Then as $T \to \infty$,

1. *For $f \geq 3$,*

$$N_f(T) = C_f T^2 + O(T), \qquad C_f = \frac{2\zeta(f)}{\zeta(f+1)} - 1,$$

2. *For $f = 2$,*

$$N_2(T) = (\frac{2\zeta(2)}{\zeta(3)} - 1)T^2 + O(T \log T),$$

3. For $f=1$,

$$N_1(T)=\frac{2}{\zeta(2)}T^2\log T+O(T^2),$$

where $\zeta(s)=\sum n^{-s}$ is the Riemann zeta function.

Remarks.

1. Recall that $\zeta(2)=\pi^2/6$, and there are similar formulas for $\zeta(2k)$.

2. The constants C_f decrease monotonically to 1 as $f\to\infty$. The first few values are $C_2=1.7368\ldots$, $C_3=1.22125\ldots$, $C_4=1.0875\ldots$, $C_5=1.0385\ldots$, etc.

3. For an application of these methods to other counting problems, see Katznelson [17, 18].

Proof. For every integer $d\le T$, let $N_f(T,d)$ be the number of solution of the system (A.1) with the extra condition $\gcd(a,b)=d$. We first estimate $N_f(T,d)$.

For such a,b write $a=Ad$, $b=Bd$ with $A,B\le T/d$ coprime. Then $ax_i=by_i$ if and only if $Ax_i=By_i$ and since A,B are *coprime*, this happens if and only if for some integer z_i,

$$x_i=Bz_i,\qquad y_i=Az_i$$

where now

$$z_i\le\min\Big(\frac{T^{1/f}}{A},\frac{T^{1/f}}{B}\Big)=T^{1/f}\min\Big(\frac{1}{A},\frac{1}{B}\Big).$$

Therefore

$$N_f(T,d)=\sum_{\substack{A,B\le T/d\\ \gcd(A,B)=1}}\left(T^{1/f}\min\Big(\frac{1}{A},\frac{1}{B}\Big)\right)^f=T\sum_{\substack{A,B\le T/d\\ \gcd(A,B)=1}}\min\Big(\frac{1}{A},\frac{1}{B}\Big)^f. \tag{A.2}$$

We rearrange the sum over A,B as

$$\sum_{\substack{A,B\le T/d\\ \gcd(A,B)=1}}\min\Big(\frac{1}{A},\frac{1}{B}\Big)^f=2\sum_{A=1}^{T/d}\sum_{\substack{B\le A\\ \gcd(A,B)=1}}\frac{1}{A^f}-1$$

by taking the cases $A\le B$ and $B\le A$ separately, with identical contributions, and the pair $A=B=1$ is counted twice so needs to be subtracted once. Thus

$$N_f(T,d) = 2T\sum_{A=1}^{T/d} \sum_{\substack{B\leq A \\ \gcd(A,B)=1}} \frac{1}{A^f} - T = 2T\sum_{A=1}^{T/d}\frac{\phi(A)}{A^f} - T,$$

where $\phi(n)$ is the number of $B \leq n$ coprime to n (Euler's function). Therefore

$$N_f(T) = \sum_{d=1}^{T} N_f(T,d) = 2T\sum_{d=1}\sum_{A=1}^{T/d}\frac{\phi(A)}{A^f} - T^2. \tag{A.3}$$

Now change order of summation to evaluate the contribution of the first term:

$$\begin{aligned}
\sum_{d=1}^{T}\sum_{A=1}^{T/d}\frac{\phi(A)}{A^f} &= \sum_{A=1}^{T}\sum_{D\leq T/A}\frac{\phi(A)}{A^f} \\
&= \sum_{A=1}^{T}\frac{\phi(A)}{A^f}\sum_{D\leq T/A} 1 \\
&= \sum_{A=1}^{T}\frac{\phi(A)}{A^f}(\frac{T}{A} + O(1)) \\
&= T\sum_{A=1}^{T}\frac{\phi(A)}{A^{f+1}} + O\left(\sum_{A=1}^{T}\frac{\phi(A)}{A^f}\right).
\end{aligned}$$

We thus have

$$N_f(T) = 2T^2\sum_{A=1}^{T}\frac{\phi(A)}{A^{f+1}} + O\left(T\sum_{A=1}^{T}\frac{\phi(A)}{A^f}\right) - T^2.$$

Note that for $f > 2$, the series $\sum \phi(A)/A^f$ converges, in fact for $s > 1$:

$$\sum_{n=1}^{\infty}\frac{\phi(n)}{n^s} = \frac{\zeta(s-1)}{\zeta(s)} \tag{A.4}$$

and

$$\sum_{A=1}^{T}\frac{\phi(A)}{A^{f+1}} = \frac{\zeta(f)}{\zeta(f+1)} + O(T^{1-f}), \qquad f > 1. \tag{A.5}$$

Therefore for $f \geq 3$ we have

$$N_f(T) = (\frac{2\zeta(f)}{\zeta(f+1)} - 1)T^2 + O(T). \tag{A.6}$$

For the remaining cases, we use (see e.g. Hardy and Wright [14] for a similar formula for $\sum_{n\leq x}\phi(n)$)

$$\sum_{n\leq T}\frac{\phi(n)}{n^2}=\frac{\log T}{\zeta(2)}+O(1).$$

Thus for $f=2$

$$\begin{aligned}N_2(T)&=2T^2\sum_{A=1}^{T}\frac{\phi(A)}{A^3}+O\left(T\sum_{A=1}^{T}\frac{\phi(A)}{A^2}\right)-T^2\\&=(\frac{2\zeta(2)}{\zeta(3)}-1)T^2+O(T\log T)\end{aligned}\tag{A.7}$$

and for $f=1$ use $\phi(n)\leq n$ and so $\sum_{n\leq T}\frac{\phi(n)}{n}\ll T$. Thus

$$\begin{aligned}N_1(T)&=2T^2\sum_{A=1}^{T}\frac{\phi(A)}{A^2}+O\left(T\sum_{A=1}^{T}\frac{\phi(A)}{A}\right)-T^2\\&=\frac{2}{\zeta(2)}T^2\log T+O(T^2).\end{aligned}\tag{A.8}$$

This concludes the proof. □

References

[1] M.V. Berry and M. Tabor, Level clustering in the regular spectrum, *Proc. Roy. Soc.* A **356** (1977) 375-394.

[2] P.M. Bleher, The energy level spacing for two harmonic oscillators with golden mean ratio of frequencies, *J. Statist. Phys.* **61** (1990) 869-876.

[3] P.M. Bleher, The energy level spacing for two harmonic oscillators with generic ratio of frequencies, *J. Statist. Phys.* **63** (1991) 261-283.

[4] P.M. Bleher, Trace formula for quantum integrable systems, lattice point problem, and small divisors, in: D. Hejhal et al. (eds.), *Emerging Applications of Number Theory*, IMA Vol. Math. Appl. **109** (Springer, New York, 1999) 1-38.

[5] J.W.S. Cassels, *An Introduction to Diophantine Approximation* (Cambridge University Press, 1957).

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

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

[8] A. Eskin, Counting problems and semisimple groups, Proc. ICM, Vol. II, *Doc. Math.* (1998) 539-552.

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

[10] A.L. Figotin and L.A. Pastur, An exactly solvable model of a multidimensional incommensurate structure, *Comm. Math. Phys.* **95** (1984) 401-425.

[11] J.F. Geelen and R.J. Simpson, A two dimensional Steinhaus theorem, *Australas. J. Comb.* **8** (1993) 169-197.

[12] C. Greenman, The generic spacing distribution of the two-dimensional harmonic oscillator, *J. Phys.* A **29** (1996) 4065-4081.

[13] C. Greenman, Is the level spacing distribution of the infinite-dimensional harmonic oscillator that of a Poisson process? *J. Phys. A* **30** (1997) 927-936.

[14] G.H. Hardy and E.M. Wright, *An Introduction to the Theory of Numbers*, 5th ed. (Oxford Clarendon, 1979).

[15] D. Hejhal, On value distribution properties of automorphic functions along closed horocycles, *XVIth Rolf Nevanlinna Colloquium (Joensuu, 1995)*, 39-52 (de Gruyter, 1996).

[16] N.M. Katz and P. Sarnak, *Random Matrices, Frobenius Eigenvalues, and Monodromy*, AMS Colloq. Publ. 45 (AMS, Providence, RI, 1999).

[17] Y.R. Katznelson, Singular matrices and a uniform bound for congruence groups of $\mathrm{SL}_n(\mathbb{Z})$, *Duke Math. J.* **69** (1993) 121-136.

[18] Y.R. Katznelson, Integral matrices of fixed rank, *Proc. Amer. Math. Soc.* **120** (1994) 667-675.

[19] D.Y. Kleinbock and G.A. Margulis, Bounded orbits of nonquasiunipotent flows on homogeneous spaces, in: L.A. Bunimovich et al. (eds.), *Sinai's Moscow Seminar on Dynamical Systems*, 141-172, AMS Transl. (2) **171**, Amer. Math. Soc., Providence, RI, 1996.

[20] P. Kurlberg and Z. Rudnick, The distribution of spacings between quadratic residues, *Duke Math. J.*, to appear.

[21] J. Marklof, Limit theorems for theta sums, *Duke Math. J.* **97** (1999) 127-153.

[22] J. Marklof, Theta sums, Eisenstein series, and the semiclassical dynamics of a precessing spin, in: D. Hejhal et al. (eds.), *Emerging Applications of Number Theory*, IMA Vol. Math. Appl. **109** (Springer, New York, 1999) 405-450.

[23] J. Marklof, Spectral form factors of rectangle billiards, *Comm. Math. Phys.* **199** (1998) 169-202.

[24] A.E. Mazel and Ya.G. Sinai, A limiting distribution connected with fractional parts of linear forms, in: S. Albeverio et al. (ed.), *Ideas and methods in mathematical analysis, stochastics and applications*, Vol. 1 (1992) 220-229.

[25] M.L. Mehta, *Random Matrices*, 2nd ed. (Academic Press,1991).

[26] C.C. Moore, Ergodicity of flows on homogeneous spaces, *Am. J. Math.* **88** (1966) 154-178.

[27] A. Pandey, O. Bohigas and M.-J. Giannoni, Level repulsion in the spectrum of two-dimensional harmonic oscillators, *J. Phys.* A **22** (1989) 4083-4088.

[28] A. Pandey and R. Ramaswamy, Level spacings for harmonic-oscillator systems, *Phys. Rev.* A **43** (1991) 4237-4243.

[29] M.S. Raghunathan, *Discrete Subgroups of Lie Groups* (Springer, 1972).

[30] Z. Rudnick and P. Sarnak, Zeros of principal L-functions and random matrix theory, *Duke Math. J.* **81** (1996) 269-322.

[31] Z. Rudnick and P. Sarnak, The pair correlation function for fractional parts of polynomials, *Comm. Math. Phys.* **194** (1998) 61-70.

[32] P. Sarnak, Asymptotic behavior of periodic orbits of the horocycle flow and Eisenstein series, *Comm. Pure Appl. Math.* **34** (1981) 719-739.

[33] P. Sarnak, Values at integers of binary quadratic forms, *Harmonic analysis and number theory* (Montreal, 1996), 181-203, CMS Conf. Proc. **21**, Amer. Math. Soc., Providence, RI, 1997.

[34] P. Sarnak, Quantum chaos, symmetry and zeta functions. Lecture I: Quantum chaos, *Curr. Dev. Math.* (1997) 84-101.

[35] P. Sarnak, Quantum chaos, symmetry and zeta functions. Lecture II: Zeta functions, *Curr. Dev. Math.* (1997) 102-115.

[36] W.M. Schmidt, *Diophantine Approximation*, Lect. Notes in Math. **785** (Springer, 1980).

[37] C.L. Siegel, *Lectures on the Geometry of Numbers* (Springer, 1989).

[38] Ya.G. Sinai, On limit theorems for Weyl multiple trigonometric sums, *Proc. Steklov Inst. Math.* **191** (1989) 131-143.

[39] N.B. Slater, Gaps and steps for the sequence $n\theta \mod 1$, *Proc. Cambridge Philos. Soc.* **63** (1967) 1115-1123.

[40] S. Swierczkowski, On successive settings of an arc on the circumference of a circle, *Fund. Math.* **46** (1958) 187-189.

[41] J.M. VanderKam, Values at integers of homogeneous polynomials, *Duke Math. J.* **97** (1999) 379-412.

[42] J.M. VanderKam, Pair correlation of four-dimensional flat tori, *Duke Math. J.* **97** (1999) 413-438.