Annals of Mathematics, **172** (2010), 1949–2033

The distribution of free path lengths in the periodic Lorentz gas and related lattice point problems

By JENS MARKLOF and ANDREAS STRÖMBERGSSON

Abstract

The periodic Lorentz gas describes the dynamics of a point particle in a periodic array of spherical scatterers, and is one of the fundamental models for chaotic diffusion. In the present paper we investigate the Boltzmann-Grad limit, where the radius of each scatterer tends to zero, and prove the existence of a limiting distribution for the free path length. We also discuss related problems, such as the statistical distribution of directions of lattice points that are visible from a fixed position.

Contents

J.M. has been supported by an EPSRC Advanced Research Fellowship and a Philip Leverhulme Prize. A.S. is a Royal Swedish Academy of Sciences Research Fellow supported by a grant from the Knut and Alice Wallenberg Foundation.

1. **Introduction**

1.1. *The periodic Lorentz gas.* The Lorentz gas, which was originally introduced by Lorentz [20] in 1905 to model the motion of electrons in a metal, describes an ensemble of noninteracting point particles in an infinite array of spherical scatterers. Lorentz was in particular interested in the stochastic properties of the dynamics that emerge in the Boltzmann-Grad limit, where the radius ρ of each scatterer tends to zero.

In the present and subsequent papers [23], [26], [25] we investigate the periodic set-up, where the scatterers are placed at the vertices of a euclidean lattice $\mathscr{L} \subset \mathbb{R}^d$ (Figure 1). We will identify a new random process that governs the macroscopic dynamics of a particle cloud in the Boltzmann-Grad limit. In the case of a Poisson-distributed (rather than periodic) configuration of scatterers, the limiting process is described by the linear Boltzmann equation, see Gallavotti [16],

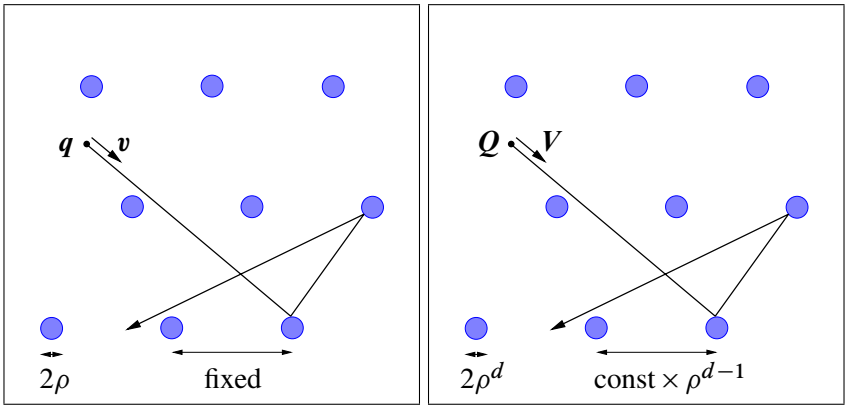

Figure 1. Left: The periodic Lorentz gas in "microscopic" coordinates—the lattice $\mathscr{L}$ remains fixed as the radius ρ of the scatterer tends to zero. Right: The periodic Lorentz gas in "macroscopic" coordinates —both the lattice constant and the radius of each scatter tend to zero, in such a way that the mean free path length remains finite.

Spohn [35], and Boldrighini, Bunimovich and Sinai [8]. It already follows from the estimates in [9], [19] that the linear Boltzmann equation does not hold in the periodic set-up; this was pointed out recently by Golse [18].

The first step towards the proof of the existence of a limiting process for the periodic Lorentz gas is the understanding of the distribution of the free path length in the limit $\rho \to 0$, which is the key result of the present paper. The distribution of the free path lengths in the periodic Lorentz gas was already investigated by Pólya, who rephrased the problem in terms of the visibility in a (periodic) forest [29]. We complete the analysis of the limiting process in [23], [26] and [25], where we establish a Markov property, and provide explicit formulas and asymptotic estimates for the limiting distributions.

Our results complement classical studies in ergodic theory, where one is interested in the stochastic properties in the limit of long times, with the radius of each scatterer being *fixed*. Here Bunimovich and Sinai [10] proved, in the case of a finite horizon and in dimension $d = 2$, that the dynamics is diffusive in the limit of large times, and satisfies a central limit theorem. "Finite horizon" means that the scatterers are sufficiently large so that the path length between consecutive collisions is bounded; this hypothesis was recently removed by Szász and Varju [37] after initial work by Bleher [2]. For related recent studies of statistical properties of the two-dimensional periodic Lorentz gas, see also [13], [27], [28]. There is at present no proof of the central limit theorem for higher dimensions, even in the case of finite horizon [12], [1].

Since the point particles of the Lorentz gas are noninteracting, we can reduce the problem to the study of the billiard flow

$$\varphi_t : \mathrm{T}^1(\mathcal{H}_\rho) \to \mathrm{T}^1(\mathcal{H}_\rho), \qquad (\boldsymbol{q}_0, \boldsymbol{v}_0) \mapsto (\boldsymbol{q}(t), \boldsymbol{v}(t)), \tag{1.1}$$

where $\mathcal{H}_\rho \subset \mathbb{R}^d$ is the complement of the set $\mathcal{B}_\rho^d + \mathcal{L}$ (the "billiard domain"), and $\mathrm{T}^1(\mathcal{H}_\rho) = \mathcal{H}_\rho \times \mathrm{S}_1^{d-1}$ is its unit tangent bundle (the "phase space"). $\mathcal{B}_\rho^d$ denotes the open ball of radius ρ, centered at the origin. A point in $\mathrm{T}^1(\mathcal{H}_\rho)$ is parametrized by $(\boldsymbol{q}, \boldsymbol{v})$, with $\boldsymbol{q} \in \mathcal{H}_\rho$ denoting the position and $\boldsymbol{v} \in \mathrm{S}_1^{d-1}$ the velocity of the particle. The Liouville measure of φ_t is

$$d\nu(\boldsymbol{q}, \boldsymbol{v}) = d\,\mathrm{vol}_{\mathbb{R}^d}(\boldsymbol{q})\, d\,\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\boldsymbol{v}) \tag{1.2}$$

where $\mathrm{vol}_{\mathbb{R}^d}$ and $\mathrm{vol}_{\mathrm{S}_1^{d-1}}$ refer to the Lebesgue measures on $\mathbb{R}^d$ (restricted to $\mathcal{H}_\rho$) and S_1^{d-1}, respectively.

The free path length for the initial condition $(\boldsymbol{q}, \boldsymbol{v}) \in \mathrm{T}^1(\mathcal{H}_\rho)$ is defined as

$$\tau_1(\boldsymbol{q}, \boldsymbol{v}; \rho) = \inf\{t > 0 : \boldsymbol{q} + t\boldsymbol{v} \notin \mathcal{H}_\rho\}. \tag{1.3}$$

That is, $\tau_1(\boldsymbol{q}, \boldsymbol{v}; \rho)$ is the first time at which a particle with initial data $(\boldsymbol{q}, \boldsymbol{v})$ hits a scatterer.

From now on we will assume, without loss of generality, that $\mathcal{L}$ has covolume one.

THEOREM 1.1. *Fix a lattice $\mathcal{L}$ of covolume one, let $\boldsymbol{q} \in \mathbb{R}^d \setminus \mathcal{L}$, and let λ be a Borel probability measure on S_1^{d-1} absolutely continuous with respect to Lebesgue measure.*[1] *Then there exists a continuous probability density $\Phi_{\mathcal{L},\boldsymbol{q}}$ on $\mathbb{R}_{>0}$ such that, for every $\xi \geq 0$,*

$$\lim_{\rho \to 0} \lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} : \rho^{d-1}\tau_1(\boldsymbol{q}, \boldsymbol{v}; \rho) \geq \xi\}) = \int_\xi^\infty \Phi_{\mathcal{L},\boldsymbol{q}}(\xi')d\xi'. \tag{1.4}$$

The limiting density is in fact "universal" for generic $\boldsymbol{q}$; i.e.,

$$\Phi(\xi) := \Phi_{\mathcal{L},\boldsymbol{q}}(\xi) \tag{1.5}$$

is independent of $\mathcal{L}$ and $\boldsymbol{q}$, for Lebesgue-almost every $\boldsymbol{q}$. Theorem 1.1 is proved in Section 4, it is closely related to the lattice point problem studied in Section 3. The asymptotic tails of the limiting distribution $\Phi_{\mathcal{L},\boldsymbol{q}}(\xi)$ are calculated in [25]. In Section 4 we generalize Theorem 1.1 in several ways. We consider for instance the distribution of free paths that hit a given point on the scatterer, which will be crucial in the characterization of the limiting random process in [23].

Theorem 1.1 shows that the free path length scales like $\rho^{-(d-1)}$. This suggests to re-define position and time and use the "macroscopic" coordinates

$$(\boldsymbol{Q}(t), \boldsymbol{V}(t)) = (\rho^{d-1}\boldsymbol{q}(\rho^{-(d-1)}t), \boldsymbol{v}(\rho^{-(d-1)}t)). \tag{1.6}$$

[1] The condition $\boldsymbol{q} \in \mathbb{R}^d \setminus \mathcal{L}$ ensures that τ_1 is defined for ρ sufficiently small. In Section 4 we also consider variants of Theorem 1.1 where the initial position is near $\mathcal{L}$, e.g., $\boldsymbol{q} \in \partial\mathcal{H}_\rho$.

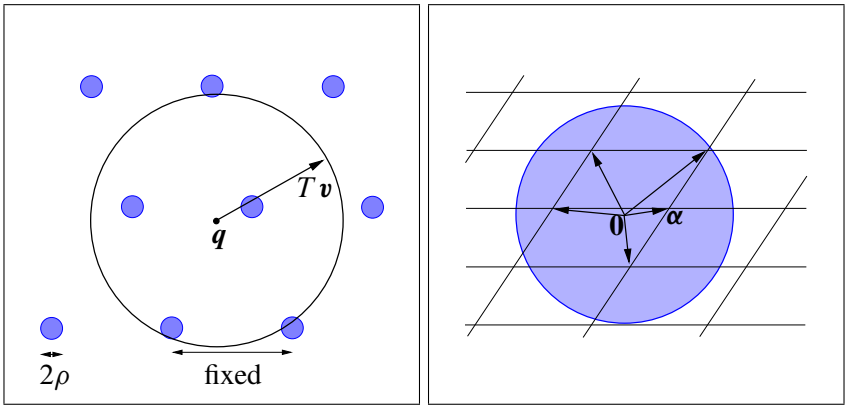

Figure 2. Left: How many lattice balls of radius ρ does a random ray of length $T = \text{const} \times \rho^{-(d-1)}$ intersect? Right: What are the statistical properties of the directions of the affine lattice points $\mathscr{L} + \boldsymbol{\alpha}$ inside a large ball?

We now state a macroscopic version of Theorem 1.1, which is a corollary of the proof of Theorem 1.1 (see §9.2). Here

$$\mathcal{T}_1(\boldsymbol{Q}, \boldsymbol{V}; \rho) = \rho^{d-1} \tau_1(\rho^{-(d-1)} \boldsymbol{Q}, \boldsymbol{V}; \rho) \tag{1.7}$$

is the corresponding macroscopic free path length.

THEOREM 1.2. *Fix a lattice $\mathscr{L}$ of covolume one and let Λ be a Borel probability measure on $\mathrm{T}^1(\mathbb{R}^d)$ absolutely continuous with respect to Lebesgue measure. Then, for every $\xi \geq 0$,*

$$\lim_{\rho \to 0} \Lambda(\{(\boldsymbol{Q}, \boldsymbol{V}) \in \mathrm{T}^1(\rho^{d-1} \mathscr{H}_\rho) \,:\, \mathcal{T}_1(\boldsymbol{Q}, \boldsymbol{V}; \rho) \geq \xi\}) = \int_\xi^\infty \Phi(\xi')\, d\xi' \tag{1.8}$$

with $\Phi(\xi)$ as in (1.5).

Variants of Theorem 1.2 were recently established by Boca and Zaharescu [7] in dimension $d = 2$, using methods from analytic number theory; cf. also their earlier work with Gologan [4], and the paper by Caglioti and Golse [11]. Our approach uses dynamics and equidistribution of flows on homogeneous spaces (the details are developed in §5), and works in arbitrary dimension. Previous work in higher dimension $d > 2$ includes the papers by Bourgain, Golse and Wennberg [9], [19] who provide tail estimates of possible limiting distributions of converging subsequences. More details on the existing literature can be found in the survey [17].

1.2. *Related lattice point problems.* The key to the understanding of the Boltzmann-Grad limit of the periodic Lorentz gas are lattice point problems for thinly stretched domains, which are randomly rotated or sheared. In Sections 2 and 3

we discuss two problems of independent interest that fall into this category: the distribution of spheres that intersect a randomly directed ray, and the statistical properties of the directions of lattice points (Figure 2). Section 6 discusses the general class of problems of this type.

Let us for example consider the affine lattice $\mathbb{Z}^2+\boldsymbol{\alpha}$, with the observer located at the origin. The directions of all lattice points with distance $< T$ are represented by points on the unit circle,

$$\frac{\boldsymbol{m}+\boldsymbol{\alpha}}{\|\boldsymbol{m}+\boldsymbol{\alpha}\|}\in \mathrm{S}^1_1, \qquad \text{for } \boldsymbol{m}\in\mathbb{Z}^2\setminus\{-\boldsymbol{\alpha}\}, \quad \|\boldsymbol{m}+\boldsymbol{\alpha}\|<T. \tag{1.9}$$

We identify the circle with the unit interval via the map $(x,y)\mapsto(2\pi)^{-1}\arg(x+\mathrm{i}y)$, and therefore the distribution of directions is reformulated as a problem of distribution mod 1 of the numbers

$$\tfrac{1}{2\pi}\arg(m+\alpha_1+\mathrm{i}(n+\alpha_2)), \quad \text{for } (m,n)\in\mathbb{Z}^2\setminus\{-\boldsymbol{\alpha}\},\ \ (m+\alpha_1)^2+(n+\alpha_2)^2<T^2. \tag{1.10}$$

We label these $N=N(T)$ numbers in order by

$$-\tfrac{1}{2}<\xi_{N,1}\le\xi_{N,2}\le\cdots\le\xi_{N,N}\le\tfrac{1}{2} \tag{1.11}$$

and define in addition $\xi_{N,0}=\xi_{N,N}-1$. It is not hard to see that this sequence (or rather: this sequence of sequences) is uniformly distributed mod 1; i.e., for every $-\frac{1}{2}\le a<b\le\frac{1}{2}$,

$$\lim_{N\to\infty}\frac{\#\{1\le j\le N:\xi_{N,j}\in[a,b)\}}{N}=b-a. \tag{1.12}$$

This (classical) equidistribution statement follows from the fact that the asymptotic number of lattice points in a fixed sector of a large disc is proportional to the area of the sector.

A popular way to characterize the "randomness" of a uniformly distributed sequence is the statistics of gaps. The following theorem, which is a corollary of more general results in Section 2, shows that there is a limiting gap distribution when $N\to\infty$.

THEOREM 1.3. *For every $\boldsymbol{\alpha}\in\mathbb{R}^2$ there exists a distribution function $P_{\boldsymbol{\alpha}}(s)$ on $\mathbb{R}_{\ge 0}$ (continuous except possibly at $s=0$) such that for every $s\ge 0$,*

$$\lim_{N\to\infty}\frac{\#\{1\le j\le N:N(\xi_{N,j}-\xi_{N,j-1})\ge s\}}{N}=P_{\boldsymbol{\alpha}}(s). \tag{1.13}$$

We will provide explicit formulas for $P_{\boldsymbol{\alpha}}(s)$, which clearly deviate from the statistics of independent random variables from a Poisson process, where $P(s)=\exp(-s)$. It is remarkable that, for $\boldsymbol{\alpha}\notin\mathbb{Q}^2$, the limiting distribution $P_{\boldsymbol{\alpha}}(s)$ is independent of $\boldsymbol{\alpha}$ and coincides with the gap distribution for the fractional parts of $\sqrt{n}$ calculated by Elkies and McMullen [14]; cf. Figure 3. There is a deep reason for this apparent coincidence, which we will return to in the next section.

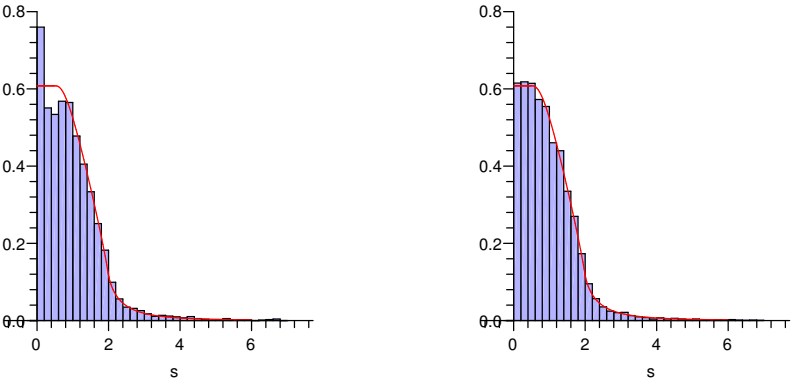

Figure 3. Left: The distribution of gaps in the sequence $\sqrt{n}$ mod 1, $n = 1, \ldots, 7765$, vs. the Elkies-McMullen distribution. Right: Gap distribution for the directions of the vectors $(m - \sqrt{2}, n) \in \mathbb{R}^2$ with $m \in \mathbb{Z}$, $n \in \mathbb{Z}_{\geq 0}$, $(m - \sqrt{2})^2 + n^2 < 4900$. The continuous curve is the Elkies-McMullen distribution.

The statistics are different for $\boldsymbol{\alpha} \in \mathbb{Q}^2$. In particular, $P_{\boldsymbol{\alpha}}(s)$ has a jump discontinuity at $s = 0$ for every $\boldsymbol{\alpha} \in \mathbb{Q}^2$, which exactly accounts for the multiplicities in the sequence (1.11); removing all repetitions from that sequence results in a limiting gap distribution which is continuous on all $\mathbb{R}_{\geq 0}$; see Corollary 2.7 below. In the particular case $\boldsymbol{\alpha} = \mathbf{0}$, this recovers a result of Boca, Cobeli and Zaharescu [3], which is closely related to the statistical distribution of Farey fractions (see also Boca and Zaharescu [5]).

The only previously known result for nonzero values of $\boldsymbol{\alpha}$ is by Boca and Zaharescu [6], who calculated the limit of the pair correlation function *on average* over $\boldsymbol{\alpha}$. (The pair correlation function is essentially the variance of the probability $E_{0,\boldsymbol{\alpha}}(r, \sigma)$ studied in §2.) Contrary to the behavior of the gap probability $P_{\boldsymbol{\alpha}}(s)$, the limiting pair correlation function is the same as for random variables from a Poisson process.[2]

1.3. *Outline of the paper.* Sections 2–4 give a detailed account of the main results of this paper. Section 2 discusses the statistical properties of affine lattice points inside a large sphere that are projected onto the unit sphere. A dual problem is the question of the probability that a ray of length T pointing in a random direction intersects exactly r lattice spheres whose radius scales as $T^{-1/(d-1)}$. The solution of the latter problem is provided in Section 3, and applied in Section 4 to the distribution of the free path lengths of the Lorentz gas. Both of the above

[2]Boca and Zaharescu consider a slightly different sequence of directions, which is obtained by replacing the last condition in (1.10) with $\max(|m + \alpha_1|, |n + \alpha_2|) < T$. This sequence is however not uniformly distributed modulo one, which explains the discrepancy with the Poisson pair correlation function observed in [6].

lattice point problems fall into a general class of lattice point problems in randomly sheared or rotated domains, which are discussed in Section 6. The central idea for the solution of such questions is to exploit equidistribution results for flows on the homogeneous spaces $\mathrm{SL}(d,\mathbb{Z})\backslash \mathrm{SL}(d,\mathbb{R})$ and $\mathrm{ASL}(d,\mathbb{Z})\backslash \mathrm{ASL}(d,\mathbb{R})$, which represent the space of lattices (resp. affine lattices) of covolume one. We establish the required ergodic-theoretic results in Section 5. The key ingredient is Ratner's theorem [30] on the classification of ergodic measures invariant under a unipotent flow. We provide useful integration formulas on $\mathrm{SL}(d,\mathbb{Z})\backslash \mathrm{SL}(d,\mathbb{R})$ and $\mathrm{ASL}(d,\mathbb{Z})\backslash \mathrm{ASL}(d,\mathbb{R})$ in Section 7 and in Section 8 we apply these to our limit functions. Detailed proofs of the main limit theorems in Sections 2–4 are given in Section 9. The proofs for Section 2 are virtually identical to those of the corresponding theorems in Section 3.

2. **Distribution of visible lattice points**

2.1. *Lattices.* Let $\mathscr{L} \subset \mathbb{R}^d$ be a euclidean lattice of covolume one. Recall that $\mathscr{L} = \mathbb{Z}^d M$ for some $M \in \mathrm{SL}(d,\mathbb{R})$ and that therefore the homogeneous space $X_1 = \mathrm{SL}(d,\mathbb{Z})\backslash \mathrm{SL}(d,\mathbb{R})$ parametrizes the space of lattices of covolume one.

Let $\mathrm{ASL}(d,\mathbb{R}) = \mathrm{SL}(d,\mathbb{R}) \ltimes \mathbb{R}^d$ be the semidirect product group with multiplication law

$$(M,\boldsymbol{\xi})(M',\boldsymbol{\xi}') = (MM', \boldsymbol{\xi} M' + \boldsymbol{\xi}'). \tag{2.1}$$

An action of $\mathrm{ASL}(d,\mathbb{R})$ on $\mathbb{R}^d$ can be defined as

$$\boldsymbol{y} \mapsto \boldsymbol{y}(M,\boldsymbol{\xi}) := \boldsymbol{y} M + \boldsymbol{\xi}. \tag{2.2}$$

Each *affine* lattice (i.e. translate of a lattice) of covolume one in $\mathbb{R}^d$ can then be expressed as $\mathbb{Z}^d g$ for some $g \in \mathrm{ASL}(d,\mathbb{R})$, and the space of affine lattices is then represented by $X = \mathrm{ASL}(d,\mathbb{Z})\backslash \mathrm{ASL}(d,\mathbb{R})$ where $\mathrm{ASL}(d,\mathbb{Z}) = \mathrm{SL}(d,\mathbb{Z}) \ltimes \mathbb{Z}^d$. We denote by μ_1 and μ the Haar measure on $\mathrm{SL}(d,\mathbb{R})$ and $\mathrm{ASL}(d,\mathbb{R})$, respectively, normalized in such a way that they represent probability measures on X_1 and X.

If $\boldsymbol{\alpha} \in \mathbb{Q}^d$, say $\boldsymbol{\alpha} = \boldsymbol{p}/q$ for $\boldsymbol{p} \in \mathbb{Z}^d$, $q \in \mathbb{Z}_{>0}$, we see that

$$\left(\mathbb{Z}^d + \frac{\boldsymbol{p}}{q}\right)\gamma M = \left(\mathbb{Z}^d + \frac{\boldsymbol{p}}{q}\right) M \tag{2.3}$$

for all

$$\gamma \in \Gamma(q) := \{\gamma \in \mathrm{SL}(d,\mathbb{Z}) \,:\, \gamma \equiv 1_d \bmod q\}, \tag{2.4}$$

the principal congruence subgroup. This means that the space of affine lattices with $\boldsymbol{\alpha} = \boldsymbol{p}/q$ can be parametrized by the homogeneous space $X_q = \Gamma(q)\backslash \mathrm{SL}(d,\mathbb{R})$ (this is not necessarily one-to-one). We denote by μ_q the Haar measure on $\mathrm{SL}(d,\mathbb{R})$ which is normalized as a probability measure on X_q.

2.2. *Basic set-up.* We fix a lattice $\mathscr{L} \subset \mathbb{R}^d$ of covolume one, and fix, once and for all, a choice of $M_0 \in \mathrm{SL}(d,\mathbb{R})$ such that $\mathscr{L} = \mathbb{Z}^d M_0$. Given $\boldsymbol{\alpha} \in \mathbb{R}^d$ we

then define the affine lattice

$$\mathscr{L}_{\boldsymbol{\alpha}} := (\mathbb{Z}^d + \boldsymbol{\alpha})M_0 = \mathbb{Z}^d(1, \boldsymbol{\alpha})(M_0, \mathbf{0}). \tag{2.5}$$

Consider the set $\mathscr{P}_T$ of lattice points $\boldsymbol{y} \in \mathscr{L}_{\boldsymbol{\alpha}}$ inside the ball $\mathscr{B}_T^d$ of radius T, or, more generally, the spherical shell

$$\mathscr{B}_T^d(c) = \{\boldsymbol{x} \in \mathbb{R}^d : cT \le \|\boldsymbol{x}\| < T\}, \qquad 0 \le c < 1. \tag{2.6}$$

For T large there are asymptotically $(1-c^d)\,\mathrm{vol}(\mathscr{B}_1^d)T^d$ such points, where $\mathrm{vol}(\mathscr{B}_1^d) = \pi^{d/2}/\Gamma(\frac{d+2}{2})$ is the volume of the unit ball. For each T, we study the corresponding *directions,*

$$\|\boldsymbol{y}\|^{-1}\boldsymbol{y} \in \mathrm{S}_1^{d-1}, \qquad \text{for } \boldsymbol{y} \in \mathscr{P}_T = \mathscr{L}_{\boldsymbol{\alpha}} \cap \mathscr{B}_T^d(c) \setminus \{\mathbf{0}\}, \tag{2.7}$$

where $\mathrm{S}_\rho^{d-1} \subset \mathbb{R}^d$ denotes the $(d-1)$-sphere of radius ρ. It is well known that, as $T \to \infty$, these points become uniformly distributed on S_1^{d-1}: For any set $\mathfrak{U} \subset \mathrm{S}_1^{d-1}$ with boundary of measure zero (with respect to the volume element $\mathrm{vol}_{\mathrm{S}_1^{d-1}}$ on S_1^{d-1}) we have

$$\lim_{T\to\infty} \frac{\#\{\boldsymbol{y} \in \mathscr{P}_T \,:\, \|\boldsymbol{y}\|^{-1}\boldsymbol{y} \in \mathfrak{U}\}}{\#\mathscr{P}_T} = \frac{\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\mathfrak{U})}{\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\mathrm{S}_1^{d-1})}. \tag{2.8}$$

Recall that $\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\mathrm{S}_1^{d-1}) = d\,\mathrm{vol}(\mathscr{B}_1^d)$.

2.3. *Distribution in small discs.* We are interested in the fine-scale distribution of the directions to points in $\mathscr{P}_T$, e.g., in the probability of finding r directions in a small disc with random center $\boldsymbol{v} \in \mathrm{S}_1^{d-1}$. We define $\mathfrak{D}_T(\sigma, \boldsymbol{v}) \subset \mathrm{S}_1^{d-1}$ to be the open disc with center $\boldsymbol{v}$ and volume

$$\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\mathfrak{D}_T(\sigma, \boldsymbol{v})) = \frac{\sigma d}{1-c^d}T^{-d}. \tag{2.9}$$

The radius of $\mathfrak{D}_T(\sigma, \boldsymbol{v})$ is thus $\asymp T^{-d/(d-1)}$ (if $\sigma > 0$). We introduce the counting function

$$\mathscr{N}_{c,T}(\sigma, \boldsymbol{v}) = \#\{\boldsymbol{y} \in \mathscr{P}_T \,:\, \|\boldsymbol{y}\|^{-1}\boldsymbol{y} \in \mathfrak{D}_T(\sigma, \boldsymbol{v})\} \tag{2.10}$$

for the number of points in $\mathfrak{D}_T(\sigma, \boldsymbol{v})$. The motivation for the definition (2.9) is that it implies, via (2.8), that the expectation value for the counting function is asymptotically equal to σ (for $T \to \infty$ and σ fixed):

$$\begin{aligned}\int_{\mathrm{S}_1^{d-1}} \mathscr{N}_{c,T}(\sigma, \boldsymbol{v})\, d\lambda(\boldsymbol{v}) \\ \sim \mathrm{vol}(\mathscr{B}_T^d(c))\frac{\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\mathfrak{D}_T(\sigma))}{\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\mathrm{S}_1^{d-1})} = \frac{1-c^d}{d}\,\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\mathfrak{D}_T(\sigma))T^d = \sigma,\end{aligned} \tag{2.11}$$

where λ is any probability measure on S_1^{d-1} with continuous density, and $\mathfrak{D}_T(\sigma) = \mathfrak{D}_T(\sigma, \boldsymbol{e}_1)$ is the disc centered at $\boldsymbol{e}_1 = (1, 0, \dots, 0)$.

THEOREM 2.1. *Let λ be a Borel probability measure on S_1^{d-1} absolutely continuous with respect to Lebesgue measure. Then, for every $\sigma \geq 0$ and $r \in \mathbb{Z}_{\geq 0}$, the limit*

$$E_{c,\boldsymbol{\alpha}}(r,\sigma) := \lim_{T\to\infty} \lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} : \mathcal{N}_{c,T}(\sigma, \boldsymbol{v}) = r\}) \tag{2.12}$$

exists, and for fixed $c, \boldsymbol{\alpha}, r$ the convergence is uniform with respect to σ in any compact subset of $\mathbb{R}_{\geq 0}$. The limit function is given by

(2.13)

$$E_{c,\boldsymbol{\alpha}}(r,\sigma) = \begin{cases} \mu_1(\{M \in X_1 : \#(\mathbb{Z}^d_* M \cap \mathfrak{C}(c,\sigma)) = r\}) & \text{if } \boldsymbol{\alpha} \in \mathbb{Z}^d \\ \mu_q(\{M \in X_q : \#((\mathbb{Z}^d + \frac{\boldsymbol{p}}{q})M \cap \mathfrak{C}(c,\sigma)) = r\}) & \text{if } \boldsymbol{\alpha} = \frac{\boldsymbol{p}}{q} \in \mathbb{Q}^d \setminus \mathbb{Z}^d \\ \mu(\{(M,\boldsymbol{\xi}) \in X : \#((\mathbb{Z}^d M + \boldsymbol{\xi}) \cap \mathfrak{C}(c,\sigma)) = r\}) & \text{if } \boldsymbol{\alpha} \notin \mathbb{Q}^d, \end{cases}$$

where

$$\mathfrak{C}(c,\sigma) = \big\{(x_1, \dots, x_d) \in \mathbb{R}^d : c < x_1 < 1,\ \|(x_2, \dots, x_d)\| \leq x_1 A(c,\sigma)\big\}, \tag{2.14}$$

$$A(c,\sigma) = \left(\frac{\sigma d}{(1-c^d)\,\mathrm{vol}(\mathcal{B}_1^{d-1})}\right)^{\frac{1}{d-1}}, \qquad \mathrm{vol}(\mathcal{B}_1^{d-1}) = \frac{\pi^{(d-1)/2}}{\Gamma(\frac{d+1}{2})}. \tag{2.15}$$

In particular, $E_{c,\boldsymbol{\alpha}}(r,\sigma)$ is continuous in σ and independent of $\mathfrak{L}$ and λ.

In the above, we use the notation $\mathbb{Z}^d_* := \mathbb{Z}^d \setminus \{\mathbf{0}\}$. Although the use of $\mathbb{Z}^d_*$ is superfluous at this point (since $\mathfrak{C}(c,\sigma)$ does not contain zero), it appears as the natural object in the proof. This subtlety is due to the fact that for generic M we have $\mathbb{Z}^d M \cap \mathfrak{C}(0,\sigma) \neq \mathbb{Z}^d M \cap \overline{\mathfrak{C}(0,\sigma)}$ but $\mathbb{Z}^d_* M \cap \mathfrak{C}(0,\sigma) = \mathbb{Z}^d_* M \cap \overline{\mathfrak{C}(0,\sigma)}$.

Theorem 2.1 says that the limiting distribution $E_{c,\mathbf{0}}(r,\sigma)$ is given by the probability that there are r points of a random lattice in the cone $\mathfrak{C}(c,\sigma)$, and $E_{c,\boldsymbol{\alpha}}(r,\sigma)$ for $\boldsymbol{\alpha} \notin \mathbb{Q}^d$ is the corresponding probability for a random affine lattice. Hence in particular $E_{c,\boldsymbol{\alpha}}(r,\sigma)$ is independent of $\boldsymbol{\alpha}$ when $\boldsymbol{\alpha} \notin \mathbb{Q}^d$.

Remark 2.2. We will furthermore prove that when $c = 0$ the function $E_{c,\boldsymbol{\alpha}}(r,\sigma)$ is C^1 with respect to $\sigma > 0$; see Section 8.5. We expect that the same statement should also be true for any fixed $0 < c < 1$.

Remark 2.3. In the case $c = 0$, $d = 2$ and $\boldsymbol{\alpha} \notin \mathbb{Q}^2$ our distribution coincides with Elkies and McMullen's limiting distribution [14] for the probability of finding r elements of the sequence $\sqrt{n} \bmod 1$ ($n = 1, \dots, N$) in a randomly shifted interval of length σ/N ($N \to \infty$). Although the two problems are seemingly unrelated, the reason for this coincidence is that both results use equidistribution of translates of different orbits on the space of affine lattices X with respect to the same test functions.

Remark 2.4. By a general statistical argument, cf. e.g. [14], [22], Theorem 1.3 is an immediate corollary of Theorem 2.1 in the case $d = 2$, $r = 0$, with the limit function $P_{\boldsymbol{\alpha}}(s)$ explicitly given by

$$P_{\boldsymbol{\alpha}}(s) := -\frac{d}{ds} E_{0,\boldsymbol{\alpha}}(0, s) \quad (s > 0); \qquad P_{\boldsymbol{\alpha}}(0) := 1. \tag{2.16}$$

The continuity of $P_{\boldsymbol{\alpha}}(s)$ for $s > 0$ follows from Remark 2.2.

To exhibit explicitly the group action which will play a central role in the proof of the above statements, it is convenient to realize S_1^{d-1} as the homogeneous space $\mathrm{SO}(d-1)\backslash \mathrm{SO}(d)$ by setting $\boldsymbol{v} = \boldsymbol{e}_1 K$ with $\boldsymbol{e}_1 = (1, 0, \ldots, 0)$ and $K \in \mathrm{SO}(d)$. The stabilizer of $\boldsymbol{e}_1$ is isomorphic to $\mathrm{SO}(d-1)$ (acting from the right), where $\mathrm{SO}(d-1)$ is identified with the subgroup

$$\begin{pmatrix} 1 & \mathbf{0} \\ {}^{\mathrm{t}}\mathbf{0} & \mathrm{SO}(d-1) \end{pmatrix} \subset \mathrm{SO}(d). \tag{2.17}$$

Then

$$\mathfrak{D}_T(\sigma, \boldsymbol{v}) = \mathfrak{D}_T(\sigma)K = \{\boldsymbol{x} : \boldsymbol{x}K^{-1} \in \mathfrak{D}_T(\sigma)\} \tag{2.18}$$

and

$$\mathcal{N}_{c,T}(\sigma, K) = \#(\mathcal{P}_T \cap \mathfrak{D}_T(\sigma)K) \tag{2.19}$$

is the number of points in $\mathfrak{D}_T(\sigma)K$. Note that $\mathcal{N}_{c,T}(\sigma, K)$ is left-invariant under the action of $\mathrm{SO}(d-1)$ and thus may be viewed as a function on $\mathrm{SO}(d-1)\backslash \mathrm{SO}(d)$. The statement equivalent to Theorem 2.1 is now that, if λ is a Borel probability measure on $\mathrm{SO}(d)$ absolutely continuous with respect to Lebesque measure, then

$$\lim_{T\to\infty} \lambda(\{K \in \mathrm{SO}(d) : \mathcal{N}_{c,T}(\sigma, K) = r\}) = E_{c,\boldsymbol{\alpha}}(r, \sigma). \tag{2.20}$$

2.4. *Visible lattice points.* In the study of directions of affine lattice points it is natural to restrict our attention to those points that are visible from the origin. That is, we consider the set of directions without counting multiplicities. Nontrivial multiplicities are only obtained when the $\mathbb{Q}$-linear span of 1 and the components of $\boldsymbol{\alpha}$ has dimension ≤ 2. If $\boldsymbol{\alpha} \notin \mathbb{Q}^d$ then the multiplicities are statistically insignificant; in fact they can only occur along at most a single line through the origin, and thus restricting to considering only the visible lattice points still yields the same limit distribution as in Theorem 2.1.

Hence from now on we will assume $\boldsymbol{\alpha} \in \mathbb{Q}^d$. If $\boldsymbol{\alpha} = \mathbf{0}$ then the visible lattice points are exactly the *primitive* lattice points, i.e. those points $\boldsymbol{m}M_0 \in \mathcal{L}$ for which $\boldsymbol{m} \in \mathbb{Z}_*^d$, $\gcd(\boldsymbol{m}) = 1$. In the general case $\boldsymbol{\alpha} = \frac{\boldsymbol{p}}{q} \in \mathbb{Q}^d$ ($q \in \mathbb{Z}_{>0}$, $\boldsymbol{p} \in \mathbb{Z}^d$), the set of visible lattice points is:

$$\widehat{\mathcal{L}}_{\boldsymbol{\alpha}} = \widehat{\mathbb{Z}}^d_{\boldsymbol{\alpha}} M_0, \qquad \widehat{\mathbb{Z}}^d_{\boldsymbol{\alpha}} := \{\boldsymbol{x} \in (\mathbb{Z}^d + \boldsymbol{\alpha}) \setminus \{\mathbf{0}\} : \gcd(q\boldsymbol{x}) \leq q\}. \tag{2.21}$$

From now on in this section we will assume that $q \in \mathbb{Z}_{>0}$ is the *minimal* integer which gives $q\boldsymbol{\alpha} \in \mathbb{Z}^d$. Given $0 \leq c < 1$ we set $\widehat{\mathcal{P}}_T = \widehat{\mathcal{L}}_{\boldsymbol{\alpha}} \cap \mathcal{B}_T^d(c)$; then by a

sieving argument using (2.21) and (2.8) one shows that for any set $\mathfrak{U} \subset \mathrm{S}_1^{d-1}$ with boundary of measure zero,

(2.22)
$$\lim_{T\to\infty} \frac{\#\{\boldsymbol{y} \in \widehat{\mathcal{P}}_T \cap \mathcal{B}_T^d(c) \,:\, \|\boldsymbol{y}\|^{-1}\boldsymbol{y} \in \mathfrak{U}\}}{\operatorname{vol} \mathcal{B}_T^d(c)} = \kappa_q \frac{\operatorname{vol}_{\mathrm{S}_1^{d-1}}(\mathfrak{U})}{\operatorname{vol}_{\mathrm{S}_1^{d-1}}(\mathrm{S}_1^{d-1})},$$
$$\text{with} \qquad \kappa_q := \Big(\sum_{\substack{n\ge 1\\(n,q)=1}} \mu(n) n^{-d}\Big) \sum_{\substack{1\le t\le q\\(t,q)=1}} t^{-d} = \Big(\sum_{\substack{n\ge 1\\(n,q)=1}} n^{-d}\Big)^{-1} \sum_{\substack{1\le t\le q\\(t,q)=1}} t^{-d}.$$

When $\boldsymbol{\alpha} \in \mathbb{Z}^d$ this specializes to the well-known fact that the asymptotic density of the primitive points in $\mathbb{Z}^d$ is $\zeta(d)^{-1}$. It follows from (2.22) that if we introduce the following analogue of (2.10) for visible lattice points:

(2.23)
$$\widehat{\mathcal{N}}_{c,T}(\sigma, \boldsymbol{v}) = \#\{\boldsymbol{y} \in \widehat{\mathcal{P}}_T \,:\, \|\boldsymbol{y}\|^{-1}\boldsymbol{y} \in \mathfrak{D}_T(\kappa_q^{-1}\sigma, \boldsymbol{v})\}, \qquad \widehat{\mathcal{P}}_T = \widehat{\mathcal{L}}_{\boldsymbol{\alpha}} \cap \mathcal{B}_T^d(c);$$

then the expectation value for $\widehat{\mathcal{N}}$ is again asymptotically equal to σ:

$$\lim_{T\to\infty} \int_{\mathrm{S}_1^{d-1}} \widehat{\mathcal{N}}_{c,T}(\sigma, \boldsymbol{v})\, d\lambda(\boldsymbol{v}) = \sigma, \tag{2.24}$$

for any fixed $\sigma \ge 0$, $0 \le c < 1$ and λ as in (2.11).

THEOREM 2.5. *Let λ be a Borel probability measure on S_1^{d-1} absolutely continuous with respect to Lebesgue measure. Then, for every $\sigma \ge 0$ and $r \in \mathbb{Z}_{\ge 0}$, the limit*

$$\widehat{E}_{c,\boldsymbol{\alpha}}(r, \sigma) := \lim_{T\to\infty} \lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} : \widehat{\mathcal{N}}_{c,T}(\sigma, \boldsymbol{v}) = r\}) \tag{2.25}$$

exists, and for fixed c, r the convergence is uniform with respect to σ in any compact subset of $\mathbb{R}_{\ge 0}$. The limit function is given by

$$\mu_q(\{M \in X_q \,:\, \#(\widehat{\mathbb{Z}}^d_{\boldsymbol{\alpha}} M \cap \mathfrak{C}(c, \kappa_q^{-1}\sigma)) = r\}) \qquad (\boldsymbol{\alpha} = \tfrac{\boldsymbol{p}}{q} \in \mathbb{Q}^d). \tag{2.26}$$

In particular, $\widehat{E}_{c,\boldsymbol{\alpha}}(r, \sigma)$ is continuous in σ and independent of $\mathcal{L}$ and λ.

Remark 2.6. The function $\widehat{E}_{0,\boldsymbol{\alpha}}(r, \sigma)$ is C^1 with respect to $\sigma > 0$. This is proved by adapting the arguments of Sections 7.1 and 8.5 to the setting of visible lattice points.

In dimension $d = 2$, considering only visible lattice points gives a variant of Theorem 1.3 with an everywhere continuous distribution function: Take $\boldsymbol{\alpha} \in \mathbb{Q}^2$, and consider the set of rescaled directions

$$\{\tfrac{1}{2\pi} \arg(x_1 + \mathrm{i}x_2) \,:\, \boldsymbol{x} = (x_1, x_2) \in \widehat{\mathbb{Z}}^2_{\boldsymbol{\alpha}}, \quad x_1^2 + x_2^2 < T^2\}. \tag{2.27}$$

Let us label these $M = M(T)$ numbers in order by

$$-\tfrac{1}{2} < \hat{\xi}_{M,1} < \hat{\xi}_{M,2} < \cdots < \hat{\xi}_{M,M} \le \tfrac{1}{2} \tag{2.28}$$

and define in addition $\hat{\xi}_{M,0} = \hat{\xi}_{M,M} - 1$. Note that this is exactly the sequence which is obtained from (1.11) by removing all repetitions. We now have:

COROLLARY 2.7. *There exists a distribution function* $\widehat{P}_{\boldsymbol{\alpha}}(s)$ *on* $\mathbb{R}_{\geq 0}$, *continuous on all of* $\mathbb{R}_{\geq 0}$, *such that for every* $s \geq 0$,

$$\lim_{M\to\infty} M^{-1}\#\big\{1 \leq j \leq M : M(\hat{\xi}_{M,j} - \hat{\xi}_{M,j-1}) \geq s\big\} = \widehat{P}_{\boldsymbol{\alpha}}(s). \tag{2.29}$$

Proof. Just as in Remark 2.4, the limit relation (2.29) follows from Theorem 2.5 together with the fact $M \sim \kappa_q \pi T^2$ as $T \to \infty$ (cf. (2.22)), and $\widehat{P}_{\boldsymbol{\alpha}}(s)$ is explicitly given by

$$\widehat{P}_{\boldsymbol{\alpha}}(s) := -\frac{d}{ds}\widehat{E}_{0,\boldsymbol{\alpha}}(0,s) \quad (s>0); \qquad \widehat{P}_{\boldsymbol{\alpha}}(0) := 1. \tag{2.30}$$

Note that $\widehat{E}_{0,\boldsymbol{\alpha}}(0,s) = E_{0,\boldsymbol{\alpha}}(0,\kappa_q^{-1}s)$ for all $s \geq 0$, since $\mathfrak{C}(0,\kappa_q^{-1}s)$ is star shaped. Hence

$$\widehat{P}_{\boldsymbol{\alpha}}(s) = \kappa_q^{-1} P_{\boldsymbol{\alpha}}(\kappa_q^{-1}s) \qquad \text{for } s>0. \tag{2.31}$$

The continuity of $\widehat{P}_{\boldsymbol{\alpha}}(s)$ for $s>0$ follows from Remark 2.2, or Remark 2.6. Furthermore, in Section 8.5 we will prove that (for $d=2$),

$$E_{0,\boldsymbol{\alpha}}(0,\sigma) = 1 - \kappa_q\sigma, \qquad \text{for all } \sigma \in \big[0,(2q)^{-1}\big], \tag{2.32}$$

and this implies that $\widehat{P}_{\boldsymbol{\alpha}}(s)$ is also continuous at $s=0$. □

When $\boldsymbol{\alpha} = \mathbf{0}$, Corollary 2.7 specializes to give the limiting gap distribution for directions of primitive lattice points in $\mathbb{Z}^2$, which was proved earlier by Boca, Cobeli and Zaharescu [3].

The proofs of Theorems 2.1 and 2.5 are virtually identical to those of Theorems 3.1 and 3.7; we will therefore only outline the differences in Section 9.4. In [24] we carry out a more detailed statistical analysis of the distribution of visible lattice points, which yields generalizations of Theorems 2.1 and 2.5, and also provide explicit formulas and tail estimates of the limiting distributions.

3. **The number of spheres in a random direction**

We now turn to a lattice point problem that is in some sense dual to the one studied in the previous Section 2. Its solution will also answer the question of the distribution of free path lengths in the periodic Lorentz gas, see Section 4 below for details.

3.1. *Spheres centered at lattice points.* We place at each lattice point $\boldsymbol{y} \in \mathscr{L}_{\boldsymbol{\alpha}}$ a ball of small radius ρ and consider the set $\mathscr{B}_\rho^d + \mathscr{L}_{\boldsymbol{\alpha}}$. The set of balls with centers inside the shell (2.6) is

$$\{\boldsymbol{x} \in \mathscr{B}_\rho^d + \boldsymbol{y} \,:\, \boldsymbol{y} \in \mathscr{L}_{\boldsymbol{\alpha}} \cap \mathscr{B}_T^d(c) \setminus \{\mathbf{0}\}\}. \tag{3.1}$$

Note that we remove any ball at $\boldsymbol{y} = \boldsymbol{0}$ (this is only relevant in the case $\boldsymbol{\alpha} \in \mathbb{Z}^d$). Furthermore, we will always keep $\rho \le m(\mathcal{L}_{\boldsymbol{\alpha}}) := \min\{\|\boldsymbol{y}\| : \boldsymbol{y} \in \mathcal{L}_{\boldsymbol{\alpha}} \setminus \{\boldsymbol{0}\}\}$, so that $\boldsymbol{0}$ lies outside each of the balls in our set. We are interested in the number $\mathcal{N}_{c,T}(\rho, \boldsymbol{v})$ of intersections of this set with a ray starting at the origin $\boldsymbol{0}$ that points in the random direction $\boldsymbol{v} \in \mathrm{S}_1^{d-1}$ distributed according to the probability measure λ. That is,

$$\mathcal{N}_{c,T}(\rho, \boldsymbol{v}) := \#\big\{\boldsymbol{y} \in \mathcal{L}_{\boldsymbol{\alpha}} \cap \mathcal{B}_T^d(c) \setminus \{\boldsymbol{0}\} \,:\, \mathbb{R}_{>0}\boldsymbol{v} \cap (\mathcal{B}_\rho^d + \boldsymbol{y}) \neq \varnothing\big\}. \tag{3.2}$$

If $\rho \le \|\boldsymbol{y}\|$, then a ray in direction $\boldsymbol{v}$ hits the ball $\mathcal{B}_\rho^d + \boldsymbol{y}$ if and only if

$$\|\boldsymbol{y}\|^{-1}\boldsymbol{y} \in \mathfrak{D}(\|\boldsymbol{y}\|^{-1}\rho, \boldsymbol{v}) \tag{3.3}$$

with the disc

$$\begin{aligned} &\mathfrak{D}(\epsilon, \boldsymbol{v}) = (\mathcal{B}_\epsilon^d + \boldsymbol{v})(1-\epsilon^2)^{-1/2} \cap \mathrm{S}_1^{d-1} \qquad (0 < \epsilon < 1); \\ &\mathfrak{D}(1, \boldsymbol{v}) = \{\boldsymbol{w} \in \mathrm{S}_1^{d-1} \,:\, \boldsymbol{w} \cdot \boldsymbol{v} > 0\}. \end{aligned} \tag{3.4}$$

We will again use the shorthand $\mathfrak{D}(\epsilon) = \mathfrak{D}(\epsilon, \boldsymbol{e}_1)$. The radius of this disc is $\sim \epsilon$, for $\epsilon \to 0$. Hence the number of balls hit by a ray in direction $\boldsymbol{v}$ is

$$\mathcal{N}_{c,T}(\rho, \boldsymbol{v}) = \#\left\{\boldsymbol{y} \in \mathcal{L}_{\boldsymbol{\alpha}} \cap \mathcal{B}_T^d(c) \setminus \{\boldsymbol{0}\} \,:\, \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} \in \mathfrak{D}(\|\boldsymbol{y}\|^{-1}\rho, \boldsymbol{v})\right\}; \tag{3.5}$$

compare (2.19).

For any λ as in (2.11), one finds for the expectation value as $T \to \infty$, $\rho \to 0$

$$\begin{aligned} \int_{\mathrm{S}_1^{d-1}} \mathcal{N}_{c,T}(\rho, \boldsymbol{v})\, d\lambda(\boldsymbol{v}) &\sim \int_{\boldsymbol{y} \in \mathcal{B}_T^d(c),\, \|\boldsymbol{y}\| > \rho} \frac{\mathrm{vol}(\mathfrak{D}(\|\boldsymbol{y}\|^{-1}\rho))}{\mathrm{vol}(\mathrm{S}_1^{d-1})}\, d\,\mathrm{vol}(\boldsymbol{y}) \\ &\sim \frac{\mathrm{vol}(\mathcal{B}_1^{d-1})}{\mathrm{vol}(\mathrm{S}_1^{d-1})}\rho^{d-1} \int_{\mathcal{B}_T^d(c)} \frac{d\,\mathrm{vol}(\boldsymbol{y})}{\|\boldsymbol{y}\|^{d-1}} \\ &= \mathrm{vol}(\mathcal{B}_1^{d-1})(1-c)\rho^{d-1}T. \end{aligned} \tag{3.6}$$

This suggests the scaling $\rho = \sigma T^{-1/(d-1)}$ with $\sigma \ge 0$ fixed.

THEOREM 3.1. *Let λ be a Borel probability measure on S_1^{d-1} absolutely continuous with respect to $\mathrm{vol}_{\mathrm{S}_1^{d-1}}$. Then, for every $\sigma \ge 0$ and $r \in \mathbb{Z}_{\ge 0}$, the limit*

$$F_{c,\boldsymbol{\alpha}}(r, \sigma) := \lim_{T \to \infty} \lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} : \mathcal{N}_{c,T}(\sigma T^{-1/(d-1)}, \boldsymbol{v}) = r\}) \tag{3.7}$$

exists, and for fixed $\boldsymbol{\alpha}, r$ the convergence is uniform with respect to σ in any compact subset of $\mathbb{R}_{\ge 0}$, and with respect to $c \in [0,1]$. The limit function is given by

$$F_{c,\boldsymbol{\alpha}}(r, \sigma) = \begin{cases} \mu_1(\{M \in X_1 : \#(\mathbb{Z}_*^d M \cap \mathfrak{Z}(c, \sigma)) = r\}) & \text{if } \boldsymbol{\alpha} \in \mathbb{Z}^d \\ \mu_q(\{M \in X_q : \#((\mathbb{Z}^d + \frac{\boldsymbol{p}}{q})M \cap \mathfrak{Z}(c, \sigma)) = r\}) & \text{if } \boldsymbol{\alpha} = \frac{\boldsymbol{p}}{q} \in \mathbb{Q}^d \setminus \mathbb{Z}^d \\ \mu(\{(M, \boldsymbol{\xi}) \in X : \#((\mathbb{Z}^d M + \boldsymbol{\xi}) \cap \mathfrak{Z}(c, \sigma)) = r\}) & \text{if } \boldsymbol{\alpha} \notin \mathbb{Q}^d, \end{cases} \tag{3.8}$$

where

$$\mathfrak{Z}(c,\sigma)=\big\{(x_1,\dots,x_d)\in\mathbb{R}^d : c<x_1<1, \|(x_2,\dots,x_d)\|<\sigma\big\}. \tag{3.9}$$

In particular, $F_{c,\boldsymbol{\alpha}}(r,\sigma)$ is continuous in σ and independent of $\mathscr{L}$ and λ.

Remark 3.2. In the case $c=0$ the function $F_{c,\boldsymbol{\alpha}}(r,\sigma)$ is C^1 with respect to $\sigma>0$; we will prove this in Section 8.3. (We expect the same should be true also for any fixed $0<c<1$.) If $\boldsymbol{\alpha}\notin\mathbb{Q}^d$ then $F_{c,\boldsymbol{\alpha}}(r,\sigma)$ is independent of $\boldsymbol{\alpha}$; we denote this "universal" limit function simply by $F_c(r,\sigma)$. We prove in Section 8.3 that $F_c(r,\sigma)$ is C^2 with respect to $\sigma>0$, for any fixed $0\le c<1$.

Remark 3.3. We give tail estimates for $F_0(0,\sigma)$ and $F_{0,\mathbf{0}}(0,\sigma)$ for general dimension d in [25]. In the special case $d=2$, explicit formulas for $F_0(r,\sigma)$ and $F_{0,\mathbf{0}}(r,\sigma)$ were given in [36], where these limit functions came up in a different set of problems. Specifically, $F_0(r,\sigma)=f_r^{\mathrm{box,ASL}_2}(2\sigma)$ and $F_{0,\mathbf{0}}(r,\sigma)=f_{2r+1}^{\mathrm{box,SL}_2}(4\sigma)$ in the notation of [36, §7].

3.2. *A variation.* Instead of rays emerging from the origin we consider now the family of rays starting at the points $\rho\boldsymbol{\beta}(\boldsymbol{v})$ in direction $\boldsymbol{v}$, where $\boldsymbol{\beta}:\mathrm{S}_1^{d-1}\to\mathbb{R}^d$ is some fixed continuous function. We will keep ρ so small that, for all $\boldsymbol{y}\in\mathscr{L}_{\boldsymbol{\alpha}}\setminus\{\mathbf{0}\}$ and all $\boldsymbol{v}\in\mathrm{S}_1^{d-1}$, the point $\rho\boldsymbol{\beta}(\boldsymbol{v})$ lies outside the ball $\mathscr{B}_\rho^d+\boldsymbol{y}$. Then the ray $\rho\boldsymbol{\beta}(\boldsymbol{v})+\mathbb{R}_{>0}\boldsymbol{v}$ hits the ball $\mathscr{B}_\rho^d+\boldsymbol{y}$ if and only if

$$\frac{\boldsymbol{y}-\rho\boldsymbol{\beta}(\boldsymbol{v})}{\|\boldsymbol{y}-\rho\boldsymbol{\beta}(\boldsymbol{v})\|}\in\mathfrak{D}(\|\boldsymbol{y}-\rho\boldsymbol{\beta}(\boldsymbol{v})\|^{-1}\rho,\boldsymbol{v}), \tag{3.10}$$

compare the analogous argument in the previous section. Hence the number of balls in (3.1) intersecting this ray is $\mathscr{N}_{c,T}(\rho,\boldsymbol{v},\boldsymbol{\beta}(\boldsymbol{v}))$, where

$$\mathscr{N}_{c,T}(\rho,\boldsymbol{v},\boldsymbol{w}):=\#\left\{\boldsymbol{y}\in(\mathscr{L}_{\boldsymbol{\alpha}}\cap\mathscr{B}_T^d(c)\setminus\{\mathbf{0}\})-\rho\boldsymbol{w}:\frac{\boldsymbol{y}}{\|\boldsymbol{y}\|}\in\mathfrak{D}(\|\boldsymbol{y}\|^{-1}\rho,\boldsymbol{v})\right\}. \tag{3.11}$$

THEOREM 3.4. *Let λ be a Borel probability measure on S_1^{d-1} absolutely continuous with respect to Lebesgue measure. Then, for every $\sigma\ge0$ and $r\in\mathbb{Z}_{\ge0}$, the limit*

$$F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma):=\lim_{T\to\infty}\lambda(\{\boldsymbol{v}\in\mathrm{S}_1^{d-1}:\mathscr{N}_{c,T}(\sigma T^{-1/(d-1)},\boldsymbol{v},\boldsymbol{\beta}(\boldsymbol{v}))=r\}) \tag{3.12}$$

exists, and for fixed $\boldsymbol{\alpha},\boldsymbol{\beta},\lambda,r$ the convergence is uniform with respect to σ in any compact subset of $\mathbb{R}_{\ge0}$, and with respect to $c\in[0,1]$. The limit function is given by

(3.13)

$$F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma)=\begin{cases}(\mu_1\times\lambda)(\{(M,\boldsymbol{v})\in X_1\times\mathrm{S}_1^{d-1}:\#(\mathbb{Z}_*^dM\cap\mathfrak{Z}_{\boldsymbol{v}}(c,\sigma))=r\}) & \text{if } \boldsymbol{\alpha}\in\mathbb{Z}^d\\ (\mu_q\times\lambda)(\{(M,\boldsymbol{v})\in X_q\times\mathrm{S}_1^{d-1}:\#((\mathbb{Z}^d+\frac{\boldsymbol{p}}{q})M\cap\mathfrak{Z}_{\boldsymbol{v}}(c,\sigma))=r\}) & \\ \qquad\qquad\text{if } \boldsymbol{\alpha}=\frac{\boldsymbol{p}}{q}\in\mathbb{Q}^d\setminus\mathbb{Z}^d & \\ \mu(\{(M,\boldsymbol{\xi})\in X:\#((\mathbb{Z}^dM+\boldsymbol{\xi})\cap\mathfrak{Z}(c,\sigma))=r\}) \text{ if } \boldsymbol{\alpha}\notin\mathbb{Q}^d, & \end{cases}$$

where

$$\mathfrak{Z}_{\boldsymbol{v}}(c,\sigma) = \mathfrak{Z}(c,\sigma) + \sigma \left\| \mathrm{Proj}_{\{\boldsymbol{v}\}^\perp} \boldsymbol{\beta}(\boldsymbol{v}) \right\| \cdot \boldsymbol{e}_2. \tag{3.14}$$

($\mathrm{Proj}_{\{\boldsymbol{v}\}^\perp}$ *denotes the orthogonal projection from* $\mathbb{R}^d$ *onto the orthogonal complement of* $\boldsymbol{v}$, *and* $\boldsymbol{e}_2 = (0,1,0,\dots,0)$.) *In particular,* $F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma)$ *is continuous in* σ *and independent of* $\mathscr{L}$, *and if* $\boldsymbol{\alpha} \notin \mathbb{Q}^d$ *then* $F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma) = F_c(r,\sigma)$, *independently of* $\boldsymbol{\beta}$ *and* λ.

Remark 3.5. Again, we prove in the case $c = 0$ that the function $F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma)$ is C^1 with respect to $\sigma > 0$; see Section 8.3.

Remark 3.6. It will be useful for several of the results in Section 4 below, as well as in the proofs in [23], to know that

$$\lim_{\sigma\to 0} F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(0,\sigma) = 1 \quad \text{and} \quad \lim_{\sigma\to\infty} F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma) = 0,$$

and that this holds uniformly with respect to the various parameters. This follows from the following two basic bounds, which we prove in Section 8.4. More exact asymptotic formulas will be given in [25].

Let $v_d := \mathrm{vol}(\mathcal{B}_1^{d-1}) = \pi^{(d-1)/2}/\Gamma(\frac{d+1}{2})$. Then for all $\sigma > 0$ we have

$$F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(0,\sigma) \ge 1 - v_d(1-c)\sigma^{d-1} \tag{3.15}$$

$$\text{and thus} \qquad \sum_{r=1}^{\infty} F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma) \le v_d(1-c)\sigma^{d-1}.$$

Furthermore, there exists a constant $C > 0$ which only depends on r, d (thus C is independent of c, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, λ) such that for all $\sigma > 0$ we have

$$F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma) \le C(1-c)^{-1}\sigma^{1-d}. \tag{3.16}$$

3.3. *Spheres centered at visible lattice points.* Now assume $\boldsymbol{\alpha} = \frac{\boldsymbol{p}}{q} \in \mathbb{Q}^d$ and set

$$\widehat{\mathcal{N}}_{c,T}(\rho,\boldsymbol{v},\boldsymbol{w}) := \#\left\{ \boldsymbol{y} \in (\widehat{\mathscr{L}}_{\boldsymbol{\alpha}} \cap \mathcal{B}_T^d(c)) - \rho\boldsymbol{w} \,:\, \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} \in \mathfrak{D}(\|\boldsymbol{y}\|^{-1}\rho, \boldsymbol{v}) \right\}. \tag{3.17}$$

THEOREM 3.7. *Let* λ *be a Borel probability measure on* S_1^{d-1} *absolutely continuous with respect to Lebesgue measure. Then, for every* $\sigma \ge 0$ *and* $r \in \mathbb{Z}_{\ge 0}$, *the limit*

$$\widehat{F}_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma) := \lim_{T\to\infty} \lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} : \widehat{\mathcal{N}}_{c,T}(\sigma T^{-1/(d-1)}, \boldsymbol{v}, \boldsymbol{\beta}(\boldsymbol{v})) = r\}) \tag{3.18}$$

exists, and for fixed $\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda, r$ *the convergence is uniform with respect to* σ *in any compact subset of* $\mathbb{R}_{\ge 0}$, *and with respect to* $c \in [0,1]$. *The limiting function is given by*

$$\widehat{F}_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma) = (\mu_q \times \lambda)(\{(M,\boldsymbol{v}) \in X_q \times \mathrm{S}_1^{d-1} : \#(\widehat{\mathbb{Z}}_{\boldsymbol{\alpha}}^d M \cap \mathfrak{Z}_{\boldsymbol{v}}(c,\sigma)) = r\}). \tag{3.19}$$

In particular, $\widehat{F}_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma)$ *is continuous in* σ *and independent of* $\mathscr{L}$.

Remark 3.8. The function $\widehat{F}_{0,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma)$ is C^1 with respect to $\sigma > 0$. This is proved by adapting the arguments of Sections 7.1, 8.1 and 8.3 to the setting of visible lattice points.

3.4. *Nonspherical objects.* Instead of balls we now consider more general objects

$$\mathcal{Q}_T = T^{-1/(d-1)}\mathcal{Q} = \{\boldsymbol{x} \in \mathbb{R}^d \,:\, T^{1/(d-1)}\boldsymbol{x} \in \mathcal{Q}\}, \tag{3.20}$$

where $\mathcal{Q}$ is a bounded open subset of $\mathbb{R}^d$ which satisfies the technical condition that, for Lebesgue-almost every $\boldsymbol{v} \in \mathrm{S}_1^{d-1}$, the subset $\mathrm{Proj}_{\{\boldsymbol{v}\}^\perp}\mathcal{Q} \subset \{\boldsymbol{v}\}^\perp$ has boundary of ($(d-1)$-dimensional) volume measure zero. This assumption is readily verified to hold for any "nice" set $\mathcal{Q}$; for instance it certainly holds whenever $\mathcal{Q}$ is convex, but also for much more general sets $\mathcal{Q}$.

As before we place translates of $\mathcal{Q}$ at lattice points, and consider the set

$$\{\boldsymbol{x} \in \mathcal{Q}_T + \boldsymbol{y} \,:\, \boldsymbol{y} \in \mathcal{L}_{\boldsymbol{\alpha}} \cap \mathcal{B}_T^d(c) \setminus \{\boldsymbol{0}\}\}. \tag{3.21}$$

The number of intersections with a ray starting at the origin in direction $\boldsymbol{v}$ is

$$\mathcal{N}_{c,T}(\mathcal{Q},\boldsymbol{v}) := \#\big\{\boldsymbol{y} \in \mathcal{L}_{\boldsymbol{\alpha}} \cap \mathcal{B}_T^d(c) \setminus \{\boldsymbol{0}\} \,:\, \mathbb{R}_{>0}\boldsymbol{v} \cap (\mathcal{Q}_T + \boldsymbol{y}) \neq \varnothing\big\}. \tag{3.22}$$

THEOREM 3.9. *Let λ be a Borel probability measure on S_1^{d-1} absolutely continuous with respect to $\mathrm{vol}_{\mathrm{S}_1^{d-1}}$. Then, for every $r \in \mathbb{Z}_{\geq 0}$, the limit*

$$F_{c,\boldsymbol{\alpha}}(r,\mathcal{Q}) := \lim_{T\to\infty} \lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} \,:\, \mathcal{N}_{c,T}(\mathcal{Q},\boldsymbol{v}) = r\}) \tag{3.23}$$

exists, and is given by

(3.24)
$$\begin{cases} (\lambda\times\mu_1)(\{(\boldsymbol{v},M) \in \mathrm{S}_1^{d-1}\times X_1 \,:\, \#(\mathbb{Z}_*^d M \cap \mathfrak{Z}(c,\mathcal{Q},\boldsymbol{v})) = r\}) & \text{if } \boldsymbol{\alpha} \in \mathbb{Z}^d \\ (\lambda\times\mu_q)(\{(\boldsymbol{v},M) \in \mathrm{S}_1^{d-1}\times X_q \,:\, \#((\mathbb{Z}^d + \frac{\boldsymbol{p}}{q})M \cap \mathfrak{Z}(c,\mathcal{Q},\boldsymbol{v})) = r\}) & \text{if } \boldsymbol{\alpha} = \frac{\boldsymbol{p}}{q} \in \mathbb{Q}^d \setminus \mathbb{Z}^d \\ (\lambda\times\mu)(\{(\boldsymbol{v},g) \in \mathrm{S}_1^{d-1}\times X \,:\, \#(\mathbb{Z}^d g \cap \mathfrak{Z}(c,\mathcal{Q},\boldsymbol{v})) = r\}) & \text{if } \boldsymbol{\alpha} \notin \mathbb{Q}^d, \end{cases}$$

where

$$\mathfrak{Z}(c,\mathcal{Q},\boldsymbol{v}) = \big\{\boldsymbol{x} \in \mathbb{R}^d \,:\, c < \boldsymbol{x}\cdot\boldsymbol{v} < 1,\ \mathbb{R}\boldsymbol{v} \cap (\mathcal{Q} + \boldsymbol{x}) \neq \varnothing\big\}. \tag{3.25}$$

In particular, $F_{c,\boldsymbol{\alpha}}(r,\mathcal{Q})$ is independent of $\mathcal{L}$.

The analogous statement holds for visible lattice points. Assume $\boldsymbol{\alpha} = \frac{\boldsymbol{p}}{q} \in \mathbb{Q}^d$ and set

$$\widehat{\mathcal{N}}_{c,T}(\mathcal{Q},\boldsymbol{v}) := \#\big\{\boldsymbol{y} \in \widehat{\mathcal{L}}_{\boldsymbol{\alpha}} \cap \mathcal{B}_T^d(c) \,:\, \mathbb{R}_{>0}\boldsymbol{v} \cap (\mathcal{Q}_T + \boldsymbol{y}) \neq \varnothing\big\}. \tag{3.26}$$

THEOREM 3.10. *Let λ be a Borel probability measure on S_1^{d-1} absolutely continuous with respect to Haar measure. Then, for every $\sigma > 0$ and $r \in \mathbb{Z}_{\geq 0}$, the limit*

$$\widehat{F}_{c,\boldsymbol{\alpha}}(r,\mathcal{Q}) := \lim_{T\to\infty} \lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} \,:\, \widehat{\mathcal{N}}_{c,T}(\mathcal{Q},\boldsymbol{v}) = r\}) \tag{3.27}$$

exists, and is given by

$$(\lambda \times \mu_q)(\{(\boldsymbol{v}, M) \in \mathrm{S}_1^{d-1} \times X_q \,:\, \#(\widehat{\mathbb{Z}}^d_{\boldsymbol{\alpha}} M \cap \mathfrak{Z}(c, \mathfrak{Q}, \boldsymbol{v})) = r\}). \tag{3.28}$$

In particular, $\widehat{F}_{c,\boldsymbol{\alpha}}(r, \mathfrak{Q})$ *is independent of* $\mathfrak{L}$.

All statements in this section are proved in Section 9.

4. **The periodic Lorentz gas**

We now show how the results of the previous Section 3 can be applied to the distribution of free path lengths (§4.1). We will then generalize these results to provide joint distributions of free path lengths and exact location of impact on the scatterer (§4.2), and the distribution of the velocity vector after the first hit (§4.3).

4.1. *Free path lengths.* Recall that the free path length for the initial condition $(\boldsymbol{q}, \boldsymbol{v}) \in \mathrm{T}^1(\mathcal{H}_\rho)$ is defined as

$$\tau_1(\boldsymbol{q}, \boldsymbol{v}; \rho) = \inf\{t > 0 \,:\, \boldsymbol{q} + t\boldsymbol{v} \notin \mathcal{H}_\rho\}. \tag{4.1}$$

The crucial observation is that if λ is any given probability measure on S_1^{d-1} and $0 < \rho < T$, $(\boldsymbol{q}, \boldsymbol{v}) \in \mathrm{T}^1(\mathcal{H}_\rho)$, then we have

$$\begin{aligned} \lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} \,:\, \mathcal{N}_{0,T+\rho}(\rho, \boldsymbol{v}) = 0\}) \\ \le \lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} \,:\, \tau_1(\boldsymbol{q}, \boldsymbol{v}; \rho) \ge T\}) \\ \le \lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} \,:\, \mathcal{N}_{0,T-\rho}(\rho, \boldsymbol{v}) = 0\}), \end{aligned} \tag{4.2}$$

where $\mathcal{N}_{0,T}$ is as defined in (3.5) with affine lattice $\mathcal{L}_{\boldsymbol{\alpha}} = \mathcal{L} - \boldsymbol{q}$ (thus $\boldsymbol{\alpha} \equiv -\boldsymbol{q} M_0^{-1} \bmod \mathbb{Z}^d$).

Let

$$\Phi_{\boldsymbol{\alpha}}(\xi) = -\frac{d}{d\xi} F_{0,\boldsymbol{\alpha}}(0, \xi^{1/(d-1)}). \tag{4.3}$$

This defines a continuous probability density on $\mathbb{R}_{>0}$ (cf. Remark 3.2). If $\boldsymbol{\alpha} \notin \mathbb{Q}^d$ then $\Phi_{\boldsymbol{\alpha}}(\xi)$ is independent of $\boldsymbol{\alpha}$ and we write $\Phi(\xi)$ for this function (as in (1.5)).

The following is a restatement of Theorem 1.1.

COROLLARY 4.1. *Fix a lattice* $\mathcal{L} = \mathbb{Z}^d M_0$. *Let* $\boldsymbol{q} \in \mathbb{R}^d \setminus \mathcal{L}$ *and* $\boldsymbol{\alpha} = -\boldsymbol{q} M_0^{-1}$, *and let* λ *be a Borel probability measure on* S_1^{d-1} *absolutely continuous with respect to Lebesgue measure. Then, for every* $\xi \ge 0$,

$$\lim_{\rho \to 0} \lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} \,:\, \rho^{d-1} \tau_1(\boldsymbol{q}, \boldsymbol{v}; \rho) \ge \xi\}) = \int_\xi^\infty \Phi_{\boldsymbol{\alpha}}(\xi') d\xi'. \tag{4.4}$$

Note here that the condition $\boldsymbol{q} \notin \mathcal{L}$ is ensures that $\tau_1(\boldsymbol{q}, \boldsymbol{v}; \rho)$ is defined for all sufficiently small ρ. Corollary 4.1 follows directly from (4.2) and Theorem 3.1; cf. the proof of Corollary 4.2 below.

The analogous result corresponding to the set-up of Section 3.2 is as follows. As in that section we let $\boldsymbol{\beta} : \mathrm{S}_1^{d-1} \to \mathbb{R}^d$ be a continuous function, and again let

λ be a Borel probability measure on S_1^{d-1} absolutely continuous with respect to Lebesgue measure.

If $\boldsymbol{q} \in \mathscr{L}$, it is possible that the trajectory $\boldsymbol{q} + \rho\boldsymbol{\beta}(\boldsymbol{v}) + \mathbb{R}_{>0}\boldsymbol{v}$ starts inside the scatterer (if $\|\boldsymbol{\beta}(\boldsymbol{v})\| < 1$), or will hit the scatterer at $\boldsymbol{q}$ (if $\|\boldsymbol{\beta}(\boldsymbol{v})\| \geq 1$ and $\boldsymbol{v}$ is suitably chosen). In the first case the corresponding free path length is undefined; in the second case $\tau_1(\boldsymbol{q} + \rho\boldsymbol{\beta}(\boldsymbol{v}), \boldsymbol{v}; \rho) = O(\rho)$. The measure of directions with short free path lengths,

$$\lambda\big(\big\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} \,:\, \tau_1(\boldsymbol{q} + \rho\boldsymbol{\beta}(\boldsymbol{v}), \boldsymbol{v}; \rho) \leq \tfrac{1}{2}m(\mathscr{L}_{\boldsymbol{\alpha}})\big\}\big) \tag{4.5}$$

is independent of ρ, for ρ sufficiently small.

In order to avoid these pathological cases we will from now on assume that $\boldsymbol{\beta}$ is such that *if $\boldsymbol{q} \in \mathscr{L}$, then the ray $\boldsymbol{\beta}(\boldsymbol{v}) + \mathbb{R}_{>0}\boldsymbol{v}$ lies completely outside $\mathscr{B}_1^d$, for each* $\boldsymbol{v} \in \mathrm{S}_1^{d-1}$. This assumption will be in force throughout the remainder of Section 4.

Set

$$\Phi_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\xi) = -\frac{d}{d\xi}F_{0,\boldsymbol{\alpha},\boldsymbol{\beta}}(0, \xi^{1/(d-1)}), \tag{4.6}$$

which, unlike $\Phi_{\boldsymbol{\alpha}}$, depends on the choice of the measure λ; cf. (3.13). The function $\Phi_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\xi)$ again defines a continuous probability density on $\mathbb{R}_{>0}$; see Remark 3.5.

COROLLARY 4.2. *For every* $\xi \geq 0$,

$$\lim_{\rho\to 0} \lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} \,:\, \rho^{d-1}\tau_1(\boldsymbol{q} + \rho\boldsymbol{\beta}(\boldsymbol{v}), \boldsymbol{v}; \rho) \geq \xi\}) = \int_\xi^\infty \Phi_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\xi')d\xi'. \tag{4.7}$$

In this statement, $\tau_1(\boldsymbol{q} + \rho\boldsymbol{\beta}(\boldsymbol{v}), \boldsymbol{v}; \rho)$ is well-defined for all $\boldsymbol{v} \in \mathrm{S}_1^{d-1}$ so long as ρ is sufficiently small. (For if $\boldsymbol{q} \in \mathscr{L}$ then, by our assumptions on $\boldsymbol{\beta}$, we have in particular $\|\boldsymbol{\beta}(\boldsymbol{v})\| \geq 1$ for all $\boldsymbol{v}$.)

Proof of Corollary 4.2. Set $C = 1 + \sup_{\mathrm{S}_1^{d-1}} \|\boldsymbol{\beta}\|$. Generalizing (4.2) we note that when ρ is sufficiently small and T is sufficiently large, we have

$$\begin{aligned}\lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} \,:\, \mathscr{N}_{0,T+C\rho}(\rho, \boldsymbol{v}, \boldsymbol{\beta}(\boldsymbol{v})) = 0\}) \\ \leq \lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} \,:\, \tau_1(\boldsymbol{q} + \rho\boldsymbol{\beta}(\boldsymbol{v}), \boldsymbol{v}; \rho) \geq T\}) \\ \leq \lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} \,:\, \mathscr{N}_{0,T-C\rho}(\rho, \boldsymbol{v}, \boldsymbol{\beta}(\boldsymbol{v})) = 0\}),\end{aligned} \tag{4.8}$$

where $\mathscr{N}_{0,T}$ is as defined in (3.11) with affine lattice $\mathscr{L}_{\boldsymbol{\alpha}} = \mathscr{L} - \boldsymbol{q}$ (in (4.8) we used our assumption that if $\boldsymbol{q} \in \mathscr{L}$ then $(\boldsymbol{\beta}(\boldsymbol{v}) + \mathbb{R}_{>0}\boldsymbol{v}) \cap \mathscr{B}_1^d = \varnothing$ for all $\boldsymbol{v} \in \mathrm{S}_1^{d-1}$). In particular, writing $T_1 = \xi\rho^{1-d} + C\rho$ and $\sigma(\rho) = T_1^{\frac{1}{d-1}}\rho$ we have, for any $\rho > 0$ sufficiently small,

$$\begin{aligned}\lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} \,:\, \rho^{d-1}\tau_1(\boldsymbol{q} + \rho\boldsymbol{\beta}(\boldsymbol{v}), \boldsymbol{v}; \rho) \geq \xi\}) \\ \geq \lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} \,:\, \mathscr{N}_{0,T_1}(\sigma(\rho)T_1^{-\frac{1}{d-1}}, \boldsymbol{v}, \boldsymbol{\beta}(\boldsymbol{v})) = 0\}).\end{aligned} \tag{4.9}$$

But $T_1 \to \infty$ and $\sigma(\rho) \to \xi^{1/(d-1)}$ as $\rho \to 0^+$; hence by Theorem 3.4 the right-hand side above tends to $F_{0,\boldsymbol{\alpha},\boldsymbol{\beta}}(0, \xi^{1/(d-1)})$. This equals $\int_\xi^\infty \Phi_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\xi')\,d\xi'$, because of

(4.6) and $\lim_{\sigma\to\infty} F_{0,\boldsymbol{\alpha},\boldsymbol{\beta}}(0,\sigma) = 0$ (see Remark 3.6). Hence we have proved

$$\liminf_{\rho\to 0} \lambda(\{\boldsymbol{v}\in \mathrm{S}_1^{d-1} : \rho^{d-1}\tau_1(\boldsymbol{q}+\rho\boldsymbol{\beta}(\boldsymbol{v}),\boldsymbol{v};\rho)\ge\xi\}) \ge \int_\xi^\infty \Phi_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\xi')\,d\xi'. \tag{4.10}$$

But using the last inequality in (4.8) we obtain the same upper bound for the corresponding lim sup, and hence (4.7) is proved. □

Remark 4.3. When $\mathscr{L}=\mathbb{Z}^2$, $\boldsymbol{q}=\boldsymbol{0}$, $\boldsymbol{\beta}(\boldsymbol{v})=\boldsymbol{v}$ (say) and λ = uniform measure on S_1^1, Corollary 4.2 specializes to the limit result proved in Boca, Gologan and Zaharescu [4]. Similarly for $\mathscr{L}=\mathbb{Z}^2$, Theorem 1.2 (which is basically a $\boldsymbol{q}$-averaged version of Corollary 4.1; cf. also Corollary 9.4 below) specializes to the limit result proved in Boca and Zaharescu [7]. The known explicit formulas for the volumes $F_{0,\boldsymbol{0}}(0,\sigma)$ and $F_{\boldsymbol{0}}(0,\sigma)$ in (3.8) in the case $d=2$ (cf. [36] and Remark 3.3) indeed agree, via (4.3) and (4.6), with the limit formulas obtained in [4] and [7] using methods of analytic number theory.

Analogous results are valid for nonspherical scatterers, as direct corollaries of Theorem 3.9.

4.2. *Location of the first collision.* The position of the particle when hitting the first scatterer is

$$\boldsymbol{q}_1(\boldsymbol{q},\boldsymbol{v};\rho) := \boldsymbol{q}+\tau_1(\boldsymbol{q},\boldsymbol{v};\rho)\boldsymbol{v}. \tag{4.11}$$

We are now interested in the joint distribution of the free path length (considered in the previous section), and the precise location *on* the scatterer where the particle hits.

By definition there is a unique $\boldsymbol{m}\in\mathscr{L}$ such that $\boldsymbol{q}_1(\boldsymbol{q},\boldsymbol{v};\rho)\in \mathrm{S}_\rho^{d-1}+\boldsymbol{m}$; hence there is a unique point $\boldsymbol{w}_1=\boldsymbol{w}_1(\boldsymbol{q},\boldsymbol{v};\rho)\in\mathrm{S}_1^{d-1}$ such that $\boldsymbol{q}_1(\boldsymbol{q},\boldsymbol{v};\rho)=\rho\boldsymbol{w}_1+\boldsymbol{m}$. Let us fix a map $K:\mathrm{S}_1^{d-1}\to\mathrm{SO}(d)$ such that $\boldsymbol{v}K(\boldsymbol{v})=\boldsymbol{e}_1$ for all $\boldsymbol{v}\in\mathrm{S}_1^{d-1}$; we assume that K is smooth when restricted to S_1^{d-1} minus one point.[3] It is evident that $-\boldsymbol{w}_1K(\boldsymbol{v})\in {\mathrm{S}'_1}^{d-1}$, with the hemisphere

$${\mathrm{S}'_1}^{d-1} = \{\boldsymbol{v}=(v_1,\dots,v_d)\in\mathrm{S}_1^{d-1} : v_1>0\}.$$

Recall that we are assuming that $\boldsymbol{\beta}$ is a continuous function $\mathrm{S}_1^{d-1}\to\mathbb{R}^d$ such that if $\boldsymbol{q}\in\mathscr{L}$ then $(\boldsymbol{\beta}(\boldsymbol{v})+\mathbb{R}_{>0}\boldsymbol{v})\cap\mathscr{B}_1^d=\varnothing$ for all $\boldsymbol{v}\in\mathrm{S}_1^{d-1}$. We will use the shorthand $\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{v})=\boldsymbol{q}+\rho\boldsymbol{\beta}(\boldsymbol{v})$ for the initial position. For the statement of the theorem below, we define the following submanifolds of X_q and X, respectively:

(4.12)

$$\begin{aligned} X_q(\boldsymbol{y}) &:= \{M\in X_q : \boldsymbol{y}\in(\mathbb{Z}^d+\boldsymbol{\alpha})M\} && (\text{for } \boldsymbol{y}\in\mathbb{R}^d\setminus\{\boldsymbol{0}\} \text{ and fixed } \boldsymbol{\alpha}\in q^{-1}\mathbb{Z}^d);\\ X(\boldsymbol{y}) &:= \{g\in X : \boldsymbol{y}\in\mathbb{Z}^d g\} && (\text{for } \boldsymbol{y}\in\mathbb{R}^d). \end{aligned}$$

[3] For example, we may choose K as $K(\boldsymbol{e}_1)=I$, $K(-\boldsymbol{e}_1)=-I$ and $K(\boldsymbol{v})=E\Big(-\frac{2\arcsin\left(\|\boldsymbol{v}-\boldsymbol{e}_1\|/2\right)}{\|\boldsymbol{v}_\perp\|}\boldsymbol{v}_\perp\Big)$ for $\boldsymbol{v}\in\mathrm{S}_1^{d-1}\setminus\{\boldsymbol{e}_1,-\boldsymbol{e}_1\}$, where $\boldsymbol{v}_\perp:=(v_2,\dots,v_d)\in\mathbb{R}^{d-1}$ and $E(\boldsymbol{w})=\exp\left(\begin{smallmatrix}0 & \boldsymbol{w}\\ -{}^t\boldsymbol{w} & 0\end{smallmatrix}\right)\in\mathrm{SO}(d)$. Then K is smooth when restricted to $\mathrm{S}_1^{d-1}\setminus\{-\boldsymbol{e}_1\}$.

These submanifolds will be studied in Section 7, where we will introduce a natural Borel probability measure $\nu_{\boldsymbol{y}}$ on each of them.

We will also use the notation $\boldsymbol{x}_\perp = \boldsymbol{x} - (\boldsymbol{x}\cdot\boldsymbol{e}_1)\boldsymbol{e}_1$ for $\boldsymbol{x}\in\mathbb{R}^d$.

THEOREM 4.4. *Fix a lattice $\mathcal{L} = \mathbb{Z}^d M_0$. Let $\boldsymbol{q}\in\mathbb{R}^d$ and $\boldsymbol{\alpha} = -\boldsymbol{q}M_0^{-1}$. There exists a function $\Phi_{\boldsymbol{\alpha}} : \mathbb{R}_{>0}\times(\{0\}\times\mathcal{B}_1^{d-1})\times(\{0\}\times\mathbb{R}^{d-1})\to\mathbb{R}_{\geq 0}$ such that for any Borel probability measure λ on S_1^{d-1} absolutely continuous with respect to $\mathrm{vol}_{\mathrm{S}_1^{d-1}}$, any subset $\mathfrak{U}\subset {\mathrm{S}'_1}^{d-1}$ with $\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\partial\mathfrak{U}) = 0$, and any $0\leq\xi_1<\xi_2$, we have*

(4.13)
$$\lim_{\rho\to 0}\lambda\big(\big\{\boldsymbol{v}\in\mathrm{S}_1^{d-1}:\rho^{d-1}\tau_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{v}),\boldsymbol{v};\rho)\in[\xi_1,\xi_2),\,-\boldsymbol{w}_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{v}),\boldsymbol{v};\rho)K(\boldsymbol{v})\in\mathfrak{U}\big\}\big)$$
$$=\int_{\xi_1}^{\xi_2}\int_{\mathfrak{U}_\perp}\int_{\mathrm{S}_1^{d-1}}\Phi_{\boldsymbol{\alpha}}\big(\xi,\boldsymbol{w},(\boldsymbol{\beta}(\boldsymbol{v})K(\boldsymbol{v}))_\perp\big)\,d\lambda(\boldsymbol{v})d\boldsymbol{w}\,d\xi,$$

where $d\boldsymbol{w}$ denotes the $(d-1)$-dimensional Lebesgue volume measure on $\{0\}\times\mathbb{R}^{d-1}$. The function $\Phi_{\boldsymbol{\alpha}}$ is explicitly given by

(4.14)
$$\Phi_{\boldsymbol{\alpha}}(\xi,\boldsymbol{w},\boldsymbol{z})$$
$$=\begin{cases}\nu_{\boldsymbol{y}}\big(\{M\in X_q(\boldsymbol{y}):(\mathbb{Z}^d+\boldsymbol{\alpha})M\cap(\mathfrak{Z}(0,\xi,1)+\boldsymbol{z})=\varnothing\}\big) & \text{if } \boldsymbol{\alpha}\in q^{-1}\mathbb{Z}^d\\ \nu_{\boldsymbol{y}}\big(\{g\in X(\boldsymbol{y}):\mathbb{Z}^d g\cap(\mathfrak{Z}(0,\xi,1)+\boldsymbol{z})=\varnothing\}\big) & \text{if } \boldsymbol{\alpha}\notin\mathbb{Q}^d,\end{cases}$$

where $\boldsymbol{y} = \xi\boldsymbol{e}_1+\boldsymbol{w}+\boldsymbol{z}$, and

$$\mathfrak{Z}(c_1,c_2,\sigma) = \big\{(x_1,\dots,x_d)\in\mathbb{R}^d : c_1<x_1<c_2,\,\|(x_2,\dots,x_d)\|<\sigma\big\}. \tag{4.15}$$

Remark 4.5. Note that $\Phi_{\boldsymbol{\alpha}}(\xi,\boldsymbol{w},\boldsymbol{z})$ is independent of $\boldsymbol{\beta}$. For $\boldsymbol{\alpha}\in\mathbb{Q}^d$ the function $\Phi_{\boldsymbol{\alpha}}(\xi,\boldsymbol{w},\boldsymbol{z})$ is Borel measurable, and in fact only depends on ($\boldsymbol{\alpha}$ and) the four real numbers $\xi,\|\boldsymbol{z}\|,\|\boldsymbol{w}\|,\boldsymbol{z}\cdot\boldsymbol{w}$. Also for $\boldsymbol{\alpha}\in\mathbb{Q}^d$, if we restrict to $\|\boldsymbol{z}\|\leq 1$ [and if $d=2$: $\boldsymbol{z}+\boldsymbol{w}\neq\boldsymbol{0}$], then $\Phi_{\boldsymbol{\alpha}}(\xi,\boldsymbol{w},\boldsymbol{z})$ is jointly continuous in the three variables $\xi,\boldsymbol{w},\boldsymbol{z}$. If $\boldsymbol{\alpha}\notin\mathbb{Q}^d$ then $\Phi_{\boldsymbol{\alpha}}(\xi,\boldsymbol{w},\boldsymbol{z})$ is everywhere continuous in the three variables, and it is independent of both $\boldsymbol{\alpha}$ and $\boldsymbol{z}$; in fact it only depends on ξ and $\|\boldsymbol{w}\|$. All these statements will be proved in Sections 8.1 and 8.2. In particular, if $\boldsymbol{\alpha}\notin\mathbb{Q}^d$ then the limit in (4.13) is independent of $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, λ.

Remark 4.6. It follows from (4.13) that

$$\int_0^\infty\int_{\{0\}\times\mathcal{B}_1^{d-1}}\Phi_{\boldsymbol{\alpha}}(\xi,\boldsymbol{w},\boldsymbol{z})\,d\boldsymbol{w}\,d\xi = 1 \tag{4.16}$$

holds for almost all $\boldsymbol{z}\in\{0\}\times\mathbb{R}^{d-1}$, and from (4.13) and Corollary 4.2 that

$$\int_{\{0\}\times\mathcal{B}_1^{d-1}}\int_{\mathrm{S}_1^{d-1}}\Phi_{\boldsymbol{\alpha}}\big(\xi,\boldsymbol{w},(\boldsymbol{\beta}(\boldsymbol{v})K(\boldsymbol{v}))_\perp\big)\,d\lambda(\boldsymbol{v})d\boldsymbol{w} = \Phi_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\xi) \tag{4.17}$$

holds for almost all $\xi > 0$. As a consistency check we derive in Section 8.3 (see Remark 8.12) the relations (4.16) and (4.17) directly from the explicit formula (4.14). In fact it turns out that (4.16) holds for *all* $\boldsymbol{z} \in \{0\} \times \mathbb{R}^{d-1}$ and (4.17) holds for *all* $\xi > 0$.

As a preparation for Theorem 4.8 below and for the results in [23], we also state a version of Theorem 4.4 involving an arbitrary continuous test function.

COROLLARY 4.7. *Let λ be a Borel probability measure on S_1^{d-1} absolutely continuous with respect to $\mathrm{vol}_{\mathrm{S}_1^{d-1}}$. For any bounded continuous function $f : \mathrm{S}_1^{d-1} \times \mathbb{R}_{>0} \times \mathrm{S}_1^{d-1} \to \mathbb{R}$,*

$$\begin{aligned}(4.18)\quad \lim_{\rho\to 0} \int_{\mathrm{S}_1^{d-1}} & f\big(\boldsymbol{v}, \rho^{d-1}\tau_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{v}), \boldsymbol{v}; \rho), \boldsymbol{w}_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{v}), \boldsymbol{v}; \rho)\big) d\lambda(\boldsymbol{v}) \\ &= \int_{\mathrm{S}_1'^{\,d-1}} \int_{\mathbb{R}_{>0}} \int_{S_1^{d-1}} f\big(\boldsymbol{v}, \xi, -\boldsymbol{\omega} K(\boldsymbol{v})^{-1}\big) \\ &\quad \cdot \Phi_{\boldsymbol{\alpha}}\big(\xi, \boldsymbol{\omega}_\perp, (\boldsymbol{\beta}(\boldsymbol{v})K(\boldsymbol{v}))_\perp\big)\, \omega_1\, d\lambda(\boldsymbol{v})\, d\xi\, d\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\boldsymbol{\omega}),\end{aligned}$$

where $\boldsymbol{\omega} = (\omega_1, \dots, \omega_d)$.

Proof. For f with compact support the result follows in a standard way by approximating f from above and below by linear combinations of characteristic functions and applying Theorem 4.4. When extending to arbitrary bounded continuous functions f one uses (4.17), (4.6) and Remark 3.6. □

4.3. *Velocity after the first collision.* If a particle moving with velocity $\boldsymbol{v}_0$ hits a spherical scatterer at the point $\boldsymbol{q}_1$ and is elastically reflected, its velocity changes to

$$\boldsymbol{v}_1 = \boldsymbol{v}_0 - 2(\boldsymbol{v}_0 \cdot \boldsymbol{w}_1)\boldsymbol{w}_1, \tag{4.19}$$

where $\boldsymbol{w}_1 \in \mathrm{S}_1^{d-1}$ is the location of the hit relative to the center of the sphere, as defined in Section 4.2. This implies

$$\boldsymbol{w}_1 = \frac{\boldsymbol{v}_1 - \boldsymbol{v}_0}{\|\boldsymbol{v}_1 - \boldsymbol{v}_0\|}. \tag{4.20}$$

THEOREM 4.8. *Let λ be a Borel probability measure on S_1^{d-1} absolutely continuous with respect to $\mathrm{vol}_{\mathrm{S}_1^{d-1}}$. For any bounded continuous function $f : \mathrm{S}_1^{d-1} \times \mathbb{R}_{>0} \times \mathrm{S}_1^{d-1} \to \mathbb{R}$,*

$$\begin{aligned}(4.21)\quad \lim_{\rho\to 0} \int_{\mathrm{S}_1^{d-1}} & f\big(\boldsymbol{v}_0, \rho^{d-1}\tau_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{v}_0), \boldsymbol{v}_0; \rho), \boldsymbol{v}_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{v}_0), \boldsymbol{v}_0; \rho)\big) d\lambda(\boldsymbol{v}_0) \\ &= \int_{\mathrm{S}_1^{d-1}} \int_{\mathbb{R}_{>0}} \int_{\mathrm{S}_1^{d-1}} f\big(\boldsymbol{v}_0, \xi, \boldsymbol{v}_1\big)\, p_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{v}_0, \xi, \boldsymbol{v}_1)\, d\lambda(\boldsymbol{v}_0)\, d\xi\, d\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\boldsymbol{v}_1),\end{aligned}$$

with the probability density $p_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ *defined by*

(4.22)
$$p_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{v}_0,\xi,\boldsymbol{v}_1)\,d\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\boldsymbol{v}_1)=\Phi_{\boldsymbol{\alpha}}\big(\xi,\boldsymbol{\omega}_\perp,(\boldsymbol{\beta}(\boldsymbol{v}_0)K(\boldsymbol{v}_0))_\perp\big)\,\omega_1\,d\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\boldsymbol{\omega})$$

where

$$\boldsymbol{v}_1=(\boldsymbol{e}_1-2(\boldsymbol{e}_1\cdot\boldsymbol{\omega})\boldsymbol{\omega})K(\boldsymbol{v}_0)^{-1},\qquad \boldsymbol{\omega}\in {\mathrm{S}_1'}^{d-1}. \tag{4.23}$$

Remark 4.9. The relationship between $p_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{v}_0,\xi,\boldsymbol{v}_1)$ and $\Phi_{\boldsymbol{\alpha}}(\xi,\boldsymbol{w},\boldsymbol{z})$ can be expressed more explicitly as

(4.24)
$$p_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{v}_0,\xi,\boldsymbol{v}_1)=\frac{1}{4}\,\|\boldsymbol{v}_1-\boldsymbol{v}_0\|^{3-d}\,\Phi_{\boldsymbol{\alpha}}\Big(\xi,-\frac{(\boldsymbol{v}_1K(\boldsymbol{v}_0))_\perp}{\|\boldsymbol{v}_1-\boldsymbol{v}_0\|},(\boldsymbol{\beta}(\boldsymbol{v}_0)K(\boldsymbol{v}_0))_\perp\Big).$$

The function $p_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{v}_0,\xi,\boldsymbol{v}_1)$ is independent of the choice of the function $K:\mathrm{S}_1^{d-1}\to\mathrm{SO}(d)$, since $\Phi_{\boldsymbol{\alpha}}(\xi,\boldsymbol{w},\boldsymbol{z})$ only depends on the four real numbers (cf. Remark 4.5) ξ, $\|\boldsymbol{w}\|$, $\|\boldsymbol{z}\|$, $\boldsymbol{w}\cdot\boldsymbol{z}$, which in (4.24) can be expressed as ξ, $\frac{\sqrt{1-(\boldsymbol{v}_0\cdot\boldsymbol{v}_1)^2}}{\|\boldsymbol{v}_1-\boldsymbol{v}_0\|}$, $\sqrt{1-(\boldsymbol{\beta}(\boldsymbol{v}_0)\cdot\boldsymbol{v}_0)^2}$, $\frac{(\boldsymbol{v}_1\cdot\boldsymbol{v}_0)(\boldsymbol{\beta}(\boldsymbol{v}_0)\cdot\boldsymbol{v}_0)-\boldsymbol{v}_1\cdot\boldsymbol{\beta}(\boldsymbol{v}_0)}{\|\boldsymbol{v}_1-\boldsymbol{v}_0\|}$, respectively.

5. **Equidistribution in homogeneous spaces**

This section provides the ergodic-theoretic results, which are the key ingredients in the proofs of the main theorems. These equidistribution theorems are consequences of Ratner's classification of measures that are invariant under the action of a unipotent flow [30], and may in particular be viewed as variants of Shah's Theorem 1.4 in [31].

5.1. *Translates of expanding unipotent orbits.* The following is a special case of Shah's Theorem 1.4 in [31]. Let G be a connected Lie group and let Γ be a lattice in G.

THEOREM 5.1. *Suppose G contains a Lie subgroup H isomorphic to* $\mathrm{SL}(d,\mathbb{R})$ (*we denote the corresponding embedding by* $\varphi:\mathrm{SL}(d,\mathbb{R})\to G$), *such that the set* $\Gamma\backslash\Gamma H$ *is dense in* $\Gamma\backslash G$. *Let* λ *be a Borel probability measure on* $\mathbb{R}^{d-1}$ *which is absolutely continuous with respect to Lebesgue measure, and let* $f:\Gamma\backslash G\to\mathbb{R}$ *be bounded continuous. Then*

(5.1)
$$\lim_{t\to\infty}\int_{\mathbb{R}^{d-1}} f\left(\varphi\left(\begin{pmatrix}1 & \boldsymbol{x}\\ {}^{\mathrm{t}}\mathbf{0} & 1_{d-1}\end{pmatrix}\begin{pmatrix}\mathrm{e}^{-(d-1)t} & \mathbf{0}\\ {}^{\mathrm{t}}\mathbf{0} & \mathrm{e}^{t}1_{d-1}\end{pmatrix}\right)\right)d\lambda(\boldsymbol{x})=\int_{\Gamma\backslash G} f\,d\mu,$$

where μ *is the unique G-right-invariant probability measure on* $\Gamma\backslash G$.

Let us set

$$n_-(\boldsymbol{x})=\left(\begin{pmatrix}1 & \boldsymbol{x}\\ {}^{\mathrm{t}}\mathbf{0} & 1_{d-1}\end{pmatrix},\mathbf{0}\right)\in\mathrm{ASL}(d,\mathbb{R}) \tag{5.2}$$

and

$$\Phi^t = \left(\begin{pmatrix} \mathrm{e}^{-(d-1)t} & \mathbf{0} \\ {}^{\mathrm{t}}\mathbf{0} & \mathrm{e}^{t} 1_{d-1} \end{pmatrix}, \mathbf{0} \right) \in \mathrm{ASL}(d, \mathbb{R}). \tag{5.3}$$

Theorem 5.1 implies the following.

THEOREM 5.2. *Let λ be a Borel probability measure on $\mathbb{R}^{d-1}$ which is absolutely continuous with respect to Lebesgue measure, and let $f : X \to \mathbb{R}$ be bounded continuous. Then, for every $\boldsymbol{\alpha} \in \mathbb{R}^d \setminus \mathbb{Q}^d$ and every $M \in \mathrm{SL}(d, \mathbb{R})$,*

$$\lim_{t\to\infty} \int_{\mathbb{R}^{d-1}} f\left((1_d, \boldsymbol{\alpha})(M, \mathbf{0}) n_-(\boldsymbol{x}) \Phi^t\right) d\lambda(\boldsymbol{x}) = \int_X f(g)\, d\mu(g). \tag{5.4}$$

Proof. Let $G = \mathrm{ASL}(d, \mathbb{R})$, $\Gamma = \mathrm{ASL}(d, \mathbb{Z})$ and define the embedding

$$\varphi : \mathrm{SL}(d, \mathbb{R}) \to G, \quad \widetilde{M} \mapsto (1_d, \boldsymbol{\alpha})(M \widetilde{M} M^{-1}, \mathbf{0})(1_d, -\boldsymbol{\alpha}). \tag{5.5}$$

We now wish to establish that $\Gamma\backslash\Gamma H$ with $H = \varphi(\mathrm{SL}(d, \mathbb{R}))$ is dense in $\Gamma\backslash G$. To this end it suffices to show that

$$(\gamma, \boldsymbol{m})(1_d, \boldsymbol{\alpha})(M \widetilde{M}, \mathbf{0}) = (\gamma M \widetilde{M}, (\boldsymbol{\alpha} + \boldsymbol{m}) M \widetilde{M}) \tag{5.6}$$

are dense in $\mathrm{ASL}(d, \mathbb{R})$, as γ, $\boldsymbol{m}$ and $\widetilde{M}$ vary over $\mathrm{SL}(d, \mathbb{Z})$, $\mathbb{Z}^d$ and $\mathrm{SL}(d, \mathbb{R})$, respectively. It is evident that this in turn is equivalent to showing that $\{(\boldsymbol{\alpha} + \boldsymbol{m})\gamma^{-1}\}$ is dense in $\mathbb{R}^d$.

Letting $C \subset \mathbb{R}^d/\mathbb{Z}^d$ be the closure of the image of $\boldsymbol{\alpha}\, \mathrm{SL}(d, \mathbb{Z}) \subset \mathbb{R}^d$ under the natural projection $\mathbb{R}^d \to \mathbb{R}^d/\mathbb{Z}^d$, our task is to show $C = \mathbb{R}^d/\mathbb{Z}^d$. Since $\boldsymbol{\alpha} \notin \mathbb{Q}^d$ there is a choice of $\gamma \in \mathrm{SL}(d, \mathbb{Z})$ either a permutation matrix or $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ which gives $\boldsymbol{w} = (w_1, \ldots, w_d) := \boldsymbol{\alpha}\gamma \in C$ with $w_1 \notin \mathbb{Q}$. Then by choosing $\gamma' = \left(\begin{smallmatrix} 1 & \boldsymbol{a} \\ {}^{\mathrm{t}}\mathbf{0} & 1_{d-1} \end{smallmatrix}\right) \in \mathrm{SL}(d, \mathbb{Z})$ with appropriate $\boldsymbol{a} \in \mathbb{Z}^{d-1}$, the point $\boldsymbol{w}\gamma'$ can be made to lie arbitrarily close to $(w_1, 0, \ldots, 0)$ in $\mathbb{R}^d/\mathbb{Z}^d$. Hence since C is closed we have $(w_1, 0, \ldots, 0) \in C$. Now let $\boldsymbol{y} = (y_1, \ldots, y_d) \in \mathbb{R}^d$ and $\varepsilon > 0$ be given. Then there is $m \in \mathbb{Z}\setminus\{0\}$ such that $\|mw_1 - y_1\| < \varepsilon$ (where $\|x\| = \inf_{n\in\mathbb{Z}} |x-n|$ as usual). Letting γ'' be any matrix in $\mathrm{SL}(d, \mathbb{Z})$ with top left entry m we have $(mw_1, *, \ldots, *) = (w_1, 0, \ldots, 0)\gamma'' \in C$, and hence since C is right $\mathrm{SL}(d, \mathbb{Z})$ invariant and $mw_1 \notin \mathbb{Q}$, an argument as above shows $(mw_1, 0, \ldots, 0) \in C$. Finally by choosing (again) $\gamma''' = \left(\begin{smallmatrix} 1 & \boldsymbol{a} \\ {}^{\mathrm{t}}\mathbf{0} & 1_{d-1} \end{smallmatrix}\right) \in \mathrm{SL}(d, \mathbb{Z})$ with appropriate $\boldsymbol{a} \in \mathbb{Z}^{d-1}$, the point $(mw_1, 0, \ldots, 0)\gamma''' \in C$ can be made to lie arbitrarily close to $(mw_1, y_2, \ldots, y_d)$. Since ε is arbitrary and C is closed we obtain $\boldsymbol{y} \in C$. Hence $C = \mathbb{R}^d/\mathbb{Z}^d$, as desired.

Having established the required density, Theorem 5.1 implies that for any bounded continuous $\tilde{f} : X \to \mathbb{R}$

$$\lim_{t\to\infty} \int_{\mathbb{R}^{d-1}} \tilde{f}((M, \boldsymbol{\alpha}M) n_-(\boldsymbol{x}) \Phi^t (M, \boldsymbol{\alpha}M)^{-1}) d\lambda(\boldsymbol{x}) = \int_X \tilde{f}(g) d\mu(g). \tag{5.7}$$

Choosing the test function $\tilde{f}(g) = f(g(M, \boldsymbol{\alpha}M))$ completes the proof. □

We now extend Theorem 5.2 by considering sequences of test functions with additional parameter dependence.

THEOREM 5.3. *Let λ be a Borel probability measure on $\mathbb{R}^{d-1}$ which is absolutely continuous with respect to Lebesgue measure. Let $f : \mathbb{R}^{d-1} \times X \to \mathbb{R}$ be bounded continuous and $f_t : \mathbb{R}^{d-1} \times X \to \mathbb{R}$ a family of uniformly bounded (i.e., $|f_t| < K$ for some absolute constant K), continuous functions such that $f_t \to f$ as $t \to \infty$, uniformly on compacta. Then, for every $\boldsymbol{\alpha} \in \mathbb{R}^d \setminus \mathbb{Q}^d$, $M \in \mathrm{SL}(d, \mathbb{R})$,*

$$\lim_{t\to\infty} \int_{\mathbb{R}^{d-1}} f_t\big(\boldsymbol{x}, (1_d, \boldsymbol{\alpha})(M, \mathbf{0})n_-(\boldsymbol{x})\Phi^t\big)\, d\lambda(\boldsymbol{x}) = \int_{\mathbb{R}^{d-1}\times X} f(\boldsymbol{x}, g)\, d\mu(g) d\lambda(\boldsymbol{x}). \tag{5.8}$$

Proof. Let us first assume that f_t and f have support in the compact set $\mathcal{K} \subset \mathbb{R}^{d-1} \times X$. Hence the convergence $f_t \to f$ is uniform and all functions are uniformly continuous. Therefore, given $\delta > 0$ there exist $\epsilon > 0, t_0 > 0$ such that

$$f(\boldsymbol{x}_0, g) - \delta \le f(\boldsymbol{x}, g) \le f(\boldsymbol{x}_0, g) + \delta \tag{5.9}$$

and

$$f(\boldsymbol{x}_0, g) - \delta \le f_t(\boldsymbol{x}, g) \le f(\boldsymbol{x}_0, g) + \delta \tag{5.10}$$

for all $\boldsymbol{x} \in \boldsymbol{x}_0 + [0, \epsilon)^{d-1}$, $t > t_0$. Now

$$\begin{aligned} &\int_{\mathbb{R}^{d-1}} f_t(\boldsymbol{x}, (1_d, \boldsymbol{\alpha})(M, \mathbf{0})n_-(\boldsymbol{x})\Phi^t) d\lambda(\boldsymbol{x}) \\ &\quad = \sum_{\boldsymbol{k}\in\mathbb{Z}^{d-1}} \int_{\epsilon\boldsymbol{k}+[0,\epsilon)^{d-1}} f_t(\boldsymbol{x}, (1_d, \boldsymbol{\alpha})(M, \mathbf{0})n_-(\boldsymbol{x})\Phi^t) d\lambda(\boldsymbol{x}) \\ &\quad \le \sum_{\boldsymbol{k}\in\mathbb{Z}^{d-1}} \int_{\epsilon\boldsymbol{k}+[0,\epsilon)^{d-1}} f(\epsilon\boldsymbol{k}, (1_d, \boldsymbol{\alpha})(M, \mathbf{0})n_-(\boldsymbol{x})\Phi^t) d\lambda(\boldsymbol{x}) + \delta. \end{aligned} \tag{5.11}$$

By Theorem 5.2,

$$\begin{aligned} &\lim_{t\to\infty} \int_{\epsilon\boldsymbol{k}+[0,\epsilon)^{d-1}} f(\epsilon\boldsymbol{k}, (1_d, \boldsymbol{\alpha})(M, \mathbf{0})n_-(\boldsymbol{x})\Phi^t) d\lambda(\boldsymbol{x}) \\ &\quad = \int_X f(\epsilon\boldsymbol{k}, g) d\mu(g) \int_{\epsilon\boldsymbol{k}+[0,\epsilon)^{d-1}} d\lambda(\boldsymbol{x}) \\ &\quad \le \int_X \int_{\epsilon\boldsymbol{k}+[0,\epsilon)^{d-1}} [f(\boldsymbol{x}, g) + \delta] d\lambda(\boldsymbol{x}) d\mu(g), \end{aligned} \tag{5.12}$$

and so

$$\limsup_{t\to\infty} \int_{\mathbb{R}^{d-1}} f_t(\boldsymbol{x}, (1_d, \boldsymbol{\alpha})(M, \mathbf{0})n_-(\boldsymbol{x})\Phi^t) d\lambda(\boldsymbol{x}) \le \int_X \int_{\mathbb{R}^{d-1}} f(\boldsymbol{x}, g) d\lambda(\boldsymbol{x}) d\mu(g) + 2\delta. \tag{5.13}$$

An analogous argument shows

$$\liminf_{t\to\infty}\int_{\mathbb{R}^{d-1}} f_t(\boldsymbol{x},(1_d,\boldsymbol{\alpha})(M,\mathbf{0})n_-(\boldsymbol{x})\Phi^t)d\lambda(\boldsymbol{x}) \geq \int_X\int_{\mathbb{R}^{d-1}} f(\boldsymbol{x},g)d\lambda(\boldsymbol{x})d\mu(g)-2\delta. \tag{5.14}$$

It therefore follows that the limit exists and

$$\lim_{t\to\infty}\int_{\mathbb{R}^{d-1}} f_t(\boldsymbol{x},(1_d,\boldsymbol{\alpha})(M,\mathbf{0})n_-(\boldsymbol{x})\Phi^t)d\lambda(\boldsymbol{x}) = \int_X\int_{\mathbb{R}^{d-1}} f(\boldsymbol{x},g)d\lambda(\boldsymbol{x})d\mu(g). \tag{5.15}$$

We now extend the result to bounded continuous test functions f_t, uniformly bounded by $|f_t|<K$. Given $\delta>0$ we choose compact sets $\mathcal{K}_1\subset\mathbb{R}^{d-1}$ and $\mathcal{K}_2\subset X$ so large that $(1-\lambda(\mathcal{K}_1))+(1-\mu(\mathcal{K}_2))\leq\delta/K$. Let $c_1:\mathbb{R}^{d-1}\to[0,1]$ and $c_2:X\to[0,1]$ be continuous functions which have compact support and satisfy $\chi_{\mathcal{K}_1}\leq c_1$ and $\chi_{\mathcal{K}_2}\leq c_2$, respectively. Write

$$f_t=f_t^1+f_t^2,\qquad \text{with}\quad f_t^1(\boldsymbol{x},g)=c_1(\boldsymbol{x})c_2(g)\,f_t(\boldsymbol{x},g),\quad f_t^2=f_t-f_t^1. \tag{5.16}$$

Then f_t^1 is compactly supported as in the previous paragraph. For f_t^2 we have, using Theorem 5.2,

$$\begin{aligned}\limsup_{t\to\infty}\int_{\mathbb{R}^{d-1}}\big|f_t^2(\boldsymbol{x},(1_d,\boldsymbol{\alpha})(M,\mathbf{0})n_-(\boldsymbol{x})\Phi^t)\big|\,d\lambda(\boldsymbol{x}) &\leq K(1-\lambda(\mathcal{K}_1))+\limsup_{t\to\infty}\int_{\mathcal{K}_1}K\big(1-c_2\big((1_d,\boldsymbol{\alpha})(M,\mathbf{0})n_-(\boldsymbol{x})\Phi^t\big)\big)\,d\lambda(\boldsymbol{x})\\ &= K(1-\lambda(\mathcal{K}_1))+K\int_X\big(1-c_2(g)\big)\,d\mu(g)\\ &\leq K\big(1-\lambda(\mathcal{K}_1)\big)+K\big(1-\mu(\mathcal{K}_2)\big)\leq\delta.\end{aligned} \tag{5.17}$$

This upper bound shows that the statement of the theorem can be extended from compactly supported to bounded test functions. □

5.2. *Spherical averages.* We will now show that the statement of Theorem 5.3 (and thus of Theorem 5.2) holds when $n_-(\boldsymbol{x})$ is replaced by

$$(E(\boldsymbol{x}),\mathbf{0})=\left(\exp\begin{pmatrix}0 & \boldsymbol{x}\\ -{}^{\mathrm{t}}\boldsymbol{x} & 0_{d-1}\end{pmatrix},\mathbf{0}\right). \tag{5.18}$$

In fact we can prove a more general fact with almost no extra effort:

COROLLARY 5.4. *Let $D\subset\mathbb{R}^{d-1}$ be an open subset and let $E_1:D\to\mathrm{SO}(d)$ be a smooth map such that the map $D\ni\boldsymbol{x}\mapsto\boldsymbol{e}_1E_1(\boldsymbol{x})^{-1}\in\mathrm{S}_1^{d-1}$ has nonsingular differential at (Lebesgue-)almost all $\boldsymbol{x}\in D$. Let λ be a Borel probability measure on D, absolutely continuous with respect to Lebesgue measure. We then have, for any bounded continuous function $f:D\times X\to\mathbb{R}$ and any family of uniformly*

bounded continuous functions $f_t : D \times X \to \mathbb{R}$ *such that* $f_t \to f$ *as* $t \to \infty$, *uniformly on compacta, and for every* $\boldsymbol{\alpha} \in \mathbb{R}^d \setminus \mathbb{Q}^d$, $M \in \mathrm{SL}(d, \mathbb{R})$,

(5.19)
$$\lim_{t\to\infty} \int_D f_t(\boldsymbol{x}, (1_d, \boldsymbol{\alpha})(M, \mathbf{0})(E_1(\boldsymbol{x}), \mathbf{0})\Phi^t)\, d\lambda(\boldsymbol{x}) = \int_{D\times X} f(\boldsymbol{x}, g)\, d\mu(g) d\lambda(\boldsymbol{x}).$$

Remark 5.5. Taking $E_1(\boldsymbol{x}) = E(\boldsymbol{x})$ as in (5.18) is indeed a valid choice in Corollary 5.4, for note that $\boldsymbol{e}_1 E(\boldsymbol{x})^{-1} = \big(\cos\|\boldsymbol{x}\|, -\frac{\sin\|\boldsymbol{x}\|}{\|\boldsymbol{x}\|}\boldsymbol{x}\big)$, and one checks that this map has nonsingular differential except when $\|\boldsymbol{x}\| \in \{\pi, 2\pi, 3\pi, \dots\}$.

Proof of Corollary 5.4. We first prove that if $\boldsymbol{x}_0 \in D$ is any point where the map $\boldsymbol{x} \mapsto \boldsymbol{e}_1 E_1(\boldsymbol{x})^{-1}$ has nonsingular differential, then there is some open neighborhood $D_0 \subset D$ of $\boldsymbol{x}_0$ such that (5.19) holds when D is replaced by D_0 or by any Borel subset of D_0.

To see this, write $E_0 = E_1(\boldsymbol{x}_0)$ and

(5.20)
$$E_2(\boldsymbol{x}) := E_0^{-1} E_1(\boldsymbol{x}) = \begin{pmatrix} c & \boldsymbol{w} \\ {}^{\mathrm{t}}\boldsymbol{v} & A \end{pmatrix} = \begin{pmatrix} c(\boldsymbol{x}) & \boldsymbol{w}(\boldsymbol{x}) \\ {}^{\mathrm{t}}\boldsymbol{v}(\boldsymbol{x}) & A(\boldsymbol{x}) \end{pmatrix}, \qquad c \in \mathbb{R},\ \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^{d-1}.$$

Then $E_2(\boldsymbol{x}_0) = 1_d$ and thus $c(\boldsymbol{x}_0) = 1$ and $\boldsymbol{v}(\boldsymbol{x}_0) = \mathbf{0}$. Furthermore, the map $\boldsymbol{x} \mapsto (c(\boldsymbol{x}), \boldsymbol{v}(\boldsymbol{x})) \in \mathrm{S}_1^{d-1}$ has nonsingular differential at $\boldsymbol{x} = \boldsymbol{x}_0$, since $(c(\boldsymbol{x}), \boldsymbol{v}(\boldsymbol{x})) = \boldsymbol{e}_1 E_2(\boldsymbol{x})^{-1} = (\boldsymbol{e}_1 E_1(\boldsymbol{x})^{-1}) E_0$, and thus also the map $\boldsymbol{x} \mapsto \tilde{\boldsymbol{x}} := -c(\boldsymbol{x})^{-1}\boldsymbol{v}(\boldsymbol{x}) \in \mathbb{R}^{d-1}$ must have nonsingular differential at $\boldsymbol{x} = \boldsymbol{x}_0$. Hence there exists some bounded open neighborhood D_0' of $\boldsymbol{x}_0$ with $\overline{D_0'} \subset D$ such that $c(\boldsymbol{x}) > 1/2$ for all $\boldsymbol{x} \in D_0'$ and such that $\boldsymbol{x} \mapsto \tilde{\boldsymbol{x}}$ is a diffeomorphism of D_0' onto a bounded open subset $\widetilde{D}_0' \subset \mathbb{R}^{d-1}$. Now for each $\boldsymbol{x} \in D_0'$ we have

(5.21)
$$\Big(E_2(\boldsymbol{x}), \mathbf{0}\Big) = n_-(\tilde{\boldsymbol{x}}) \left(\begin{pmatrix} 1 & -\tilde{\boldsymbol{x}} \\ {}^t\mathbf{0} & 1_{d-1} \end{pmatrix} \begin{pmatrix} c & \boldsymbol{w} \\ {}^{\mathrm{t}}\boldsymbol{v} & A \end{pmatrix}, \mathbf{0} \right) = n_-(\tilde{\boldsymbol{x}}) \left(\begin{pmatrix} c^{-1} & \mathbf{0} \\ {}^{\mathrm{t}}\boldsymbol{v} & A \end{pmatrix}, \mathbf{0} \right),$$

since writing out ${}^{\mathrm{t}}E_2(\boldsymbol{x})E_2(\boldsymbol{x}) = 1_d$ one gets the relations $c\boldsymbol{w} + \boldsymbol{v}A = \mathbf{0}$ and $c^2 + \boldsymbol{v}\,{}^{\mathrm{t}}\boldsymbol{v} = 1$, viz. $\boldsymbol{w} - \tilde{\boldsymbol{x}}A = \mathbf{0}$ and $c - \tilde{\boldsymbol{x}}\,{}^{\mathrm{t}}\boldsymbol{v} = c^{-1}$.

Hence also

$$\Big(E_1(\boldsymbol{x}), \mathbf{0}\Big)\Phi^t = \big(E_0, \mathbf{0}\big) n_-(\tilde{\boldsymbol{x}}) \Phi^t \left(\begin{pmatrix} c(\boldsymbol{x})^{-1} & \mathbf{0} \\ {}^{\mathrm{t}}\boldsymbol{v}(\boldsymbol{x}) e^{-dt} & A(\boldsymbol{x}) \end{pmatrix}, \mathbf{0} \right). \tag{5.22}$$

Now fix D_0 as an open neighborhood of $\boldsymbol{x}_0$ such that $\bar{D}_0 \subset D_0'$, and consider any Borel subset B of D_0. Write $\widetilde{B} \subset \widetilde{D}_0 \subset \widetilde{D}_0'$ for the images of B and D_0 under $\boldsymbol{x} \mapsto \tilde{\boldsymbol{x}}$. Let us assume $\lambda(B) > 0$, and let $\widetilde{\lambda}$ be the measure on $\mathbb{R}^{d-1}$ which corresponds to $\lambda(B)^{-1}\lambda|_B$ under the diffeomorphism $\boldsymbol{x} \mapsto \tilde{\boldsymbol{x}}$; then $\widetilde{\lambda}$ is a Borel probability measure with bounded support and absolutely continuous with respect to Lebesgue measure. Since $\overline{\widetilde{D}}_0 \subset \widetilde{D}_0'$, we may choose a continuous cutoff function $h : \mathbb{R}^{d-1} \to [0, 1]$ such that $\chi_{\widetilde{D}_0} \le h \le \chi_{\widetilde{D}_0'}$.

If f_t and f are given as in the statement of the corollary, we may define a family of continuous functions $\widetilde{f}_t : \mathbb{R}^{d-1} \times X \to \mathbb{R}$ and a continuous function $\widetilde{f} : \mathbb{R}^{d-1} \times X \to \mathbb{R}$ through

$$\widetilde{f}_t(\tilde{\boldsymbol{x}}, g) := h(\tilde{\boldsymbol{x}}) f_t\left(\boldsymbol{x}, g\left(\begin{pmatrix} c(\boldsymbol{x})^{-1} & \mathbf{0} \\ {}^{\mathrm{t}}\boldsymbol{v}(\boldsymbol{x}) e^{-dt} & A(\boldsymbol{x}) \end{pmatrix}, \mathbf{0}\right)\right) \quad \text{if } \tilde{\boldsymbol{x}} \in \widetilde{D}_0'; \tag{5.23}$$

$$\widetilde{f}(\tilde{\boldsymbol{x}}, g) := h(\tilde{\boldsymbol{x}}) f\left(\boldsymbol{x}, g\left(\begin{pmatrix} c(\boldsymbol{x})^{-1} & \mathbf{0} \\ {}^{\mathrm{t}}\mathbf{0} & A(\boldsymbol{x}) \end{pmatrix}, \mathbf{0}\right)\right) \quad \text{if } \tilde{\boldsymbol{x}} \in \widetilde{D}_0';$$

$$\widetilde{f}_t(\tilde{\boldsymbol{x}}, g) = \widetilde{f}(\tilde{\boldsymbol{x}}, g) := 0 \quad \text{if } \tilde{\boldsymbol{x}} \notin \widetilde{D}_0'.$$

(We here view $\boldsymbol{x} \in D_0'$ as a function of $\tilde{\boldsymbol{x}} \in \widetilde{D}_0'$.) We then have $\widetilde{f}_t(\tilde{\boldsymbol{x}}, g) \to \widetilde{f}(\tilde{\boldsymbol{x}}, g)$ as $t \to \infty$, uniformly on compacta. Applying Theorem 5.3 for $\widetilde{\lambda}$, $\widetilde{f}_t$, $\widetilde{f}$, and with M replaced by ME_0, we get

$$\lim_{t\to\infty} \int_{\mathbb{R}^{d-1}} \widetilde{f}_t\Big(\tilde{\boldsymbol{x}}, (1_d, \boldsymbol{\alpha})(ME_0, \mathbf{0}) n_-(\tilde{\boldsymbol{x}}) \Phi^t\Big)\, d\widetilde{\lambda}(\tilde{\boldsymbol{x}}) = \int_{\mathbb{R}^{d-1}\times X} \widetilde{f}(\tilde{\boldsymbol{x}}, g)\, d\mu(g) d\widetilde{\lambda}(\tilde{\boldsymbol{x}}). \tag{5.24}$$

Here the left-hand side equals, using $\widetilde{\lambda} = \widetilde{\lambda}|_{\widetilde{B}}$ and (5.22),

$$\lim_{t\to\infty} \int_{\widetilde{B}} f_t\Big(\boldsymbol{x}, (1_d, \boldsymbol{\alpha})(M, \mathbf{0})\big(E_1(\boldsymbol{x}), \mathbf{0}\big)\Phi^t\Big)\, d\widetilde{\lambda}(\tilde{\boldsymbol{x}}) = \lambda(B)^{-1} \lim_{t\to\infty} \int_B f_t\Big(\boldsymbol{x}, (1_d, \boldsymbol{\alpha})(M, \mathbf{0})\big(E_1(\boldsymbol{x}), \mathbf{0}\big)\Phi^t\Big)\, d\lambda(\boldsymbol{x}), \tag{5.25}$$

and the right-hand side equals (using the right invariance of μ)

$$\int_{\widetilde{B}\times X} f(\boldsymbol{x}, g)\, d\mu(g) d\widetilde{\lambda}(\tilde{\boldsymbol{x}}) = \lambda(B)^{-1} \int_{B\times X} f\ (\boldsymbol{x}, g)\ d\mu(g) d\lambda(\boldsymbol{x}). \tag{5.26}$$

This proves our claim: (5.19) holds when D is replaced by any Borel subset B of D_0. We have proved this under the assumption $\lambda(B) > 0$, but it is trivially true also in the case $\lambda(B) = 0$.

Now the proof of Corollary 5.4 is completed by a simple covering argument: Given $\varepsilon > 0$ there is some compact subset $K \subset D$ such that $\lambda(K) > 1 - \varepsilon$ and the map $D \ni \boldsymbol{x} \mapsto \boldsymbol{e}_1 E_1(\boldsymbol{x})^{-1} \in \mathrm{S}_1^{d-1}$ has nonsingular differential at every $\boldsymbol{x} \in K$. Then by what we have proved and since K is compact, there exists a finite family $D_1, \dots, D_n$ of open subsets of D which cover K and which have the same property as D_0 above. Set $B_1 := D_1 \cap K$ and, recursively, $B_j := (D_j \cap K) \setminus (B_1 \cup \cdots \cup B_{j-1})$ for $j = 2, \dots, n$. Then each B_j is a Borel subset of D_j so that (5.19) holds when D is replaced by B_j. Furthermore, K is the disjoint union of $B_1, \dots, B_n$; hence by adding we obtain that (5.19) holds when D is replaced by K. Using $\lambda(K) > 1 - \varepsilon$ and our assumption that the family f_t is uniformly bounded, we obtain (5.19) upon letting $\varepsilon \to 0$. □

5.3. *Characteristic functions.* We recall the definition of limits of a family of sets $\{\mathscr{E}_t\}_{t\geq t_0}$, where t_0 is a fixed real constant:

$$\liminf \mathscr{E}_t := \bigcup_{t\geq t_0}\bigcap_{s\geq t}\mathscr{E}_s, \qquad \limsup \mathscr{E}_t := \bigcap_{t\geq t_0}\bigcup_{s\geq t}\mathscr{E}_s. \tag{5.27}$$

We will also use the notation

$$\lim(\inf \mathscr{E}_t)^\circ := \bigcup_{t\geq t_0}\left(\bigcap_{s\geq t}\mathscr{E}_s\right)^\circ, \qquad \lim\overline{\sup \mathscr{E}_t} := \bigcap_{t\geq t_0}\overline{\bigcup_{s\geq t}\mathscr{E}_s}. \tag{5.28}$$

Note that $\lim(\inf \mathscr{E}_t)^\circ$ is open and $\lim\overline{\sup \mathscr{E}_t}$ is closed.

If $\{\mathscr{E}_t\}_{t\geq t_0}$ is a decreasing family and $\mathscr{E} = \bigcap_{t\geq t_0}\mathscr{E}_t$ we write $\mathscr{E}_t \downarrow \mathscr{E}$; if $\{\mathscr{E}_t\}_{t\geq t_0}$ is an increasing family and $\mathscr{E} = \bigcup_{t\geq t_0}\mathscr{E}_t$ we write $\mathscr{E}_t \uparrow \mathscr{E}$.

THEOREM 5.6. *Let λ be a Borel probability measure on $\mathbb{R}^{d-1}$ which is absolutely continuous with respect to Lebesgue measure, and let $\mathscr{E}_t$ be a family of subsets of $\mathbb{R}^{d-1}\times X$. Then, for $\boldsymbol{\alpha}\in\mathbb{R}^d\setminus\mathbb{Q}^d$ and $M\in \mathrm{SL}(d,\mathbb{R})$,*

$$\liminf_{t\to\infty}\int_{\mathbb{R}^{d-1}}\chi_{\mathscr{E}_t}(\boldsymbol{x},(1_d,\boldsymbol{\alpha})(M,\mathbf{0})n_-(\boldsymbol{x})\Phi^t)d\lambda(\boldsymbol{x}) \geq \int_{\lim(\inf\mathscr{E}_t)^\circ} d\mu(g)d\lambda(\boldsymbol{x}), \tag{5.29}$$

and

$$\limsup_{t\to\infty}\int_{\mathbb{R}^{d-1}}\chi_{\mathscr{E}_t}(\boldsymbol{x},(1_d,\boldsymbol{\alpha})(M,\mathbf{0})n_-(\boldsymbol{x})\Phi^t)d\lambda(\boldsymbol{x}) \leq \int_{\lim\overline{\sup\mathscr{E}_t}} d\mu(g)d\lambda(\boldsymbol{x}). \tag{5.30}$$

If, furthermore, the set $\lim\overline{\sup\mathscr{E}_t}\setminus\lim(\inf\mathscr{E}_t)^\circ$ has measure zero, then

$$\lim_{t\to\infty}\int_{\mathbb{R}^{d-1}}\chi_{\mathscr{E}_t}(\boldsymbol{x},(1_d,\boldsymbol{\alpha})(M,\mathbf{0})n_-(\boldsymbol{x})\Phi^t)d\lambda(\boldsymbol{x}) = \int_{\limsup\mathscr{E}_t} d\mu(g)d\lambda(\boldsymbol{x}). \tag{5.31}$$

Proof. We begin with the proof of (5.30). Define the closed set

$$\widetilde{\mathscr{E}}_t := \overline{\bigcup_{s\geq t}\mathscr{E}_s}. \tag{5.32}$$

Clearly $\mathscr{E}_t\subset\widetilde{\mathscr{E}}_t\subset\widetilde{\mathscr{E}}_{t_1}$ for $t\geq t_1$. So

$$\begin{aligned}\limsup_{t\to\infty}\int_{\mathbb{R}^{d-1}}\chi_{\mathscr{E}_t}(\boldsymbol{x},(1_d,\boldsymbol{\alpha})(M,\mathbf{0})n_-(\boldsymbol{x})\Phi^t)d\lambda(\boldsymbol{x})\\ \leq \limsup_{t_1\to\infty}\limsup_{t\to\infty}\int_{\mathbb{R}^{d-1}}\chi_{\widetilde{\mathscr{E}}_{t_1}}(\boldsymbol{x},(1_d,\boldsymbol{\alpha})(M,\mathbf{0})n_-(\boldsymbol{x})\Phi^t)d\lambda(\boldsymbol{x}).\end{aligned} \tag{5.33}$$

It follows from Theorem 5.3 (for a constant family of test functions $f = f_t$) by a standard probabilistic argument in which characteristic functions are approximated

by bounded continuous functions f (see e.g., [33, Chap. III]) that

(5.34)
$$\limsup_{t\to\infty}\int_{\mathbb{R}^{d-1}}\chi_{\widetilde{\mathscr{E}}_{t_1}}(\boldsymbol{x},(1_d,\boldsymbol{\alpha})(M,\mathbf{0})n_-(\boldsymbol{x})\Phi^t)\,d\lambda(\boldsymbol{x})\le\int_{\widetilde{\mathscr{E}}_{t_1}}d\mu(g)\,d\lambda(\boldsymbol{x}).$$

Since $\widetilde{\mathscr{E}}_{t_1}\downarrow\lim\overline{\sup\mathscr{E}_t}$,

$$\limsup_{t_1\to\infty}\int_{\widetilde{\mathscr{E}}_{t_1}}d\mu(g)\,d\lambda(\boldsymbol{x})=\int_{\lim\overline{\sup\mathscr{E}_t}}d\mu(g)\,d\lambda(\boldsymbol{x}),\tag{5.35}$$

and (5.30) follows. Relation (5.29) is established by taking complements, and (5.31) then follows from (5.29) and (5.30). □

Remark 5.7. Let $E_1 : D\to\mathrm{SO}(d)$ be any map as in Corollary 5.4; then the assertions of Theorem 5.6 also hold with $n_-(\boldsymbol{x})$ replaced by $(E_1(\boldsymbol{x}),\mathbf{0})$: Let λ be a Borel probability measure on D, absolutely continuous with respect to Lebesgue measure, and let $\mathscr{E}_t$ be a family of subsets of $\mathbb{R}^{d-1}\times X$. Then, for $\boldsymbol{\alpha}\in\mathbb{R}^d\setminus\mathbb{Q}^d$ and $M\in\mathrm{SL}(d,\mathbb{R})$,

(5.36)
$$\liminf_{t\to\infty}\int_D\chi_{\mathscr{E}_t}\big(\boldsymbol{x},(1_d,\boldsymbol{\alpha})(M,\mathbf{0})(E_1(\boldsymbol{x}),\mathbf{0})\Phi^t\big)\,d\lambda(\boldsymbol{x})\ge\int_{\lim(\inf\mathscr{E}_t)^\circ}d\mu(g)\,d\lambda(\boldsymbol{x}),$$

and we have corresponding analogues of (5.30) and (5.31). The proof is exactly as the proof of Theorem 5.6, except that Corollary 5.4 is used in place of Theorem 5.3.

5.4. *Corresponding results for* $\mathrm{SL}(d,\mathbb{R})$. By following the same line of arguments as for $\mathrm{ASL}(d,\mathbb{R})$, one can prove the analogous equidistribution results for any homogeneous space $\Gamma\backslash\mathrm{SL}(d,\mathbb{R})$ with Γ a lattice in $\mathrm{SL}(d,\mathbb{R})$. The lattices relevant for our application are the congruence subgroups $\Gamma=\Gamma(q)$. The main results are as follows (cf. Theorem 5.3, Corollary 5.4, Theorem 5.6 and Remark 5.7 above).

THEOREM 5.8. *Let λ be a Borel probability measure on $\mathbb{R}^{d-1}$ which is absolutely continuous with respect to Lebesgue measure. Let $f:\mathbb{R}^{d-1}\times X_q\to\mathbb{R}$ be bounded continuous and $f_t:\mathbb{R}^{d-1}\times X_q\to\mathbb{R}$ a family of uniformly bounded, continuous functions such that $f_t\to f$ as $t\to\infty$, uniformly on compacta. Then, for every $M\in\mathrm{SL}(d,\mathbb{R})$,*

$$\lim_{t\to\infty}\int_{\mathbb{R}^{d-1}}f_t\left(\boldsymbol{x},M\begin{pmatrix}1&\boldsymbol{x}\\ {}^{\mathrm{t}}\mathbf{0}&1_{d-1}\end{pmatrix}\begin{pmatrix}\mathrm{e}^{-(d-1)t}&\mathbf{0}\\ {}^{\mathrm{t}}\mathbf{0}&\mathrm{e}^{t}1_{d-1}\end{pmatrix}\right)d\lambda(\boldsymbol{x})=\int_{\mathbb{R}^{d-1}\times X_q}f(\boldsymbol{x},M)\,d\mu_q(M)\,d\lambda(\boldsymbol{x}).\tag{5.37}$$

COROLLARY 5.9. *Let $E_1 : D\to\mathrm{SO}(d)$ be any map as in Corollary* 5.4, *let λ be a Borel probability measure on D, absolutely continuous with respect to Lebesgue measure, and let $f : D\times X_q\to\mathbb{R}$ and $f_t : D\times X_q\to\mathbb{R}$ be uniformly bounded continuous functions such that $f_t\to f$ as $t\to\infty$, uniformly on compacta.*

Then, for every $M \in \mathrm{SL}(d,\mathbb{R})$,

$$\lim_{t\to\infty}\int_D f_t\left(\boldsymbol{x}, ME_1(\boldsymbol{x})\begin{pmatrix}\mathrm{e}^{-(d-1)t} & \mathbf{0}\\ {}^{\mathrm{t}}\mathbf{0} & \mathrm{e}^{t}1_{d-1}\end{pmatrix}\right)d\lambda(\boldsymbol{x}) = \int_{D\times X_q} f(\boldsymbol{x},M)\,d\mu_q(M)d\lambda(\boldsymbol{x}). \tag{5.38}$$

THEOREM 5.10. *Let* λ *be a Borel probability measure on* $\mathbb{R}^{d-1}$ *which is absolutely continuous with respect to Lebesgue measure, and let* $\mathscr{E}_t$ *be a family of subsets of* $\mathbb{R}^{d-1}\times X_q$. *Then, for every* $M\in\mathrm{SL}(d,\mathbb{R})$,

$$\liminf_{t\to\infty}\int_{\mathbb{R}^{d-1}}\chi_{\mathscr{E}_t}\left(\boldsymbol{x}, M\begin{pmatrix}1 & \boldsymbol{x}\\ {}^{\mathrm{t}}\mathbf{0} & 1_{d-1}\end{pmatrix}\begin{pmatrix}\mathrm{e}^{-(d-1)t} & \mathbf{0}\\ {}^{\mathrm{t}}\mathbf{0} & \mathrm{e}^{t}1_{d-1}\end{pmatrix}\right)d\lambda(\boldsymbol{x}) \geq \int_{\lim(\inf\mathscr{E}_t)^\circ} d\mu_q(M)d\lambda(\boldsymbol{x}), \tag{5.39}$$

and

$$\limsup_{t\to\infty}\int_{\mathbb{R}^{d-1}}\chi_{\mathscr{E}_t}\left(\boldsymbol{x}, M\begin{pmatrix}1 & \boldsymbol{x}\\ {}^{\mathrm{t}}\mathbf{0} & 1_{d-1}\end{pmatrix}\begin{pmatrix}\mathrm{e}^{-(d-1)t} & \mathbf{0}\\ {}^{\mathrm{t}}\mathbf{0} & \mathrm{e}^{t}1_{d-1}\end{pmatrix}\right)d\lambda(\boldsymbol{x}) \leq \int_{\overline{\lim\sup\mathscr{E}_t}} d\mu_q(M)d\lambda(\boldsymbol{x}). \tag{5.40}$$

If, furthermore, the set $\overline{\lim\sup\mathscr{E}_t}\setminus\lim(\inf\mathscr{E}_t)^\circ$ *has measure zero, then*

$$\lim_{t\to\infty}\int_{\mathbb{R}^{d-1}}\chi_{\mathscr{E}_t}\left(\boldsymbol{x}, M\begin{pmatrix}1 & \boldsymbol{x}\\ {}^{\mathrm{t}}\mathbf{0} & 1_{d-1}\end{pmatrix}\begin{pmatrix}\mathrm{e}^{-(d-1)t} & \mathbf{0}\\ {}^{\mathrm{t}}\mathbf{0} & \mathrm{e}^{t}1_{d-1}\end{pmatrix}\right)d\lambda(\boldsymbol{x}) = \int_{\lim\sup\mathscr{E}_t} d\mu_q(M)d\lambda(\boldsymbol{x}). \tag{5.41}$$

Remark 5.11. Let $E_1 : D\to\mathrm{SO}(d)$ be any map as in Corollary 5.4; then the assertions of Theorem 5.10 hold with $\left(\begin{smallmatrix}1 & \boldsymbol{x}\\ {}^{\mathrm{t}}\mathbf{0} & 1_{d-1}\end{smallmatrix}\right)$ replaced with $E_1(\boldsymbol{x})$: Let λ be a Borel probability measure on D, absolutely continuous with respect to Lebesgue measure, let $\mathscr{E}_t$ be a family of subsets of $\mathbb{R}^{d-1}\times X_q$, and let $M\in\mathrm{SL}(d,\mathbb{R})$. Then

$$\liminf_{t\to\infty}\int_D\chi_{\mathscr{E}_t}\left(\boldsymbol{x}, ME_1(\boldsymbol{x})\begin{pmatrix}\mathrm{e}^{-(d-1)t} & \mathbf{0}\\ {}^{\mathrm{t}}\mathbf{0} & \mathrm{e}^{t}1_{d-1}\end{pmatrix}\right)d\lambda(\boldsymbol{x}) \geq \int_{\lim(\inf\mathscr{E}_t)^\circ} d\mu_q(M)d\lambda(\boldsymbol{x}), \tag{5.42}$$

and we have corresponding analogues of (5.40) and (5.41).

Note that these statements for $\mathrm{SL}(d,\mathbb{R})$ are in fact consequences of the mixing property of diagonal one-parameter subgroups of $\mathrm{SL}(d,\mathbb{R})$ on $\Gamma\backslash\mathrm{SL}(d,\mathbb{R})$ (cf. the arguments used in [15], [21]), and do not require an application of Ratner's theory.

6. **Lattice points in thin sets**

6.1. *Affine lattices with irrational α.* In the following we consider subsets $\mathfrak{B}$ of $\mathbb{R}^{d-1}\times\mathbb{R}^d$; we use the notation

$$\mathfrak{B}|_{\boldsymbol{x}} = (\{\boldsymbol{x}\}\times\mathbb{R}^d)\cap\mathfrak{B} \tag{6.1}$$

which we identify with a subset of $\mathbb{R}^d$ by projection onto the $\mathbb{R}^d$ component. Our goal in this section is to study, for a given affine lattice, the limit distribution of the number of lattice points contained in such a set $\mathfrak{B}|_{\boldsymbol{x}}$ after it has been deformed, thinly stretched, and then sheared (or rotated) by a random amount. As we will see in Section 9, the problems discussed in Sections 2 and 3 correspond to special cases of the present question.

THEOREM 6.1. *Let λ be a Borel probability measure on $\mathbb{R}^{d-1}$ which is absolutely continuous with respect to Lebesgue measure, and let $\mathfrak{B}_t$ be a family of subsets of $\mathbb{R}^{d-1}\times\mathbb{R}^d$ such that $\cup_t\mathfrak{B}_t$ is bounded. Then, for $r\in\mathbb{Z}_{\geq 0}$, $\boldsymbol{\alpha}\in\mathbb{R}^d\setminus\mathbb{Q}^d$ and $M\in\mathrm{SL}(d,\mathbb{R})$,*

$$\begin{aligned}\liminf_{t\to\infty}\lambda\Big(\Big\{\boldsymbol{x}\in\mathbb{R}^{d-1}\,:\,\#\big(\mathfrak{B}_t|_{\boldsymbol{x}}\Phi^{-t}n_-(\boldsymbol{x})\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M\big)\geq r\Big\}\Big)\\ \geq(\lambda\times\mu)\Big(\Big\{(\boldsymbol{x},g)\in\mathbb{R}^{d-1}\times X\,:\,\#\big((\lim(\inf\mathfrak{B}_t)^\circ)|_{\boldsymbol{x}}\cap\mathbb{Z}^d g\big)\geq r\Big\}\Big),\end{aligned} \tag{6.2}$$

and

$$\begin{aligned}\limsup_{t\to\infty}\lambda\Big(\Big\{\boldsymbol{x}\in\mathbb{R}^{d-1}:\#\big(\mathfrak{B}_t|_{\boldsymbol{x}}\Phi^{-t}n_-(\boldsymbol{x})\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M\big)\geq r\Big\}\Big)\\ \leq(\lambda\times\mu)\Big(\Big\{(\boldsymbol{x},g)\in\mathbb{R}^{d-1}\times X:\#\big((\lim\overline{\sup}\,\mathfrak{B}_t)|_{\boldsymbol{x}}\cap\mathbb{Z}^d g\big)\geq r\Big\}\Big).\end{aligned} \tag{6.3}$$

If, furthermore, the set $\lim\overline{\sup}\,\mathfrak{B}_t\setminus\lim(\inf\mathfrak{B}_t)^\circ$ has Lebesgue-measure zero, then

$$\begin{aligned}\lim_{t\to\infty}\lambda\Big(\Big\{\boldsymbol{x}\in\mathbb{R}^{d-1}\,:\,\#\big(\mathfrak{B}_t|_{\boldsymbol{x}}\Phi^{-t}n_-(\boldsymbol{x})\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M\big\}\big)\geq r\Big\}\Big)\\ =(\lambda\times\mu)\Big(\Big\{(\boldsymbol{x},g)\in\mathbb{R}^{d-1}\times X\,:\,\#\big((\limsup\mathfrak{B}_t)|_{\boldsymbol{x}}\cap\mathbb{Z}^d g\big)\geq r\Big\}\Big).\end{aligned} \tag{6.4}$$

We will require the following lemma for the proof of Theorem 6.1. Given a set $\mathfrak{B}\subset\mathbb{R}^{d-1}\times\mathbb{R}^d$ and an integer $r\in\mathbb{Z}_{>0}$, we define the subset

$$\mathscr{E}(\mathfrak{B},r)=\Big\{(\boldsymbol{x},g)\in\mathbb{R}^{d-1}\times X\,:\,\#\big(\mathfrak{B}|_{\boldsymbol{x}}\cap\mathbb{Z}^d g\big)\geq r\Big\}. \tag{6.5}$$

LEMMA 6.2. *Fix $r\in\mathbb{Z}_{>0}$. Then the following statements hold*:

(i) *If $\mathfrak{A}\subset\mathfrak{B}$, then $\mathscr{E}(\mathfrak{A},r)\subset\mathscr{E}(\mathfrak{B},r)$.*

(ii) *If $\mathfrak{B}_t$ is a decreasing family of bounded sets, then $\cap_t\mathscr{E}(\mathfrak{B}_t,r)=\mathscr{E}(\cap_t\mathfrak{B}_t,r)$.*

(iii) *If $\mathfrak{B}_t$ is an increasing family of sets then $\cup_t\mathscr{E}(\mathfrak{B}_t,r)=\mathscr{E}(\cup_t\mathfrak{B}_t,r)$.*

(iv) *If $\mathfrak{B}$ is open, then $\mathscr{E}(\mathfrak{B},r)$ is open.*

(v) *If $\mathfrak{B}$ is closed and bounded, then $\mathscr{E}(\mathfrak{B},r)$ is closed.*

(vi) *If $\mathfrak{B}$ has zero Lebesgue measure, then $\mathscr{E}(\mathfrak{B}, r)$ has zero measure with respect to* $\mathrm{vol}_{\mathbb{R}^{d-1}} \times \mu$.

Proof of (i). Clear. □

Proof of (ii). By (i), $\cap_t \mathscr{E}(\mathfrak{B}_t, r) \supset \mathscr{E}(\cap_t \mathfrak{B}_t, r)$. To prove the opposite inclusion, let $(\boldsymbol{x}, g) \in \mathbb{R}^{d-1} \times X$ be an arbitrary point outside $\mathscr{E}(\cap_t \mathfrak{B}_t, r)$, where $g \in \mathrm{ASL}(d, \mathbb{R})$ is a fixed representative for a point in X. Then $\#\big((\cap_t \mathfrak{B}_t)|_{\boldsymbol{x}} \cap \mathbb{Z}^d g\big) < r$. Because of our assumptions there is a bounded set $\mathfrak{C} \subset \mathbb{R}^d$ such that $\mathfrak{B}_t|_{\boldsymbol{x}} \subset \mathfrak{C}$ for all $t \geq t_0$ (for some constant $t_0 \in \mathbb{R}$). Let F be the finite set $F := \{\boldsymbol{m} \in \mathbb{Z}^d : \boldsymbol{m} g \in \mathfrak{C}\}$, and let $F' := \{\boldsymbol{m} \in \mathbb{Z}^d : \boldsymbol{m} g \in (\cap_t \mathfrak{B}_t)|_{\boldsymbol{x}}\} \subset F$. Then $\#F' < r$. For each $\boldsymbol{m} \in F \setminus F'$ there is some $t \geq t_0$ such that $\boldsymbol{m} g \notin \mathfrak{B}_t|_{\boldsymbol{x}}$; thus for all sufficiently large t we have $\boldsymbol{m} g \notin \mathfrak{B}_t|_{\boldsymbol{x}}$ for all $\boldsymbol{m} \in F \setminus F'$. Hence for these t we have $\#\big(\mathfrak{B}_t|_{\boldsymbol{x}} \cap \mathbb{Z}^d g\big) \leq \#F' < r$. Hence $(\boldsymbol{x}, g) \notin \cap_t \mathscr{E}(\mathfrak{B}_t, r)$. □

Proof of (iii). It follows from (i) that $\cup_t \mathscr{E}(\mathfrak{B}_t, r) \subset \mathscr{E}(\cup_t \mathfrak{B}_t, r)$. To prove the other inclusion, take an arbitrary point $(\boldsymbol{x}, g) \in \mathscr{E}(\cup_t \mathfrak{B}_t, r)$. Then there are r distinct vectors $\boldsymbol{m}_1, \ldots, \boldsymbol{m}_r \in \mathbb{Z}^d$ with $\boldsymbol{m}_j g \in (\cup_t \mathfrak{B}_t)|_{\boldsymbol{x}} = \cup_t (\mathfrak{B}_t|_{\boldsymbol{x}})$. Hence for t sufficiently large we have $\boldsymbol{m}_j g \in \mathfrak{B}_t|_{\boldsymbol{x}}$ for all $j = 1, \ldots, r$. Hence $(\boldsymbol{x}, g) \in \cup_t \mathscr{E}(\mathfrak{B}_t, r)$. □

Proof of (iv). Assume that $\mathfrak{B}$ is open. Take $(\boldsymbol{x}_0, g_0) \in \mathscr{E}(\mathfrak{B}, r)$, where $g_0 \in \mathrm{ASL}(d, \mathbb{R})$ is a fixed representative for a point in X. Then there exist r distinct points $\boldsymbol{m}_1, \ldots, \boldsymbol{m}_r \in \mathbb{Z}^d$ satisfying $\boldsymbol{m}_j g_0 \in \mathfrak{B}|_{\boldsymbol{x}_0}$, i.e. $(\boldsymbol{x}_0, \boldsymbol{m}_j g_0) \in \mathfrak{B}$. Writing $\Omega = \cap_{j=1}^r f_j^{-1}(\mathfrak{B})$ where $f_j : \mathbb{R}^{d-1} \times \mathrm{ASL}(d, \mathbb{R}) \ni (\boldsymbol{x}, g) \mapsto (\boldsymbol{x}, \boldsymbol{m}_j g) \in \mathbb{R}^{d-1} \times \mathbb{R}^d$, we have $(\boldsymbol{x}_0, g_0) \in \Omega$, and each $(\boldsymbol{x}, g) \in \Omega$ projects to a point in $\mathscr{E}(\mathfrak{B}, r)$. Also Ω is an open subset of $\mathbb{R}^{d-1} \times \mathrm{ASL}(d, \mathbb{R})$, each f_j being continuous. Since $(\boldsymbol{x}_0, g_0)$ was arbitrary in $\mathscr{E}(\mathfrak{B}, r)$ we conclude that $\mathscr{E}(\mathfrak{B}, r)$ is open. □

Proof of (v). Assume that $\mathfrak{B}$ is closed and bounded. Take $(\boldsymbol{x}_0, g_0) \in \mathbb{R}^{d-1} \times X$ outside $\mathscr{E}(\mathfrak{B}, r)$, where again $g_0 \in \mathrm{ASL}(d, \mathbb{R})$ is a fixed representative for a point in X. Then $\#(\mathfrak{B}|_{\boldsymbol{x}_0} \cap \mathbb{Z}^d g_0) < r$.

Let U_1 be a neighborhood of the identity in $\mathrm{SL}(d, \mathbb{R})$ such that $\|\boldsymbol{y} M - \boldsymbol{y}\| \leq \frac{1}{2}\|\boldsymbol{y}\|$ for all $\boldsymbol{y} \in \mathbb{R}^d$, $M \in U_1$. Let $R = \sup\{\|\boldsymbol{y}\| : \boldsymbol{y} \in \cup_{\boldsymbol{x} \in \mathbb{R}^{d-1}} \mathfrak{B}|_{\boldsymbol{x}}\}$. Then $U = U_1 \times \mathcal{B}_R^d$ is a neighborhood of the identity in $\mathrm{ASL}(d, \mathbb{R}) = \mathrm{SL}(d, \mathbb{R}) \ltimes \mathbb{R}^d$, and for each $\boldsymbol{y} \in \mathbb{R}^d$ with $\|\boldsymbol{y}\| > 4R$ and $g = (M, \boldsymbol{\xi}) \in U$ we have

$$\|\boldsymbol{y} g\| = \|\boldsymbol{y} M + \boldsymbol{\xi}\| \geq \|\boldsymbol{y}\| - \|\boldsymbol{y} M - \boldsymbol{y}\| - \|\boldsymbol{\xi}\| > \tfrac{1}{2}\|\boldsymbol{y}\| - R > R. \tag{6.6}$$

Hence $\boldsymbol{y} g \notin \mathfrak{B}|_{\boldsymbol{x}}$ holds automatically for all $g \in U$, $\boldsymbol{x} \in \mathbb{R}^{d-1}$ and all $\boldsymbol{y} \in \mathbb{R}^d$ with $\|\boldsymbol{y}\| > 4R$. Let F be the finite set of points $\boldsymbol{m} \in \mathbb{Z}^d$ which satisfy $\|\boldsymbol{m} g_0\| \leq 4R$ and $\boldsymbol{m} g_0 \notin \mathfrak{B}|_{\boldsymbol{x}_0}$. For each $\boldsymbol{m} \in F$ we choose some open sets $V_{\boldsymbol{m}} \subset \mathbb{R}^{d-1}$ and $V'_{\boldsymbol{m}} \subset \mathbb{R}^d$ such that $(\boldsymbol{x}_0, \boldsymbol{m} g_0) \in V_{\boldsymbol{m}} \times V'_{\boldsymbol{m}} \subset \complement\mathfrak{B}$. Now set

$$U' = (g_0 U) \cap \Big(\bigcap_{\boldsymbol{m} \in F} \{g \in \mathrm{ASL}(d, \mathbb{R}) : \boldsymbol{m} g \in V'_{\boldsymbol{m}}\} \Big); \qquad V = \bigcap_{\boldsymbol{m} \in F} V_{\boldsymbol{m}}. \tag{6.7}$$

These are open subsets of $\mathrm{ASL}(d,\mathbb{R})$ and $\mathbb{R}^d$, respectively, and $(\boldsymbol{x}_0, g_0) \in V \times U'$. Furthermore, if $(\boldsymbol{x}, g) \in V \times U'$ then by construction $\boldsymbol{m}g \notin \mathfrak{B}|_{\boldsymbol{x}}$ for each $\boldsymbol{m} \in \mathbb{Z}^d$ with $\boldsymbol{m}g_0 \notin \mathfrak{B}|_{\boldsymbol{x}_0}$, and thus $\#(\mathbb{Z}^d g \cap \mathfrak{B}|_{\boldsymbol{x}}) < r$ since $\#(\mathbb{Z}^d g_0 \cap \mathfrak{B}|_{\boldsymbol{x}_0}) < r$. Hence each $(\boldsymbol{x}, g) \in V \times U'$ projects to a point in $\mathbb{R}^{d-1} \times X$ *outside* $\mathscr{E}(\mathfrak{B}, r)$.

Since $(\boldsymbol{x}_0, g_0)$ was an arbitrary point outside $\mathscr{E}(\mathfrak{B}, r)$ we conclude that $\mathscr{E}(\mathfrak{B}, r)$ is closed. □

Proof of (vi). Assume that $\mathfrak{B}$ has Lebesgue measure zero. Note that

$$\begin{aligned}(6.8)\quad &(\mathrm{vol}_{\mathbb{R}^{d-1}} \times \mu)\Big(\big\{(\boldsymbol{x}, g) \in \mathbb{R}^{d-1} \times \mathrm{ASL}(d,\mathbb{R}) \,:\, \mathfrak{B}|_{\boldsymbol{x}} \cap \mathbb{Z}^d g \neq \varnothing\big\}\Big)\\ &\le \sum_{\boldsymbol{m} \in \mathbb{Z}^d} (\mathrm{vol}_{\mathbb{R}^{d-1}} \times \mu)\Big(\big\{(\boldsymbol{x}, g) \in \mathbb{R}^{d-1} \times \mathrm{ASL}(d,\mathbb{R}) \,:\, \boldsymbol{m}g \in \mathfrak{B}|_{\boldsymbol{x}}\big\}\Big)\\ &= \sum_{\boldsymbol{m} \in \mathbb{Z}^d} \int_{\mathbb{R}^{d-1}} \int_{\mathrm{SL}(d,\mathbb{R})} \int_{\mathbb{R}^d} I\Big(\boldsymbol{m}(M,\boldsymbol{\xi}) \in \mathfrak{B}|_{\boldsymbol{x}}\Big) d\,\mathrm{vol}_{\mathbb{R}^d}(\boldsymbol{\xi})\, d\mu_1(M)\, d\,\mathrm{vol}_{\mathbb{R}^{d-1}}(\boldsymbol{x}),\end{aligned}$$

where I is the indicator function. The innermost integral equals $\mathrm{vol}_{\mathbb{R}^d}(\mathfrak{B}|_{\boldsymbol{x}})$, since $\boldsymbol{m}(M,\boldsymbol{\xi}) = \boldsymbol{m}M + \boldsymbol{\xi}$. But $\mathrm{vol}(\mathfrak{B}|_{\boldsymbol{x}}) = 0$ holds for almost every $\boldsymbol{x} \in \mathbb{R}^{d-1}$, and thus the total integral is zero. Hence, a fortiori, $\mathscr{E}(\mathfrak{B}, r)$ has measure zero. □

Proof of Theorem 6.1. If $r = 0$ then the statements are trivial; thus from now on we may assume $r > 0$. Define the decreasing family of sets

$$(6.9)\qquad \widehat{\mathscr{E}}_t := \mathscr{E}\Big(\overline{\bigcup_{s \ge t} \mathfrak{B}_s}, r\Big).$$

These sets are clearly closed (cf. Lemma 6.2(v)). Then

$$\begin{aligned}(6.10)\qquad &\limsup_{t\to\infty} \lambda\Big(\Big\{\boldsymbol{x} \in \mathbb{R}^{d-1} : \#\big(\mathfrak{B}_t|_{\boldsymbol{x}} \Phi^{-t} n_-(\boldsymbol{x}) \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M\big) \ge r\Big\}\Big)\\ &\qquad \le \limsup_{t\to\infty} \int_{\mathbb{R}^{d-1}} \chi_{\widehat{\mathscr{E}}_t}\big(\boldsymbol{x}, (1_d, \boldsymbol{\alpha})(M, \boldsymbol{0}) n_-(-\boldsymbol{x})\Phi^t\big)\, d\lambda(\boldsymbol{x})\\ &\qquad \le \int_{\limsup \widehat{\mathscr{E}}_t} d\mu(g) d\lambda(\boldsymbol{x}),\end{aligned}$$

due to Theorem 5.6. (To be precise, to treat "$n_-(-\boldsymbol{x})$" as above, one applies Theorem 5.6 to λ' and $\widehat{\mathscr{E}}'_t$, defined through $\lambda'(B) = \lambda(-B)$ for $B \subset \mathbb{R}^{d-1}$, and $\widehat{\mathscr{E}}'_t = \{(-\boldsymbol{x}, g) : (\boldsymbol{x}, g) \in \widehat{\mathscr{E}}_t\}$.) In view of Lemma 6.2(ii),

$$(6.11)\qquad \limsup \widehat{\mathscr{E}}_t = \bigcap_t \widehat{\mathscr{E}}_t = \mathscr{E}\Big(\bigcap_t \overline{\bigcup_{s \ge t} \mathfrak{B}_s}, r\Big) = \mathscr{E}(\limsup \mathfrak{B}_t, r),$$

and hence

$$\begin{aligned}(6.12)\qquad &\int_{\limsup \widehat{\mathscr{E}}_t} d\mu(g) d\lambda(\boldsymbol{x}) = \int_{\mathscr{E}(\limsup \mathfrak{B}_t, r)} d\mu(g) d\lambda(\boldsymbol{x})\\ &\qquad = (\lambda \times \mu)\Big(\Big\{(\boldsymbol{x}, g) \in \mathbb{R}^{d-1} \times X : \#\big((\limsup \mathfrak{B}_t)|_{\boldsymbol{x}} \cap \mathbb{Z}^d g\big) \ge r\Big\}\Big),\end{aligned}$$

which proves (6.3). The proof of (6.2) is analogous, using Lemma 6.2(iii) and (iv). Finally (6.4) follows using Lemma 6.2(vi) for $r = 1$, since λ is absolutely continuous with respect to $\mathrm{vol}_{\mathbb{R}^{d-1}}$, and

$$(6.13)\quad \mathscr{E}\big(\limsup \overline{\mathfrak{B}_t}, r\big) \setminus \mathscr{E}\big(\lim(\inf \mathfrak{B}_t)^\circ, r\big) \subset \mathscr{E}\big(\limsup \overline{\mathfrak{B}_t} \setminus \lim(\inf \mathfrak{B}_t)^\circ, 1\big). \quad \square$$

Theorem 6.1 is easily generalized to multiple families of sets:

THEOREM 6.3. *Let λ be a Borel probability measure on $\mathbb{R}^{d-1}$ which is absolutely continuous with respect to Lebesgue measure. For each $j = 1, \dots, m$, let $\mathfrak{B}_t^{(j)}$ be a family of subsets of $\mathbb{R}^{d-1} \times \mathbb{R}^d$ such that $\cup_t \mathfrak{B}_t^{(j)}$ is bounded. Then, for any $r_1, \dots, r_m \in \mathbb{Z}_{\geq 0}$, $\boldsymbol{\alpha} \in \mathbb{R}^d \setminus \mathbb{Q}^d$ and $M \in \mathrm{SL}(d, \mathbb{R})$,*

(6.14)
$$\liminf_{t\to\infty} \lambda\Big(\Big\{\boldsymbol{x} \in \mathbb{R}^{d-1} : \#\big(\mathfrak{B}_t^{(j)}|_{\boldsymbol{x}} \Phi^{-t} n_-(\boldsymbol{x}) \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M\big) \geq r_j,\ j = 1, \dots, m\Big\}\Big)$$
$$\geq (\lambda \times \mu)\Big(\Big\{(\boldsymbol{x}, g) \in \mathbb{R}^{d-1} \times X : \#\big((\lim(\inf \mathfrak{B}_t^{(j)})^\circ)|_{\boldsymbol{x}} \cap \mathbb{Z}^d g\big) \geq r_j,\ j = 1, \dots, m\Big\}\Big),$$

and

(6.15)
$$\limsup_{t\to\infty} \lambda\Big(\Big\{\boldsymbol{x} \in \mathbb{R}^{d-1} : \#\big(\mathfrak{B}_t^{(j)}|_{\boldsymbol{x}} \Phi^{-t} n_-(\boldsymbol{x}) \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M\big) \geq r_j,\ j = 1, \dots, m\Big\}\Big)$$
$$\leq (\lambda \times \mu)\Big(\Big\{(\boldsymbol{x}, g) \in \mathbb{R}^{d-1} \times X : \#\big((\limsup \overline{\mathfrak{B}_t^{(j)}})|_{\boldsymbol{x}} \cap \mathbb{Z}^d g\big) \geq r_j,\ j = 1, \dots, m\Big\}\Big).$$

If, furthermore, each set $\limsup \overline{\mathfrak{B}_t^{(j)}} \setminus \lim(\inf \mathfrak{B}_t^{(j)})^\circ$ $(j = 1, \dots, m)$ has Lebesgue-measure zero, then

(6.16)
$$\lim_{t\to\infty} \lambda\Big(\Big\{\boldsymbol{x} \in \mathbb{R}^{d-1} : \#\big(\mathfrak{B}_t^{(j)}|_{\boldsymbol{x}} \Phi^{-t} n_-(\boldsymbol{x}) \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M\big\}\big) \geq r_j,\ j = 1, \dots, m\Big\}\Big)$$
$$= (\lambda \times \mu)\Big(\Big\{(\boldsymbol{x}, g) \in \mathbb{R}^{d-1} \times X : \#\big((\limsup \mathfrak{B}_t^{(j)})|_{\boldsymbol{x}} \cap \mathbb{Z}^d g\big) \geq r_j,\ j = 1, \dots, m\Big\}\Big).$$

Proof. We may throw away each j for which $r_j = 0$. Thus from now on $r_j > 0$ for each j. Define the decreasing family of sets

$$(6.17)\qquad \widehat{\mathscr{E}}_t := \bigcap_{j=1}^{m} \mathscr{E}\Big(\overline{\bigcup_{s\geq t} \mathfrak{B}_s^{(j)}}, r_j\Big).$$

These sets are clearly closed (cf. Lemma 6.2(v)). Now (6.10) generalizes in the obvious way. In view of Lemma 6.2(ii),

(6.18)
$$\limsup \widehat{\mathscr{E}}_t = \bigcap_t \widehat{\mathscr{E}}_t = \bigcap_{j=1}^m \bigcap_t \mathscr{E}\Big(\overline{\bigcup_{s\ge t} \mathfrak{B}_s^{(j)}}, r_j\Big) = \bigcap_{j=1}^m \mathscr{E}\Big(\bigcap_t \overline{\bigcup_{s\ge t} \mathfrak{B}_s^{(j)}}, r_j\Big)$$
$$= \bigcap_{j=1}^m \mathscr{E}\Big(\overline{\limsup \mathfrak{B}_t^{(j)}}, r_j\Big),$$

and hence (6.12) carries over to give a proof of (6.15). The proof of (6.14) is analogous, using Lemma 6.2(iii) and (iv), and noticing that

$$\bigcup_t \bigcap_{j=1}^m \mathscr{E}\Big(\Big(\bigcap_{s\ge t} \mathfrak{B}_s^{(j)}\Big)^\circ, r_j\Big) = \bigcap_{j=1}^m \bigcup_t \mathscr{E}\Big(\Big(\bigcap_{s\ge t} \mathfrak{B}_s^{(j)}\Big)^\circ, r_j\Big). \tag{6.19}$$

Finally (6.16) follows using Lemma 6.2(vi) for $r = 1$, since λ is absolutely continuous with respect to $\mathrm{vol}_{\mathbb{R}^{d-1}}$, and

$$\Big(\bigcap_{j=1}^m \mathscr{E}\big(\overline{\limsup \mathfrak{B}_t^{(j)}}, r_j\big)\Big) \setminus \Big(\bigcap_{j=1}^m \mathscr{E}\big(\lim(\inf \mathfrak{B}_t^{(j)})^\circ, r_j\big)\Big)$$
$$\subset \bigcup_{j=1}^m \mathscr{E}\big(\overline{\limsup \mathfrak{B}_t^{(j)}} \setminus \lim(\inf \mathfrak{B}_t^{(j)})^\circ, 1\big). \tag{6.20}$$

This concludes the proof. □

Remark 6.4. The assertions of Theorem 6.1 and Theorem 6.3 also hold if $n_-(\boldsymbol{x})$ is replaced by $(E_1(\boldsymbol{x}), \mathbf{0})^{-1}$ where $E_1 : D \to \mathrm{SO}(d)$ is any map as in Corollary 5.4. Specifically, if λ is any Borel probability measure on D, absolutely continuous with respect to Lebesgue measure, then for any given families $\mathfrak{B}_t^{(j)} \subset \mathbb{R}^{d-1} \times \mathbb{R}^d$ as above, and any $r_j \in \mathbb{Z}_{\ge 0}$, $\boldsymbol{\alpha} \in \mathbb{R}^d \setminus \mathbb{Q}^d$ and $M \in \mathrm{SL}(d, \mathbb{R})$, we have

$$\liminf_{t\to\infty} \lambda\Big(\Big\{\boldsymbol{x} \in D : \#\big(\mathfrak{B}_t^{(j)}|_{\boldsymbol{x}}\, \Phi^{-t}(E_1(\boldsymbol{x}), \mathbf{0})^{-1} \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M\big) \ge r_j,\ j = 1, \dots, m\Big\}\Big)$$
$$\ge (\lambda \times \mu)\Big(\Big\{(\boldsymbol{x}, g) \in D \times X : \#\big((\lim(\inf \mathfrak{B}_t^{(j)})^\circ)|_{\boldsymbol{x}} \cap \mathbb{Z}^d g\big) \ge r_j,\ j = 1, \dots, m\Big\}\Big), \tag{6.21}$$

and corresponding relations for the lim sup and the limes; cf. (6.15) and (6.16). The proof is exactly as the proofs of Theorem 6.1 and Theorem 6.3, except that Remark 5.7 is used in place of Theorem 5.6.

6.2. *The case of rational* $\boldsymbol{\alpha}$. Using the same strategy of proof, the above results can be readily established for $\boldsymbol{\alpha} \in \mathbb{Q}^d$, if the space X is replaced by X_q and the measure μ by μ_q, for some q with $\boldsymbol{\alpha} \in q^{-1}\mathbb{Z}^d$. In the proofs one uses the

following analogue of (6.5):

$$\mathscr{E}_q(\mathfrak{B},r)=\Big\{(\boldsymbol{x},M)\in\mathbb{R}^{d-1}\times X_q\ :\ \#\big(\mathfrak{B}|_{\boldsymbol{x}}\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M\setminus\{\mathbf{0}\}\big)\ge r\Big\}. \tag{6.22}$$

The reason for removing the point $\mathbf{0}$ is so as to make all of Lemma 6.2 valid in the present setting. (Specifically, in the proof of the analogue of Lemma 6.2(vi) we need to note that $\int_{\mathrm{SL}(d,\mathbb{R})} I\big(\boldsymbol{m}M\in\mathfrak{C}\big)\,d\mu_1(M)=0$ holds for each subset $\mathfrak{C}\subset\mathbb{R}^d$ of measure 0. This is true for each $\boldsymbol{m}\in\mathbb{R}^d$ except $\boldsymbol{m}=\mathbf{0}$.)

We thus have the following.

THEOREM 6.5. *Let λ be a Borel probability measure on $\mathbb{R}^{d-1}$ which is absolutely continuous with respect to Lebesgue measure. For each $j=1,\dots,m$, let $\mathfrak{B}_t^{(j)}$ be a family of subsets of $\mathbb{R}^{d-1}\times\mathbb{R}^d$ such that $\cup_t\mathfrak{B}_t^{(j)}$ is bounded. Then, for any $r_1,\dots,r_m\in\mathbb{Z}_{\ge0}$, $\boldsymbol{\alpha}=\frac{\boldsymbol{p}}{q}\in\mathbb{Q}^d$ and $M\in\mathrm{SL}(d,\mathbb{R})$,*

$$\begin{aligned}&\liminf_{t\to\infty}\lambda\Big(\Big\{\boldsymbol{x}\in\mathbb{R}^{d-1}:\#\big(\mathfrak{B}_t^{(j)}|_{\boldsymbol{x}}\Phi^{-t}n_-(\boldsymbol{x})\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M\setminus\{\mathbf{0}\}\big)\ge r_j,\ j=1,\dots,m\Big\}\Big)\\&\quad\ge(\lambda\times\mu_q)\Big(\Big\{(\boldsymbol{x},M')\in\mathbb{R}^{d-1}\times X_q:\\&\qquad\qquad\#\big((\lim(\inf\mathfrak{B}_t^{(j)})^\circ)|_{\boldsymbol{x}}\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M'\setminus\{\mathbf{0}\}\big)\ge r_j,\ j=1,\dots,m\Big\}\Big),\end{aligned} \tag{6.23}$$

and

$$\begin{aligned}&\limsup_{t\to\infty}\lambda\Big(\Big\{\boldsymbol{x}\in\mathbb{R}^{d-1}:\#\big(\mathfrak{B}_t^{(j)}|_{\boldsymbol{x}}\Phi^{-t}n_-(\boldsymbol{x})\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M\setminus\{\mathbf{0}\}\big)\ge r_j,\ j=1,\dots,m\Big\}\Big)\\&\quad\le(\lambda\times\mu_q)\Big(\Big\{(\boldsymbol{x},M')\in\mathbb{R}^{d-1}\times X_q:\\&\qquad\qquad\#\big((\overline{\limsup}\,\mathfrak{B}_t^{(j)})|_{\boldsymbol{x}}\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M'\setminus\{\mathbf{0}\}\big)\ge r_j,\ j=1,\dots,m\Big\}\Big).\end{aligned} \tag{6.24}$$

If, furthermore, each set $\overline{\limsup}\,\mathfrak{B}_t^{(j)}\setminus\lim(\inf\mathfrak{B}_t^{(j)})^\circ$ $(j=1,\dots,m)$ has Lebesgue-measure zero, then

$$\begin{aligned}&\lim_{t\to\infty}\lambda\Big(\Big\{\boldsymbol{x}\in\mathbb{R}^{d-1}:\#\big(\mathfrak{B}_t^{(j)}|_{\boldsymbol{x}}\Phi^{-t}n_-(\boldsymbol{x})\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M\setminus\{\mathbf{0}\}\big)\ge r_j,\ j=1,\dots,m\Big\}\Big)\\&\quad=(\lambda\times\mu_q)\Big(\Big\{(\boldsymbol{x},M')\in\mathbb{R}^{d-1}\times X_q:\\&\qquad\qquad\#\big((\limsup\mathfrak{B}_t^{(j)})|_{\boldsymbol{x}}\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M'\setminus\{\mathbf{0}\}\big)\ge r_j,\ j=1,\dots,m\Big\}\Big).\end{aligned} \tag{6.25}$$

Remark 6.6. The assertion of Theorem 6.5 holds if $n_-(\boldsymbol{x})$ is replaced by $(E_1(\boldsymbol{x}),\mathbf{0})^{-1}$, where $E_1:D\to\mathrm{SO}(d)$ is any map as in Corollary 5.4. Compare Remark 6.4.

6.3. *Visible lattice points.* In the case of rational $\boldsymbol{\alpha}$, all results are equally valid for $\widehat{\mathbb{Z}}^d_{\boldsymbol{\alpha}}$ in place of $(\mathbb{Z}^d+\boldsymbol{\alpha})\setminus\{\mathbf{0}\}$.

THEOREM 6.7. *Let λ be a Borel probability measure on $\mathbb{R}^{d-1}$ which is absolutely continuous with respect to Lebesgue measure. For each $j = 1, \dots, m$, let $\mathfrak{B}_t^{(j)}$ be a family of subsets of $\mathbb{R}^{d-1} \times \mathbb{R}^d$ such that $\cup_t \mathfrak{B}_t^{(j)}$ is bounded. Then, for any $r_1, \dots, r_m \in \mathbb{Z}_{\geq 0}$, $\boldsymbol{\alpha} = \frac{\boldsymbol{p}}{q} \in \mathbb{Q}^d$ and $M \in \mathrm{SL}(d, \mathbb{R})$,*

(6.26)

$$\begin{aligned}\liminf_{t\to\infty} \lambda(\{\boldsymbol{x} \in \mathbb{R}^{d-1} : \#(\mathfrak{B}_t^{(j)}|_{\boldsymbol{x}} \Phi^{-t} n_-(\boldsymbol{x}) \cap \widehat{\mathbb{Z}}^d_{\boldsymbol{\alpha}} M) \geq r_j,\ j = 1, \dots, m\}) \\ \geq (\lambda \times \mu_q)(\{(\boldsymbol{x}, M') \in \mathbb{R}^{d-1} \times X_q : \\ \#((\liminf \mathfrak{B}_t^{(j)})^\circ)|_{\boldsymbol{x}} \cap \widehat{\mathbb{Z}}^d_{\boldsymbol{\alpha}} M') \geq r_j,\ j = 1, \dots, m\}),\end{aligned}$$

and

(6.27)

$$\begin{aligned}\limsup_{t\to\infty} \lambda(\{\boldsymbol{x} \in \mathbb{R}^{d-1} : \#(\mathfrak{B}_t^{(j)}|_{\boldsymbol{x}} \Phi^{-t} n_-(\boldsymbol{x}) \cap \widehat{\mathbb{Z}}^d_{\boldsymbol{\alpha}} M) \geq r_j,\ j = 1, \dots, m\}) \\ \leq (\lambda \times \mu_q)(\{(\boldsymbol{x}, M') \in \mathbb{R}^{d-1} \times X_q : \\ \#((\overline{\limsup \mathfrak{B}_t^{(j)}})|_{\boldsymbol{x}} \cap \widehat{\mathbb{Z}}^d_{\boldsymbol{\alpha}} M') \geq r_j,\ j = 1, \dots, m\}).\end{aligned}$$

If, furthermore, each set $\overline{\limsup \mathfrak{B}_t^{(j)}} \setminus \liminf(\mathfrak{B}_t^{(j)})^\circ$ $(j = 1, \dots, m)$ has Lebesgue-measure zero, then

$$\begin{aligned}\lim_{t\to\infty} \lambda(\{\boldsymbol{x} \in \mathbb{R}^{d-1} : \#(\mathfrak{B}_t^{(j)}|_{\boldsymbol{x}} \Phi^{-t} n_-(\boldsymbol{x}) \cap \widehat{\mathbb{Z}}^d_{\boldsymbol{\alpha}} M) \geq r_j,\ j = 1, \dots, m\}) \\ = (\lambda \times \mu_q)(\{(\boldsymbol{x}, M') \in \mathbb{R}^{d-1} \times X_q : \\ \#((\limsup \mathfrak{B}_t^{(j)})|_{\boldsymbol{x}} \cap \widehat{\mathbb{Z}}^d_{\boldsymbol{\alpha}} M') \geq r_j,\ j = 1, \dots, m\}).\end{aligned} \tag{6.28}$$

Remark 6.8. Just as in previous remarks, the assertion of Theorem 6.7 holds if $n_-(\boldsymbol{x})$ is replaced by $(E_1(\boldsymbol{x}), \boldsymbol{0})^{-1}$, where $E_1 : D \to \mathrm{SO}(d)$ is any map as in Corollary 5.4.

7. **Integration formulas on X and X_q**

In this section we prove some formulas for integrals and volumes in the spaces (X, μ) and (X_q, μ_q), which we will need to be able to generalize a technique which was introduced in Elkies and McMullen [14] in the case of $d = 2$ and (X, μ) (cf. also [36]). The goal is to obtain a more precise understanding of the explicit limit functions described in our main theorems; we will achieve this in Section 8.

Recall that we have fixed μ_q as the Haar measure on $\mathrm{SL}(d, \mathbb{R})$ normalized to be a probability measure on $X_q = \Gamma(q) \backslash \mathrm{SL}(d, \mathbb{R})$. This implies, via a well-known volume formula by Siegel, that μ_1 can be explicitly given as the measure on $\mathrm{SL}(d, \mathbb{R})$ which satisfies

$$d\mu_1(M) \frac{dt}{t} = \Big(\zeta(2)\zeta(3) \dots \zeta(d)\Big)^{-1} \big(\det(x_{ij})\big)^{-d}\, dx_{11} dx_{12} \cdots\cdots dx_{dd} \tag{7.1}$$

when parametrizing $\mathrm{GL}^+(d,\mathbb{R})$ as $(x_{ij}) = t^{1/d}M \in \mathrm{GL}^+(d,\mathbb{R})$, with $M \in \mathrm{SL}(d,\mathbb{R})$, $t > 0$; cf., e.g., [34], [21]. From this it follows that

$$\mu_q = I_q^{-1}\mu_1; \qquad \text{where } I_q := [\Gamma(q) : \Gamma(1)], \tag{7.2}$$

and also that the Haar measure μ on $\mathrm{ASL}(d,\mathbb{R})$ which we have normalized by $\mu(X) = 1$, is explicitly given by

$$d\mu(M,\boldsymbol{\xi}) = d\mu_1(M)d\boldsymbol{\xi}, \qquad (M,\boldsymbol{\xi}) \in \mathrm{ASL}(d,\mathbb{R}), \tag{7.3}$$

where $d\boldsymbol{\xi} = d\xi_1 \cdot\cdots\cdot d\xi_d$ is the standard Lebesgue measure on $\mathbb{R}^d$.

The following lattice average formula is also due to Siegel (at least on X_1). Recall that we always keep $d \geq 2$.

PROPOSITION 7.1. *Let $F \in \mathrm{L}^1(\mathbb{R}^d)$, $q \in \mathbb{Z}_{>0}$ and $\boldsymbol{\alpha} \in q^{-1}\mathbb{Z}^d$. Then*

$$\int_{X_q} \sum_{\boldsymbol{k}\in\mathbb{Z}^d+\boldsymbol{\alpha}\setminus\{\mathbf{0}\}} F(\boldsymbol{k}M)d\mu_q(M) = \int_{\mathbb{R}^d} F(\boldsymbol{x})d\boldsymbol{x}. \tag{7.4}$$

Proof. If $\boldsymbol{\alpha} \in \mathbb{Z}^d$ then one easily reduces to the case $q = 1$, and then the formula is proved in Siegel, [34]. (Cf. also [21, §3.7].)

From now on we assume $\boldsymbol{\alpha} \notin \mathbb{Z}^d$ (and thus $\mathbb{Z}^d + \boldsymbol{\alpha}\setminus\{\mathbf{0}\} = \mathbb{Z}^d + \boldsymbol{\alpha}$). Write $\boldsymbol{\alpha} = \frac{\boldsymbol{p}}{q}$ with $\boldsymbol{p} \in \mathbb{Z}^d$. Let $\mathcal{F} \subset \mathrm{SL}(d,\mathbb{R})$ be a fundamental domain for $\mathrm{SL}(d,\mathbb{Z})\backslash\mathrm{SL}(d,\mathbb{R})$ and choose representatives $T_j \in \mathrm{SL}(d,\mathbb{Z})$ so that $\mathrm{SL}(d,\mathbb{Z}) = \bigsqcup_{j=1}^{I_q} \Gamma(q)T_j$ (with $\bigsqcup$ denoting disjoint union); then $\bigsqcup_{j=1}^{I_q} T_j\mathcal{F}$ is a fundamental domain for $\Gamma(q)$. Hence

$$\begin{aligned}\int_{X_q} \sum_{\boldsymbol{k}\in\mathbb{Z}^d+\boldsymbol{\alpha}} F(\boldsymbol{k}M)\,d\mu_q(M) &= I_q^{-1}\sum_{j=1}^{I_q}\int_{\mathcal{F}} \sum_{\boldsymbol{k}\in\mathbb{Z}^d+\boldsymbol{\alpha}} F(\boldsymbol{k}T_jM)\,d\mu_1(M)\\ &= I_q^{-1}\int_{\mathcal{F}} \sum_{\ell=1}^{\infty} F(q^{-1}\boldsymbol{n}_\ell M)\,d\mu_1(M),\end{aligned} \tag{7.5}$$

where $\boldsymbol{n}_1, \boldsymbol{n}_2, \cdots \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$ is an enumeration (with multiplicities taken into account) of the points $\boldsymbol{m}T_j$, for $\boldsymbol{m} \in \boldsymbol{p} + q\mathbb{Z}^d$, $j \in \{1,\dots,I_q\}$. For every $\gamma \in \mathrm{SL}(d,\mathbb{Z})$, the list $\boldsymbol{n}_1\gamma, \boldsymbol{n}_2\gamma, \dots$ can be obtained as a permutation of $\boldsymbol{n}_1, \boldsymbol{n}_2, \dots$. (To see this, note that given $\gamma \in \mathrm{SL}(d,\mathbb{Z})$ there are elements $\gamma_1,\dots,\gamma_{I_q} \in \Gamma(q)$ and a permutation ρ of $\{1,\dots,I_q\}$ such that $T_j\gamma = \gamma_j T_{\rho(j)}$ for all $j \in \{1,\dots,I_q\}$. Also note $(\boldsymbol{p} + q\mathbb{Z}^d)\gamma_j = \boldsymbol{p} + q\mathbb{Z}^d$.) Recall that each orbit for the action of $\mathrm{SL}(d,\mathbb{Z})$ on $\mathbb{Z}^d \setminus \{\mathbf{0}\}$ equals $t\widehat{\mathbb{Z}}^d$ for some $t \in \mathbb{Z}_{>0}$, where $\widehat{\mathbb{Z}}^d$ as before denotes the set of primitive lattice points in $\mathbb{Z}^d$. It follows that for each $t \in \mathbb{Z}_{>0}$ there is some multiplicity $m_t \in \mathbb{Z}_{\geq 0}$ such that the sequence $\boldsymbol{n}_1, \boldsymbol{n}_2, \dots$ visits each point in $t\widehat{\mathbb{Z}}^d$ exactly m_t times. Now the above integral may be rewritten as

$$I_q^{-1}\int_{X_1} \sum_{t=1}^{\infty} m_t \sum_{\boldsymbol{c}\in\widehat{\mathbb{Z}}^d} F(q^{-1}t\boldsymbol{c}M)\,d\mu_1(M). \tag{7.6}$$

Arguing as in [21, pp. 1150–1151] we find that this equals

$$\frac{1}{I_q\zeta(d)}\sum_{t=1}^{\infty}\frac{m_t}{t^d}\int_{\mathbb{R}^d}F(q^{-1}\boldsymbol{x})\,d\boldsymbol{x}=\frac{q^d}{I_q\zeta(d)}\sum_{t=1}^{\infty}\frac{m_t}{t^d}\int_{\mathbb{R}^d}F(\boldsymbol{x})\,d\boldsymbol{x}. \tag{7.7}$$

Finally an argument as in [21, p. 1152 (top)] shows that the constant in front of the integral must actually be 1, i.e. $\sum_{t=1}^{\infty}\frac{m_t}{t^d}=q^{-d}I_q\zeta(d)$, and the proof is complete. □

The identity $\sum_{t=1}^{\infty}\frac{m_t}{t^d}=q^{-d}I_q\zeta(d)$ can of course also be proved by a more explicit computation: One easily reduces the situation to the case where q is the minimal denominator of the given $\boldsymbol{\alpha}\in\mathbb{Q}^d$; in other words $\boldsymbol{\alpha}=\frac{\boldsymbol{p}}{q}$ where $\boldsymbol{p}=(p_1,\dots,p_d)\in\mathbb{Z}^d$ has $\gcd(q,p_1,\dots,p_d)=1$. Then note that the $\mathrm{SL}(d,\mathbb{Z}/q\mathbb{Z})$-orbit of $\boldsymbol{p}+q\mathbb{Z}^d$ in $\mathbb{Z}^d/q\mathbb{Z}^d$ equals

$$V=\big\{\boldsymbol{x}+q\mathbb{Z}^d:\boldsymbol{x}=(x_1,\dots,x_d)\in\mathbb{Z}^d,\gcd(q,x_1,\dots,x_d)=1\big\}\subset\mathbb{Z}^d/q\mathbb{Z}^d, \tag{7.8}$$

and since $\#\,\mathrm{SL}(d,\mathbb{Z}/q\mathbb{Z})=I_q$ we see that the sequence $\boldsymbol{n}_1,\boldsymbol{n}_2,\dots$ visits exactly those points in $\mathbb{Z}^d$ which belong to the preimage of V, and each such point is visited exactly $I_q/\#V$ times. Hence

$$\sum_{t=1}^{\infty}\frac{m_t}{t^d}=\frac{I_q}{\#V}\sum_{t\ge1,\,(t,q)=1}t^{-d}=\frac{I_q}{\#V}\zeta(d)\sum_{e|q}\mu(e)e^{-d}. \tag{7.9}$$

On the other hand $\#V=q^d\sum_{e|q}\mu(e)e^{-d}$, and the identity follows.

7.1. *The submanifolds $X_q(\boldsymbol{y})$ of X_q.* Fix $\boldsymbol{\alpha}=\boldsymbol{p}/q\in\mathbb{Q}^d$. Given any $\boldsymbol{y}\in\mathbb{R}^d\setminus\{\boldsymbol{0}\}$ we define

$$X_q(\boldsymbol{y}):=\big\{M\in X_q\,:\,\boldsymbol{y}\in(\mathbb{Z}^d+\boldsymbol{\alpha})M\big\}. \tag{7.10}$$

This set can be given the structure of an embedded submanifold in X_q of codimension d, and with a countably infinite number of connected components. To see this we first note that $X_q(\boldsymbol{y})=\bigcup_{\boldsymbol{k}\in\mathbb{Z}^d+\boldsymbol{\alpha}\setminus\{\boldsymbol{0}\}}X_q(\boldsymbol{k},\boldsymbol{y})$, where

$$X_q(\boldsymbol{k},\boldsymbol{y}):=\big\{\Gamma(q)M\in X_q\,:\,M\in\mathrm{SL}(d,\mathbb{R}),\,\boldsymbol{k}M=\boldsymbol{y}\big\}. \tag{7.11}$$

One checks that for any $\boldsymbol{k},\boldsymbol{k}'\in\mathbb{Z}^d+\boldsymbol{\alpha}\setminus\{\boldsymbol{0}\}$, we have $X_q(\boldsymbol{k},\boldsymbol{y})=X_q(\boldsymbol{k}',\boldsymbol{y})$ if $\boldsymbol{k}'\in\boldsymbol{k}\Gamma(q)$; whereas $X_q(\boldsymbol{k},\boldsymbol{y})\cap X_q(\boldsymbol{k}',\boldsymbol{y})=\varnothing$ whenever $\boldsymbol{k}'\notin\boldsymbol{k}\Gamma(q)$. Hence if S is any subset of $\mathbb{Z}^d+\boldsymbol{\alpha}\setminus\{\boldsymbol{0}\}$ containing exactly one representative from each orbit of the right action of $\Gamma(q)$ on $\mathbb{Z}^d+\boldsymbol{\alpha}\setminus\{\boldsymbol{0}\}$, then we can express $X_q(\boldsymbol{y})$ as a disjoint union

$$X_q(\boldsymbol{y})=\bigsqcup_{\boldsymbol{k}\in S}X_q(\boldsymbol{k},\boldsymbol{y}). \tag{7.12}$$

To describe each $X_q(\boldsymbol{k}, \boldsymbol{y})$ further we set

(7.13)
$$H = \{g \in \mathrm{SL}(d,\mathbb{R}) : \boldsymbol{e}_1 g = \boldsymbol{e}_1\} = \left\{\begin{pmatrix} 1 & \mathbf{0} \\ {}^t\boldsymbol{v} & A \end{pmatrix} : \boldsymbol{v} \in \mathbb{R}^{d-1},\ A \in \mathrm{SL}(d-1,\mathbb{R})\right\}.$$

This is a closed subgroup of $\mathrm{SL}(d,\mathbb{R})$ which is isomorphic with $\mathrm{ASL}(d-1,\mathbb{R})$ (as defined in (2.1)) through

$$H \ni \begin{pmatrix} 1 & \mathbf{0} \\ {}^t\boldsymbol{v} & A \end{pmatrix} \mapsto ({}^tA^{-1}, \boldsymbol{v}\,{}^tA^{-1}) \in \mathrm{ASL}(d-1,\mathbb{R}). \tag{7.14}$$

We let μ_H be the Haar measure on H given by $d\mu_H(g) = d\mu_1^{(d-1)}(A)\,d\boldsymbol{v}$, with $A, \boldsymbol{v}$ as in (7.13), $\mu_1^{(d-1)}$ the Haar measure on $\mathrm{SL}(d-1,\mathbb{R})$ from (7.1), and $d\boldsymbol{v}$ the standard Lebesgue measure on $\mathbb{R}^{d-1}$. In particular, in dimension $d = 2$ we have $H = \left\{\left(\begin{smallmatrix} 1 & 0 \\ v & 1 \end{smallmatrix}\right) : v \in \mathbb{R}\right\}$ and we set $d\mu_H = dv$.

Now fix some $M_{\boldsymbol{k}}$, $M_{\boldsymbol{y}} \in \mathrm{SL}(d,\mathbb{R})$ such that $\boldsymbol{k} = \boldsymbol{e}_1 M_{\boldsymbol{k}}$, $\boldsymbol{y} = \boldsymbol{e}_1 M_{\boldsymbol{y}}$. Then $X_q(\boldsymbol{k}, \boldsymbol{y})$ is the image of $M_{\boldsymbol{k}}^{-1} H M_{\boldsymbol{y}} \subset \mathrm{SL}(d,\mathbb{R})$ under the standard projection $\pi : \mathrm{SL}(d,\mathbb{R}) \to X_q$, and $h_1, h_2 \in H$ give the same point $\pi(M_{\boldsymbol{k}}^{-1} h_1 M_{\boldsymbol{y}}) = \pi(M_{\boldsymbol{k}}^{-1} h_2 M_{\boldsymbol{y}})$ if and only if h_1 and h_2 belong to the same left $(M_{\boldsymbol{k}}\Gamma(q)M_{\boldsymbol{k}}^{-1} \cap H)$-coset. This gives an identification of sets

$$X_q(\boldsymbol{k}, \boldsymbol{y}) = M_{\boldsymbol{k}}^{-1}\Big(\big(M_{\boldsymbol{k}}\Gamma(q)M_{\boldsymbol{k}}^{-1} \cap H\big)\backslash H\Big)M_{\boldsymbol{y}}. \tag{7.15}$$

Since $M_{\boldsymbol{k}}\Gamma(q)M_{\boldsymbol{k}}^{-1} \cap H$ is a discrete subgroup of H, the quotient space

$$\big(M_{\boldsymbol{k}}\Gamma(q)M_{\boldsymbol{k}}^{-1} \cap H\big)\backslash H$$

is a connected (d^2-d-1)-dimensional manifold, and hence $X_q(\boldsymbol{k}, \boldsymbol{y})$ inherits a natural structure as a connected (d^2-d-1)-dimensional manifold. One verifies that this structure does not depend on the choice of $M_{\boldsymbol{y}}$ or $M_{\boldsymbol{k}}$ (since left or right multiplication by any fixed H-element gives a diffeomorphism of H). Since the map $H \ni h \mapsto M_{\boldsymbol{k}}^{-1} h M_{\boldsymbol{y}} \in \mathrm{SL}(d,\mathbb{R})$ is an immersion we see that $X_q(\boldsymbol{k}, \boldsymbol{y})$ is a connected immersed submanifold of X_q. Hence since the union (7.12) is disjoint we have now given $X_q(\boldsymbol{y})$ a structure as an immersed submanifold of X_q with a countably infinite number of connected components. ($X_q(\boldsymbol{y})$ is even an embedded submanifold of X_q, but we will not need this fact.)

Note that the measure μ_H induces a Borel measure on each quotient space $\big(M_{\boldsymbol{k}}\Gamma(q)M_{\boldsymbol{k}}^{-1} \cap H\big)\backslash H$, which we also call μ_H. *We endow $X_q(\boldsymbol{y})$ with the Borel measure $\nu_{\boldsymbol{y}}$ defined on each $X_q(\boldsymbol{k}, \boldsymbol{y})$ as coming from $(I_q\zeta(d))^{-1}\mu_H$ under the identification* (7.15). This measure $\nu_{\boldsymbol{y}}$ is independent of the choices of S and matrices $M_{\boldsymbol{k}}$, $M_{\boldsymbol{y}}$, as is easily seen from the fact that μ_H is both left and right invariant.

LEMMA 7.2. *For any $\boldsymbol{y} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, $T \in \mathrm{SL}(d,\mathbb{R})$ and any Borel subset $\mathscr{E} \subset X_q(\boldsymbol{y})$ we have $\nu_{\boldsymbol{y}}(\mathscr{E}) = \nu_{\boldsymbol{y}T}(\mathscr{E}T)$.*

Proof. For any given subset $\mathscr{E}' \subset X_q(\boldsymbol{k}, \boldsymbol{y})$ we have $\mathscr{E}'T \subset X_q(\boldsymbol{k}, \boldsymbol{y}T)$, and if we choose $M_{\boldsymbol{y}T} = M_{\boldsymbol{y}}T$ in the above definitions then these two subsets correspond to exactly the same subset of $(M_{\boldsymbol{k}}\Gamma(q)M_{\boldsymbol{k}}^{-1} \cap H)\backslash H$ under the identification(s) (7.15). The lemma follows trivially from this. □

PROPOSITION 7.3. *Let $\mathscr{E} \subset X_q$ be any Borel set; then $\boldsymbol{y} \mapsto \nu_{\boldsymbol{y}}(\mathscr{E} \cap X_q(\boldsymbol{y}))$ is a measurable function of $\boldsymbol{y} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$. If $U \subset \mathbb{R}^d \setminus \{\mathbf{0}\}$ is any Borel set such that $\mathscr{E} \subset \bigcup_{\boldsymbol{y}\in U} X_q(\boldsymbol{y})$, then*

$$\mu_q(\mathscr{E}) \le \int_U \nu_{\boldsymbol{y}}(\mathscr{E} \cap X_q(\boldsymbol{y}))\, d\boldsymbol{y}. \tag{7.16}$$

Furthermore, if for all $\boldsymbol{y}_1 \ne \boldsymbol{y}_2 \in U : X_q(\boldsymbol{y}_1) \cap X_q(\boldsymbol{y}_2) \cap \mathscr{E} = \varnothing$, then equality holds in (7.16).

The following lemma will be required for the proof.

LEMMA 7.4. *For each $\boldsymbol{y} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, choose some $M_{\boldsymbol{y}} \in \mathrm{SL}(d,\mathbb{R})$ with $\boldsymbol{e}_1 M_{\boldsymbol{y}} = \boldsymbol{y}$. Then for every $f \in \mathrm{L}^1(\mathrm{SL}(d,\mathbb{R}), \mu_q)$ we have*

$$\int_{\mathrm{SL}(d,\mathbb{R})} f(M)\, d\mu_q(M) = \frac{1}{I_q\zeta(d)} \int_{\mathbb{R}^d\setminus\{\mathbf{0}\}} \Big(\int_H f(hM_{\boldsymbol{y}})\, d\mu_H(h) \Big)\, d\boldsymbol{y}. \tag{7.17}$$

Proof. First of all the integral $\int_H f(hM_{\boldsymbol{y}})\, d\mu_H(h)$ (if it exists) only depends on f and $\boldsymbol{y}$, and not on the choice of $M_{\boldsymbol{y}}$, since for a given $\boldsymbol{y}$ the matrix $M_{\boldsymbol{y}}$ is uniquely determined up to left multiplication by an element from H, and μ_H is right H-invariant. Hence we may fix the following specific choices of $M_{\boldsymbol{y}}$, for $\boldsymbol{y} = (y_1, \dots, y_d)$ with $y_1 > 0$:

$$M_{\boldsymbol{y}} = M_{\boldsymbol{y}}^{(0)} := \begin{pmatrix} y_1 & \boldsymbol{y}' \\ {}^{\mathrm{t}}\mathbf{0} & y_1^{-\frac{1}{d-1}} 1_{d-1} \end{pmatrix} \quad \text{with } \boldsymbol{y}' = (y_2, \dots, y_d); \tag{7.18}$$

and for $\boldsymbol{y} = (y_1, \dots, y_d)$ with $y_1 < 0$:

$$M_{\boldsymbol{y}} = M_{\boldsymbol{y}K_0}^{(0)} K_0, \qquad \text{where } K_0 = \mathrm{diag}[-1, -1, 1, \dots, 1] \in \mathrm{SL}(d,\mathbb{R}). \tag{7.19}$$

We may leave $M_{\boldsymbol{y}}$ unspecified when $y_1 = 0$, since these $\boldsymbol{y}$'s form a subset of $\mathbb{R}^d \setminus \{\mathbf{0}\}$ of Lebesgue measure zero.

Write $G = \mathrm{SL}(d,\mathbb{R})$, $G^+ = \{(m_{jk}) \in G : m_{11} > 0\}$ and $G^- = \{(m_{jk}) \in G : m_{11} < 0\}$. Then the map $\langle h, \boldsymbol{y}\rangle \mapsto M = hM_{\boldsymbol{y}}^{(0)}$ gives a diffeomorphism from $H \times \{\boldsymbol{y} \in \mathbb{R}^d : y_1 > 0\}$ onto G^+ (indeed, the inverse is easily computed explicitly and seen to be smooth). Furthermore, in this parametrization we have, via a standard computation similar to, e.g., [21, (3.70), case $r = 1$], $d\mu_q(M) = (I_q\zeta(d))^{-1} d\mu_H(h) d\boldsymbol{y}$. Hence

$$\int_{G^+} f(M)\, d\mu_q(M) = (I_q\zeta(d))^{-1} \int_{\{\boldsymbol{y}\in\mathbb{R}^d\,:\,y_1>0\}} \int_H f(hM_{\boldsymbol{y}})\, d\mu_H(h)\, d\boldsymbol{y}.$$

Similarly one verifies

$$\int_{G^-} f(M)\,d\mu_q(M) = (I_q\zeta(d))^{-1}\int_{\{\boldsymbol{y}\in\mathbb{R}^d\,:\,y_1<0\}}\int_H f(hM_{\boldsymbol{y}})\,d\mu_H(h)\,d\boldsymbol{y},$$

and (7.17) follows by addition of these two. □

Proof of Proposition 7.3. Let $\mathcal{F}\subset \mathrm{SL}(d,\mathbb{R})$ be a (measurable) fundamental domain for $\Gamma(q)\backslash \mathrm{SL}(d,\mathbb{R})$, in the set theoretic sense. That is, we assume $\mathcal{F}\cap\gamma\mathcal{F} = \varnothing$ for all $\gamma\in\Gamma(q)$, and $\bigcup_{\gamma\in\Gamma(q)}\gamma\mathcal{F} = \mathrm{SL}(d,\mathbb{R})$. For each $\boldsymbol{y}\in\mathbb{R}^d\setminus\{\boldsymbol{0}\}$, fix some $M_{\boldsymbol{y}}\in\mathrm{SL}(d,\mathbb{R})$ with $\boldsymbol{e}_1 M_{\boldsymbol{y}} = \boldsymbol{y}$. Now for any $\boldsymbol{y}\in\mathbb{R}^d\setminus\{\boldsymbol{0}\}$ we have, via (7.12), (7.15) and the definition of $\nu_{\boldsymbol{y}}$,

$$\nu_{\boldsymbol{y}}(\mathscr{E}\cap X_q(\boldsymbol{y})) = (I_q\zeta(d))^{-1}\sum_{\boldsymbol{k}\in S}\int_{\mathcal{F}_1}\chi_{\mathscr{E}_0}\big(M_{\boldsymbol{k}}^{-1}hM_{\boldsymbol{y}}\big)\,d\mu_H(h), \tag{7.20}$$

where $\mathcal{F}_1\subset H$ is any fixed fundamental domain for $(M_{\boldsymbol{k}}\Gamma(q)M_{\boldsymbol{k}}^{-1}\cap H)\backslash H$, $\mathscr{E}_0$ denotes the pre-image in $\mathrm{SL}(d,\mathbb{R})$ of $\mathscr{E}\subset X_q$, and $\chi_{\mathscr{E}_0}$ is its characteristic function. We may choose $\mathcal{F}_1 = H\cap\mathcal{F}_2$ where $\mathcal{F}_2\subset\mathrm{SL}(d,\mathbb{R})$ is any fixed fundamental domain for $(M_{\boldsymbol{k}}\Gamma(q)M_{\boldsymbol{k}}^{-1}\cap H)\backslash\mathrm{SL}(d,\mathbb{R})$, and such an $\mathcal{F}_2$ may be fixed as $\mathcal{F}_2 = M_{\boldsymbol{k}}\big(\bigsqcup_{\gamma\in S^{(\boldsymbol{k})}}\gamma\mathcal{F}\big)M_{\boldsymbol{y}}^{-1}$, where $S^{(\boldsymbol{k})}\subset\Gamma(q)$ is any set of coset representatives for $(\Gamma(q)\cap M_{\boldsymbol{k}}^{-1}HM_{\boldsymbol{k}})\backslash\Gamma(q)$. Hence, since $\mathscr{E}_0\subset\mathrm{SL}(d,\mathbb{R})$ is left $\Gamma(q)$ invariant,

$$\nu_{\boldsymbol{y}}(\mathscr{E}\cap X_q(\boldsymbol{y})) = (I_q\zeta(d))^{-1}\sum_{\boldsymbol{k}\in S}\sum_{\gamma\in S^{(\boldsymbol{k})}}\int_H\chi_{\mathcal{F}\cap\mathscr{E}_0}\big(\gamma^{-1}M_{\boldsymbol{k}}^{-1}hM_{\boldsymbol{y}}\big)\,d\mu_H(h). \tag{7.21}$$

But for each $\boldsymbol{k}\in S$ and $\gamma\in S^{(\boldsymbol{k})}$ we have $\boldsymbol{e}_1M_{\boldsymbol{k}}\gamma = \boldsymbol{k}\gamma = \boldsymbol{e}_1M_{\boldsymbol{k}\gamma}$ and thus $M_{\boldsymbol{k}}\gamma = h_0M_{\boldsymbol{k}\gamma}$ for some $h_0\in H$; hence using the invariance of μ_H we see that we may replace $\gamma^{-1}M_{\boldsymbol{k}}^{-1}$ with $M_{\boldsymbol{k}\gamma}^{-1}$ inside the integrand. Furthermore, given $\gamma,\gamma'\in\Gamma(q)$ we have the following chain of equivalent statements:

$$\begin{aligned}(\Gamma(q)\cap M_{\boldsymbol{k}}^{-1}HM_{\boldsymbol{k}})\gamma = (\Gamma(q)\cap M_{\boldsymbol{k}}^{-1}HM_{\boldsymbol{k}})\gamma' &\iff \gamma'\gamma^{-1}\in M_{\boldsymbol{k}}^{-1}HM_{\boldsymbol{k}}\\ &\iff \boldsymbol{e}_1M_{\boldsymbol{k}}\gamma'\gamma^{-1}M_{\boldsymbol{k}}^{-1} = \boldsymbol{e}_1 \iff \boldsymbol{k}\gamma' = \boldsymbol{k}\gamma.\end{aligned} \tag{7.22}$$

Hence by the definition of S and $S^{(\boldsymbol{k})}$, as $\boldsymbol{k}$ and γ run through the double sum in (7.21), $\boldsymbol{k}\gamma$ visits each vector in $\mathbb{Z}^d+\boldsymbol{\alpha}\setminus\{\boldsymbol{0}\}$ exactly once. Hence

$$\nu_{\boldsymbol{y}}(\mathscr{E}\cap X_q(\boldsymbol{y})) = (I_q\zeta(d))^{-1}\sum_{\boldsymbol{k}\in\mathbb{Z}^d+\boldsymbol{\alpha}\setminus\{\boldsymbol{0}\}}\int_H\chi_{\mathcal{F}\cap\mathscr{E}_0}\big(M_{\boldsymbol{k}}^{-1}hM_{\boldsymbol{y}}\big)\,d\mu_H(h). \tag{7.23}$$

Here for each $\boldsymbol{k}$ the function $\mathbb{R}^d\setminus\{\boldsymbol{0}\}\ni\boldsymbol{y}\mapsto\int_H\chi_{M_{\boldsymbol{k}}(\mathcal{F}\cap\mathscr{E}_0)}(hM_{\boldsymbol{y}})\,d\mu_H(h)$ is measurable (this is implicit in Lemma 7.4); hence also the above sum (7.23) is measurable as a function from $\boldsymbol{y}\in\mathbb{R}^d\setminus\{\boldsymbol{0}\}$ into $\mathbb{R}_{\geq 0}\cup\{\infty\}$.

Now to prove (7.16), note that the assumption on U implies

$$\mathscr{E}_0 = \bigcup_{\boldsymbol{k}\in\mathbb{Z}^d+\boldsymbol{\alpha}\setminus\{\mathbf{0}\}} \mathscr{E}_{\boldsymbol{k}},$$

where $\mathscr{E}_{\boldsymbol{k}} := \{M \in \mathscr{E}_0 : \boldsymbol{k}M \in U\}$. We have by (7.23),

$$\begin{aligned}(7.24)\quad &\int_U \nu_{\boldsymbol{y}}(\mathscr{E}\cap X_q(\boldsymbol{y}))\,d\boldsymbol{y}\\ &\qquad = (I_q\zeta(d))^{-1}\sum_{\boldsymbol{k}\in\mathbb{Z}^d+\boldsymbol{\alpha}\setminus\{\mathbf{0}\}}\int_U\int_H \chi_{\mathscr{F}\cap\mathscr{E}_0}\big(M_{\boldsymbol{k}}^{-1}hM_{\boldsymbol{y}}\big)\,d\mu_H(h)\,d\boldsymbol{y},\end{aligned}$$

and for any $\boldsymbol{k}, \boldsymbol{y}, h$ appearing in the above expression we have $\boldsymbol{k}(M_{\boldsymbol{k}}^{-1}hM_{\boldsymbol{y}}) = \boldsymbol{y} \in U$, so that $M_{\boldsymbol{k}}^{-1}hM_{\boldsymbol{y}} \in \mathscr{E}_{\boldsymbol{k}}$ must hold whenever $M_{\boldsymbol{k}}^{-1}hM_{\boldsymbol{y}} \in \mathscr{E}_0$. Also for every $\boldsymbol{y} \in \mathbb{R}^d\setminus(U\cup\{\mathbf{0}\})$ we have $\boldsymbol{k}(M_{\boldsymbol{k}}^{-1}hM_{\boldsymbol{y}}) = \boldsymbol{y} \notin U$, so that $M_{\boldsymbol{k}}^{-1}hM_{\boldsymbol{y}} \notin \mathscr{E}_{\boldsymbol{k}}$. Hence

$$\begin{aligned}(7.25)\quad &\int_U \nu_{\boldsymbol{y}}(\mathscr{E}\cap X_q(\boldsymbol{y}))\,d\boldsymbol{y}\\ &= (I_q\zeta(d))^{-1}\sum_{\boldsymbol{k}\in\mathbb{Z}^d+\boldsymbol{\alpha}\setminus\{\mathbf{0}\}}\int_{\mathbb{R}^d\setminus\{\mathbf{0}\}}\int_H \chi_{\mathscr{F}\cap\mathscr{E}_{\boldsymbol{k}}}\big(M_{\boldsymbol{k}}^{-1}hM_{\boldsymbol{y}}\big)\,d\mu_H(h)\,d\boldsymbol{y}\\ &= \sum_{\boldsymbol{k}\in\mathbb{Z}^d+\boldsymbol{\alpha}\setminus\{\mathbf{0}\}}\mu_q\big(M_{\boldsymbol{k}}(\mathscr{F}\cap\mathscr{E}_{\boldsymbol{k}})\big)\\ &= \sum_{\boldsymbol{k}\in\mathbb{Z}^d+\boldsymbol{\alpha}\setminus\{\mathbf{0}\}}\mu_q\big(\mathscr{F}\cap\mathscr{E}_{\boldsymbol{k}}\big) \ge \mu_q\big(\mathscr{F}\cap\mathscr{E}_0\big) = \mu_q(\mathscr{E}),\end{aligned}$$

where we used Lemma 7.4, the invariance of μ_q, and $\mathscr{E}_0 = \bigcup_{\boldsymbol{k}\in\mathbb{Z}^d+\boldsymbol{\alpha}\setminus\{\mathbf{0}\}}\mathscr{E}_{\boldsymbol{k}}$. (To avoid any confusion in the last step: Recall that we use μ_q to denote both a Haar measure on $\mathrm{SL}(d,\mathbb{R})$ and the corresponding probability measure on X_q.) Hence (7.16) is proved. To prove the final statement about equality, note that the condition $\forall\,\boldsymbol{y}_1 \ne \boldsymbol{y}_2 \in U : X_q(\boldsymbol{y}_1)\cap X_q(\boldsymbol{y}_2)\cap\mathscr{E} = \varnothing$ implies that the sets $\mathscr{E}_{\boldsymbol{k}}$ are pairwise disjoint, and thus $\sum_{\boldsymbol{k}\in\mathbb{Z}^d+\boldsymbol{\alpha}\setminus\{\mathbf{0}\}}\mu_q\big(\mathscr{F}\cap\mathscr{E}_{\boldsymbol{k}}\big) = \mu_q(\mathscr{E})$. □

PROPOSITION 7.5. *For each $\boldsymbol{y} \in \mathbb{R}^d\setminus\{\mathbf{0}\}$ we have $\nu_{\boldsymbol{y}}(X_q(\boldsymbol{y})) = 1$.*

Proof. Let us write $\boldsymbol{\alpha} = \frac{\boldsymbol{p}}{q}$ with $\boldsymbol{p} = (p_1,\dots,p_d) \in \mathbb{Z}^d$. We first show that without loss of generality we may assume $\gcd(q, p_1, \dots, p_d) = 1$, i.e. that q is the minimal denominator of the given vector $\boldsymbol{\alpha} \in \mathbb{Q}^d$. Indeed, any *other* denominator of $\boldsymbol{\alpha}$ can be written as $q' = qq_1$, with $q_1 \in \mathbb{Z}_{\ge 2}$; for each such q' there is a canonical covering map $\pi : X_{q'} \to X_q$ of index $[\Gamma(q') : \Gamma(q)] = I_{q'}/I_q$, and it follows straight from the definition (7.10) that $X_{q'}(\boldsymbol{y}) = \pi^{-1}(X_q(\boldsymbol{y}))$, i.e. $X_{q'}(\boldsymbol{y})$ is a covering of the manifold $X_q(\boldsymbol{y})$ of index $[\Gamma(q') : \Gamma(q)]$. Furthermore, the measure $\nu_{\boldsymbol{y}}^{(q)}$ on $X_q(\boldsymbol{y})$ lifts to $[\Gamma(q') : \Gamma(q)]\nu_{\boldsymbol{y}}^{(q')}$ on $X_{q'}(\boldsymbol{y})$ (in an obvious notation). Hence if $\nu_{\boldsymbol{y}}^{(q)}(X_q(\boldsymbol{y})) = 1$ then also $\nu_{\boldsymbol{y}}^{(q')}(X_{q'}(\boldsymbol{y})) = 1$, as desired.

Thus from now on we assume $\gcd(q, p_1, \dots, p_d) = 1$. By (7.12) and (7.15),

$$\nu_{\boldsymbol{y}}(X_q(\boldsymbol{y})) = (I_q\zeta(d))^{-1} \sum_{\boldsymbol{k}\in S} \mu_H\big((M_{\boldsymbol{k}}\Gamma(q)M_{\boldsymbol{k}}^{-1} \cap H)\backslash H\big). \tag{7.26}$$

Given $\boldsymbol{k} = (k_1, \dots, k_d) \in \mathbb{Z}^d + \boldsymbol{\alpha} \setminus \{\boldsymbol{0}\}$, set $t_{\boldsymbol{k}} := \gcd(qk_1, qk_2, \dots, qk_d) \in \mathbb{Z}_{>0}$. Then $(q/t_{\boldsymbol{k}})\boldsymbol{k}$ is a primitive vector in $\mathbb{Z}^d$, and hence there is some $\gamma \in \Gamma(1)$ so that $(q/t_{\boldsymbol{k}})\boldsymbol{k} = \boldsymbol{e}_1\gamma$. For each $\delta > 0$ we define $g_\delta = \operatorname{diag}[\delta, \delta^{-1}, 1, \dots, 1] \in \mathrm{SL}(d, \mathbb{R})$. Then we may choose $M_{\boldsymbol{k}}$ as $M_{\boldsymbol{k}} := g_{t_{\boldsymbol{k}}/q}\gamma$ (since this gives $\boldsymbol{e}_1 M_{\boldsymbol{k}} = \boldsymbol{k}$). With this choice we have $M_{\boldsymbol{k}}\Gamma(q)M_{\boldsymbol{k}}^{-1} = g_{t_{\boldsymbol{k}}/q}\Gamma(q)g_{t_{\boldsymbol{k}}/q}^{-1}$, since $\Gamma(q)$ is normal in $\Gamma(1)$. Note that $\alpha : H \ni h \mapsto g_{t_{\boldsymbol{k}}/q}hg_{t_{\boldsymbol{k}}/q}^{-1} \in H$ gives an automorphism of H, and hence $M_{\boldsymbol{k}}\Gamma(q)M_{\boldsymbol{k}}^{-1} \cap H = \alpha(\Gamma(q) \cap H)$. Furthermore, one verifies by a quick computation that α scales the Haar measure with a factor $(q/t_{\boldsymbol{k}})^d$, i.e. $\mu_H(\alpha(A)) = (q/t_{\boldsymbol{k}})^d\mu_H(A)$ for any measurable $A \subset H$. Hence

$$\nu_{\boldsymbol{y}}(X_q(\boldsymbol{y})) = \frac{q^d}{I_q\zeta(d)} \sum_{\boldsymbol{k}\in S} t_{\boldsymbol{k}}^{-d}\mu_H\big((\Gamma(q) \cap H)\backslash H\big). \tag{7.27}$$

For each $\boldsymbol{k} = (k_1, \dots, k_d) \in \mathbb{Z}^d + \boldsymbol{\alpha} \setminus \{\boldsymbol{0}\}$ we have $(t_{\boldsymbol{k}}, q) = 1$, since $q\boldsymbol{k} \in \boldsymbol{p} + q\mathbb{Z}^d$ and $\gcd(q, p_1, \dots, p_d) = 1$. On the other hand, given any $t \in \mathbb{Z}_{>0}$ with $(t, q) = 1$ we may choose $t^* \in \mathbb{Z}$ so that $t^*t \equiv 1 \bmod q$; then $\gcd(q, t^*p_1, \dots, t^*p_d) = 1$ since $(q, t^*) = 1$, and thus there exists some primitive vector $\boldsymbol{m}$ in $t^*\boldsymbol{p} + q\mathbb{Z}^d$, [4] and then $\boldsymbol{k} = (t/q)\boldsymbol{m} \in \mathbb{Z}^d + \boldsymbol{\alpha} \setminus \{\boldsymbol{0}\}$ has $t_{\boldsymbol{k}} = t$. Furthermore, we claim that $t_{\boldsymbol{k}_1} = t_{\boldsymbol{k}_2}$ holds if and only if $\boldsymbol{k}_1\Gamma(q) = \boldsymbol{k}_2\Gamma(q)$. To prove the nontrivial direction of this claim we assume that $\boldsymbol{k}_1, \boldsymbol{k}_2 \in \mathbb{Z}^d + \boldsymbol{\alpha} \setminus \{\boldsymbol{0}\}$ have $t := t_{\boldsymbol{k}_1} = t_{\boldsymbol{k}_2}$. Then $(q/t)\boldsymbol{k}_j$ is a primitive vector in $\mathbb{Z}^d$ and hence there are some $\gamma_1, \gamma_2 \in \Gamma(1)$ with $(q/t)\boldsymbol{k}_j = \boldsymbol{e}_1\gamma_j$. Now both vectors $\boldsymbol{e}_1\gamma_j$ belong to $t^*\boldsymbol{p} + q\mathbb{Z}^d$ with t^* as before; hence $\boldsymbol{e}_1\gamma_1\gamma_2^{-1} \equiv \boldsymbol{e}_1 \bmod q\mathbb{Z}^d$, so that $\gamma_1\gamma_2^{-1} = \left(\begin{smallmatrix} x_1 & \boldsymbol{x}' \\ {}^t\boldsymbol{v} & A \end{smallmatrix}\right)$ with $x_1 \equiv 1 \bmod q$ and $\boldsymbol{x}' \in q\mathbb{Z}^{d-1}$. Reducing mod q we also see that $A \bmod q$ lies in $\mathrm{SL}(d-1, \mathbb{Z}/q\mathbb{Z})$; hence there is some $A' \in \mathrm{SL}(d-1, \mathbb{Z})$ so that $A' \equiv A \bmod q$ [32, p. 21]. Now $\left(\begin{smallmatrix} 1 & \boldsymbol{0} \\ {}^t\boldsymbol{v} & A' \end{smallmatrix}\right) \in \Gamma(1)$, and this matrix has the same projection as $\gamma_1\gamma_2^{-1}$ in $\mathrm{SL}(d, \mathbb{Z}/q\mathbb{Z}) \cong \Gamma(1)/\Gamma(q)$. Hence $\gamma_0 := \gamma_1^{-1}\left(\begin{smallmatrix} 1 & \boldsymbol{0} \\ {}^t\boldsymbol{v} & A' \end{smallmatrix}\right)\gamma_2$ belongs to $\Gamma(q)$, and we see that $\boldsymbol{e}_1\gamma_1\gamma_0 = \boldsymbol{e}_1\gamma_2$, and thus $\boldsymbol{k}_1\Gamma(q) = \boldsymbol{k}_2\Gamma(q)$, as desired.

It follows that (7.27) may be rewritten as

$$\nu_{\boldsymbol{y}}(X_q(\boldsymbol{y})) = \frac{q^d\,\mu_H\big((\Gamma(q) \cap H)\backslash H\big)}{I_q\zeta(d)} \sum_{\substack{t\geq 1\\(t,q)=1}} t^{-d}. \tag{7.28}$$

But note $\mu_H\big((\Gamma(q)\cap H)\backslash H\big) = \#\big((\Gamma(q)\cap H)\backslash(\Gamma(1)\cap H)\big)\cdot\mu_H\big((\Gamma(1)\cap H)\backslash H\big)$, where the second factor equals one by the definition of μ_H, and the first factor is seen to equal $\#H(\mathbb{Z}/q\mathbb{Z})$ with $H(\mathbb{Z}/q\mathbb{Z}) := \left\{\left(\begin{smallmatrix} 1 & \boldsymbol{0} \\ {}^t\boldsymbol{v} & A \end{smallmatrix}\right) \in \mathrm{SL}(d, \mathbb{Z}/q\mathbb{Z})\right\}$ (for this

[4] This can for example be shown using Dirichlet's theorem on arithmetic progressions, for by that theorem we may find $m_j \in t^*p_j + q\mathbb{Z}$, $j = 1, \dots, d$ such that $m_j = r_j \gcd(p_j, q)$ with prime numbers $q < r_1 < r_2 < \cdots < r_d$; then $\boldsymbol{m} = (m_1, \dots, m_d)$ lies in $t^*\boldsymbol{p} + q\mathbb{Z}^d$ and is primitive.

one uses the surjectivity of the map $\mathrm{SL}(d-1,\mathbb{Z}) \to \mathrm{SL}(d-1,\mathbb{Z}/q\mathbb{Z})$; cf. [32, p. 21]). Also note that we have a decomposition of $\mathrm{SL}(d,\mathbb{Z}/q\mathbb{Z})$ analogous to the decomposition of $\mathrm{SL}(d,\mathbb{R})$ in the proof of Lemma 7.4: Take $V \subset \mathbb{Z}^d/q\mathbb{Z}^d$ to be as in (7.8). For each $\boldsymbol{y} \in V$ we fix a matrix $M_{\boldsymbol{y}} \in \mathrm{SL}(d,\mathbb{Z}/q\mathbb{Z})$ whose first row equals $\boldsymbol{y}$. (Such a matrix exists, for we may lift $y_1,\dots,y_d$ to integers satisfying $\gcd(y_1,\dots,y_d)=1$, cf. footnote 4 above, and then apply [34, VIII.1–2].) One then verifies that the map $H(\mathbb{Z}/q\mathbb{Z}) \times V \ni \langle h, \boldsymbol{y}\rangle \mapsto hM_{\boldsymbol{y}} \in \mathrm{SL}(d,\mathbb{Z}/q\mathbb{Z})$ is a bijection. Hence $I_q = \#\mathrm{SL}(d,\mathbb{Z}/q\mathbb{Z}) = \#H(\mathbb{Z}/q\mathbb{Z}) \cdot \#V$. Finally recall that $\#V = q^d \sum_{e|q} \mu(e) e^{-d}$ and $\sum_{t\ge 1,\,(t,q)=1} t^{-d} = \zeta(d) \sum_{e|q} \mu(e) e^{-d}$. Hence we obtain

$$\nu_{\boldsymbol{y}}(X_q(\boldsymbol{y})) = \frac{q^d \,\#H(\mathbb{Z}/q\mathbb{Z})}{\#V \cdot \#H(\mathbb{Z}/q\mathbb{Z})} \sum_{e|q} \mu(e) e^{-d} = 1. \tag{7.29}$$ □

We next prove an analogue for $X_q(\boldsymbol{y})$ of Siegel's lattice average formula, Proposition 7.1.

PROPOSITION 7.6. *Assume $d \ge 3$ and $\boldsymbol{\alpha} = \frac{\boldsymbol{p}}{q}$ with $\boldsymbol{p} = (p_1,\dots,p_d) \in \mathbb{Z}^d$ and $\gcd(q, p_1,\dots,p_d) = 1$. If $F : \mathbb{R}^d \to \mathbb{R}$ is a bounded measurable function of compact support, then for any $\boldsymbol{y} \in \mathbb{R}^d \setminus \{\boldsymbol{0}\}$ we have*

$$\begin{aligned} \int_{X_q(\boldsymbol{y})} \sum_{\boldsymbol{m}\in\mathbb{Z}^d} F((\boldsymbol{m}+\boldsymbol{\alpha})M)\, d\nu_{\boldsymbol{y}}(M) \\ = \int_{\boldsymbol{x}\in\mathbb{R}^d} F(\boldsymbol{x})\, d\boldsymbol{x} + \sum_{\substack{t\ge 1\\(t,q)=1}} t^{-d} \sum_{\substack{a\in t+q\mathbb{Z}\\(a,t)=1}} F\Big(\frac{a}{t}\boldsymbol{y}\Big), \end{aligned} \tag{7.30}$$

where all sums and integrals are absolutely convergent.

We require the following lemma.

LEMMA 7.7. *Let $F : \mathbb{R}^d \to \mathbb{R}$ be a bounded measurable function of compact support. If $d = 2$ then we furthermore require that $F(x\boldsymbol{e}_1 + m\boldsymbol{e}_2)$ is a measurable function of $x \in \mathbb{R}$ for each fixed $m \in \mathbb{Z}$. Let $\boldsymbol{\alpha} = (\alpha_1,\dots,\alpha_d)$ with $\alpha_1 \in \mathbb{R}$ and $\alpha_2,\dots,\alpha_d \in \mathbb{Z}$. Then*

$$\begin{aligned} &\int_{(\Gamma(q)\cap H)\backslash H} \sum_{\boldsymbol{m}\in\mathbb{Z}^d} F((\boldsymbol{m}+\boldsymbol{\alpha})M)\, d\mu_H(M) \\ &= q^{d-1} \begin{cases} I_q^{(d-1)}\big(\sum_{\ell\in\mathbb{Z}} F((\ell+\alpha_1)\boldsymbol{e}_1) + \int_{\boldsymbol{x}\in\mathbb{R}^d} F(\boldsymbol{x})\, d\boldsymbol{x}\big) & d \ge 3 \\ \sum_{\ell\in\mathbb{Z}} F((\ell+\alpha_1)\boldsymbol{e}_1) + \sum_{m\in\mathbb{Z}\setminus\{0\}} \int_{\mathbb{R}} F(x\boldsymbol{e}_1 + m\boldsymbol{e}_2)\, dx & d = 2. \end{cases} \end{aligned} \tag{7.31}$$

Proof of Lemma 7.7. The right-hand side in (7.31) is clearly absolutely convergent; it will be clear from the proof that the left-hand side is also absolutely convergent.

We first give the proof in the case $d \geq 3$. Write $\boldsymbol{\alpha} = (\alpha_1, \boldsymbol{\alpha}') \in \mathbb{R} \times \mathbb{Z}^{d-1}$ and express $M \in H$ as $M = \begin{pmatrix} 1 & \mathbf{0} \\ {}^{\mathrm{t}}\boldsymbol{v} & M_1 \end{pmatrix}$; then

$$(\mathbb{Z}^d + \boldsymbol{\alpha})M = \bigsqcup_{\ell \in \mathbb{Z}} \bigsqcup_{\boldsymbol{m} \in \mathbb{Z}^{d-1}} \Big((\ell + \alpha_1) + (\boldsymbol{m} + \boldsymbol{\alpha}')\,{}^{\mathrm{t}}\boldsymbol{v}, (\boldsymbol{m} + \boldsymbol{\alpha}')M_1 \Big), \tag{7.32}$$

and a fundamental domain for $(\Gamma(q) \cap H)\backslash H$ is given by

$$\big\{ M \in H \,:\, \boldsymbol{v} \in [0,q)^{d-1}, M_1 \in \mathscr{F} \big\},$$

where $\mathscr{F}$ is any fixed fundamental domain for $\Gamma(q)\backslash \mathrm{SL}(d-1,\mathbb{R})$. Set $F_1(x, \boldsymbol{y}) = \sum_{\ell \in \mathbb{Z}} F(x + \ell, \boldsymbol{y})$ where in the right-hand side we identify $\mathbb{R}^d$ with $\mathbb{R} \times \mathbb{R}^{d-1}$ in the obvious way. Since $\boldsymbol{\alpha}' \in \mathbb{Z}^{d-1}$, the integral in the left-hand side of (7.31) can now be expressed as

$$\int_{\mathscr{F}} \int_{[0,q)^{d-1}} \sum_{\boldsymbol{m} \in \mathbb{Z}^{d-1}} F_1(\alpha_1 + \boldsymbol{m}\,{}^{\mathrm{t}}\boldsymbol{v}, \boldsymbol{m}M_1)\, d\mu_1(M_1) d\boldsymbol{v}. \tag{7.33}$$

Note that $\int_{[0,q)} F_1(a + bx, \boldsymbol{y})\, dx = q \int_{\mathbb{R}/\mathbb{Z}} F_1(x, \boldsymbol{y})\, dx$ for any $a \in \mathbb{R}$ and $b \in \mathbb{Z}_{\neq 0}$. Thus, defining $F_2(\boldsymbol{y}) := \int_{\mathbb{R}/\mathbb{Z}} F_1(x, \boldsymbol{y})\, dx = \int_{\mathbb{R}} F(x, \boldsymbol{y})\, dx$ (so that $F_2(\boldsymbol{y})$ is defined for almost every $\boldsymbol{y} \in \mathbb{R}^{d-1}$, and the function F_2 is measurable, by Fubini's theorem), we have $\int_{\boldsymbol{v} \in [0,q)^{d-1}} F_1(\alpha_1 + \boldsymbol{m}\,{}^{\mathrm{t}}\boldsymbol{v}, \boldsymbol{y})\, d\boldsymbol{v} = q^{d-1} F_2(\boldsymbol{y})$ for each $\boldsymbol{m} \in \mathbb{Z}^{d-1} \setminus \{\mathbf{0}\}$, and hence (7.33) equals

$$q^{d-1} \int_{\mathscr{F}} \Big(F_1(\alpha_1, \mathbf{0}) + \sum_{\boldsymbol{m} \in \mathbb{Z}^{d-1} \setminus \{\mathbf{0}\}} F_2(\boldsymbol{m}M_1) \Big)\, d\mu_1(M_1). \tag{7.34}$$

The integrand in (7.34) only depends on the $\Gamma(1)$-coset of M_1, i.e. the integration over $\mathscr{F}$ may be replaced by $I_q^{(d-1)}$ times an integral over $\Gamma(1)\backslash \mathrm{SL}(d-1,\mathbb{R}) = X_1^{(d-1)}$; hence by Proposition 7.1 (applied for "$d-1$" and "$q = 1$") we get

$$= q^{d-1} I_q^{(d-1)} \Big(F_1(\alpha_1, \mathbf{0}) + \int_{\boldsymbol{y} \in \mathbb{R}^{d-1}} F_2(\boldsymbol{y})\, d\boldsymbol{y} \Big), \tag{7.35}$$

which gives the formula (7.31).

In the remaining case $d = 2$ we obtain as before (7.33) and (7.34), but with the inner integration sign removed and instead just taking $M_1 = 1$ in the formulas. Now (7.34) agrees with (7.31), and we are done. □

Proof of Proposition 7.6. We first prove the absolute convergence. It will be clear from the proof below that it suffices to prove that the right-hand side of (7.30) is absolutely convergent. This is clear for the integral; thus we turn to the double sum. Assume $|F(\boldsymbol{x})| \leq B$ for all $\boldsymbol{x} \in \mathbb{R}^d$ and take $C > 0$ such that $F(r\boldsymbol{y}) = 0$

whenever $r \le -C$ or $r \ge C$ (for our given $\boldsymbol{y} \in \mathbb{R}^d \setminus \{\boldsymbol{0}\}$). Then

$$\sum_{\substack{t\ge 1\\(t,q)=1}} t^{-d} \sum_{\substack{a\in t+q\mathbb{Z}\\(a,t)=1}} \Big| F\Big(\frac{a}{t}\boldsymbol{y}\Big)\Big| \le \sum_{t\ge 1} t^{-d}(1+2Cq^{-1}t)B. \tag{7.36}$$

This is obviously absolutely convergent, since $d \ge 3$.

We now prove the identity. In view of (7.12) and (7.15) the left-hand side of (7.30) decomposes as

$$(I_q\zeta(d))^{-1} \sum_{\boldsymbol{k}\in S} \int_{(M_{\boldsymbol{k}}\Gamma(q)\,\mathrm{M}_{\boldsymbol{k}}^{-1}\cap H)\backslash H} \sum_{\boldsymbol{m}\in\mathbb{Z}^d} F((\boldsymbol{m}+\boldsymbol{\alpha})M_{\boldsymbol{k}}^{-1}hM_{\boldsymbol{y}})\,d\mu_H(h). \tag{7.37}$$

For each fixed $\boldsymbol{k} \in S$ we now perform the same manipulations as in the proof of Proposition 7.5, just before (7.27); we thus get (since $\mathbb{Z}^d\gamma^{-1} = \mathbb{Z}^d$ for every $\gamma \in \Gamma(1)$)

$$= \frac{q^d}{I_q\zeta(d)} \sum_{\boldsymbol{k}\in S} t_{\boldsymbol{k}}^{-d} \int_{(\Gamma(q)\cap H)\backslash H} \sum_{\boldsymbol{m}\in\mathbb{Z}^d} F((\boldsymbol{m}+\boldsymbol{\alpha}\gamma_{\boldsymbol{k}}^{-1})hg_{t_{\boldsymbol{k}}/q}^{-1}M_{\boldsymbol{y}})\,d\mu_H(h), \tag{7.38}$$

where $\gamma_{\boldsymbol{k}}$ is any map in $\Gamma(1)$ with $(q/t_{\boldsymbol{k}})\boldsymbol{k} = \boldsymbol{e}_1\gamma_{\boldsymbol{k}}$. Now note (for each $\boldsymbol{k} \in S \subset \mathbb{Z}^d + \boldsymbol{\alpha} \setminus \{\boldsymbol{0}\}$) that $\boldsymbol{\alpha}\gamma_{\boldsymbol{k}}^{-1} \in (\boldsymbol{k}+\mathbb{Z}^d)\gamma_{\boldsymbol{k}}^{-1} = (t_{\boldsymbol{k}}/q)\boldsymbol{e}_1 + \mathbb{Z}^d$. Hence Lemma 7.7 applies, giving

$$\begin{aligned} &= \frac{q^{2d-1}I_q^{(d-1)}}{I_q^{(d)}\zeta(d)} \sum_{\boldsymbol{k}\in S} t_{\boldsymbol{k}}^{-d} \Big(\sum_{\ell\in\mathbb{Z}} F\Big(\Big(\ell+\frac{t_{\boldsymbol{k}}}{q}\Big)\boldsymbol{e}_1 g_{t_{\boldsymbol{k}}/q}^{-1}M_{\boldsymbol{y}}\Big) \\ &\qquad\qquad + \int_{\boldsymbol{x}\in\mathbb{R}^d} F\big(\boldsymbol{x}\,g_{t_{\boldsymbol{k}}/q}^{-1}M_{\boldsymbol{y}}\big)\,d\boldsymbol{x}\Big) \\ &= \frac{q^{2d-1}I_q^{(d-1)}}{I_q^{(d)}\zeta(d)} \sum_{\boldsymbol{k}\in S} t_{\boldsymbol{k}}^{-d} \Big(\sum_{\ell\in\mathbb{Z}} F\Big(\Big(\frac{\ell q}{t_{\boldsymbol{k}}}+1\Big)\boldsymbol{y}\Big) + \int_{\boldsymbol{x}\in\mathbb{R}^d} F\big(\boldsymbol{x}\big)\,d\boldsymbol{x}\Big). \end{aligned} \tag{7.39}$$

But we saw in the proof of Proposition 7.5 that when $\boldsymbol{k}$ runs through S then $t_{\boldsymbol{k}}$ visits each $t \in \mathbb{Z}_{>0}$ with $(t,q) = 1$ exactly once, and no other numbers. Also from that proof we have $I_q^{(d)} = \#H(\mathbb{Z}/q\mathbb{Z})\cdot\#V = \big(q^{d-1}I_q^{(d-1)}\big)\cdot q^d\sum_{e|q}\mu(e)e^{-d}$, and recall $\sum_{t\ge1,\,(t,q)=1} t^{-d} = \zeta(d)\sum_{e|q}\mu(e)e^{-d}$. Hence we get

$$= \int_{\boldsymbol{x}\in\mathbb{R}^d} F\big(\boldsymbol{x}\big)\,d\boldsymbol{x} + \Big(\sum_{t\ge1,\,(t,q)=1} t^{-d}\Big)^{-1} \sum_{t\ge1,\,(t,q)=1} t^{-d} \sum_{\ell\in\mathbb{Z}} F\Big(\big(\frac{\ell q}{t}+1\big)\boldsymbol{y}\Big). \tag{7.40}$$

In the last double sum we substitute $e = (\ell, t)$; thus $\ell = e\ell_1$, $t = et_1$ with $(\ell_1, t_1) = 1$, and the double sum becomes

$$\sum_{\substack{e\ge 1\\(e,q)=1}} \sum_{\substack{t_1\ge 1\\(t_1,q)=1}} (et_1)^{-d} \sum_{\substack{\ell_1\in\mathbb{Z}\\(\ell_1,t_1)=1}} F\Big(\big(\frac{\ell_1 q}{t_1}+1\big)\boldsymbol{y}\Big) = \Big(\sum_{\substack{e\ge 1\\(e,q)=1}} e^{-d}\Big) \sum_{\substack{t_1\ge 1\\(t_1,q)=1}} t_1^{-d} \sum_{\substack{a\in t_1+q\mathbb{Z}\\(a,t_1)=1}} F\Big(\frac{a}{t_1}\boldsymbol{y}\Big). \tag{7.41}$$

Hence we obtain the desired formula. □

We next turn to the case $d = 2$. In this case the integral in the left-hand side of (7.30) typically *diverges*. This is e.g. true for every continuous function $F \ge 0$ which is not identically zero along the line $\mathbb{R}\boldsymbol{y}$, as is seen by following the proof of Proposition 7.6 and noting that the sum $\sum_{\substack{t\ge 1\\(t,q)=1}} t^{-2} \sum_{\substack{a\in t+q\mathbb{Z}\\(a,t)=1}} F\Big(\frac{a}{t}\boldsymbol{y}\Big)$ now diverges. However we can prove that the integral in the left-hand side of (7.30) is finite if $X_q(\boldsymbol{y})$ is replaced by any subset

$$X_q^{(t_0)}(\boldsymbol{y}) := \bigsqcup_{\boldsymbol{k}\in S;\, t_{\boldsymbol{k}}\le t_0} X_q(\boldsymbol{k}, \boldsymbol{y}), \qquad (t_0 \in \mathbb{Z}_{>0}). \tag{7.42}$$

PROPOSITION 7.8. *Let $\boldsymbol{\alpha} = \boldsymbol{p}/q$ with $\boldsymbol{p} = (p_1, p_2) \in \mathbb{Z}^2$ and $\gcd(q, p_1, p_2) = 1$. Let $\boldsymbol{y} \in \mathbb{R}^2 \setminus \{\boldsymbol{0}\}$ and let $\tilde{\boldsymbol{y}} \in \mathbb{R}^2$ be any of the two vectors orthogonal to $\boldsymbol{y}$ with $\|\tilde{\boldsymbol{y}}\| = \|\boldsymbol{y}\|^{-1}$. Let $F : \mathbb{R}^2 \to \mathbb{R}$ be a nonnegative, bounded measurable function of compact support such that $F(x\boldsymbol{y} + u\tilde{\boldsymbol{y}})$ is a measurable function of $x \in \mathbb{R}$ for each fixed $u \in \mathbb{R}$. Then for any $t_0 \in \mathbb{Z}_{>0}$ we have*

$$\int_{X_q^{(t_0)}(\boldsymbol{y})} \sum_{\boldsymbol{m}\in\mathbb{Z}^2} F((\boldsymbol{m}+\boldsymbol{\alpha})M)\, d\nu_{\boldsymbol{y}}(M) \le \sum_{\substack{1\le t\le t_0\\(t,q)=1}} t^{-2} \sum_{\substack{a\in t+q\mathbb{Z}\\(a,t)=1}} F\Big(\frac{a}{t}\boldsymbol{y}\Big) + q^{-1} \sum_{v\in\mathbb{Z}\setminus\{0\}} \Big(\sum_{\substack{t\mid v\\(t,q)=1}} t^{-1}\Big) \int_{\mathbb{R}} F\Big(x\boldsymbol{y} + \frac{v}{q}\tilde{\boldsymbol{y}}\Big)\, dx. \tag{7.43}$$

Proof. This follows by imitating the proof of Proposition 7.6 but noting the special form of Lemma 7.7 when $d = 2$, and using the restriction $t_{\boldsymbol{k}} \le t_0$ from (7.42) in the treatment of the ℓ-sum from (7.31). When treating the constant factor in front of the ("new") second term, one uses the fact that $\sum_{(t,q)=1} t^{-2} > 1$. □

7.2. *The submanifolds $X(\boldsymbol{y})$ of X.* These are analogous to the submanifolds $X_q(\boldsymbol{y})$ of X_q, but we will see that many details are quite a bit simpler. Given any $\boldsymbol{y} \in \mathbb{R}^d$ we define

$$X(\boldsymbol{y}) := \big\{g \in X \,:\, \boldsymbol{y} \in \mathbb{Z}^d g\big\}. \tag{7.44}$$

We will write $\Gamma = \mathrm{ASL}(d,\mathbb{Z})$ throughout this section. Since $\mathbb{Z}^d = \mathbf{0}\Gamma$ we actually have

$$X(\boldsymbol{y}) = \big\{\Gamma g \,:\, g \in \mathrm{ASL}(d,\mathbb{R}),\ \mathbf{0}g = \boldsymbol{y}\big\} = \big\{\Gamma(M,\boldsymbol{y}) \,:\, M \in \mathrm{SL}(d,\mathbb{R})\big\}. \tag{7.45}$$

Furthermore, one checks that $M_1, M_2 \in \mathrm{SL}(d,\mathbb{R})$ give the same point $\Gamma(M_1,\boldsymbol{y}) = \Gamma(M_2,\boldsymbol{y})$ in X if and only if $\mathrm{SL}(d,\mathbb{Z})M_1 = \mathrm{SL}(d,\mathbb{Z})M_2$. Hence we get an identification of the sets $X(\boldsymbol{y})$ and $X_1 = \mathrm{SL}(d,\mathbb{Z})\backslash \mathrm{SL}(d,\mathbb{R})$, through

$$X(\boldsymbol{y}) = \big\{(M,\boldsymbol{y}) \,:\, M \in X_1\big\}. \tag{7.46}$$

This gives $X(\boldsymbol{y})$ the structure of an embedded submanifold of X, of dimension d^2-1. *We endow $X(\boldsymbol{y})$ with the Borel probability measure $\nu_{\boldsymbol{y}}$ which comes from μ_1 on X_1 under the identification* (7.46). Hence, automatically, $\nu_{\boldsymbol{y}}(X(\boldsymbol{y})) = 1$.

LEMMA 7.9. *For any $\boldsymbol{y} \in \mathbb{R}^d$, $h \in \mathrm{ASL}(d,\mathbb{R})$ and any Borel subset $\mathscr{E} \subset X(\boldsymbol{y})$ we have $\nu_{\boldsymbol{y}}(\mathscr{E}) = \nu_{\boldsymbol{y}h}(\mathscr{E}h)$.*

Proof. This follows easily using the fact that μ_1 is invariant under the (right) action of $\mathrm{SL}(d,\mathbb{R})$ on X_1. □

PROPOSITION 7.10. *Let $\mathscr{E} \subset X$ be any Borel set; then $\boldsymbol{y} \mapsto \nu_{\boldsymbol{y}}(\mathscr{E} \cap X(\boldsymbol{y}))$ is a measurable function from $\mathbb{R}^d$ to $\mathbb{R}$. If $U \subset \mathbb{R}^d$ is any Borel set such that $\mathscr{E} \subset \bigcup_{\boldsymbol{y}\in U} X(\boldsymbol{y})$, then*

$$\mu(\mathscr{E}) \le \int_U \nu_{\boldsymbol{y}}(\mathscr{E} \cap X(\boldsymbol{y}))\, d\boldsymbol{y}. \tag{7.47}$$

Furthermore, if for all $\boldsymbol{y}_1 \ne \boldsymbol{y}_2 \in U : X(\boldsymbol{y}_1) \cap X(\boldsymbol{y}_2) \cap \mathscr{E} = \varnothing$, then equality holds in (7.16).

Proof. Let us denote by $\mathscr{F}_1 \subset \mathrm{SL}(d,\mathbb{R})$ a (measurable) fundamental domain for $\Gamma(1)\backslash \mathrm{SL}(d,\mathbb{R})$, in the set theoretic sense. Then

$$\mathscr{F} := \big\{(M,\boldsymbol{\xi}) \,:\, M \in \mathscr{F}_1,\ \boldsymbol{\xi} \in [0,1)^d M\big\} \tag{7.48}$$

is a fundamental domain for $\Gamma\backslash \mathrm{ASL}(d,\mathbb{R})$. Now by the definition of $\nu_{\boldsymbol{y}}$ we have for each $\boldsymbol{y} \in \mathbb{R}^d$,

$$\nu_{\boldsymbol{y}}(\mathscr{E} \cap X(\boldsymbol{y})) = \int_{\mathscr{F}_1} \chi_{\mathscr{E}_0}(M,\boldsymbol{y})\, d\mu_1(M) = \int_{\mathrm{SL}(d,\mathbb{R})} \chi_{(\mathscr{F}_1\times\mathbb{R}^d)\cap\mathscr{E}_0}(M,\boldsymbol{y})\, d\mu_1(M), \tag{7.49}$$

where $\mathscr{E}_0$ denotes the pre-image in $\mathrm{ASL}(d,\mathbb{R})$ of $\mathscr{E} \subset X$. But the set $(\mathscr{F}_1\times\mathbb{R}^d)\cap\mathscr{E}_0$ is μ-measurable, and recall from (7.3) that $d\mu(M,\boldsymbol{\xi}) = d\mu_1(M)d\boldsymbol{\xi}$; hence by Fubini's theorem, (7.49) shows that $\boldsymbol{y} \mapsto \nu_{\boldsymbol{y}}(\mathscr{E} \cap X(\boldsymbol{y}))$ is a measurable function with respect to the Lebesgue measure on $\mathbb{R}^d$.

Next, to prove (7.47) we note that (7.49) implies

$$\int_U \nu_{\boldsymbol{y}}(\mathscr{E}\cap X(\boldsymbol{y}))\,d\boldsymbol{y} = \int_{\mathbb{R}^d}\int_{\mathrm{SL}(d,\mathbb{R})} \chi_{(\mathscr{F}_1\times U)\cap\mathscr{E}_0}(M,\boldsymbol{y})\,d\mu_1(M)\,d\boldsymbol{y} = \mu\big((\mathscr{F}_1\times U)\cap\mathscr{E}_0\big), \tag{7.50}$$

where we again used (7.3) in the last step. But it follows from our assumption $\mathscr{E}\subset\bigcup_{\boldsymbol{y}\in U} X(\boldsymbol{y})$ that each point in $\mathscr{E}\subset X$ has at least one representative in $(\mathscr{F}_1\times U)\cap\mathscr{E}_0\subset \mathrm{SL}(d,\mathbb{R})$. Hence $\mu\big((\mathscr{F}_1\times U)\cap\mathscr{E}_0\big)\ge\mu(\mathscr{E})$ and (7.47) is proved. To prove the final statement about equality, note that the condition $\forall\,\boldsymbol{y}_1\ne\boldsymbol{y}_2\in U: X(\boldsymbol{y}_1)\cap X(\boldsymbol{y}_2)\cap\mathscr{E}=\varnothing$ implies that each point in $\mathscr{E}$ has *exactly* one representative in $(\mathscr{F}_1\times U)\cap\mathscr{E}_0$, and thus $\mu\big((\mathscr{F}_1\times U)\cap\mathscr{E}_0\big)=\mu(\mathscr{E})$. □

PROPOSITION 7.11. *If $F\in \mathrm{L}^1(\mathbb{R}^d)$ and $\boldsymbol{y}\in\mathbb{R}^d$ then*

$$\int_{X(\boldsymbol{y})}\sum_{\boldsymbol{m}\in\mathbb{Z}^d} F(\boldsymbol{m}g)\,d\nu_{\boldsymbol{y}}(g) = F(\boldsymbol{y}) + \int_{\mathbb{R}^d} F(\boldsymbol{x})\,d\boldsymbol{x}. \tag{7.51}$$

Proof. This follows directly from (7.46) and Proposition 7.1 (with $\boldsymbol{\alpha}=\boldsymbol{0}$, $q=1$). □

7.3. *A thin region seldom contains an extra lattice point.* It will be important for our applications of Proposition 7.3 and Proposition 7.10 to know that if a bounded set $U\subset\mathbb{R}^d$ is thin in at least one direction (i.e. contained between two parallel hyperplanes close together) then a random lattice with a vertex in U is unlikely to have *another* vertex in U. Precisely, we will need an upper bound on the integral in (7.52) below. Since this integral is obviously monotone with respect to the set U, it suffices to consider the case when U is a translate of a cylinder $\mathfrak{Z}(c_1,c_2,\sigma)$ (cf. (4.15)) with c_2-c_1 small.

LEMMA 7.12. *Assume $\boldsymbol{\alpha}\in q^{-1}\mathbb{Z}^d$, fix $C>1$ and write $U=\boldsymbol{z}+\mathfrak{Z}(c_1,c_2,C)$. Then if $d\ge 3$ we have*

$$\int_U \nu_{\boldsymbol{y}}\Big(\Big\{M\in X_q(\boldsymbol{y}) : \#\big(U\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M\big)\ge 2\Big\}\Big)\,d\boldsymbol{y} \ll (c_2-c_1)^2, \tag{7.52}$$

uniformly over all $\boldsymbol{z}\in\{0\}\times\mathbb{R}^{d-1}$ and $C^{-1}\le c_1<c_2$. If $d=2$ then the same integral is

$$\ll (c_2-c_1)^2\log\big(2+(c_2-c_1)^{-1}\big), \tag{7.53}$$

uniformly over all $\boldsymbol{z}\in\{0\}\times[-C,C]$ and $C^{-1}\le c_1<c_2\le C$. (In the first bound the implied constant depends only on C,d; in the second bound it depends only on C,q.)

Proof. As in the proof of Proposition 7.5, we may assume $\gcd(q,p_1,\dots,p_d)=1$, without loss of generality. Take $\boldsymbol{z}\in\{0\}\times\mathbb{R}^{d-1}$, $C^{-1}\le c_1<c_2$ and let $U=\boldsymbol{z}+\mathfrak{Z}(c_1,c_2,C)$. For each $\boldsymbol{y}\in U$ and $M\in X_q(\boldsymbol{y})$ we have $\sum_{\boldsymbol{m}\in\mathbb{Z}^d}\chi_U((\boldsymbol{m}+\boldsymbol{\alpha})M)\ge 1$

by the definition of $X_q(\boldsymbol{y})$, and the same sum is ≥ 2 whenever $\#(U \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M) \geq 2$. Hence, using also $\nu_{\boldsymbol{y}}(X_q(\boldsymbol{y})) = 1$ (see Proposition 7.5), we have for each $\boldsymbol{y} \in U$,

$$\nu_{\boldsymbol{y}}\Big(\Big\{M \in X_q(\boldsymbol{y}) : \#\big(U \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M\big) \geq 2\Big\}\Big) \leq -1 + \int_{X_q(\boldsymbol{y})} \sum_{\boldsymbol{m} \in \mathbb{Z}^d} \chi_U((\boldsymbol{m} + \boldsymbol{\alpha})M)\, d\nu_{\boldsymbol{y}}(M). \tag{7.54}$$

If $d \geq 3$ then this is, by Proposition 7.6,

$$= -1 + \operatorname{vol}(U) + \sum_{\substack{t \geq 1 \\ (t,q)=1}} t^{-d} \sum_{\substack{a \in t + q\mathbb{Z} \\ (a,t)=1}} \chi_U\Big(\frac{a}{t}\boldsymbol{y}\Big) \leq O\big(c_2 - c_1\big) + \sum_{t=2}^{\infty} t^{-d} \sum_{\substack{a \in \mathbb{Z} \\ (a,t)=1}} \chi_U\Big(\frac{a}{t}\boldsymbol{y}\Big), \tag{7.55}$$

where in the second step we used $C^{-1} \leq c_1 < c_2$ to get $\sum_{a \in \mathbb{Z}} \chi_U(a\boldsymbol{y}) \leq 1 + O(c_2 - c_1)$. If some $t \geq 2$ gives nonvanishing contribution to the last sum then we must have $\frac{a}{t}\boldsymbol{y} \in U$ either for $a = t+1$ or $a = t-1$. In the first case it follows that $\frac{t+1}{t} c_1 < \frac{t+1}{t} y_1 < c_2$ so that $t > \frac{c_1}{c_2 - c_1}$; in the second case it follows that $\frac{t-1}{t} c_2 > \frac{t-1}{t} y_1 > c_1$ so that $t > \frac{c_2}{c_2 - c_1}$. Hence all t-values which contribute to the sum must satisfy $t > \frac{c_1}{c_2 - c_1}$. For each such t, a given $a \in \mathbb{Z}$ gives nonvanishing contribution only if $\frac{a}{t} y_1 < c_2$ (implying $a < t(c_2/c_1)$) and $\frac{a}{t} y_1 > c_1$ (implying $a > t(c_1/c_2)$); the number of such a's is $\leq \#\big(\mathbb{Z} \cap \big(t\frac{c_1}{c_2}, t\frac{c_2}{c_1}\big)\big) \leq 1 + t(\frac{c_2}{c_1} - \frac{c_1}{c_2})$; hence the sum in (7.55) is

$$\leq \sum_{t \geq \max(2, \frac{c_1}{c_2 - c_1})} t^{-d}\Big(1 + t\big(\frac{c_2}{c_1} - \frac{c_1}{c_2}\big)\Big) = \sum_{t \geq \max(2, x^{-1})} t^{-d}\Big(1 + tx\frac{2+x}{1+x}\Big), \tag{7.56}$$

where $x = \frac{c_2}{c_1} - 1$. Hence if $x \leq \frac{1}{2}$ then the full expression in (7.55) is (when $d \geq 3$)

$$\leq O(c_2 - c_1) + O\big(x^{d-1}\big) + O\big(x^{d-1}\big) = O\Big(c_2 - c_1 + \Big(\frac{c_2 - c_1}{c_1}\Big)^{d-1}\Big), \tag{7.57}$$

whereas if $x > \frac{1}{2}$ we get

$$\leq O(c_2 - c_1) + O(1 + x) = O\Big(c_2 - c_1 + \frac{c_2}{c_1}\Big). \tag{7.58}$$

Using $C^{-1} \leq c_1 < c_2$ the above is $\leq O(c_2 - c_1)$, in both cases. Hence we have proved

$$\nu_{\boldsymbol{y}}\Big(\Big\{M \in X_q(\boldsymbol{y}) : \#\big(U \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M\big) \geq 2\Big\}\Big) \leq O(c_2 - c_1), \qquad \forall\, \boldsymbol{y} \in U, \tag{7.59}$$

where the implied constant depends only on C and d. Since this bound is uniform over $\boldsymbol{y} \in U$ we obtain (7.52) by integration.

We now turn to the case $d = 2$. We take $\boldsymbol{z} \in \{0\} \times [-C, C]$ and $C^{-1} \le c_1 < c_2 \le C$. Take $t_0 \in \mathbb{Z}_{\ge 10}$, to be fixed later. Recall the definition of $X_q^{(t_0)}(\boldsymbol{y})$, (7.42). The left-hand side in (7.54) is

(7.60)
$$\le \int_{X_q^{(t_0)}(\boldsymbol{y})} \Big(-1 + \sum_{\boldsymbol{m} \in \mathbb{Z}^2} \chi_U((\boldsymbol{m}+\boldsymbol{\alpha})M)\Big)\, d\nu_{\boldsymbol{y}}(M) + \mathrm{vol}\Big(X_q(\boldsymbol{y}) \setminus X_q^{(t_0)}(\boldsymbol{y})\Big).$$

Imitating the proof of Proposition 7.5 one shows that the last volume is $\ll t_0^{-1}$. Hence by Proposition 7.8 the above is

$$\le -1 + O(t_0^{-1}) + \sum_{1 \le t \le t_0} t^{-2} \sum_{\substack{a \in \mathbb{Z} \\ (a,t)=1}} \chi_U\Big(\frac{a}{t}\boldsymbol{y}\Big) + q^{-1} \sum_{v \in \mathbb{Z}\setminus\{0\}} \Big(\sum_{t \mid v} t^{-1}\Big) \int_{\mathbb{R}} \chi_U\Big(x\boldsymbol{y} + \frac{v}{q}\tilde{\boldsymbol{y}}\Big)\, dx. \tag{7.61}$$

Arguing along the same lines as before we find, with $x = \frac{c_2}{c_1} - 1$,

$$-1 + \sum_{1 \le t \le t_0} t^{-2} \sum_{\substack{a \in \mathbb{Z} \\ (a,t)=1}} \chi_U\Big(\frac{a}{t}\boldsymbol{y}\Big) \le O(c_2 - c_1) + \sum_{\max(2, x^{-1}) \le t \le t_0} t^{-2}\Big(1 + tx\frac{2+x}{1+x}\Big) \le O((c_2 - c_1)\log t_0). \tag{7.62}$$

Finally we treat the last sum in (7.61). Let $L = \sqrt{4C^2 + (c_2 - c_1)^2}$, the length of the diagonal of U. If $\|\frac{v}{q}\tilde{\boldsymbol{y}}\| > L$ then $\int_{\mathbb{R}} \chi_U\Big(x\boldsymbol{y} + \frac{v}{q}\tilde{\boldsymbol{y}}\Big)\, dx = 0$ for all $\boldsymbol{y} \in U$. Hence only $v \in \mathbb{Z} \setminus \{0\}$ with $|v| \le Lq\|\tilde{\boldsymbol{y}}\|^{-1} = Lq\|\boldsymbol{y}\|$ give contributions to the last sum in (7.61), and since $\|\boldsymbol{y}\| < c_2 + \|\boldsymbol{z}\| + C \le 3C$ and $L \le \sqrt{5}C$ it follows that these v's are bounded in absolute value by a constant which only depends on C, q. Hence the last sum in (7.61) is

$$\le O(1) \sum_{v \in \mathbb{Z}\setminus\{0\}} \int_{\mathbb{R}} \chi_U\Big(x\boldsymbol{y} + \frac{v}{q}\tilde{\boldsymbol{y}}\Big)\, dx. \tag{7.63}$$

Now for each $v \in \mathbb{Z}_{>0}$ for which the integral is nonzero, there exists some $x' \in \mathbb{R}$ such that $x'\boldsymbol{y} + \frac{v-1}{q}\tilde{\boldsymbol{y}} \in U$ (since $\boldsymbol{y} \in U$ and U is convex); hence if the contribution from our v equals $\int_{x_1}^{x_2} dx$ then U must contain the triangle with vertices $x'\boldsymbol{y} + \frac{v-1}{q}\tilde{\boldsymbol{y}}$, $x_1\boldsymbol{y} + \frac{v}{q}\tilde{\boldsymbol{y}}$ and $x_2\boldsymbol{y} + \frac{v}{q}\tilde{\boldsymbol{y}}$, which has area

$$\frac{1}{2}(x_2 - x_1)\|\boldsymbol{y}\| \cdot \frac{1}{q}\|\tilde{\boldsymbol{y}}\| = \frac{1}{2q}(x_2 - x_1).$$

Note also that distinct v's lead to pairwise disjoint triangles inside U; hence the total contribution in (7.63) from positive v's is $\le 2q\mathrm{Area}(U)$. Similarly for the negative v's. Combining our bounds we have now proved that for each $\boldsymbol{y} \in U$ the left-hand side in (7.54) is

$$\le O\big(t_0^{-1} + (c_2 - c_1)\log t_0\big). \tag{7.64}$$

Choosing $t_0 = \max(10, [(c_2 - c_1)^{-1}])$ and integrating over all $\boldsymbol{y} \in U$ we obtain the bound (7.53). □

The corresponding bound in the case $\boldsymbol{\alpha} \notin \mathbb{Q}^d$ is as follows:

LEMMA 7.13. *Let $d \ge 2$ and $C > 1$ and write $U = \boldsymbol{z} + \mathfrak{Z}(c_1, c_2, C)$. Then*

$$\int_U \nu_{\boldsymbol{y}}\Big(\Big\{g \in X(\boldsymbol{y}) : \#\big(U \cap \mathbb{Z}^d g\big) \ge 2\Big\}\Big)\, d\,\boldsymbol{y} \ll (c_2 - c_1)^2, \tag{7.65}$$

uniformly over all $\boldsymbol{z} \in \{0\} \times \mathbb{R}^{d-1}$ and $c_1 < c_2$. (The implied constant depends only on C, d.)

Proof. This follows by arguing as in the first part of the proof of Lemma 7.12 (up until (7.55)) but using Proposition 7.11 in place of Proposition 7.6. □

8. **Properties of the limit functions**

8.1. *An important volume function, for $\boldsymbol{\alpha} \in \mathbb{Q}^d$.* In this section we will prove some "quasi-continuity" properties of the limit function $\Phi_{\boldsymbol{\alpha}}(\xi, \boldsymbol{w}, \boldsymbol{z})$ in Theorem 4.4, and for some more general functions. These considerations will be of importance for the proof of Theorem 4.4.

Given $r \in \mathbb{Z}_{\ge 0}$ and $\boldsymbol{\alpha} \in q^{-1}\mathbb{Z}^d$ we introduce the function

$$\begin{aligned} &f_r(c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}) \\ &\qquad := \nu_{\boldsymbol{y}}\big(\big\{M \in X_q(\boldsymbol{y}) : \#\big((\mathbb{Z}^d + \boldsymbol{\alpha})M \cap (\mathfrak{Z}(c_1, c_2, \sigma) + \boldsymbol{z})\big) = r\big\}\big) \end{aligned} \tag{8.1}$$

in the domain

$$\begin{aligned} \Omega = \big\{\langle c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}\rangle \in \mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\big(\{0\}\times\mathbb{R}^{d-1}\big)\times\mathbb{R}^d \\ : 0 \le c_1 < c_2 \le y_1,\ 0 \le \sigma\big\}. \end{aligned} \tag{8.2}$$

Arguing as in the first paragraph of the proof of Proposition 7.5 we see that although the function f_r depends on the given vector $\boldsymbol{\alpha} \in \mathbb{Q}^d$, it does not depend on the choice of denominator q of $\boldsymbol{\alpha}$; hence from now on in this section we will always assume that q is the minimal denominator of $\boldsymbol{\alpha}$, so that Propositions 7.6 and 7.8 apply. We also write, for $\xi > 0$ and $\boldsymbol{z}, \boldsymbol{w} \in \{0\} \times \mathbb{R}^{d-1}$,

$$F_r(\xi, \boldsymbol{w}, \boldsymbol{z}) := f_r(0, \xi, 1, \boldsymbol{z}, \xi\boldsymbol{e}_1 + \boldsymbol{w} + \boldsymbol{z}). \tag{8.3}$$

Thus the function $\Phi_{\boldsymbol{\alpha}}(\xi, \boldsymbol{w}, \boldsymbol{z})$ in Theorem 4.4 is the same as $F_0(\xi, \boldsymbol{w}, \boldsymbol{z})$.

LEMMA 8.1. *For any* $K = \left(\begin{smallmatrix} 1 & \mathbf{0} \\ {}^t\mathbf{0} & K_1 \end{smallmatrix}\right) \in \mathrm{O}(d)$ *we have*

$$f_r(c_1, c_2, \sigma, \boldsymbol{z}K, \boldsymbol{y}K) = f_r(c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}), \tag{8.4}$$

and for any $\delta > 0$ *we have*

$$f_r(c_1\delta^{d-1}, c_2\delta^{d-1}, \sigma\delta^{-1}, \delta^{-1}\boldsymbol{z}, \delta^{-1}\boldsymbol{y}) = f_r(c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}). \tag{8.5}$$

Proof. If $K_1 \in \mathrm{SO}(d-1)$ then the first claim follows immediately from Lemma 7.2 with $T = K$, using $\mathfrak{Z}(c_1, c_2, \sigma)K^{-1} = \mathfrak{Z}(c_1, c_2, \sigma)$. Similarly the second claim follows from Lemma 7.2 using

$$\mathfrak{Z}(c_1, c_2, \sigma)T_\delta^{-1} = \mathfrak{Z}(c_1\delta^{d-1}, c_2\delta^{d-1}, \sigma\delta^{-1}), \quad \text{for } T_\delta := \mathrm{diag}[\delta^{1-d}, \delta, \dots, \delta]. \tag{8.6}$$

To extend the first claim to general $K_1 \in \mathrm{O}(d-1)$ it now suffices to treat the single case $K = K_0 := \mathrm{diag}[1, \dots, 1, -1]$. Fix some $\gamma \in \mathrm{SL}(d, \mathbb{Z})$ such that $\boldsymbol{\alpha}\gamma K_0 = \boldsymbol{\alpha}$, and thus $(\mathbb{Z}^d + \boldsymbol{\alpha})\gamma K_0 = \mathbb{Z}^d + \boldsymbol{\alpha}$. Then $a : M \mapsto \gamma K_0 M K_0$ gives a well-defined diffeomorphism from X_q onto X_q, and one checks by a straightforward computation that for any Borel subset $\mathscr{E} \subset X_q(\boldsymbol{y})$ we have $a(\mathscr{E}) \subset X_q(\boldsymbol{y}K_0)$ and $\nu_{\boldsymbol{y}}(\mathscr{E}) = \nu_{\boldsymbol{y}K_0}(a(\mathscr{E}))$. Applying this with

$$\mathscr{E} = \{M \in X_q(\boldsymbol{y}) : \#((\mathbb{Z}^d + \boldsymbol{\alpha})M \cap (\mathfrak{Z}(c_1, c_2, \sigma) + \boldsymbol{z})) = r\}$$

we get $f_r(c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}) = f_r(c_1, c_2, \sigma, \boldsymbol{z}K_0, \boldsymbol{y}K_0)$, as desired. □

Remark 8.2. It follows that $F_r(\xi, \boldsymbol{w}K, \boldsymbol{z}K) = F_r(\xi, \boldsymbol{w}, \boldsymbol{z})$ for all K as in the lemma, and hence $F_r(\xi, \boldsymbol{w}, \boldsymbol{z})$ only depends on the four real numbers ξ, $\|\boldsymbol{z}\|$, $\|\boldsymbol{w}\|$, $\boldsymbol{z} \cdot \boldsymbol{w}$.

We will now prove our main technical result about $f_r(c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y})$ being not too far from continuous. For $N \in \mathbb{Z}_{\geq 2}$ we let $\mathfrak{F}_N$ be the set of rational numbers strictly between 0 and 1 and with denominator $\leq N$; that is,

$$\mathfrak{F}_N = \left\{\tfrac{h}{k} : h, k \in \mathbb{Z},\ 0 < h < k \leq N\right\}. \tag{8.7}$$

Given $\langle c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}\rangle \in \Omega$ and $\delta \in \mathfrak{F}_N$ we define

$$s(\delta) = s_{\langle c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}\rangle}(\delta) = \begin{cases} 1 & \text{if } \boldsymbol{y} \in \delta^{-1}(\boldsymbol{z} + \mathfrak{Z}(c_1, c_2, \sigma)) \\ 0 & \text{if } \boldsymbol{y} \notin \delta^{-1}(\boldsymbol{z} + \mathfrak{Z}(c_1, c_2, \sigma)). \end{cases} \tag{8.8}$$

For $C > 1$ we write

$$\Omega_C := \begin{cases} \{\langle c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}\rangle \in \Omega : \sigma, \|\boldsymbol{z}\|, \|\boldsymbol{y}\| \leq C;\ C^{-1} \leq |y_1|, |y_2|\} & \text{if } d = 2 \\ \{\langle c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}\rangle \in \Omega : \sigma, \|\boldsymbol{z}\|, \|\boldsymbol{y}\| \leq C;\ C^{-1} \leq \|\boldsymbol{y}\|\} & \text{if } d \geq 3. \end{cases} \tag{8.9}$$

PROPOSITION 8.3. *Fix* $d \geq 2$ *and* $r \in \mathbb{Z}_{\geq 0}$. *Given* $C > 1$ *and* $\varepsilon > 0$ *there exist some* $\eta > 0$ *and* $N \in \mathbb{Z}_{\geq 2}$ *such that*

$$\left| f_r(c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}) - f_r(c_1', c_2', \sigma', \boldsymbol{z}', \boldsymbol{y}') \right| \leq \varepsilon \tag{8.10}$$

holds for all $\langle c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}\rangle, \langle c_1', c_2', \sigma', \boldsymbol{z}', \boldsymbol{y}'\rangle \in \Omega_C$ *satisfying* $|c_1 - c_1'| \le \eta$, $|c_2 - c_2'| \le \eta$, $|\sigma - \sigma'| \le \eta$, $\|\boldsymbol{z} - \boldsymbol{z}'\| \le \eta$, $\|\boldsymbol{y} - \boldsymbol{y}'\| \le \eta$ *and* $s_{\langle c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}\rangle}(\delta) = s_{\langle c_1', c_2', \sigma', \boldsymbol{z}', \boldsymbol{y}'\rangle}(\delta)$ *for all* $\delta \in \mathfrak{F}_N$.

Proof. For certain technical statements in the following proof to be correct we need to introduce the notation $\widetilde{\mathfrak{Z}}(c_1, c_2, \sigma) := \{x_1 \boldsymbol{e}_1 : c_1 < x_1 < c_2\}$ when $\sigma = 0$, but $:= \mathfrak{Z}(c_1, c_2, \sigma)$ when $\sigma > 0$. Let $C > 1$ and $\varepsilon > 0$ be given. If $d \ge 3$ then we choose $0 < \eta_1 < 1$ so small that $\mathrm{vol}(\overline{\mathcal{B}^d_{\eta_1}} + \partial\widetilde{\mathfrak{Z}}(c_1, c_2, \sigma)) \le \frac{\varepsilon}{2}$ for all $\langle c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}\rangle \in \Omega_C$ (this is possible since $\langle c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}\rangle \in \Omega_C$ implies $0 \le c_1 < c_2 \le C$ and $\sigma \le C$); if $d = 2$ then we instead set $\eta_1 = \min\big(1, \varepsilon/(20C \sum_{1\le|v|\le 4C^2 q} \sum_{t|v} t^{-1})\big)$. We will denote by $\|A\|$ the operator norm of any $d \times d$ matrix A, viz. $\|A\| = \sup_{\boldsymbol{v}\in S_1^{d-1}} \|\boldsymbol{v}A\|$. Take $\eta \in \big(0, \min(\frac{\eta_1}{10}, \frac{1}{C})\big)$ so small that $\|M^{(0)}_{\boldsymbol{w}} - I\| \le \frac{\eta_1}{40C}$ for all $\boldsymbol{w} \in \boldsymbol{e}_1 + \overline{\mathcal{B}^d_{C\eta}}$, where $M^{(0)}_{\boldsymbol{w}}$ is as in (7.18). If $d \ge 3$ we take N so large that $\sum_{t\ge N} t^{1-d} < \frac{\varepsilon}{2}$; if $d = 2$ we take N so large that $\nu_{\boldsymbol{y}}\big(X_q(\boldsymbol{y}) \setminus X_q^{(N)}(\boldsymbol{y})\big) \le \frac{\varepsilon}{2}$ (cf. (7.42)).

Let $\langle c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}\rangle, \langle c_1', c_2', \sigma', \boldsymbol{z}', \boldsymbol{y}'\rangle$ be any two points satisfying all assumptions in the proposition, for our fixed η, N. Then $\|\boldsymbol{y} - \boldsymbol{y}'\| \le \eta \le C\eta\|\boldsymbol{y}\|$, and hence by our choice of η we can find some $T \in \mathrm{SL}(d, \mathbb{R})$ such that

$$\boldsymbol{y}' = \boldsymbol{y}T \quad \text{and} \quad \|T - I\| \le \tfrac{\eta_1}{40C} \qquad (< \tfrac{1}{40}) \tag{8.11}$$

(namely: let $T = K^{-1} M^{(0)}_{\|\boldsymbol{y}\|^{-1}\boldsymbol{y}'K^{-1}} K$ for some $K \in \mathrm{SO}(d)$ with $\boldsymbol{y} = \|\boldsymbol{y}\|\boldsymbol{e}_1 K$). Then also $\|T^{-1} - I\| \le \frac{\|T-I\|}{1-\|T-I\|} < \frac{\eta_1}{39C}$. Hence, since the constraints in Ω_C imply that $\boldsymbol{z} + \mathfrak{Z}(c_1, c_2, \sigma)$ is contained in $\mathcal{B}^d_{3C}$, we have:

$$\|\boldsymbol{x}T - \boldsymbol{x}\| < \tfrac{\eta_1}{10}; \qquad \|\boldsymbol{x}T^{-1} - \boldsymbol{x}\| < \tfrac{\eta_1}{10}, \qquad \forall \boldsymbol{x} \in \boldsymbol{z} + \mathfrak{Z}(c_1, c_2, \sigma) \tag{8.12}$$

(and similarly for $\boldsymbol{z}' + \mathfrak{Z}(c_1', c_2', \sigma')$).

Now by Lemma 7.2 we have

$$\begin{aligned} &f_r(c_1', c_2', \sigma', \boldsymbol{z}', \boldsymbol{y}') \\ &\quad = \nu_{\boldsymbol{y}}\big(\{M \in X_q(\boldsymbol{y}) : \#\big((\mathbb{Z}^d + \boldsymbol{\alpha})MT \cap (\boldsymbol{z}' + \mathfrak{Z}(c_1', c_2', \sigma'))\big) = r\}\big) \\ &\quad = \nu_{\boldsymbol{y}}\big(\{M \in X_q(\boldsymbol{y}) : \#\big((\mathbb{Z}^d + \boldsymbol{\alpha})M \cap (\boldsymbol{z}' + \mathfrak{Z}(c_1', c_2', \sigma'))T^{-1}\big) = r\}\big), \end{aligned} \tag{8.13}$$

and hence

$$\begin{aligned} &\Big|f_r(c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}) - f_r(c_1', c_2', \sigma', \boldsymbol{z}', \boldsymbol{y}')\Big| \\ &\qquad\qquad \le \nu_{\boldsymbol{y}}\big(\{M \in X_q(\boldsymbol{y}) : \#\big((\mathbb{Z}^d + \boldsymbol{\alpha})M \cap U\big) \ge 1\}\big), \end{aligned} \tag{8.14}$$

where U is the symmetric set difference

$$U = (\boldsymbol{z}' + \mathfrak{Z}(c_1', c_2', \sigma'))T^{-1} \triangle (\boldsymbol{z} + \mathfrak{Z}(c_1, c_2, \sigma)). \tag{8.15}$$

But (8.14) is $\le \int_{X_q(\boldsymbol{y})} \sum_{\boldsymbol{m}\in\mathbb{Z}^d} \chi_U\big((\boldsymbol{m}+\boldsymbol{\alpha})M\big)\, dv_{\boldsymbol{y}}$, and by Propositions 7.6 and 7.8 this is

$$\text{(8.16)} \qquad \text{if } d\ge 3: \quad \le \operatorname{vol}(U) + \sum_{t=1}^{\infty} t^{-d} \sum_{a\in\mathbb{Z}} \chi_U\Big(\frac{a}{t}\boldsymbol{y}\Big);$$

$$\text{(8.17)} \qquad \text{if } d=2: \quad \le \frac{\varepsilon}{2} + \sum_{t=1}^{N} t^{-d} \sum_{a\in\mathbb{Z}} \chi_U\Big(\frac{a}{t}\boldsymbol{y}\Big) + q^{-1} \sum_{v\in\mathbb{Z}\setminus\{0\}} \Big(\sum_{t\mid v} t^{-1}\Big) \int_{\mathbb{R}} \chi_U\Big(x\boldsymbol{y} + \frac{v}{q}\tilde{\boldsymbol{y}}\Big)\, dx.$$

We now claim that

$$\text{(8.18)} \qquad U \subset \overline{\mathcal{B}^d_{\eta_1}} + \partial\big(\boldsymbol{z} + \widetilde{\mathfrak{Z}}(c_1,c_2,\sigma)\big).$$

Indeed, using $|c_1-c_1'|$, $|c_2-c_2'|$, $|\sigma-\sigma'|$, $\|\boldsymbol{z}-\boldsymbol{z}'\| \le \eta$ one verifies

$$\text{(8.19)} \qquad \begin{aligned} &\boldsymbol{z} + \mathfrak{Z}(c_1,c_2,\sigma) \subset \big(\boldsymbol{z}' + \widetilde{\mathfrak{Z}}(c_1',c_2',\sigma')\big) + \mathcal{B}^d_{3\eta} \qquad \text{and} \\ &\boldsymbol{z}' + \mathfrak{Z}(c_1',c_2',\sigma') \subset \big(\boldsymbol{z} + \widetilde{\mathfrak{Z}}(c_1,c_2,\sigma)\big) + \mathcal{B}^d_{3\eta}. \end{aligned}$$

Hence using (8.12) and $\eta < \frac{\eta_1}{10}$ we have

$$\text{(8.20)} \qquad \begin{aligned} &\big(\boldsymbol{z} + \mathfrak{Z}(c_1,c_2,\sigma)\big)T \subset \big(\boldsymbol{z}' + \widetilde{\mathfrak{Z}}(c_1',c_2',\sigma')\big) + \mathcal{B}^d_{\eta_1/2} \qquad \text{and} \\ &\big(\boldsymbol{z}' + \mathfrak{Z}(c_1',c_2',\sigma')\big)T^{-1} \subset \big(\boldsymbol{z} + \widetilde{\mathfrak{Z}}(c_1,c_2,\sigma)\big) + \mathcal{B}^d_{\eta_1/2}, \end{aligned}$$

and since $\|T-I\| < \frac{1}{40}$ implies $\mathcal{B}^d_{\eta_1/2}T^{-1} \subset \mathcal{B}^d_{\eta_1}$ we also get

$$\text{(8.21)} \qquad \big(\boldsymbol{z} + \mathfrak{Z}(c_1,c_2,\sigma)\big) \subset \big(\boldsymbol{z}' + \widetilde{\mathfrak{Z}}(c_1',c_2',\sigma')\big)T^{-1} + \mathcal{B}^d_{\eta_1}.$$

Our claim (8.18) follows easily from (8.20) and (8.21), using also the convexity of the set $\big(\boldsymbol{z}' + \widetilde{\mathfrak{Z}}(c_1',c_2',\sigma')\big)T^{-1}$.

To see this take $\boldsymbol{x} \in U$. Then *either* $\boldsymbol{x} \in \big(\boldsymbol{z}' + \mathfrak{Z}(c_1',c_2',\sigma')\big)T^{-1}$ and $\boldsymbol{x} \notin \boldsymbol{z} + \mathfrak{Z}(c_1,c_2,\sigma)$; and in this case (8.20) shows that there exists a point $\boldsymbol{x}' \in \boldsymbol{z} + \widetilde{\mathfrak{Z}}(c_1,c_2,\sigma)$ with $\|\boldsymbol{x}'-\boldsymbol{x}\| < \eta_1/2$. Then some point on the line segment between $\boldsymbol{x}$ and $\boldsymbol{x}'$ must lie in $\partial\big(\boldsymbol{z} + \widetilde{\mathfrak{Z}}(c_1,c_2,\sigma)\big)$—q.e.d. *Or else* we have $\boldsymbol{x} \notin \big(\boldsymbol{z}' + \mathfrak{Z}(c_1',c_2',\sigma')\big)T^{-1}$ and $\boldsymbol{x} \in \boldsymbol{z} + \mathfrak{Z}(c_1,c_2,\sigma)$. (Thus $\sigma > 0$ and $\mathfrak{Z}(c_1,c_2,\sigma) = \widetilde{\mathfrak{Z}}(c_1,c_2,\sigma)$.) Then, since $\big(\boldsymbol{z}' + \widetilde{\mathfrak{Z}}(c_1',c_2',\sigma')\big)T^{-1}$ is convex, there is a hyperplane $\Pi \subset \mathbb{R}^d$ through $\boldsymbol{x}$ such that $\big(\boldsymbol{z}' + \widetilde{\mathfrak{Z}}(c_1',c_2',\sigma')\big)T^{-1}$ lies in one of the closed half spaces determined by Π. Let $\boldsymbol{x}'$ be that point which lies in the *other* half space, on the normal line to Π through $\boldsymbol{x}$, with $\|\boldsymbol{x}'-\boldsymbol{x}\| = \eta_1$. Then (8.21) implies $\boldsymbol{x}' \notin \boldsymbol{z} + \mathfrak{Z}(c_1,c_2,\sigma)$ and hence by our assumption on $\boldsymbol{x}$, some point on the line segment between $\boldsymbol{x}$ and $\boldsymbol{x}'$ must lie in $\partial\big(\boldsymbol{z} + \widetilde{\mathfrak{Z}}(c_1,c_2,\sigma)\big)$ — q.e.d.

If $d \ge 3$ then (8.18) implies that $\operatorname{vol}(U) \le \frac{\varepsilon}{2}$, by our choice of η_1.

Next we will show that $\delta\boldsymbol{y} \in U$ with $\delta \in \mathbb{Q}$ implies that δ has a large denominator. For each $\delta \ge 1$ we have $\delta\boldsymbol{y} \notin \boldsymbol{z} + \mathfrak{Z}(c_1,c_2,\sigma)$ since $y_1 \ge c_2$, and

also $\delta \boldsymbol{y} \notin (\boldsymbol{z}' + \mathfrak{Z}(c_1', c_2', \sigma'))T^{-1}$ since $\delta \boldsymbol{y} T = \delta \boldsymbol{y}' \notin (\boldsymbol{z}' + \mathfrak{Z}(c_1', c_2', \sigma'))$; hence $\delta \boldsymbol{y} \notin U$. Similarly $\delta \boldsymbol{y} \notin U$ for each $\delta \le 0$. Also if $\delta \in \mathfrak{F}_N$ then our assumption $s_{\langle c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}\rangle}(\delta) = s_{\langle c_1', c_2', \sigma', \boldsymbol{z}', \boldsymbol{y}'\rangle}(\delta)$ implies that the point $\delta \boldsymbol{y}$ either lies in both sets $\boldsymbol{z} + \mathfrak{Z}(c_1, c_2, \sigma)$ and $(\boldsymbol{z}' + \mathfrak{Z}(c_1', c_2', \sigma'))T^{-1}$, or else in none of them; thus $\delta \boldsymbol{y} \notin U$. Hence it follows that $\delta \boldsymbol{y} \in U$ for rational δ can only hold if $0 < \delta < 1$ and δ's denominator is larger than N.

It follows from this that if $d \ge 3$ then the sum in (8.16) is $\le \sum_{t \ge N} t^{-d} t < \frac{\varepsilon}{2}$ and hence since $\mathrm{vol}(U) \le \frac{\varepsilon}{2}$ we have now shown that (8.14) is $\le \varepsilon$, i.e. the proof of the proposition is complete.

If $d = 2$ then it follows that the first sum in (8.17) *vanishes*, and it remains to bound the second sum in (8.17). Since $U \subset \mathcal{B}_{4C}^2$ we get nonvanishing contributions in that sum only when $|v| \le 4Cq\|\boldsymbol{y}\| \le 4C^2 q$. Furthermore, it follows from (8.18) that U is contained in the union of two translates of $[0, c_2 - c_1 + 2\eta_1] \times [0, 2\eta_1]$ and two translates of $[0, 2\eta_1] \times [0, 2\sigma + 2\eta_1]$. Using now the condition $|y_2| \ge C^{-1}$ we see that for each translate B of $[0, c_2 - c_1 + 2\eta_1] \times [0, 2\eta_1]$ and any $\boldsymbol{w} \in \mathbb{R}^2$, the interval $\{x \in \mathbb{R} : x\boldsymbol{y} + \boldsymbol{w} \in B\}$ has length $\le 2\eta_1/|y_2| \le 2C\eta_1$, and hence the total contribution from B to the v-sum in (8.17) is $\le \sum_{1 \le |v| \le 4C^2 q} (\sum_{t|v} t^{-1}) 2C\eta_1$, and by our choice of η_1 this is $\le \frac{\varepsilon}{10}$. Similarly using $|y_1| \ge C^{-1}$ one finds that the total contribution from each *vertical* side is also $\le \frac{\varepsilon}{10}$. Hence in total (8.17) is $\le \frac{\varepsilon}{2} + 0 + \frac{\varepsilon}{10} + \frac{\varepsilon}{10} + \frac{\varepsilon}{10} + \frac{\varepsilon}{10} < \varepsilon$, and the proof is complete. □

We now point out several consequences of Proposition 8.3. First, the following technical lemma will be quite convenient to use in our proof of Theorem 4.4.

LEMMA 8.4. *Given any C, ε and corresponding η, N as in Proposition* 8.3, *then for all $c, \xi > 0$ and $\boldsymbol{w}, \boldsymbol{z} \in \{0\} \times \mathbb{R}^{d-1}$ satisfying $C^{-1} \le c \le \xi \le c + \min(\eta, c/N)$ and $\xi + \|\boldsymbol{w}\| + \|\boldsymbol{z}\| \le C$ [and if $d = 2$: $\|\boldsymbol{w} + \boldsymbol{z}\| \ge C^{-1}$], we have*

$$\begin{aligned} &\big| f_r(0, c, 1, \boldsymbol{z}, \xi \boldsymbol{e}_1 + \boldsymbol{w} + \boldsymbol{z}) - F_r(\xi, \boldsymbol{w}, \boldsymbol{z}) \big| \\ &\quad = \big| f_r(0, c, 1, \boldsymbol{z}, \xi \boldsymbol{e}_1 + \boldsymbol{w} + \boldsymbol{z}) - f_r(0, \xi, 1, \boldsymbol{z}, \xi \boldsymbol{e}_1 + \boldsymbol{w} + \boldsymbol{z}) \big| \le \varepsilon. \end{aligned} \tag{8.22}$$

Proof. The assumptions imply that both $\langle 0, c, 1, \boldsymbol{z}, \xi \boldsymbol{e}_1 + \boldsymbol{w} + \boldsymbol{z}\rangle$ and $\langle 0, \xi, 1, \boldsymbol{z}, \xi \boldsymbol{e}_1 + \boldsymbol{w} + \boldsymbol{z}\rangle$ belong to Ω_C, and these 5-tuples differ only in the second coordinate, by an amount $\le \eta$; hence by Proposition 8.3 we only need to check that

$$s_{\langle 0, c, 1, \boldsymbol{z}, \xi \boldsymbol{e}_1 + \boldsymbol{w} + \boldsymbol{z}\rangle}(\delta) = s_{\langle 0, \xi, 1, \boldsymbol{z}, \xi \boldsymbol{e}_1 + \boldsymbol{w} + \boldsymbol{z}\rangle}(\delta)$$

holds for every $\delta \in \mathfrak{F}_N$. Fix $\delta \in \mathfrak{F}_N$; our task is now to prove that the point $\xi \boldsymbol{e}_1 + \boldsymbol{w} + \boldsymbol{z}$ either belongs to both or none of the two sets $\delta^{-1}(\boldsymbol{z} + \mathfrak{Z}(0, c, 1))$ and $\delta^{-1}(\boldsymbol{z} + \mathfrak{Z}(0, \xi, 1))$. Note that $\delta \le \frac{N-1}{N}$; thus using $0 \le \xi - c \le c/N$ we have $\xi < \delta^{-1} c$ as well as $\xi < \delta^{-1}\xi$. Hence the two containment relations are *both* equivalent with $\|\boldsymbol{w} + \boldsymbol{z} - \delta^{-1}\boldsymbol{z}\| < \delta^{-1}$, and we are done. □

We next prove several lemmas relating directly to the function F_r.

LEMMA 8.5. *$F_r(\xi, \boldsymbol{w}, \boldsymbol{z})$ is Borel measurable.*

Proof. We first take $d \geq 3$. It suffices to prove that the restriction of F_r to any given compact subset K of $\mathbb{R}_{>0} \times (\{0\} \times \mathbb{R}^{d-1}) \times (\{0\} \times \mathbb{R}^{d-1})$ is Borel measurable. Using Proposition 8.3 we see that on K we can obtain F_r as a uniform limit of functions which take only a finite number of values, each level set being a finite union of sets of the form

$$B \cap \{\langle \xi, \boldsymbol{w}, \boldsymbol{z}\rangle \in K : s_{\langle 0,\xi,1,\boldsymbol{z},\xi \boldsymbol{e}_1+\boldsymbol{w}+\boldsymbol{z}\rangle}(\delta) = s_0(\delta),\ \forall \delta \in \mathfrak{F}_N\}, \tag{8.23}$$

with B a box region and s_0 some function from $\mathfrak{F}_N$ to $\{0,1\}$. Since each such level set is a Borel set we have thus expressed $(F_r)|_K$ as a uniform limit of Borel measurable functions, and we are done.

We now turn to the case $d = 2$. In this case an application of Proposition 8.3 as above shows that the restriction of F_r to any given compact subset K of $\{\langle \xi, w\boldsymbol{e}_2, z\boldsymbol{e}_2\rangle : w + z \neq \boldsymbol{0}\}$ is Borel measurable. Next, by a computation using the set-up from Proposition 7.5 one finds

$$\begin{aligned} F_r(\xi, -z\boldsymbol{e}_2, z\boldsymbol{e}_2) = \Big(\sum_{\substack{t\geq 1\\(t,q)=1}} t^{-2}\Big)^{-1} \sum_{\substack{t\geq 1\\(t,q)=1}} t^{-2} \\ \times \int_{\mathbb{R}/\mathbb{Z}} I\Big(\sum_{\substack{n\in\mathbb{Z}\\ q\xi(z-1)<nt<q\xi(z+1)}} \#\big(\mathbb{Z} \cap (nx, nx+\tfrac{t}{q})\big) = r\Big)\, dx. \end{aligned} \tag{8.24}$$

In particular, $F_r(\xi, -z\boldsymbol{e}_2, z\boldsymbol{e}_2)$ is constant on any set of the form

$$\begin{aligned} M_{a_1,a_2} = \big\{(\xi, z) : q\xi(z-1) \in [a_1, a_1+1), \\ q\xi(z+1) \in (a_2, a_2+1]\big\} \end{aligned} \qquad (a_1, a_2 \in \mathbb{Z}). \tag{8.25}$$

This implies that also the restriction of F_r to $\{\langle \xi, w\boldsymbol{e}_2, z\boldsymbol{e}_2\rangle : w + z = \boldsymbol{0}\}$ is Borel measurable, and we are done. □

In particular, this proves the claim about Borel measurability in Remark 4.5. This shows that we may freely change order of integration in the right-hand side of the limit formula (4.13).

Next we prove the claim about continuity in Remark 4.5.

LEMMA 8.6. *If we keep* $\|\boldsymbol{w}\| < 1$ *and* $\|\boldsymbol{z}\| \leq 1$ *[and if* $d = 2$: $\boldsymbol{z} + \boldsymbol{w} \neq \boldsymbol{0}$*] then the function* $F_r(\xi, \boldsymbol{w}, \boldsymbol{z})$ *is jointly continuous in all three variables.*

Proof. This is a simple consequence of Proposition 8.3 once we note that $s_{\langle 0,\xi,1,\boldsymbol{z},\xi\boldsymbol{e}_1+\boldsymbol{w}+\boldsymbol{z}\rangle}(\delta) = 1$ holds for any $\xi > 0$, $\boldsymbol{w} \in \{0\} \times \mathcal{B}_1^{d-1}$, $\boldsymbol{z} \in \{0\} \times \overline{\mathcal{B}_1^{d-1}}$, and any $\delta \in \mathfrak{F}_N$. This fact follows from $0 < \xi < \delta^{-1}\xi$ and $\|\boldsymbol{w} + \boldsymbol{z} - \delta^{-1}\boldsymbol{z}\| \leq \|\boldsymbol{w}\| + (\delta^{-1} - 1)\|\boldsymbol{z}\| < 1 + (\delta^{-1} - 1) = \delta^{-1}$. □

LEMMA 8.7. *For any fixed* $\boldsymbol{z}, \boldsymbol{w}$ *[if* $d = 2$: *assume* $\boldsymbol{z} + \boldsymbol{w} \neq \boldsymbol{0}$*], the function* $F_r(\xi, \boldsymbol{w}, \boldsymbol{z})$ *is continuous in the variable* $\xi > 0$.

Proof. This follows directly from Proposition 8.3 once we note that for any $\delta \in \mathfrak{F}_N$, the function $s_{\langle 0,\xi,1,\boldsymbol{z},\xi\boldsymbol{e}_1+\boldsymbol{w}+\boldsymbol{z}\rangle}(\delta)$ is independent of ξ. Indeed, since $\delta < 1$, $s_{\langle 0,\xi,1,\boldsymbol{z},\xi\boldsymbol{e}_1+\boldsymbol{w}+\boldsymbol{z}\rangle}(\delta) = 1$ holds if and only if $\|\boldsymbol{w}+\boldsymbol{z}-\delta^{-1}\boldsymbol{z}\| < \delta^{-1}$. □

LEMMA 8.8. *Let $\mathcal{W}$ be any bounded Borel subset of $\{0\}\times\mathbb{R}^{d-1}$; then the integral $\int_{\mathcal{W}} F_r(\xi,\boldsymbol{w},\boldsymbol{z})\,d\boldsymbol{w}$ exists for all $\xi > 0$, $\boldsymbol{z} \in \{0\}\times\mathbb{R}^{d-1}$, and is jointly continuous in these two variables. In fact, given any $\varepsilon > 0$ and $B > 1$ there is some $\nu > 0$ such that*

$$\int_{\mathcal{W}} \left| F_r(\xi,\boldsymbol{w},\boldsymbol{z}) - F_r(\xi',\boldsymbol{w},\boldsymbol{z}')\right| d\boldsymbol{w} < \varepsilon \tag{8.26}$$

holds for all $\xi,\xi' \in [B^{-1},B]$, $\boldsymbol{z},\boldsymbol{z}' \in \{0\}\times\mathcal{B}_B^{d-1}$ satisfying $|\xi-\xi'|<\nu$ and $\|\boldsymbol{z}-\boldsymbol{z}'\|<\nu$.

Proof. Since $0 \le F_r(\xi,\boldsymbol{w},\boldsymbol{z}) \le 1$, the existence of the integral follows from the Borel measurability proved in Lemma 8.5.

To prove (8.26), let $\varepsilon > 0$ and $B > 1$ be given. Applying Proposition 8.3 with $\varepsilon' := (2+\mathrm{vol}_{d-1}(\mathcal{W}))^{-1}\varepsilon$ in place of ε and with $C = \max(2B+\sup_{\boldsymbol{w}\in\mathcal{W}}\|\boldsymbol{w}\|, 4/\varepsilon')$, we get that there are some $\eta > 0$ and $N \in \mathbb{Z}_{\ge 2}$ such that $\left|F_r(\xi,\boldsymbol{w},\boldsymbol{z}) - F_r(\xi',\boldsymbol{w},\boldsymbol{z}')\right| \le \varepsilon'$ holds for all $\xi,\xi' \in [B^{-1},B]$, $\boldsymbol{w} \in \mathcal{W}$ and $\boldsymbol{z},\boldsymbol{z}' \in \{0\}\times\mathcal{B}_B^{d-1}$ satisfying $|\xi-\xi'| < \frac{\eta}{2}$ and $\|\boldsymbol{z}-\boldsymbol{z}'\| < \frac{\eta}{2}$ and $s_{\langle 0,\xi',1,\boldsymbol{z}',\xi'\boldsymbol{e}_1+\boldsymbol{w}+\boldsymbol{z}'\rangle}(\delta) = s_{\langle 0,\xi,1,\boldsymbol{z},\xi\boldsymbol{e}_1+\boldsymbol{w}+\boldsymbol{z}\rangle}(\delta)$, $\forall \delta \in \mathfrak{F}_N$. If $d = 2$ then we must also require $\|\boldsymbol{w}+\boldsymbol{z}\| \ge C^{-1}$ and $\|\boldsymbol{w}+\boldsymbol{z}'\| \ge C^{-1}$. The s-conditions are seen to hold if and only if, for each $\delta \in \mathfrak{F}_N$, either both or none of $\|\boldsymbol{w}-(\delta^{-1}-1)\boldsymbol{z}\| < \delta^{-1}$ and $\|\boldsymbol{w}-(\delta^{-1}-1)\boldsymbol{z}'\| < \delta^{-1}$ are true. For each $\delta \in \mathfrak{F}_N$, the set of *exceptional* $\boldsymbol{w}$'s is thus seen to lie in a union of two translates of the region $\delta^{-1}\big(\mathcal{B}_{1+\|\boldsymbol{z}-\boldsymbol{z}'\|}^{d-1} \setminus \mathcal{B}_1^{d-1}\big)$. Hence, since $\delta^{-1} \le N$ and $\mathfrak{F}_N$ is finite, there is some $\nu \in (0,\frac{\eta}{2}]$ such that the volume of the total set of exceptional $\boldsymbol{w}$'s is less than ε' whenever $\|\boldsymbol{z}-\boldsymbol{z}'\| < \nu$. For $d = 2$ we also have to consider the set of exceptional $\boldsymbol{w}$'s satisfying $\|\boldsymbol{w}+\boldsymbol{z}\| < C^{-1}$ or $\|\boldsymbol{w}+\boldsymbol{z}'\| < C^{-1}$; this set has volume $\le 4C^{-1} \le \varepsilon'$. Hence, since the integrand in (8.26) is everywhere ≤ 1, we see that for any $\xi,\xi' \in [B^{-1},B]$ and $\boldsymbol{z},\boldsymbol{z}' \in \{0\}\times\mathcal{B}_B^{d-1}$ satisfying $|\xi-\xi'| < \nu$ (or just $< \frac{\eta}{2}$) and $\|\boldsymbol{z}-\boldsymbol{z}'\| < \nu$, the integral in (8.26) is $\le (2+\mathrm{vol}_{d-1}(\mathcal{W}))\varepsilon' = \varepsilon$, as desired. □

8.2. *An important volume function, for $\boldsymbol{\alpha} \notin \mathbb{Q}^d$.* The questions treated in the last section become much simpler if we consider the submanifolds $X(\boldsymbol{y})$ in place of $X_q(\boldsymbol{y})$. Indeed, let us define, in analogy with (8.1) above:

$$f_r(c_1,c_2,\sigma,\boldsymbol{z},\boldsymbol{y}) := \nu_{\boldsymbol{y}}\big(\{g \in X(\boldsymbol{y}) : \#\big(\mathbb{Z}^d g \cap (\mathfrak{Z}(c_1,c_2,\sigma)+\boldsymbol{z})\big) = r\}\big), \tag{8.27}$$

with the same domain Ω as before, and for $\xi > 0$ and $\boldsymbol{z},\boldsymbol{w} \in \{0\}\times\mathbb{R}^{d-1}$,

$$F_r(\xi,\boldsymbol{w},\boldsymbol{z}) := f_r(0,\xi,1,\boldsymbol{z},\xi\boldsymbol{e}_1+\boldsymbol{w}+\boldsymbol{z}). \tag{8.28}$$

It will be clear from the context which case of functions f_r, F_r ((8.1), (8.3) or (8.27), (8.28)) we are referring to.

LEMMA 8.9. *$f_r(c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y})$ in* (8.27) *satisfies the same invariance relations as in the $X_q(\boldsymbol{y})$-case* (*see Lemma* 8.1), *and also*

$$f_r(c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}) = f_r(c_1, c_2, \sigma, \boldsymbol{0}, \boldsymbol{y} - \boldsymbol{z}).$$

Proof. Cf. the proof of Lemma 8.1 but use Lemma 7.9 in place of Lemma 7.2, and also use the transformation $h = (1_d, -\boldsymbol{z}) \in \mathrm{ASL}(d, \mathbb{R})$. □

Hence $F_r(\xi, \boldsymbol{w}, \boldsymbol{z})$ in fact only depends on ξ and $\|\boldsymbol{w}\|$. (In particular, this is true for $\Phi_{\boldsymbol{\alpha}}(\xi, \boldsymbol{w}, \boldsymbol{z}) = F_0(\xi, \boldsymbol{w}, \boldsymbol{z})$, as pointed out in Remark 4.5.)

PROPOSITION 8.10. *The function $f_r(c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y})$ in* (8.27) *is continuous everywhere in Ω.*

Proof. This follows by the same method of proof as in Proposition 8.3, but the details are much simpler: Using Proposition 7.11 in place of Proposition 7.6 one finds that (8.16) is now replaced by

$$\left| f_r(c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}) - f_r(c_1', c_2', \sigma', \boldsymbol{z}', \boldsymbol{y}') \right| \le \mathrm{vol}(U) + \chi_U(\boldsymbol{y}), \tag{8.29}$$

and as before one sees that $\chi_U(\boldsymbol{y}) = 0$ and that $\mathrm{vol}(U)$ can be made arbitrarily small by taking $\langle c_1', c_2', \sigma', \boldsymbol{z}', \boldsymbol{y}' \rangle$ close to $\langle c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y} \rangle$. □

We end by remarking some relations which will be useful in Proposition 8.13 below and in our discussion of explicit formulas in [26]. First, using (7.46) and the definition of $\nu_{\boldsymbol{y}}$ just below (7.46) we see that

$$f_r(c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}) = \mu_1\big(\{M \in X_1 : \#\big(\mathbb{Z}^d M \cap (\boldsymbol{z} - \boldsymbol{y} + \mathfrak{Z}(c_1, c_2, \sigma))\big) = r\}\big). \tag{8.30}$$

In particular, we have

$$F_r(\xi, \boldsymbol{w}, \boldsymbol{z}) = \mu_1\big(\{M \in X_1 : \#\big(\mathbb{Z}^d M \cap (-\xi \boldsymbol{e}_1 - \boldsymbol{w} + \mathfrak{Z}(0, \xi, 1))\big) = r\}\big). \tag{8.31}$$

Here $-\xi \boldsymbol{e}_1 - \boldsymbol{w} + \mathfrak{Z}(0, \xi, 1)$ may be replaced by its pointwise negate, $\xi \boldsymbol{e}_1 + \boldsymbol{w} - \mathfrak{Z}(0, \xi, 1)$, and since $\boldsymbol{w} \in \{0\} \times \mathbb{R}^{d-1}$ this set is seen to equal $\boldsymbol{w} + \mathfrak{Z}(0, \xi, 1)$. Hence

$$F_r(\xi, \boldsymbol{w}, \boldsymbol{z}) = \mu_1\big(\{M \in X_1 : \#\big(\mathbb{Z}^d M \cap (\boldsymbol{w} + \mathfrak{Z}(0, \xi, 1))\big) = r\}\big). \tag{8.32}$$

One may note that this volume is a special case of the limit function $F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(r, \sigma)$ obtained in Theorem 3.4 for $\boldsymbol{\alpha} = \boldsymbol{0}$. Indeed, using the relation

$$\big(\boldsymbol{w} + \mathfrak{Z}(0, \xi, 1)\big) \begin{pmatrix} \xi^{-1} & \boldsymbol{0} \\ {}^{\mathrm{t}}\boldsymbol{0} & \xi^{1/(d-1)} \end{pmatrix} = \xi^{\frac{1}{d-1}} \boldsymbol{w} + \mathfrak{Z}(0, 1, \xi^{\frac{1}{d-1}})$$

we see that $F_r(\xi, \boldsymbol{w}, \boldsymbol{z}) = F_{0,\boldsymbol{0},\boldsymbol{\beta}}(r, \xi^{\frac{1}{d-1}})$ holds for any choice of function $\boldsymbol{\beta}(\boldsymbol{v})$ such that $\big\|\mathrm{Proj}_{\{\boldsymbol{v}\}^\perp} \boldsymbol{\beta}(\boldsymbol{v})\big\| = \|\boldsymbol{w}\|$ for all $\boldsymbol{v} \in \mathrm{S}_1^{d-1}$.

8.3. *Differentiability properties.*

PROPOSITION 8.11. *For any fixed $\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda, r$ (and $c = 0$) as in Theorem* 3.4 *with $\boldsymbol{\alpha} \in \mathbb{Q}^d$, the function $F_{0,\boldsymbol{\alpha},\boldsymbol{\beta}}(r, \sigma)$ defined in* (3.13) *is* C^1 *with respect to $\sigma > 0$.*

Proof. In analogy with (3.14) we define

$$\mathfrak{Z}_{\boldsymbol{v}}(c_1,c_2,\sigma) := \mathfrak{Z}(c_1,c_2,\sigma) + \sigma\big\|\mathrm{Proj}_{\{\boldsymbol{v}\}^\perp}\boldsymbol{\beta}(\boldsymbol{v})\big\|\cdot\boldsymbol{e}_2,$$

so that $\mathfrak{Z}_{\boldsymbol{v}}(c,\sigma) = \mathfrak{Z}_{\boldsymbol{v}}(c,1,\sigma)$. Then

$$\mathfrak{Z}_{\boldsymbol{v}}(c'c,c',c'^{-\frac{1}{d-1}}\sigma) = \mathfrak{Z}_{\boldsymbol{v}}(c,1,\sigma)\begin{pmatrix} c' & \boldsymbol{0} \\ {}^{\mathrm t}\boldsymbol{0} & c'^{-1/(d-1)}1_{d-1}\end{pmatrix}$$

for all $c'>0$, and hence, using also the invariance of μ_q, we have

$$F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma) = (\mu_q\times\lambda)\big(\big\{(M,\boldsymbol{v})\in X_q\times \mathrm{S}_1^{d-1} : \#\big((\mathbb{Z}^d+\boldsymbol{\alpha})M\cap\mathfrak{Z}_{\boldsymbol{v}}(c\sigma^{d-1},\sigma^{d-1},1)\big)=r\big\}\big). \tag{8.33}$$

To simplify the notation we write $\sigma=\xi^{\frac{1}{d-1}}$. Now, for any $\xi>0$ and $h>0$,

$$\begin{aligned}&\big(F_{0,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,(\xi+h)^{\frac{1}{d-1}}) - F_{0,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\xi^{\frac{1}{d-1}})\big)/h \\ &= h^{-1}\int_{\mathrm{S}_1^{d-1}}\mu_q\Big(\Big\{M\in X_q : \#\big((\mathbb{Z}^d+\boldsymbol{\alpha})M\cap\mathfrak{Z}_{\boldsymbol{v}}(0,\xi,1)\big)<r, \\ &\qquad\qquad \#\big((\mathbb{Z}^d+\boldsymbol{\alpha})M\cap\mathfrak{Z}_{\boldsymbol{v}}(0,\xi+h,1)\big)=r\Big\}\Big)\,d\lambda(\boldsymbol{v}) \\ &\quad - h^{-1}\int_{\mathrm{S}_1^{d-1}}\mu_q\Big(\Big\{M\in X_q : \#\big((\mathbb{Z}^d+\boldsymbol{\alpha})M\cap\mathfrak{Z}_{\boldsymbol{v}}(0,\xi,1)\big)=r, \\ &\qquad\qquad \#\big((\mathbb{Z}^d+\boldsymbol{\alpha})M\cap\mathfrak{Z}_{\boldsymbol{v}}(0,\xi+h,1)\big)>r\Big\}\Big)\,d\lambda(\boldsymbol{v}).\end{aligned} \tag{8.34}$$

If $r\ge 1$, then using Proposition 7.3 and Lemma 7.12 we find (cf. the discussion of (9.53) below) that the first term in the right-hand side of (8.34) equals, as $h\to 0$,

$$\begin{aligned}&O\big(h\log(h^{-1})\big) \\ &+h^{-1}\int_{\mathrm{S}_1^{d-1}}\int_{\xi}^{\xi+h}\int_{\{0\}\times\mathcal{B}_1^{d-1}} f_{r-1}(0,\xi,1,\boldsymbol{z}_{\boldsymbol{v}},\xi'\boldsymbol{e}_1+\boldsymbol{w}+\boldsymbol{z}_{\boldsymbol{v}})\,d\boldsymbol{w}\,d\xi'\,d\lambda(\boldsymbol{v}),\end{aligned} \tag{8.35}$$

where $\boldsymbol{z}_{\boldsymbol{v}} := \big\|\mathrm{Proj}_{\{\boldsymbol{v}\}^\perp}\boldsymbol{\beta}(\boldsymbol{v})\big\|\cdot\boldsymbol{e}_2$. This tends to

$$\int_{\mathrm{S}_1^{d-1}}\int_{\{0\}\times\mathcal{B}_1^{d-1}} F_{r-1}(\xi,\boldsymbol{w},\boldsymbol{z}_{\boldsymbol{v}})\,d\boldsymbol{w}\,d\lambda(\boldsymbol{v})$$

as $h\to 0$, by Lemma 8.4. Treating the second term in (8.34) in the same way we obtain

$$\begin{aligned}\lim_{h\to 0^+}\ & h^{-1}\Big(F_{0,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,(\xi+h)^{\frac{1}{d-1}}) - F_{0,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\xi^{\frac{1}{d-1}})\Big) \\ &= \int_{\mathrm{S}_1^{d-1}}\int_{\{0\}\times\mathcal{B}_1^{d-1}}\Big(F_{r-1}(\xi,\boldsymbol{w},\boldsymbol{z}_{\boldsymbol{v}}) - F_r(\xi,\boldsymbol{w},\boldsymbol{z}_{\boldsymbol{v}})\Big)\,d\boldsymbol{w}\,d\lambda(\boldsymbol{v}).\end{aligned} \tag{8.36}$$

This is valid also for $r=0$ if we define $F_{-1}:\equiv 0$. Inspecting the proof just carried out and using the uniformity in the statements of Lemma 7.12 and Lemma 8.4 we see that the convergence in (8.36) is uniform with respect to ξ in any compact

subset of $\mathbb{R}_{>0}$. Hence the formula (8.36) is also valid in the limit $h \to 0^-$, and Lemma 8.8 gives that $F_{0,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma)$ is indeed C^1 with respect to σ. We also note that (8.36) gives an explicit formula for the derivative. □

Remark 8.12. The explicit formula for the derivative, (8.36), specializes to the formula (4.17) in Remark 4.6 in the case $r = 0$. (For recall (4.6), $F_0(\xi, \boldsymbol{w}, \boldsymbol{z}) = \Phi_{\boldsymbol{\alpha}}(\xi, \boldsymbol{w}, \boldsymbol{z})$, and Remark 8.2.)

We also note that the argument in the above proof applies without changes to the case when λ is a (not absolutely continuous) probability measure which gives mass one to a single point. Hence for each $\boldsymbol{z} \in \{0\} \times \mathbb{R}^{d-1}$ we have

$$\frac{d}{d\xi}\mu_q\big(\{M \in X_q \,:\, (\mathbb{Z}^d + \boldsymbol{\alpha})M \cap (\mathfrak{Z}(0,\xi,1) + \boldsymbol{z}) = \varnothing\}\big) = -\int_{\{0\}\times\mathcal{B}_1^{d-1}} \Phi_{\boldsymbol{\alpha}}(\xi, \boldsymbol{w}, \boldsymbol{z})\, d\boldsymbol{w}; \tag{8.37}$$

in particular the derivative in the left-hand side is a continuous function of ξ, cf. Lemma 8.8. The set in the left-hand side of (8.37) has μ_q-measure tending to 1 as $\xi \to 0^+$ and tending to 0 as $\xi \to \infty$, cf. the proof of Remark 3.6 in Section 8.4. Hence, integrating (8.37) over $\xi \in \mathbb{R}_{>0}$ we deduce the formula (4.16) in Remark 4.6, $\int_0^\infty \int_{\{0\}\times\mathcal{B}_1^{d-1}} \Phi_{\boldsymbol{\alpha}}(\xi, \boldsymbol{w}, \boldsymbol{z})\, d\boldsymbol{w}\, d\xi = 1$.

Next we turn to the case $\boldsymbol{\alpha} \notin \mathbb{Q}^d$. Recall that in this case $F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma)$ is independent of $\boldsymbol{\beta}, \lambda, \boldsymbol{\alpha}$, and we have introduced the notation $F_c(r,\sigma)$ for this function. Proposition 8.13 and the ensuing remarks carry over directly to the case $\boldsymbol{\alpha} \notin \mathbb{Q}^d$, with the usual changes of notation. However, we can say more:

PROPOSITION 8.13. *For any fixed* $0 \le c < 1$ *and* $r \in \mathbb{Z}_{\ge 0}$ *the function* $F_c(r,\sigma)$ *is* C^2 *with respect to* $\sigma > 0$.

Proof. The function $F_c(r,\sigma)$ satisfies the invariance relation

$$F_c(r,\sigma) = F_0(r, \sigma(1-c)^{\frac{1}{d-1}}),$$

which follows directly from the definition (3.8), using the right $\mathrm{ASL}(d,\mathbb{R})$-invariance of μ. Hence we may from now on assume $c = 0$.

Arguing as in the proof of Proposition 8.11 we prove that $F_0(r,\sigma)$ is C^1 with respect to σ. The explicit formula (8.36) is still valid (with $F_r(\xi, \boldsymbol{w}, \boldsymbol{z})$ now being given by (8.28), (8.27)), although the integration over S_1^{d-1} may be skipped since in this case $F_r(\xi, \boldsymbol{w}, \boldsymbol{z})$ is independent of $\boldsymbol{z}$. Rewriting (8.36) using (8.32) we get

$$\begin{aligned}\frac{d}{d\xi}F_0(r, \xi^{\frac{1}{d-1}}) &= \int_{\{0\}\times\mathcal{B}_1^{d-1}} \mu_1\big(\{M \in X_1 \,:\, \#\big(\mathbb{Z}^d M \cap (\boldsymbol{w}' + \mathfrak{Z}(0,\xi,1))\big) = r-1\}\big)\, d\boldsymbol{w}' \\ &\quad - \int_{\{0\}\times\mathcal{B}_1^{d-1}} \mu_1\big(\{M \in X_1 \,:\, \#\big(\mathbb{Z}^d M \cap (\boldsymbol{w}' + \mathfrak{Z}(0,\xi,1))\big) = r\}\big)\, d\boldsymbol{w}'.\end{aligned} \tag{8.38}$$

But here the right-hand side can again be differentiated with respect to ξ, by repeating the argument in the proof of Proposition 8.11 (with "$\boldsymbol{\alpha}=\mathbf{0}$" and letting $\boldsymbol{w}'$ play the role of $\boldsymbol{z}_{\boldsymbol{v}}$ in that proof); this leads to

(8.39)
$$\frac{d^2}{d\xi^2}F_0(r,\xi^{\frac{1}{d-1}}) = \int_{\{0\}\times\mathcal{B}_1^{d-1}}\int_{\{0\}\times\mathcal{B}_1^{d-1}}\Big(F_{r-2}^{(\boldsymbol{\alpha}=\mathbf{0})}(\xi,\boldsymbol{w},\boldsymbol{w}') - 2F_{r-1}^{(\boldsymbol{\alpha}=\mathbf{0})}(\xi,\boldsymbol{w},\boldsymbol{w}') + F_r^{(\boldsymbol{\alpha}=\mathbf{0})}(\xi,\boldsymbol{w},\boldsymbol{w}')\Big)\,d\boldsymbol{w}\,d\boldsymbol{w}',$$

where "$F_r^{(\boldsymbol{\alpha}=\mathbf{0})}$" means "$F_r$ as in (8.3), (8.1) with $\boldsymbol{\alpha}=\mathbf{0}$, $q=1$" (and we understand $F_{-2}^{(\boldsymbol{\alpha}=\mathbf{0})}:\equiv 0$ and $F_{-1}^{(\boldsymbol{\alpha}=\mathbf{0})}:\equiv 0$). Hence (for our $\boldsymbol{\alpha}\notin\mathbb{Q}^d$) $F_0(r,\sigma)$ is indeed C^2 with respect to σ; cf. Lemma 8.8. □

Remark 8.14. The formula (8.39) generalizes [36, eq. (34)] from $d=2$ to general d.

8.4. *A uniform bound.* In this section we prove the two bounds in Remark 3.6. If $\boldsymbol{\alpha}\in\mathbb{Q}^d$ we note that for each $\boldsymbol{v}\in\mathrm{S}_1^{d-1}$ we have, by Proposition 7.1,

$$\begin{aligned}(8.40)\quad &\mu_q\big(\{M\in X_q\,:\,(\mathbb{Z}^d+\boldsymbol{\alpha})M\cap\mathfrak{Z}_{\boldsymbol{v}}(c,\sigma)=\varnothing\}\big)\\ &\ge 1-\int_{X_q}\#\big((\mathbb{Z}^d+\boldsymbol{\alpha})M\cap\mathfrak{Z}_{\boldsymbol{v}}(c,\sigma)\big)\,d\mu_q(M)\\ &= 1-\mathrm{vol}\big(\mathfrak{Z}_{\boldsymbol{v}}(c,\sigma)\big) = 1-v_d(1-c)\sigma^{d-1}.\end{aligned}$$

Integrating over $\boldsymbol{v}\in\mathrm{S}_1^{d-1}$ with respect to the measure λ (cf. the definition (3.13)) we obtain the first bound in (3.15); the second one follows using $\sum_{r=0}^{\infty}F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma)=1$.

In the case $\boldsymbol{\alpha}\notin\mathbb{Q}^d$ the bound (3.15) follows using (7.3), (7.48) and a computation as in (8.40), noticing $\int_{[0,1)^d\boldsymbol{M}}\#\big((\mathbb{Z}^d\boldsymbol{M}+\boldsymbol{\xi})\cap\mathfrak{Z}_{\boldsymbol{v}}(c,\sigma)\big)\,d\boldsymbol{\xi}=\mathrm{vol}\big(\mathfrak{Z}_{\boldsymbol{v}}(c,\sigma)\big)$ for each $M\in\mathcal{F}_1$.

The bound (3.16) is a direct consequence of the following lemma.

LEMMA 8.15. *If $r\in\mathbb{Z}_{\ge 0}$ and B is any translate of a cylinder $\mathfrak{Z}(c_1,c_2,\sigma)$ (cf.* (4.15)*) in $\mathbb{R}^d$ of volume V, then*

$$\begin{aligned}(8.41)\quad &\mu_q\big(\{M\in X_q\,:\,\#((\mathbb{Z}^d+\boldsymbol{\alpha})M\cap B)\le r\}\big)\ll V^{-1},\qquad \forall\boldsymbol{\alpha}\in q^{-1}\mathbb{Z}^d;\\ &\textit{and}\quad \mu\big(\{g\in X\,:\,\#(\mathbb{Z}^d g\cap B)\le r\}\big)\ll V^{-1}.\end{aligned}$$

The implied constants depend only on r,d.

Proof. The proof uses the methods in [21, §3.6], but note that we work with a slightly different notation in the present paper. We will prove the first bound in (8.41); the proof of the second bound is quite similar. Since both sides in the inequality remain invariant if B is replaced by BM_0 for any $M_0\in\mathrm{SL}(d,\mathbb{R})$, we may assume without loss of generality that B is a translate of a cylinder $\mathfrak{Z}(c_1,c_2,\sigma)$ with $c_2-c_1=\sigma$.

Every element $M \in \mathrm{SL}(d,\mathbb{R})$ has a unique Iwasawa decomposition $M = \mathsf{nak}$, where n belongs to the group N of upper triangular matrices with 1s on the diagonal, a is diagonal with positive diagonal elements, and $\mathsf{k} \in \mathrm{SO}(d)$. We let $\mathscr{F}_N$ be the set of all matrices in N for which all entries above the diagonal lie in the interval $(-\frac{1}{2}, \frac{1}{2}]$, and introduce the following Siegel set (denoting $\mathsf{a} = \mathrm{diag}[a_1, \dots, a_d]$):

(8.42)
$$\mathscr{S} := \Big\{ \mathsf{nak} : \mathsf{n} \in \mathscr{F}_N,\ 0 < a_{j+1} \le \tfrac{2}{\sqrt{3}} a_j \ (j = 1, \dots, d-1),\ \mathsf{k} \in \mathrm{SO}(d) \Big\}.$$

It is known that $\mathscr{S}$ contains a fundamental domain for $X_1 = \mathrm{SL}(d,\mathbb{Z})\backslash \mathrm{SL}(d,\mathbb{R})$; we fix $\mathscr{F} \subset \mathscr{S}$ to be one such fundamental domain (in the set-theoretic sense). Choose representatives $T_j \in \mathrm{SL}(d,\mathbb{Z})$ so that $\mathrm{SL}(d,\mathbb{Z}) = \bigsqcup_{j=1}^m \Gamma(q) T_j$ (disjoint union); then $\bigsqcup_{j=1}^m T_j \mathscr{F}$ is a fundamental domain for $X_q = \Gamma(q)\backslash \mathrm{SL}(d,\mathbb{R})$.

Now let M be any element in $\bigsqcup_{j=1}^m T_j \mathscr{F}$. Choose j so that $T_j^{-1} M \in \mathscr{F}$, let the Iwasawa decomposition of this matrix be $T_j^{-1} M = \mathsf{nak}$, and let the row vectors of the same matrix be $\boldsymbol{b}_1, \dots, \boldsymbol{b}_d \in \mathbb{R}^d$. Then using $\mathsf{n} \in \mathscr{F}_N$ and $\mathsf{nak} \in \mathscr{S}$ we see that $\|\boldsymbol{b}_k\| \le \sum_{j=1}^d a_j \ll_d a_1$ for each $k = 1, \dots, d$. Using $T_j \in \mathrm{SL}(d,\mathbb{Z})$ we see that $\#((\mathbb{Z}^d + \boldsymbol{\alpha})M \cap B) = \#((\mathbb{Z}\boldsymbol{b}_1 + \cdots + \mathbb{Z}\boldsymbol{b}_d) \cap (B - \boldsymbol{\alpha} M))$. Choose $\xi_1, \dots, \xi_d \in \mathbb{R}$ so that $\xi_1 \boldsymbol{b}_1 + \cdots + \xi_d \boldsymbol{b}_d$ is the center of the cylinder $B - \boldsymbol{\alpha} M$, and take $m_1, \dots, m_d \in \mathbb{Z}$ so that $\xi_k - m_k \in (-\frac{1}{2}, \frac{1}{2}]$ for each k. Then the distance from $\xi_1 \boldsymbol{b}_1 + \cdots + \xi_d \boldsymbol{b}_d$ to any of the lattice points $m_1 \boldsymbol{b}_1 + \cdots + m_d \boldsymbol{b}_d + j \boldsymbol{b}_d$, for $j = 0, \dots, r$, is $\le \frac{1}{2}\big(\|\boldsymbol{b}_1\| + \cdots + \|\boldsymbol{b}_d\|\big) + r\|\boldsymbol{b}_d\| \ll_{d,r} a_1$. Hence using our assumption $c_2 - c_1 = \sigma$, we see that if $a_1 \ll_{d,r} V^{1/d}$ then all these lattice points lie in $B - \boldsymbol{\alpha} M$, so that $\#((\mathbb{Z}^d + \boldsymbol{\alpha})M \cap B) > r$. Hence the left-hand side in (8.41) is $\le \sum_j \mu_q(\{M : T_j^{-1} M = \mathsf{nak} \in \mathscr{F},\ a_1 \gg V^{1/d}\})$. Using (7.2) and the invariance of μ_1 we see that this is $\le \mu_1(\{M = \mathsf{nak} \in \mathscr{S} : a_1 \gg V^{1/d}\})$, and as in [21, §3.6] we see that this is $\ll V^{-1}$. □

8.5. *Analogous results for Section* 2. In this section we indicate how most parts of the development in Sections 8.1–8.3 carry over to the setting of Section 2.3, leading to a proof of the claim in Remark 2.2 that the function $E_{0,\boldsymbol{\alpha}}(r,\sigma)$ is C^1 with respect to σ.

For any $0 \le c_1 \le c_2$ and $\sigma \ge 0$ we let

(8.43)
$$\mathfrak{C}(c_1, c_2, \sigma) = \Big\{ (x_1, \dots, x_d) \in \mathbb{R}^d \ : \ c_1 < x_1 < c_2,\ \|(x_2, \dots, x_d)\| < \sigma x_1 \Big\},$$

so that in particular $\mathfrak{C}(c,\sigma)$ (cf. (2.14)) equals $\mathfrak{C}(c, 1, A(c,\sigma))$ up to a set of measure zero. (The reason for using "$\le$" in (2.14) is to make (9.66) below true without modification also when $\sigma_\infty = 0$.) Given $r \in \mathbb{Z}_{\ge 0}$ and $\boldsymbol{\alpha} \in q^{-1}\mathbb{Z}^d$ we now introduce

the function

$$g_r(c_1,c_2,\sigma,\boldsymbol{z},\boldsymbol{y}) := \nu_{\boldsymbol{y}}\big(\{M\in X_q(\boldsymbol{y}) : \#\big((\mathbb{Z}^d+\boldsymbol{\alpha})M\cap(\mathfrak{C}(c_1,c_2,\sigma)+\boldsymbol{z})\big)=r\}\big) \tag{8.44}$$

with domain Ω as in (8.2). Therefore, $g_r(c_1,c_2,\sigma,\boldsymbol{z},\boldsymbol{y})$ is defined exactly as $f_r(c_1,c_2,\sigma,\boldsymbol{z},\boldsymbol{y})$ in (8.1) except that we use $\mathfrak{C}(c_1,c_2,\sigma)$ in place of $\mathfrak{Z}(c_1,c_2,\sigma)$. We also write, in analogy with (8.3), for $\xi>0$ and $\boldsymbol{z},\boldsymbol{w}\in\{0\}\times\mathbb{R}^{d-1}$,

$$G_r(\xi,\boldsymbol{w},\boldsymbol{z}) = g_r(0,\xi,1,\boldsymbol{z},\xi\boldsymbol{e}_1+\boldsymbol{w}+\boldsymbol{z}). \tag{8.45}$$

Now the discussion in Section 8.1 up to and including Proposition 8.3 carries over to the case of $g_r(c_1,c_2,\sigma,\boldsymbol{z},\boldsymbol{y})$ with very minor changes. In particular, if we replace $\mathfrak{Z}(c_1,c_2,\sigma)$ by $\mathfrak{C}(c_1,c_2,\sigma)$ in the definition of $s(\delta)$, (8.8), and replace the definition of Ω_C in the case $d=2$ (cf. (8.9)) by

$$\Omega_C := \{\langle c_1,c_2,\sigma,\boldsymbol{z},\boldsymbol{y}\rangle\in\Omega \ :\ \sigma,\|\boldsymbol{z}\|,\|\boldsymbol{y}\|\le C;\ C^{-1}\le|y_1|,|y_2\pm\sigma y_1|\}, \tag{8.46}$$

then *the statement of Proposition* 8.3 *holds with* $g_r(c_1,c_2,\sigma,\boldsymbol{z},\boldsymbol{y})$ *in place of* $f_r(c_1,c_2,\sigma,\boldsymbol{z},\boldsymbol{y})$. Using this, one then also proves that *the statements of Lemmas* 8.4, 8.5 *and* 8.8 *hold with* g_r, G_r *in place of* f_r, F_r, with the only difference that *in Lemma* 8.4 *the condition "*[*and if* $d=2$: $\|\boldsymbol{w}+\boldsymbol{z}\|\ge C^{-1}$]*" must be replaced with "*[*and if* $d=2$: $\big|\|\boldsymbol{w}+\boldsymbol{z}\|-\xi\big|\ge C^{-1}$]".

Similarly, in the case $\boldsymbol{\alpha}\notin\mathbb{Q}^d$, the discussion in Section 8.2 up to and including Proposition 8.10 carries over in the obvious way to the function

$$g_r(c_1,c_2,\sigma,\boldsymbol{z},\boldsymbol{y}) := \nu_{\boldsymbol{y}}\big(\{g\in X(\boldsymbol{y}) : \#\big(\mathbb{Z}^d g\cap(\mathfrak{C}(c_1,c_2,\sigma)+\boldsymbol{z})\big)=r\}\big). \tag{8.47}$$

Also the formulas (8.30) and (8.31) carry over, but (8.32) *does not carry over*, since the cone $\mathfrak{C}(0,\xi,1)$ does not have the necessary symmetry.

Now Proposition 8.11 carries over, i.e. the function $E_{0,\boldsymbol{\alpha}}(r,\sigma)$ is C^1 with respect to $\sigma>0$ for any fixed $\boldsymbol{\alpha}\in\mathbb{R}^d$ and $r\in\mathbb{Z}_{\ge0}$, as claimed. We remark that in the proof of this we actually only need (8.44) with $\boldsymbol{z}=\boldsymbol{0}$. The analogue of the formula (8.36) is

$$\frac{d}{d\xi}E_{0,\boldsymbol{\alpha}}\Big(r,\tfrac{\mathrm{vol}(\mathcal{B}_1^{d-1})}{d}\xi^d\Big) = \int_{\{0\}\times\mathcal{B}_\xi^{d-1}}\Big(G_{r-1}(\xi,\boldsymbol{w},\boldsymbol{0})-G_r(\xi,\boldsymbol{w},\boldsymbol{0})\Big)\,d\boldsymbol{w}. \tag{8.48}$$

Finally we turn to the special case $d=2$ and $\boldsymbol{\alpha}\in\mathbb{Q}^2$ (say $\boldsymbol{\alpha}\in q^{-1}\mathbb{Z}^2$ with q minimal). We intend to prove (2.32) in Section 2.4, i.e. that $E_{0,\boldsymbol{\alpha}}(0,\sigma)=1-\kappa_q\sigma$ holds for all $\sigma\in\big[0,(2q)^{-1}\big]$. Clearly, by (8.48), it suffices to prove that if $0<\xi\le(2q)^{-1/2}$ and $|w|<\xi$ then $G_0(\xi,we_2,\boldsymbol{0})=\kappa_q$; i.e.,

$$\nu_{\boldsymbol{w}}\big(\{M\in X_q(\boldsymbol{w}) : (\mathbb{Z}^2+\boldsymbol{\alpha})M\cap\mathfrak{C}(0,\xi,1)=\varnothing\}\big)=\kappa_q, \tag{8.49}$$

where $\boldsymbol{w} = \xi \boldsymbol{e}_1 + w\boldsymbol{e}_2$. Let $M \in \mathrm{SL}(2,\mathbb{R})$ be a representative for an arbitrary element in $X_q(\boldsymbol{w})$. Then there is some $\boldsymbol{k} \in \mathbb{Z}^2 + \boldsymbol{\alpha} \setminus \{\boldsymbol{0}\}$ such that $\boldsymbol{k}M = \boldsymbol{w}$. Set $t = \gcd(q\boldsymbol{k})$; then $\frac{q}{t}\boldsymbol{k}$ is a primitive vector in $\mathbb{Z}^2$, and thus $\mathbb{Z}^2 = \mathbb{Z}\frac{q}{t}\boldsymbol{k} + \mathbb{Z}\boldsymbol{h}$ for some $\boldsymbol{h} \in \mathbb{Z}^2$. Hence $(\mathbb{Z}^2+\boldsymbol{\alpha})M = (\mathbb{Z}^2+\boldsymbol{k})M = \mathbb{Z}\frac{q}{t}\boldsymbol{k}M + \mathbb{Z}\boldsymbol{h}M + \boldsymbol{k}M \subset \mathbb{R}\boldsymbol{w} + \mathbb{Z}\boldsymbol{h}M$. Also $\mathbb{R}\boldsymbol{w} + n\boldsymbol{h}M = \mathbb{R}\boldsymbol{w} \pm n\frac{t}{q\xi}\boldsymbol{e}_2$, and from this one verifies (using $0 < \xi \le (2q)^{-1/2}$, $|w| < \xi$) that each line $\mathbb{R}\boldsymbol{w} + n\boldsymbol{h}M$ ($n \in \mathbb{Z} \setminus \{0\}$) lies outside $\mathfrak{C}(0,\xi,1)$. Hence

$$(\mathbb{Z}^2 + \boldsymbol{\alpha})M \cap \mathfrak{C}(0,\xi,1) = \mathbb{R}\boldsymbol{w} \cap (\mathbb{Z}^2 + \boldsymbol{\alpha})M \cap \mathfrak{C}(0,\xi,1) = (\boldsymbol{w} + \mathbb{Z}\tfrac{q}{t}\boldsymbol{w}) \cap \mathfrak{C}(0,\xi,1). \tag{8.50}$$

This set is empty if and only if $t \le q$. Hence by mimicking the proof of Proposition 7.5 we find that the left-hand side of (8.49) equals

$$\frac{q^2\,\mu_H\big((\Gamma(q)\cap H)\backslash H\big)}{I_q\zeta(2)} \sum_{\substack{1\le t\le q\\(t,q)=1}} t^{-2} = \Big(\sum_{\substack{n\ge 1\\(n,q)=1}} n^{-2}\Big)^{-1} \sum_{\substack{1\le t\le q\\(t,q)=1}} t^{-2} = \kappa_q, \tag{8.51}$$

and we are done.

9. **Proof of the limit theorems in Sections** 1–4

9.1. *Proofs for Section* 3. We first prove Theorem 3.4 (and thus Theorem 3.1, which is a special case). Theorem 3.4 will be derived as a direct consequence of the general limit theorems in Section 6, and our only serious task in the present section will be to compute the upper and lower limits of an appropriate family of subsets of $\mathbb{R}^{d-1} \times \mathbb{R}^d$ (see Lemma 9.2 below). In fact we will carry this out for a generalized version of Theorem 3.4, see Theorem 9.1 below. This generalization is interesting in its own right, and its proof is also a useful preparation for the demonstration of Theorem 4.4 in Section 9.3.

First let us fix a parametrization of the sphere: Let D be a bounded open subset of $\mathbb{R}^{d-1}$ and let $E_1 : D \to \mathrm{SO}(d)$ be a smooth map such that $\boldsymbol{v} = \boldsymbol{v}(\boldsymbol{x}) := \boldsymbol{e}_1 E_1(\boldsymbol{x})^{-1} \in \mathrm{S}_1^{d-1}$ gives a diffeomorphism from D to S_1^{d-1} minus one point.[5] The fact that we miss one point in S_1^{d-1} will not matter for us since the measure λ is absolutely continuous.

Now for any subset $\mathfrak{U} \subset {\mathrm{S}'_1}^{d-1} = \{\boldsymbol{z} \in \mathrm{S}_1^{d-1} : \boldsymbol{z}\cdot\boldsymbol{e}_1 > 0\}$ and any $\boldsymbol{w} \in \mathbb{R}^d$, $\rho > 0$ such that $\rho\boldsymbol{w}$ lies outside all the balls $\mathcal{B}_\rho^d + \boldsymbol{y}$ ($\boldsymbol{y} \in \mathscr{L}_{\boldsymbol{\alpha}} \setminus \{\boldsymbol{0}\}$), we define

$$\mathcal{N}_{c,T}^{(\mathfrak{U})}(\rho,\boldsymbol{x},\boldsymbol{w}) := \#\Big\{\boldsymbol{y} \in (\mathscr{L}_{\boldsymbol{\alpha}} \cap \mathcal{B}_T^d(c)\setminus\{\boldsymbol{0}\}) - \rho\boldsymbol{w} : \boldsymbol{y}\,E_1(\boldsymbol{x}) \in \mathbb{R}_{>0}\boldsymbol{e}_1 + \rho\mathfrak{U}_\perp\Big\}, \tag{9.1}$$

where we write $\mathfrak{U}_\perp := \{\boldsymbol{z}_\perp : \boldsymbol{z} \in \mathfrak{U}\}$ with $\boldsymbol{z}_\perp := \boldsymbol{z} - (\boldsymbol{z}\cdot\boldsymbol{e}_1)\boldsymbol{e}_1 = (0, z_2, \dots, z_d)$ for any $\boldsymbol{z} = (z_1,\dots,z_d) \in \mathbb{R}^d$. Note that $\mathcal{N}_{c,T}^{(\mathfrak{U})}(\rho,\boldsymbol{x},\boldsymbol{w})$ is the number of points

[5] For example, we may choose $D = \mathcal{B}_\pi^{d-1}$ and $E_1(\boldsymbol{x}) = K_0^{-1}E(-\boldsymbol{x})$ for any fixed $K_0 \in \mathrm{SO}(d)$, where $E(\boldsymbol{x}) = \exp\left(\begin{smallmatrix} 0 & \boldsymbol{x} \\ -{}^{\mathrm{t}}\boldsymbol{x} & 0_{d-1}\end{smallmatrix}\right)$.

$\boldsymbol{y} \in \mathscr{L}_{\boldsymbol{\alpha}} \cap \mathcal{B}_T^d(c) \setminus \{\boldsymbol{0}\}$ such that the ray $\rho\boldsymbol{w} + \mathbb{R}_{>0}\boldsymbol{v}$ ($\boldsymbol{v} = \boldsymbol{v}(\boldsymbol{x})$) hits the ball $\mathcal{B}_\rho^d + \boldsymbol{y}$, with the extra condition that $-\boldsymbol{w}_{\boldsymbol{y}} E_1(\boldsymbol{x}) \in \mathfrak{U}$, where $\boldsymbol{w}_{\boldsymbol{y}} = \rho^{-1}(\rho\boldsymbol{w} + \tau_{\boldsymbol{y}}\boldsymbol{v} - \boldsymbol{y}) \in \mathrm{S}_1^{d-1}$ and $\tau_{\boldsymbol{y}} = \inf\{t > 0 : \rho\boldsymbol{w} + t\boldsymbol{v} \in \mathcal{B}_\rho^d + \boldsymbol{y}\}$. Here $\boldsymbol{w}_{\boldsymbol{y}}$ is the location of the point where the ray first hits the $\boldsymbol{y}$-sphere, relative to its center $\boldsymbol{y}$. Hence, similarly as in Section 4.2, $\boldsymbol{w}_{\boldsymbol{y}}$ *always* satisfies $-\boldsymbol{w}_{\boldsymbol{y}} E_1(\boldsymbol{x}) \in {\mathrm{S}_1'}^{d-1}$. In particular, we have

$$\mathcal{N}_{c,T}(\rho, \boldsymbol{v}(\boldsymbol{x}), \boldsymbol{w}) = \mathcal{N}_{c,T}^{({\mathrm{S}_1'}^{d-1})}(\rho, \boldsymbol{x}, \boldsymbol{w}), \tag{9.2}$$

so that $\mathcal{N}_{c,T}^{(\mathfrak{U})}(\rho, \boldsymbol{x}, \boldsymbol{w})$ generalizes our notation from (3.11). We will write λ and $\boldsymbol{\beta}$ also for the lifts of λ and $\boldsymbol{\beta}$ to the variable $\boldsymbol{x}$. Thus λ is a Borel probability measure on $\mathbb{R}^{d-1}$ with bounded support (in fact $\lambda = \lambda|_D$), which is absolutely continuous with respect to Lebesgue measure. Furthermore $\boldsymbol{\beta}$ is a continuous function from D to $\mathbb{R}^d$.

THEOREM 9.1. *For every subset $\mathfrak{U} \subset {\mathrm{S}_1'}^{d-1}$ with $\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\partial\mathfrak{U}) = 0$ and for all $\sigma \geq 0$ and $r \in \mathbb{Z}_{\geq 0}$, the limit*

$$F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}^{(\mathfrak{U})}(r, \sigma) := \lim_{T\to\infty} \lambda(\{\boldsymbol{x} \in D : \mathcal{N}_{c,T}^{(\mathfrak{U})}(\sigma T^{-1/(d-1)}, \boldsymbol{x}, \boldsymbol{\beta}(\boldsymbol{x})) = r\}) \tag{9.3}$$

exists, and for fixed $\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda, r, \mathfrak{U}$ the convergence is uniform with respect to σ in any compact subset of $\mathbb{R}_{\geq 0}$ and with respect to $c \in [0, 1]$. The limit function is given by

$$F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}^{(\mathfrak{U})}(r, \sigma) \tag{9.4}$$

$$= \begin{cases} (\mu_q \times \lambda)(\{(M, \boldsymbol{x}) \in X_q \times D : \#((\mathbb{Z}^d + \frac{\boldsymbol{p}}{q})M \cap \mathfrak{Z}^{(\mathfrak{U})}(c, 1, \sigma, \boldsymbol{\beta})|_{\boldsymbol{x}}) = r\}) & \text{if } \boldsymbol{\alpha} = \frac{\boldsymbol{p}}{q} \in \mathbb{Q}^d \\ \mu(\{(M, \boldsymbol{\xi}) \in X : \#((\mathbb{Z}^d M + \boldsymbol{\xi}) \cap \mathfrak{Z}^{(\mathfrak{U})}(c, 1, \sigma)) = r\}) & \text{if } \boldsymbol{\alpha} \notin \mathbb{Q}^d, \end{cases}$$

where

$$\mathfrak{Z}^{(\mathfrak{U})}(c_1, c_2, \sigma) = \{\boldsymbol{y} = (y_1, \dots, y_d) \in \mathbb{R}^d : c_1 < y_1 < c_2,\ \boldsymbol{y}_\perp \in \sigma\mathfrak{U}_\perp\}; \tag{9.5}$$

$$\mathfrak{Z}^{(\mathfrak{U})}(c_1, c_2, \sigma, \boldsymbol{\beta}) = \{(\boldsymbol{x}, \boldsymbol{y}) \in D \times \mathbb{R}^d : \boldsymbol{y} \in \mathfrak{Z}^{(\mathfrak{U})}(c_1, c_2, \sigma) + (\sigma\boldsymbol{\beta}(\boldsymbol{x})E_1(\boldsymbol{x}))_\perp\}.$$

In particular, $F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}^{(\mathfrak{U})}(r, \sigma)$ is continuous in σ and independent of $\mathscr{L}$, and if $\boldsymbol{\alpha} \notin \mathbb{Q}$ then it is also independent of $\boldsymbol{\beta}$ and λ.

Theorem 3.4 follows from Theorem 9.1 by taking $\mathfrak{U} = {\mathrm{S}_1'}^{d-1}$. Indeed, we have $\mathfrak{Z}^{({\mathrm{S}_1'}^{d-1})}(c, 1, \sigma) = \mathfrak{Z}(c, \sigma)$ (except if $\sigma = 0$, but then both sets are of measure zero), and in the case $\boldsymbol{\alpha} \in \mathbb{Q}^d$ the volume in (9.4) equals

$$\int_D \mu_q\Big(\Big\{M \in X_q : \#\big((\mathbb{Z}^d + \boldsymbol{\alpha})M \cap (\mathfrak{Z}(c, \sigma) + (\sigma\boldsymbol{\beta}(\boldsymbol{x})E_1(\boldsymbol{x}))_\perp)\big) = r\Big\}\Big)\, d\lambda(\boldsymbol{x}). \tag{9.6}$$

Here we may replace "$(\sigma\boldsymbol{\beta}(\boldsymbol{x})E_1(\boldsymbol{x}))_\perp$" with "$\sigma \| \mathrm{Proj}_{\{\boldsymbol{v}(\boldsymbol{x})\}^\perp} \boldsymbol{\beta}(\boldsymbol{x})\| \cdot \boldsymbol{e}_2$", since (if $d \geq 3$) there is a rotation $\begin{pmatrix} 1 & \boldsymbol{0} \\ {}^t\boldsymbol{0} & K \end{pmatrix} \in \mathrm{SO}(d)$ which takes the second vector to the

first, and μ_q is invariant under the diffeomorphism $X_q \ni M \mapsto M \begin{pmatrix} 1 & \mathbf{0} \\ {}^t\mathbf{0} & K \end{pmatrix} \in X_q$. (If $d = 2$: Then either the two vectors are equal, or they correspond to each other under $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$; in the latter case one chooses $\gamma_0 \in \mathrm{SL}(2,\mathbb{Z})$ with $\boldsymbol{\alpha}\gamma_0 \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) = \boldsymbol{\alpha}$ and then uses the fact that $M \mapsto \gamma_0 \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) M \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ is a well-defined automorphism of X_q onto itself, which preserves μ_q.) Hence we obtain the volume in (3.13).

Proof of Theorem 9.1. To prove the desired uniformity, it suffices to show that, given any continuous functions $\mathbb{R}_{>0} \ni T \mapsto \sigma_T \in \mathbb{R}_{\geq 0}$ and $\mathbb{R}_{>0} \ni T \mapsto c_T \in [0,1]$ such that $\sigma_\infty = \lim_{T\to\infty} \sigma_T$ and $c_\infty = \lim_{T\to\infty} c_T$ exist, we have

(9.7)
$$\lim_{T\to\infty} \lambda(\{\boldsymbol{x} \in \mathbb{R}^{d-1} : \mathcal{N}^{(\mathfrak{U})}_{c_T,T}(\sigma_T T^{-1/(d-1)}, \boldsymbol{x}, \boldsymbol{\beta}(\boldsymbol{x})) = r\}) = F^{(\mathfrak{U})}_{c_\infty,\boldsymbol{\alpha},\boldsymbol{\beta}}(r, \sigma_\infty),$$

where the right-hand side is given by (9.4).

In the following we let $\mathrm{SL}(d,\mathbb{R})$ and $\mathrm{ASL}(d,\mathbb{R})$ act on $\mathbb{R}^{d-1} \times \mathbb{R}^d$ by leaving the first entry fixed and acting as usual on the second entry:

$$g : \mathbb{R}^{d-1} \times \mathbb{R}^d \to \mathbb{R}^{d-1} \times \mathbb{R}^d, \qquad (\boldsymbol{x}, \boldsymbol{y}) \mapsto (\boldsymbol{x}, \boldsymbol{y}g). \tag{9.8}$$

Set, for any $\sigma \geq 0$, $0 \leq c_1 \leq c_2$, $T > 0$,

$$\begin{aligned} \mathfrak{Z}^{(\mathfrak{U})}_T(c_1, c_2, \sigma, \boldsymbol{\beta}) = \Big\{ &(\boldsymbol{x}, \boldsymbol{y}) \in D \times \mathbb{R}^d \ : \ c_1 T \leq \|\boldsymbol{y}\| < c_2 T, \\ &\boldsymbol{y} - \sigma T^{-\frac{1}{d-1}} \boldsymbol{\beta}(\boldsymbol{x}) E_1(\boldsymbol{x}) \in \mathbb{R}_{>0} \boldsymbol{e}_1 + \sigma T^{-\frac{1}{d-1}} \mathfrak{U}_\perp \Big\} \begin{pmatrix} T^{-1} & \mathbf{0} \\ {}^t\mathbf{0} & T^{1/(d-1)} 1_{d-1} \end{pmatrix}. \end{aligned} \tag{9.9}$$

We then have for all $\boldsymbol{x} \in D$,

$$\begin{aligned} &\mathcal{N}^{(\mathfrak{U})}_{c,T}(\sigma T^{-1/(d-1)}, \boldsymbol{x}, \boldsymbol{\beta}(\boldsymbol{x})) \\ &\qquad = \#\big(\mathfrak{Z}^{(\mathfrak{U})}_T(c, 1, \sigma, \boldsymbol{\beta})|_{\boldsymbol{x}} \Phi^{-t}\big(E_1(\boldsymbol{x})^{-1}, \mathbf{0}\big) \cap (\mathbb{Z}^d + \boldsymbol{\alpha}) M_0 \setminus \{\mathbf{0}\}\big), \end{aligned} \tag{9.10}$$

with $T = \mathrm{e}^{(d-1)t}$, so long as T is large enough so that the left-hand side is defined.

Now taking Lemma 9.2 (with $c_{2,T} = 1$) below into account, we see that (9.7) and Theorem 9.1 follow immediately from the Theorems of Section 6. □

The flexibility of taking $c_{2,T} \not\equiv 1$ in the following lemma is not needed for the proof of Theorem 9.1, but it will be convenient later.

LEMMA 9.2. *Let $\sigma_T, c_{1,T}, c_{2,T}$ be continuous functions of $T > 0$ with $\sigma_T \geq 0$, $0 \leq c_{1,T} \leq c_{2,T}$ for all $T > 0$, and such that all three limits $\sigma_\infty = \lim_{T\to\infty} \sigma_T$, $c_{1,\infty} = \lim_{T\to\infty} c_{1,T}$ and $c_{2,\infty} = \lim_{T\to\infty} c_{2,T}$ exist. Then the union*

$$\cup_{T\geq 1} \mathfrak{Z}^{(\mathfrak{U})}_T(c_{1,T}, c_{2,T}, \sigma_T, \boldsymbol{\beta})$$

is bounded, and

$$\lim(\inf \mathfrak{Z}^{(\mathfrak{U})}_T(c_{1,T}, c_{2,T}, \sigma_T, \boldsymbol{\beta}))^\circ \supset \widetilde{\mathfrak{Z}}^{(\mathfrak{U})}(c_{1,\infty}, c_{2,\infty}, \sigma_\infty, \boldsymbol{\beta})^\circ \tag{9.11}$$

and

$$\limsup \mathfrak{Z}_T^{(\mathfrak{U})}(c_{1,T}, c_{2,T}, \sigma_T, \boldsymbol{\beta}) \subset \overline{\widetilde{\mathfrak{Z}}^{(\mathfrak{U})}(c_{1,\infty}, c_{2,\infty}, \sigma_\infty, \boldsymbol{\beta})} \tag{9.12}$$

(*closures and limits taken in* $\mathbb{R}^{d-1} \times \mathbb{R}^d$), *where*

$$\widetilde{\mathfrak{Z}}^{(\mathfrak{U})}(c_1, c_2, \sigma, \boldsymbol{\beta}) := \mathfrak{Z}^{(\mathfrak{U})}(c_1, c_2, \sigma, \boldsymbol{\beta})$$

(*cf.* (9.5)) *if* $c_1 < c_2$, *but* $\widetilde{\mathfrak{Z}}^{(\mathfrak{U})}(c_1, c_1, \sigma, \boldsymbol{\beta}) := \big\{(\boldsymbol{x}, \boldsymbol{y}) \in D \times \mathbb{R}^d : \boldsymbol{y} \in (\{c_1\} \times \sigma\mathfrak{U}_\perp) + (\sigma\boldsymbol{\beta}(\boldsymbol{x})E_1(\boldsymbol{x}))_\perp\big\}$. *Furthermore, the boundary of* $\widetilde{\mathfrak{Z}}^{(\mathfrak{U})}(c_1, c_2, \sigma, \boldsymbol{\beta})$ *intersects* $D \times \mathbb{R}^d$ *in a set of Lebesgue measure zero.*

Proof. Let $C = 1 + \sup_D \|\boldsymbol{\beta}\|$. Take $T > 0$ and consider an arbitrary point $(\boldsymbol{x}, \boldsymbol{y}) \in \mathfrak{Z}_T^{(\mathfrak{U})}(c_{1,T}, c_{2,T}, \sigma_T, \boldsymbol{\beta})$. Set $\boldsymbol{y}' = Ty_1\boldsymbol{e}_1 + T^{-\frac{1}{d-1}}\boldsymbol{y}_\perp$; then $c_{1,T}T \le \|\boldsymbol{y}'\| < c_{2,T}T$ and $\boldsymbol{y}' - \sigma_T T^{-\frac{1}{d-1}}\boldsymbol{\beta}(\boldsymbol{x})E_1(\boldsymbol{x}) \in \mathbb{R}_{>0}\boldsymbol{e}_1 + \sigma_T T^{-\frac{1}{d-1}}\mathfrak{U}_\perp$. From these we conclude

$$-\sigma_T T^{-\frac{d}{d-1}} \sup \|\boldsymbol{\beta}\| < y_1 < c_{2,T} \quad \text{and} \quad (\boldsymbol{y} - \sigma_T\boldsymbol{\beta}(\boldsymbol{x})E_1(\boldsymbol{x}))_\perp \in \sigma_T\mathfrak{U}_\perp. \tag{9.13}$$

Since $\mathfrak{U}_\perp \subset \mathcal{B}_1^d$ the last relation implies $\|\boldsymbol{y}_\perp\| \le C\sigma_T$. The first claim of the lemma follows from the inequalities noted so far.

Now let $\eta > 0$ be given, and take T_0 so large that $c_{1,T} > c_{1,\infty} - \frac{\eta}{2}$, $c_{2,T} < c_{2,\infty} + \eta$, $C\sigma_T T^{-\frac{d}{d-1}} < \frac{\eta}{8}$ and $|\sigma_T - \sigma_\infty| < \eta/C$ hold for all $T \ge T_0$. Let $T \ge T_0$ and consider any point $(\boldsymbol{x}, \boldsymbol{y}) \in \mathfrak{Z}_T^{(\mathfrak{U})}(c_{1,T}, c_{2,T}, \sigma_T, \boldsymbol{\beta})$. Then by (9.13) we have $y_1 \ge -C\sigma_T T^{-\frac{d}{d-1}} > -\frac{\eta}{8}$, but using $\|\boldsymbol{y}'\| \ge c_{1,T}T$ we also conclude $|y_1| \ge c_{1,T} - C\sigma_T T^{-\frac{d}{d-1}} > c_{1,\infty} - \frac{5\eta}{8}$. Together these two inequalities imply in particular that $y_1 > c_{1,\infty} - \eta$. Also, by (9.13), $y_1 < c_{2,T} < c_{2,\infty} + \eta$. From (9.13) we also see that there is some $\boldsymbol{w} \in \mathfrak{U}_\perp$ such that $(\boldsymbol{y} - \sigma_T\boldsymbol{\beta}(\boldsymbol{x})E_1(\boldsymbol{x}))_\perp = \sigma_T\boldsymbol{w}$. Thus

$$(\boldsymbol{y} - \sigma_\infty\boldsymbol{\beta}(\boldsymbol{x})E_1(\boldsymbol{x}))_\perp = \sigma_\infty\boldsymbol{w} + (\sigma_T - \sigma_\infty)(\boldsymbol{w} + (\boldsymbol{\beta}(\boldsymbol{x})E_1(\boldsymbol{x}))_\perp), \tag{9.14}$$

and here $\big\|\boldsymbol{w} + (\boldsymbol{\beta}(\boldsymbol{x})E_1(\boldsymbol{x}))_\perp\big\| < C$ and $|\sigma_T - \sigma_\infty| < \eta/C$, so that

$$(\boldsymbol{y} - \sigma_\infty\boldsymbol{\beta}(\boldsymbol{x})E_1(\boldsymbol{x}))_\perp \in \sigma_\infty\mathfrak{U}_\perp + \mathcal{B}_\eta^d. \tag{9.15}$$

Hence we have proved that for each $T \ge T_0$ we have

$$\begin{aligned} &\mathfrak{Z}_T^{(\mathfrak{U})}(c_{1,T}, c_{2,T}, \sigma_T, \boldsymbol{\beta}) \\ &\quad\subset \Big\{(\boldsymbol{x}, \boldsymbol{y}) \in D \times \mathbb{R}^d \ : \ c_{1,\infty} - \eta < \boldsymbol{y} \cdot \boldsymbol{e}_1 < c_{2,\infty} + \eta, \\ &\qquad\qquad (\boldsymbol{y} - \sigma_\infty\boldsymbol{\beta}(\boldsymbol{x})E_1(\boldsymbol{x}))_\perp \in \sigma_\infty\mathfrak{U}_\perp + \mathcal{B}_\eta^d\Big\}. \end{aligned} \tag{9.16}$$

We have seen that such a T_0 exists for any $\eta > 0$; this fact leads easily to (9.12).

We now turn to (9.11). Assume $(\boldsymbol{x}_0, \boldsymbol{y}_0) \in \widetilde{\mathfrak{Z}}^{(\mathfrak{U})}(c_{1,\infty}, c_{2,\infty}, \sigma_\infty, \boldsymbol{\beta})^\circ$, and take $\eta > 0$ so that

$$(\boldsymbol{x}_0 + \mathcal{B}_{2\eta}^{d-1}) \times (\boldsymbol{y}_0 + \mathcal{B}_{2\eta}^{d}) \subset \widetilde{\mathfrak{Z}}^{(\mathfrak{U})}(c_{1,\infty}, c_{2,\infty}, \sigma_\infty, \boldsymbol{\beta}). \tag{9.17}$$

Then we must have $\sigma_\infty > 0$ and $c_{1,\infty} < c_{2,\infty}$. Take T_0 so large that each of the following five inequalities hold when $T \geq T_0$:

$$\begin{aligned} &\sigma_T > 0; \qquad \Big|\frac{\sigma_\infty}{\sigma_T} - 1\Big| < \frac{\eta}{C\sigma_\infty}; \qquad \sigma_T T^{-\frac{d}{d-1}} < \frac{\eta}{C}; \\ &c_{1,T} \leq c_{1,\infty} + \eta; \qquad c_{2,T} - C\sigma_\infty T^{-\frac{d}{d-1}} > c_{2,\infty} - \eta. \end{aligned} \tag{9.18}$$

We then claim

$$(\boldsymbol{x}_0 + \mathcal{B}_{\eta}^{d-1}) \times (\boldsymbol{y}_0 + \mathcal{B}_{\eta}^{d}) \subset \mathfrak{Z}_T^{(\mathfrak{U})}(c_{1,T}, c_{2,T}, \sigma_T, \boldsymbol{\beta}), \qquad \forall T \geq T_0. \tag{9.19}$$

This implies $(\boldsymbol{x}_0, \boldsymbol{y}_0) \in \lim(\inf \mathfrak{Z}_T^{(\mathfrak{U})}(c_{1,T}, c_{2,T}, \sigma_T, \boldsymbol{\beta}))^\circ$, and hence (9.11) will be proved, since $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ was arbitrary in $\widetilde{\mathfrak{Z}}^{(\mathfrak{U})}(c_{1,\infty}, c_{2,\infty}, \sigma_\infty, \boldsymbol{\beta})^\circ$.

To prove (9.19), let $(\boldsymbol{x}, \boldsymbol{y})$ be an arbitrary point in $(\boldsymbol{x}_0 + \mathcal{B}_{\eta}^{d-1}) \times (\boldsymbol{y}_0 + \mathcal{B}_{\eta}^{d})$, and take $T \geq T_0$. Write $\boldsymbol{y}' = Ty_1\boldsymbol{e}_1 + T^{-\frac{1}{d-1}}\boldsymbol{y}_\perp$. Using $C\sigma_T T^{-\frac{d}{d-1}} < \eta$ we get

$$Ty_1 - \sigma_T T^{-\frac{1}{d-1}} \boldsymbol{\beta}(\boldsymbol{x}) E_1(\boldsymbol{x}) \cdot \boldsymbol{e}_1 \geq T\eta - \sigma_T T^{-\frac{1}{d-1}} \sup \|\boldsymbol{\beta}\| > 0. \tag{9.20}$$

Next (9.17) implies $(\boldsymbol{y} - \sigma_\infty \boldsymbol{\beta}(\boldsymbol{x}) E_1(\boldsymbol{x}))_\perp + \big(\{0\} \times \mathcal{B}_\eta^{d-1}\big) \subset \sigma_\infty \mathfrak{U}_\perp$. In particular, $\|\boldsymbol{y}_\perp\| < C\sigma_\infty$, and using $\big|\frac{\sigma_\infty}{\sigma_T} - 1\big| < \frac{\eta}{C\sigma_\infty}$ we get $\big|\frac{\sigma_\infty}{\sigma_T} - 1\big| \cdot \|\boldsymbol{y}_\perp\| < \eta$ and hence

$$(\boldsymbol{y} - \sigma_\infty \boldsymbol{\beta}(\boldsymbol{x}) E_1(\boldsymbol{x}))_\perp + \Big(\frac{\sigma_\infty}{\sigma_T} - 1\Big)\boldsymbol{y}_\perp \in \sigma_\infty \mathfrak{U}_\perp. \tag{9.21}$$

In other words $(\boldsymbol{y} - \sigma_T \boldsymbol{\beta}(\boldsymbol{x}) E_1(\boldsymbol{x}))_\perp \in \sigma_T \mathfrak{U}_\perp$, and thus

$$(\boldsymbol{y}' - \sigma_T T^{-\frac{1}{d-1}} \boldsymbol{\beta}(\boldsymbol{x}) E_1(\boldsymbol{x}))_\perp \in \sigma_T T^{-\frac{1}{d-1}} \mathfrak{U}_\perp. \tag{9.22}$$

Finally (9.17) gives $c_{1,\infty} + \eta \leq y_1 \leq c_{2,\infty} - \eta$, and using $c_{1,T} \leq c_{1,\infty} + \eta$ and $c_{2,\infty} - \eta < c_{2,T} - C\sigma_\infty T^{-\frac{d}{d-1}}$ we obtain

$$c_{1,T} T \leq \|\boldsymbol{y}'\| < c_{2,T} T. \tag{9.23}$$

But (9.23), (9.20), (9.22) imply $(\boldsymbol{x}, \boldsymbol{y}) \in \mathfrak{Z}_T^{(\mathfrak{U})}(c_{1,T}, c_{2,T}, \sigma_T, \boldsymbol{\beta})$, and hence (9.19) is proved.

Finally, the fact that $(D \times \mathbb{R}^d) \cap \partial\widetilde{\mathfrak{Z}}^{(\mathfrak{U})}(c_1, c_2, \sigma, \boldsymbol{\beta})$ has Lebesgue measure zero follows from

$$\begin{aligned} &(D \times \mathbb{R}^d) \cap \partial\widetilde{\mathfrak{Z}}^{(\mathfrak{U})}(c_1, c_2, \sigma, \boldsymbol{\beta}) \\ &\subset \big\{(\boldsymbol{x}, \boldsymbol{y}) \in D \times \mathbb{R}^d \,:\, \boldsymbol{y} \cdot \boldsymbol{e}_1 \in \{c_1, c_2\},\ \boldsymbol{y}_\perp \in \sigma(\boldsymbol{\beta}(\boldsymbol{x}) E_1(\boldsymbol{x}))_\perp + \big(\{0\} \times \overline{\mathcal{B}_\sigma^{d-1}}\big)\big\} \\ &\quad \cup \big\{(\boldsymbol{x}, \boldsymbol{y}) \in D \times \mathbb{R}^d \,:\, c_1 < \boldsymbol{y} \cdot \boldsymbol{e}_1 < c_2,\ \boldsymbol{y}_\perp \in \sigma(\boldsymbol{\beta}(\boldsymbol{x}) E_1(\boldsymbol{x}))_\perp + \partial(\sigma\mathfrak{U}_\perp)\big\}, \end{aligned} \tag{9.24}$$

using $\partial(\mathfrak{U}_\perp) = (\partial\mathfrak{U})_\perp$, and our assumption that $\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\partial\mathfrak{U}) = 0$. □

The proof of Theorem 3.7 is almost identical to the proof of Theorem 3.4, using the theorems of Section 6.3.

We proceed to the proofs of Theorems 3.9 and 3.10. To be in line with the notation used in the previous proofs, we again write $\boldsymbol{v} = \boldsymbol{e}_1 E_1(\boldsymbol{x})^{-1}$ ($\boldsymbol{x} \in D$), and write λ also for the lift of λ to the variable $\boldsymbol{x}$. Set

$$\mathfrak{Z}_T(c,\mathcal{Q}) = \big\{(\boldsymbol{x},\boldsymbol{y}) \in D \times \mathbb{R}^d \,:\, cT \le \|\boldsymbol{y}\| < T, \\ \mathbb{R}_{>0}\boldsymbol{e}_1 \cap (\mathcal{Q}_T E_1(\boldsymbol{x}) + \boldsymbol{y}) \ne \varnothing\big\} \begin{pmatrix} T^{-1} & \boldsymbol{0} \\ {}^{\mathrm{t}}\boldsymbol{0} & T^{1/(d-1)} 1_{d-1} \end{pmatrix}. \tag{9.25}$$

For the counting function defined in (3.22) we have

$$\mathcal{N}_{c,T}(\mathcal{Q}, \boldsymbol{e}_1 E_1(\boldsymbol{x})^{-1}) = \#\big(\mathfrak{Z}_T(c,\mathcal{Q})|_{\boldsymbol{x}} \Phi^{-t}(E_1(\boldsymbol{x})^{-1}, \boldsymbol{0}) \cap (\mathbb{Z}^d + \boldsymbol{\alpha}) M_0 \setminus \{\boldsymbol{0}\}\big) \tag{9.26}$$

with $T = \mathrm{e}^{(d-1)t}$. The primitive case is analogous.

Theorems 3.9 and 3.10 are again a consequence of the theorems in Section 6 and the following lemma.

LEMMA 9.3. *The union $\cup_{T\ge 1}\mathfrak{Z}_T(c,\mathcal{Q})$ is bounded, and we have*

$$\lim(\inf \mathfrak{Z}_T(c,\mathcal{Q}))^\circ \supset \mathfrak{Z}(c,\mathcal{Q})^\circ, \qquad \lim\overline{\sup \mathfrak{Z}_T(c,\mathcal{Q})} \subset \overline{\mathfrak{Z}(c,\mathcal{Q})}, \tag{9.27}$$

where

$$\mathfrak{Z}(c,\mathcal{Q}) := \big\{(\boldsymbol{x},\boldsymbol{y}) \in D \times \mathbb{R}^d \,:\, c < y_1 < 1,\ (y_2,\dots,y_d) \in -(\mathcal{Q}E_1(\boldsymbol{x}))_\perp\big\} \tag{9.28}$$

is a bounded set whose boundary intersects $D \times \mathbb{R}^d$ in a set of Lebesgue measure zero.

Proof. This is very similar to the proof of Lemma 9.2 (but slightly easier, since c and $\mathcal{Q}$ are kept fixed). To prove the last statement one first verifies that

$$(D \times \mathbb{R}^d) \cap \partial\mathfrak{Z}(c,\mathcal{Q}) \\ \subset \Big\{(\boldsymbol{x},\boldsymbol{y}) \in D \times \mathbb{R}^d \,:\, y_1 \in \{c,1\},\ \boldsymbol{y}_\perp \in -\overline{(\mathcal{Q}E_1(\boldsymbol{x}))_\perp}\Big\} \\ \cup \Big\{(\boldsymbol{x},\boldsymbol{y}) \in D \times \mathbb{R}^d \,:\, y_1 \in [c,1],\ \boldsymbol{y}_\perp \in -\partial\big((\mathcal{Q}E_1(\boldsymbol{x}))_\perp\big)\Big\}. \tag{9.29}$$

Here the first set clearly has measure zero, and the second set has measure

$$(1-c)\int_{\boldsymbol{x}\in D} \mathrm{vol}_{\mathbb{R}^{d-1}}\big(\partial\big((\mathcal{Q}E_1(\boldsymbol{x}))_\perp\big)\big)\, d\boldsymbol{x}, \tag{9.30}$$

which is zero exactly because of the technical assumption made just below (3.20). □

Lemma 9.3 is applied in the following way: If $\boldsymbol{\alpha} \in \mathbb{Q}^d$ then by (9.26), Remark 6.6 and Lemma 9.3 the limit in (3.23) exists, and equals

$$\int_D \int_{X_q} I\Big(\#\big(\mathfrak{Z}(c,\mathcal{Q})|_{\boldsymbol{x}} \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M\big) = r\Big)\, d\mu_q(M)\, d\lambda(\boldsymbol{x}). \tag{9.31}$$

But we have from (9.28), since $\boldsymbol{v} = \boldsymbol{e}_1 E_1(\boldsymbol{x})^{-1}$:

$$\begin{aligned}(9.32)\qquad \mathfrak{Z}(c,\mathfrak{Q})|_{\boldsymbol{x}} &= \{\boldsymbol{y} \in \mathbb{R}^d \,:\, c < \boldsymbol{y}\cdot\boldsymbol{e}_1 < 1,\ \mathbb{R}\boldsymbol{e}_1 \cap (\mathfrak{Q}E_1(\boldsymbol{x}) + \boldsymbol{y}) \neq \varnothing\} \\ &= \{\boldsymbol{y} \in \mathbb{R}^d \,:\, c < \boldsymbol{y}\cdot\boldsymbol{v} < 1,\ \mathbb{R}\boldsymbol{v} \cap (\mathfrak{Q} + \boldsymbol{y}) \neq \varnothing\} E_1(\boldsymbol{x}).\end{aligned}$$

Hence by substituting $M = M' E_1(\boldsymbol{x})$ in the inner integral in (9.31) we obtain the formula stated in Theorem 3.9. The proof in the case $\boldsymbol{\alpha} \notin \mathbb{Q}^d$ is entirely similar, and so is the proof of Theorem 3.10.

9.2. *Averaging over* $\boldsymbol{\alpha}$. Naturally, one can also prove $\boldsymbol{\alpha}$-averaged (or $\boldsymbol{q}$-averaged) versions of all the limit results obtained in the present paper. In this section we discuss this to the extent necessary to give a proof of Theorem 1.2.

We first give an averaged version of Corollary 4.1. Recall that if $\boldsymbol{\alpha} \notin \mathbb{Q}^d$ then $\Phi_{\boldsymbol{\alpha}}(\xi)$ is independent of $\boldsymbol{\alpha}$, and we write $\Phi(\xi)$ for this function.

COROLLARY 9.4. *Fix a lattice* $\mathfrak{L} = \mathbb{Z}^d M_0$ *and let* λ *be a Borel probability measure on* $\mathrm{T}^1(\mathbb{R}^d) = \mathbb{R}^d \times \mathrm{S}_1^{d-1}$ *which is absolutely continuous with respect to Lebesgue measure* $\mathrm{vol}_{\mathbb{R}^d} \times \mathrm{vol}_{\mathrm{S}_1^{d-1}}$. *Then, for every* $\xi \geq 0$,

$$(9.33)\qquad \lim_{\rho\to 0} \lambda(\{(\boldsymbol{q},\boldsymbol{v}) \in \mathrm{T}^1(\mathscr{H}_\rho) \,:\, \rho^{d-1}\tau_1(\boldsymbol{q},\boldsymbol{v};\rho) \geq \xi\}) = \int_\xi^\infty \Phi(\xi')\,d\xi'.$$

Proof. By the theorem of Radon-Nikodym we have

$$d\lambda(\boldsymbol{q},\boldsymbol{v}) = f(\boldsymbol{q},\boldsymbol{v})\,d\boldsymbol{q}\,d\,\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\boldsymbol{v})$$

for some nonnegative function $f \in \mathrm{L}^1(\mathbb{R}^d \times \mathrm{S}_1^{d-1})$ with $\|f\|_{\mathrm{L}^1} = 1$. By Fubini's theorem, the left-hand side of (9.33) equals

$$(9.34)\qquad \lim_{\rho\to 0} \int_{\mathbb{R}^d} \Big(\int_{\mathrm{S}_1^{d-1}} I\big(\rho^{d-1}\tau_1(\boldsymbol{q},\boldsymbol{v};\rho) \geq \xi\big) f(\boldsymbol{q},\boldsymbol{v})\,d\,\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\boldsymbol{v}) \Big)\, d\boldsymbol{q},$$

where the indicator function $I\big(\dots\big)$ is interpreted as zero whenever $\boldsymbol{q} \notin \mathscr{H}_\rho$. For almost every $\boldsymbol{q} \in \mathbb{R}^d$ we have $f(\boldsymbol{q},\cdot) \in \mathrm{L}^1(\mathrm{S}_1^{d-1})$ and $-\boldsymbol{q}M_0^{-1} \notin \mathbb{Q}^d$, and for each such (fixed) point $\boldsymbol{q}$, Corollary 4.1 implies that the inner integral in (9.34) tends to

$$(9.35)\qquad \Big(\int_{\mathrm{S}_1^{d-1}} f(\boldsymbol{q},\boldsymbol{v})\,d\,\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\boldsymbol{v}) \Big) \cdot \int_\xi^\infty \Phi(\xi')\,d\xi' \qquad \text{as } \rho \to 0.$$

By Lebesgue's Bounded Convergence Theorem (with $\boldsymbol{q} \mapsto \int_{\mathrm{S}_1^{d-1}} f(\boldsymbol{q},\boldsymbol{v})\,d\,\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\boldsymbol{v})$ as a majorant function), we may change the order between $\lim_{\rho\to 0}$ and $\int_{\mathbb{R}^d}$ in (9.34), thus obtaining (9.33). □

Proof of Theorem 1.2. Let M be the set of nonnegative functions $f \in \mathrm{L}^1(\mathbb{R}^d \times \mathrm{S}_1^{d-1})$ with $\|f\|_{\mathrm{L}^1} = 1$. By the Theorem of Radon-Nikodym and (1.7), our task is

to prove that for each $f \in M$ we have

$$\lim_{\rho\to 0} \int_{\mathcal{H}_\rho} \int_{\mathrm{S}_1^{d-1}} I\left(\rho^{d-1}\tau_1(\boldsymbol{q},\boldsymbol{v};\rho) \ge \xi\right) \cdot \rho^{d(d-1)} f(\rho^{d-1}\boldsymbol{q},\boldsymbol{v})\, d\,\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\boldsymbol{v})\, d\boldsymbol{q} = \int_\xi^\infty \Phi(\xi')\, d\xi'. \tag{9.36}$$

In fact it suffices to prove (9.36) when $f \in M$ is continuous and of compact support, since the subset of such functions is dense in M with respect to the L^1-norm.

Using the $\mathcal{L}$-periodicity of $\tau_1(\cdot,\boldsymbol{v};\rho)$, the double integral in (9.36) can be expressed as

$$\int_{F\cap\mathcal{H}_\rho} \int_{\mathrm{S}_1^{d-1}} I\left(\rho^{d-1}\tau_1(\boldsymbol{q}_0,\boldsymbol{v};\rho) \ge \xi\right) \times \Big\{\rho^{d(d-1)} \sum_{\boldsymbol{q}\in\boldsymbol{q}_0+\mathcal{L}} f(\rho^{d-1}\boldsymbol{q},\boldsymbol{v})\Big\}\, d\,\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\boldsymbol{v})\, d\boldsymbol{q}_0, \tag{9.37}$$

where $F \subset \mathbb{R}^d$ is a fundamental parallelogram for $\mathcal{L}$. But for f continuous and of compact support, the expression within the brackets in (9.37) tends to $h(\boldsymbol{v}) := \int_{\mathbb{R}^d} f(\boldsymbol{q},\boldsymbol{v})\, d\boldsymbol{q}$ as $\rho\to 0$, uniformly with respect to $\boldsymbol{v}\in \mathrm{S}_1^{d-1}$ and $\boldsymbol{q}_0 \in F$. Thus Theorem 1.2 follows from Corollary 9.4, applied with

$$d\lambda(\boldsymbol{q},\boldsymbol{v}) = \chi_F(\boldsymbol{q}) h(\boldsymbol{v})\, d\boldsymbol{q}\, d\,\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\boldsymbol{v}). \qquad \square$$

9.3. *Proofs for Section* 4.

Proof of Theorem 4.4.

As in Section 9.1 we fix a smooth map $E_1 : D \to \mathrm{SO}(d)$ such that $\boldsymbol{v} = \boldsymbol{v}(\boldsymbol{x}) = \boldsymbol{e}_1 E_1(\boldsymbol{x})^{-1} \in \mathrm{S}_1^{d-1}$ gives a diffeomorphism between the bounded open set $D \subset \mathbb{R}^{d-1}$ and S_1^{d-1} minus one point. However we now make the extra requirement that $E_1(\boldsymbol{x}) = K(\boldsymbol{v}(\boldsymbol{x}))$ for all $\boldsymbol{x}\in D$.[6]

We again write λ and $\boldsymbol{\beta}$ also for the lifts of λ and $\boldsymbol{\beta}$ to the variable $\boldsymbol{x}$. Now the measure appearing in the limit in (4.13) equals, with $\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{x}) = \boldsymbol{q} + \rho\boldsymbol{\beta}(\boldsymbol{x})$:

$$\lambda\big(\{\boldsymbol{x}\in D : \rho^{d-1}\tau_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{x}),\boldsymbol{v}(\boldsymbol{x});\rho) \in [\xi_1,\xi_2), \; -\boldsymbol{w}_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{x}),\boldsymbol{v}(\boldsymbol{x});\rho) \in \mathfrak{U}E_1(\boldsymbol{x})^{-1}\}\big). \tag{9.38}$$

This is well defined for ρ small; more specifically, if ρ is sufficiently small then $(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{x}),\boldsymbol{v}(\boldsymbol{x})) \in \mathrm{T}^1(\mathcal{H}_\rho)$ for all $\boldsymbol{x}\in D$, so that $\tau_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{x}),\boldsymbol{v}(\boldsymbol{x});\rho)$ and (if $\tau_1<\infty$) $\boldsymbol{w}_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{x}),\boldsymbol{v}(\boldsymbol{x});\rho)$ are defined. (For recall that if $\boldsymbol{q}\in\mathcal{L}$ then by our assumption on $\boldsymbol{\beta}$ we have $\|\boldsymbol{\beta}(\boldsymbol{x})\| \ge 1$ everywhere.)

For technical reasons we will prove Theorem 4.4 under the extra assumption that $\xi_1 > 0$. This is no loss of generality, for once that proof is complete, the

[6]For example, we may choose $D = \mathcal{R}_\pi^{d-1}$ and $E_1(\boldsymbol{x}) = K(\boldsymbol{e}_1 E(\boldsymbol{x}) K_0)$ where $E(\boldsymbol{x}) = \exp\left(\begin{smallmatrix} 0 & \boldsymbol{x} \\ -{}^t\boldsymbol{x} & 0_{d-1}\end{smallmatrix}\right)$ and K_0 is any fixed matrix in $\mathrm{SO}(d)$ such that $\boldsymbol{v} = -\boldsymbol{e}_1 K_0$ is the unique point where $K(\boldsymbol{v})$ is not smooth.

remaining case $\xi_1 = 0$ follows by a simple limit argument, using Corollary 4.2 in the form $\lim_{\rho\to 0} \lambda(\{\boldsymbol{v} \in \mathrm{S}_1^{d-1} : \rho^{d-1}\tau_1 < \xi\}) = 1 - F_{0,\boldsymbol{\alpha},\boldsymbol{\beta}}(0, \xi^{1/(d-1)})$ together with the fact that $\lim_{\xi\to 0} F_{0,\boldsymbol{\alpha},\boldsymbol{\beta}}(0, \xi^{1/(d-1)}) = 1$ (cf. Remark 3.6).

The measure in (9.38) can be bounded from above and below using the counting function $\mathcal{N}^{(\mathfrak{U})}_{c,T}(\rho, \boldsymbol{x}, \boldsymbol{w})$ (cf. (9.1)), taken with respect to the affine lattice $\mathscr{L}_{\boldsymbol{\alpha}} = \mathscr{L} - \boldsymbol{q}$, as follows. We will use the shorthand notation

$$\mathcal{N}_{c,T}(\rho, \boldsymbol{x}, \boldsymbol{w}) := \mathcal{N}^{(\mathrm{S}_1'^{\,d-1})}_{c,T}(\rho, \boldsymbol{x}, \boldsymbol{w}),$$

which is natural in view of (9.2). Let $C = 1 + \sup_D \|\boldsymbol{\beta}\|$. Now for any $0 < \xi_1 < \xi_2$ and any $\rho > 0$ so small that $\xi_1\rho^{1-d} - C\rho > 0$, $\xi_1\rho^{1-d} + C\rho < \xi_2\rho^{1-d} - C\rho$ and $(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{x}), \boldsymbol{v}(\boldsymbol{x})) \in \mathrm{T}^1(\mathscr{H}_\rho)$ for all $\boldsymbol{x} \in D$, we have:

$$\begin{aligned}
(9.39)\quad &\lambda\Big(\big\{\boldsymbol{x} \in D : \mathcal{N}_{0,T_1}(\rho, \boldsymbol{x}, \boldsymbol{\beta}(\boldsymbol{x})) = 0,\\
&\qquad \mathcal{N}^{(\mathfrak{U})}_{c_2,T_2}(\rho, \boldsymbol{x}, \boldsymbol{\beta}(\boldsymbol{x})) \geq 1,\ \mathcal{N}_{c_3,T_3}(\rho, \boldsymbol{x}, \boldsymbol{\beta}(\boldsymbol{x})) \leq 1\big\}\Big)\\
&\leq \lambda\Big(\big\{\boldsymbol{x} \in D : \rho^{d-1}\tau_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{x}), \boldsymbol{v}(\boldsymbol{x}); \rho) \in [\xi_1, \xi_2),\\
&\qquad -\boldsymbol{w}_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{x}), \boldsymbol{v}(\boldsymbol{x}); \rho) \in \mathfrak{U}E_1(\boldsymbol{x})^{-1}\big\}\Big)\\
&\leq \lambda\Big(\big\{\boldsymbol{x} \in D : \mathcal{N}_{0,T_4}(\rho, \boldsymbol{x}, \boldsymbol{\beta}(\boldsymbol{x})) = 0,\ \mathcal{N}^{(\mathfrak{U})}_{c_5,T_5}(\rho, \boldsymbol{x}, \boldsymbol{\beta}(\boldsymbol{x})) \geq 1\big\}\Big),
\end{aligned}$$

where $T_j > 0$, $c_j \in [0, 1]$ are defined through $T_1 = c_2T_2 = c_3T_3 = \xi_1\rho^{1-d} + C\rho$, $T_2 = \xi_2\rho^{1-d} - C\rho$, $T_3 = T_5 = \xi_2\rho^{1-d} + C\rho$, $T_4 = c_5T_5 = \xi_1\rho^{1-d} - C\rho$.

In order to prove (9.39), let $\boldsymbol{x}$ be any point in D with $\mathcal{N}_{0,T_1}(\rho, \boldsymbol{x}, \boldsymbol{\beta}(\boldsymbol{x})) = 0$, $\mathcal{N}^{(\mathfrak{U})}_{c_2,T_2}(\rho, \boldsymbol{x}, \boldsymbol{\beta}(\boldsymbol{x})) \geq 1$ and $\mathcal{N}_{c_3,T_3}(\rho, \boldsymbol{x}, \boldsymbol{\beta}(\boldsymbol{x})) \leq 1$. To show the first inequality in (9.39) it suffices to prove that these conditions imply $\tau_1 := \tau_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{x}), \boldsymbol{v}(\boldsymbol{x}); \rho) \in [\rho^{1-d}\xi_1, \rho^{1-d}\xi_2)$ and $\boldsymbol{w}_1 := \boldsymbol{w}_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{x}), \boldsymbol{v}(\boldsymbol{x}); \rho) \in -\mathfrak{U}E_1(\boldsymbol{x})^{-1}$.

It follows from $\mathcal{N}^{(\mathfrak{U})}_{c_2,T_2}(\rho, \boldsymbol{x}, \boldsymbol{\beta}(\boldsymbol{x})) \geq 1$ that there is some $\boldsymbol{y} \in \mathscr{L}_{\boldsymbol{\alpha}} \setminus \{\boldsymbol{0}\}$ with $\xi_1\rho^{1-d} + C\rho \leq \|\boldsymbol{y}\| < \xi_2\rho^{1-d} - C\rho$ and $(\boldsymbol{y} - \rho\boldsymbol{\beta}(\boldsymbol{x}))E_1(\boldsymbol{x}) \in \mathbb{R}_{>0}\boldsymbol{e}_1 + \rho\mathfrak{U}_\perp$. Since $\mathfrak{U} \subset \mathrm{S}_1'^{\,d-1}$, it follows that there exist $\boldsymbol{w} \in \mathfrak{U}$ and $t > -\rho$ such that

$$(\boldsymbol{y} - \rho\boldsymbol{\beta}(\boldsymbol{x}))E_1(\boldsymbol{x}) = t\boldsymbol{e}_1 + \rho\boldsymbol{w}.$$

This implies in particular that $\|\boldsymbol{y}\| - C\rho \leq t \leq \|\boldsymbol{y}\| + C\rho$, and thus

$$(9.40)\qquad \xi_1\rho^{1-d} \leq t < \xi_2\rho^{1-d}.$$

Set $\boldsymbol{y}' := \boldsymbol{y} + \boldsymbol{q} \in \mathscr{L} \setminus \{\boldsymbol{q}\}$ and recall $\boldsymbol{v} = \boldsymbol{v}(\boldsymbol{x}) = \boldsymbol{e}_1E_1(\boldsymbol{x})^{-1}$; then our equality says

$$(9.41)\qquad \boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{x}) + t\boldsymbol{v} = \boldsymbol{y}' - \rho\boldsymbol{w}E_1(\boldsymbol{x})^{-1}.$$

This implies $\tau_1 \leq t$. Furthermore, using $\mathcal{N}_{0,T_1}(\rho, \boldsymbol{x}, \boldsymbol{\beta}(\boldsymbol{x})) = 0$ together with our requirement that if $\boldsymbol{q} \in \mathscr{L}$ then $(\boldsymbol{\beta}(\boldsymbol{v}) + \mathbb{R}_{>0}\boldsymbol{v}) \cap \mathcal{B}_1^d = \varnothing$ for all $\boldsymbol{v} \in \mathrm{S}_1^{d-1}$, we conclude $\xi_1\rho^{1-d} \leq \tau_1$.

We claim that in fact $\tau_1 = t$ holds. Assume the opposite; then we have $\xi_1\rho^{1-d} \le \tau_1 < t < \xi_2\rho^{1-d}$. By the definition of τ_1 there exists some $\boldsymbol{y}'' \in \mathscr{L}_{\boldsymbol{\alpha}} \setminus \{\mathbf{0}\}$ such that $\rho\boldsymbol{\beta}(\boldsymbol{x}) + (\tau_1 + \varepsilon)\boldsymbol{v} \in \mathscr{B}_\rho^d + \boldsymbol{y}''$ for all sufficiently small $\varepsilon > 0$. Then $\|\boldsymbol{y}''\| \le \tau_1 + C\rho$, and also since $\mathscr{N}_{0,T_1}(\rho, \boldsymbol{x}, \boldsymbol{\beta}(\boldsymbol{x})) = 0$ we must have $\|\boldsymbol{y}''\| \ge T_1 = c_3T_3$; hence we see that $\boldsymbol{y}'' - \rho\boldsymbol{\beta}(\boldsymbol{x})$ lies in the set defining $\mathscr{N}_{c_3,T_3}(\rho, \boldsymbol{x}, \boldsymbol{\beta}(\boldsymbol{x}))$. But $\boldsymbol{y} - \rho\boldsymbol{\beta}(\boldsymbol{x})$ *also* lies in this set, and from $\tau_1 < t$ we see that $\boldsymbol{y} \ne \boldsymbol{y}''$. Hence $\mathscr{N}_{c_3,T_3}(\rho, \boldsymbol{x}, \boldsymbol{\beta}(\boldsymbol{x})) \ge 2$, contradicting our assumptions. Having thus proved $\tau_1 = t$ we obtain $\boldsymbol{w}_1 = -\boldsymbol{w}E_1(\boldsymbol{x})^{-1}$ by the definition of $\boldsymbol{w}_1$, and hence both

$$\boldsymbol{w}_1 \in -\mathfrak{U}E_1(\boldsymbol{x})^{-1} \qquad \text{and} \qquad \tau_1 \in [\xi_1\rho^{1-d}, \xi_2\rho^{1-d}).$$

Hence the proof of the first inequality in (9.39) is completed.

The proof of the second inequality in (9.39) is easier, and we leave it to the reader.

Continuing onwards, let us note the following mild generalization of (9.10). For all $\sigma \ge 0$, $c \ge 0$, $c' > 0$, $\boldsymbol{x} \in D$ and any $T > 0$ so large that the left-hand side is defined, we have

$$\begin{aligned}(9.42)\quad &\mathscr{N}_{c,T}^{(\mathfrak{U})}(\sigma T^{-\frac{1}{d-1}}, \boldsymbol{x}, \boldsymbol{\beta}(\boldsymbol{x})) \\ &= \#\Big(\mathfrak{Z}_{T/c'}^{(\mathfrak{U})}(c'c, c', c'^{-\frac{1}{d-1}}\sigma, \boldsymbol{\beta})|_{\boldsymbol{x}}\Phi^{-t}\big(E_1(\boldsymbol{x})^{-1}, \mathbf{0}\big) \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M_0 \setminus \{\mathbf{0}\}\Big),\end{aligned}$$

with $T/c' = \mathrm{e}^{(d-1)t}$. This follows directly from (9.10) combined with the invariance relation

$$(9.43)\qquad \mathfrak{Z}_{T/c'}^{(\mathfrak{U})}(c'c, c', c'^{-\frac{1}{d-1}}\sigma, \boldsymbol{\beta}) = \mathfrak{Z}_T^{(\mathfrak{U})}(c, 1, \sigma, \boldsymbol{\beta}) \begin{pmatrix} c' & \mathbf{0} \\ {}^{\mathrm{t}}\mathbf{0} & c'^{-1/(d-1)}1_{d-1} \end{pmatrix},$$

which can be verified straight from the definition (9.9).

In (9.39), introduce $\sigma_1, \dots, \sigma_5$ through $\rho = \sigma_j T_j^{-\frac{1}{d-1}}$. Using (9.42) and $T_1 = c_2T_2 = c_3T_3$ we see that when ρ is sufficiently small, the left-hand side in (9.39) can be expressed as

$$\begin{aligned}(9.44)\quad &\lambda\Big(\Big\{\boldsymbol{x} \in D : \#\big(\mathfrak{Z}_{T_1/\xi_1}(0, \xi_1, \xi_1^{-\frac{1}{d-1}}\sigma_1, \boldsymbol{\beta})|_{\boldsymbol{x}}\Phi^{-t}\big(E_1(\boldsymbol{x})^{-1}, \mathbf{0}\big) \\ &\qquad\qquad \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M_0 \setminus \{\mathbf{0}\}\big) = 0, \\ &\#\big(\mathfrak{Z}_{T_1/\xi_1}^{(\mathfrak{U})}(\xi_1, \tfrac{\xi_1}{c_2}, (\tfrac{\xi_1}{c_2})^{-\frac{1}{d-1}}\sigma_2, \boldsymbol{\beta})|_{\boldsymbol{x}}\Phi^{-t}\big(E_1(\boldsymbol{x})^{-1}, \mathbf{0}\big) \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M_0 \setminus \{\mathbf{0}\}\big) \ge 1, \\ &\#\big(\mathfrak{Z}_{T_1/\xi_1}(\xi_1, \tfrac{\xi_1}{c_3}, (\tfrac{\xi_1}{c_3})^{-\frac{1}{d-1}}\sigma_3, \boldsymbol{\beta})|_{\boldsymbol{x}}\Phi^{-t}\big(E_1(\boldsymbol{x})^{-1}, \mathbf{0}\big) \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M_0 \setminus \{\mathbf{0}\}\big) \le 1\Big\}\Big)\end{aligned}$$

with $\mathrm{e}^{(d-1)t} = T_1/\xi_1$, and using the notation

$$\mathfrak{Z}_{c,T}(c_1, c_2, \sigma, \boldsymbol{\beta}) := \mathfrak{Z}_{c,T}^{(\mathrm{S}_1'^{\,d-1})}(c_1, c_2, \sigma, \boldsymbol{\beta}).$$

Recall that all c_j, σ_j, T_j are functions of ρ, and, when $\rho \to 0$, we have $T_j \to \infty$, $\sigma_1 \to \xi_1^{\frac{1}{d-1}}$, $\sigma_2, \sigma_3 \to \xi_2^{\frac{1}{d-1}}$, and $c_2, c_3 \to \xi_1/\xi_2$. Note that $e^{(d-1)t} = T_1/\xi_1 =$

$\rho^{1-d} + \frac{C}{\xi_1}\rho$ is strictly decreasing as a function of ρ for small $\rho > 0$; hence for small ρ ($\Leftrightarrow$ large t) we may instead view ρ as a function of t; then also all c_j, σ_j, T_j are functions of t. Now, using an obvious shorthand notation, we have the following sieving type identity for (9.44):

$$\begin{aligned}(9.45)\quad &\lambda\big(\{\boldsymbol{x} \in D : \#F^{(1)}_{t,\boldsymbol{x}} = 0,\ \#F^{(2)}_{t,\boldsymbol{x}} \geq 1,\ \#F^{(3)}_{t,\boldsymbol{x}} \leq 1\}\big)\\ &\quad = \lambda\big(\{\boldsymbol{x} : \#F^{(2)}_{t,\boldsymbol{x}} \geq 1\}\big) - \lambda\big(\{\boldsymbol{x} : \#F^{(1)}_{t,\boldsymbol{x}} \geq 1,\ \#F^{(2)}_{t,\boldsymbol{x}} \geq 1\}\big)\\ &\qquad -\lambda\big(\{\boldsymbol{x} : \#F^{(2)}_{t,\boldsymbol{x}} \geq 1,\ \#F^{(3)}_{t,\boldsymbol{x}} \geq 2\}\big)\\ &\qquad +\lambda\big(\{\boldsymbol{x} : \#F^{(1)}_{t,\boldsymbol{x}} \geq 1,\ \#F^{(2)}_{t,\boldsymbol{x}} \geq 1,\ \#F^{(3)}_{t,\boldsymbol{x}} \geq 2\}\big).\end{aligned}$$

To each of the four terms in the right-hand side we can now apply the $E_1(\boldsymbol{x})$-variant of Theorem 6.3 (see Remark 6.4) and its analogue for rational $\boldsymbol{\alpha}$ (Theorem 6.5, Remark 6.6), in conjunction with Lemma 9.2. If $\boldsymbol{\alpha} \in q^{-1}\mathbb{Z}^d$ then we obtain that as $\rho \to 0$, (9.44) tends to

$$\begin{aligned}(9.46)\quad &(\lambda \times \mu_q)\Big(\Big\{(\boldsymbol{x}, M) \in D \times X_q : \#\big(\mathfrak{Z}(0, \xi_1, 1, \boldsymbol{\beta})|_{\boldsymbol{x}} \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M\big) = 0,\\ &\#\big(\mathfrak{Z}^{(\mathfrak{U})}(\xi_1, \xi_2, 1, \boldsymbol{\beta})|_{\boldsymbol{x}} \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M\big) \geq 1, \#\big(\mathfrak{Z}(\xi_1, \xi_2, 1, \boldsymbol{\beta})|_{\boldsymbol{x}} \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M\big) \leq 1\Big\}\Big).\end{aligned}$$

(Note that here we need not remove $\mathbf{0}$ from the set $\mathbb{Z}^d + \boldsymbol{\alpha}$, since $\mathbf{0}$ is anyway not contained in any of the sets $\mathfrak{Z}(0, \xi_1, 1, \boldsymbol{\beta})|_{\boldsymbol{x}}$ or $\mathfrak{Z}(\xi_1, \xi_2, 1, \boldsymbol{\beta})|_{\boldsymbol{x}}$.) In the case $\boldsymbol{\alpha} \notin \mathbb{Q}^d$ we obtain the same expression but with μ, X and $\mathbb{Z}^d$ in place of μ_q, X_q and $\mathbb{Z}^d + \boldsymbol{\alpha}$.

Similarly, the right-hand side in (9.39) can be expressed as (using also $T_4 = c_5 T_5$)

$$\begin{aligned}(9.47)\quad &\lambda\Big(\Big\{\boldsymbol{x} \in D : \#\big(\mathfrak{Z}_{T_4/\xi_1}(0, \xi_1, \xi_1^{-\frac{1}{d-1}}\sigma_4, \boldsymbol{\beta})|_{\boldsymbol{x}}\, \Phi^{-t}\big(E_1(\boldsymbol{x})^{-1}, \mathbf{0}\big)\\ &\qquad\qquad \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M_0 \setminus \{\mathbf{0}\}\big) = 0,\\ &\#\big(\mathfrak{Z}^{(\mathfrak{U})}_{T_4/\xi_1}(\xi_1, \tfrac{\xi_1}{c_5}, (\tfrac{\xi_1}{c_5})^{-\frac{1}{d-1}}\sigma_5, \boldsymbol{\beta})|_{\boldsymbol{x}}\, \Phi^{-t}\big(E_1(\boldsymbol{x})^{-1}, \mathbf{0}\big) \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M_0 \setminus \{\mathbf{0}\}\big) \geq 1\Big\}\Big)\end{aligned}$$

with $\mathrm{e}^{(d-1)t} = T_4/\xi_1$, and as $\rho \to 0$ this is seen to tend to (if $\boldsymbol{\alpha} \in q^{-1}\mathbb{Z}^d$)

$$\begin{aligned}(9.48)\quad &(\lambda \times \mu_q)\Big(\Big\{(\boldsymbol{x}, M) \in D \times X_q : \#\big(\mathfrak{Z}(0, \xi_1, 1, \boldsymbol{\beta})|_{\boldsymbol{x}} \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M\big) = 0,\\ &\qquad\qquad \#\big(\mathfrak{Z}^{(\mathfrak{U})}(\xi_1, \xi_2, 1, \boldsymbol{\beta})|_{\boldsymbol{x}} \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M\big) \geq 1\Big\}\Big).\end{aligned}$$

Hence we conclude: Given any $0 < \xi_1 < \xi_2$, the lim inf of the expression (9.38) as $\rho \to 0$ is bounded below by (9.46), and the lim sup is bounded above by (9.48) (both with the usual modifications if $\boldsymbol{\alpha} \notin \mathbb{Q}^d$). In order to get successively sharper bounds we will now *split* the original interval $[\xi_1, \xi_2)$ into many small parts, and apply the bounds just proved to each part. We will also use the results on integrals over (X, μ) and (X_q, μ_q) which we developed in Sections 7 and 8. We will give the details for the case $\boldsymbol{\alpha} \in q^{-1}\mathbb{Z}^d$, but exactly the same proof with very small

changes of notation works also in the case $\boldsymbol{\alpha} \notin \mathbb{Q}^d$; in particular all expressions below containing $f_0(\dots)$ of $F_0(\dots)$ will remain unchanged, except that they refer to the definitions (8.27), (8.28) in place of (8.1), (8.3); also some of the continuity issues below are slightly easier in the case $\boldsymbol{\alpha} \notin \mathbb{Q}^d$ since we can refer to Proposition 8.10 for all that we need.

Thus from now on we assume $\boldsymbol{\alpha} \in q^{-1}\mathbb{Z}^d$. Recall that we have defined ((8.1) with $r = 0$)

(9.49)
$$f_0(c_1, c_2, \sigma, \boldsymbol{z}, \boldsymbol{y}) = \nu_{\boldsymbol{y}}\big(\{M \in X_q(\boldsymbol{y}) : (\mathfrak{Z}(c_1, c_2, \sigma) + \boldsymbol{z}) \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M = \varnothing\}\big)$$

and $F_0(\xi, \boldsymbol{w}, \boldsymbol{z}) = f_0(0, \xi, 1, \boldsymbol{z}, \xi \boldsymbol{e}_1 + \boldsymbol{w} + \boldsymbol{z})$. (And $F_0(\xi, \boldsymbol{w}, \boldsymbol{z})$ is the same as $\Phi_{\boldsymbol{\alpha}}(\xi, \boldsymbol{w}, \boldsymbol{z})$ in (4.14) in Theorem 4.4.) Our goal now is to prove that the expression in (9.38) tends to

$$\int_{\boldsymbol{x} \in D} \int_{\xi_1}^{\xi_2} \int_{\mathfrak{U}_\perp} F_0(\xi, \boldsymbol{w}, \boldsymbol{z}_{\boldsymbol{x}})\, d\boldsymbol{w} d\xi d\lambda(\boldsymbol{x}), \tag{9.50}$$

where $\boldsymbol{z}_{\boldsymbol{x}} := (\boldsymbol{\beta}(\boldsymbol{x})E_1(\boldsymbol{x}))_\perp$. Recall that we have already seen in Lemma 8.5 that the function $F_0(\xi, \boldsymbol{w}, \boldsymbol{z})$ is Borel measurable on the $\langle \xi, \boldsymbol{w}, \boldsymbol{z}\rangle$-product space; in particular we are allowed to freely change order of integration in (9.50); hence our present aim is equivalent with proving the limit formula (4.13) in Theorem 4.4.

Let $0 < \xi_1 < \xi_2$ be given once and for all. Take $\varepsilon > 0$ arbitrary (we will take $\varepsilon \to 0$ in the end). Fix a constant C so large that $C \geq 1 + \xi_2 + \sup_D \|\boldsymbol{\beta}\|$, $C \geq \xi_1^{-1}$ and if $d = 2$ then also require $2C^{-1} \leq \varepsilon$. Next choose $\eta > 0$ and $N \in \mathbb{Z}_{\geq 2}$ as in Proposition 8.3, for $r = 0$ and our fixed C and ε; if necessary shrink η further so that $\eta < \xi_1/N$. By Lemma 7.12 we may *also* assume, after possibly shrinking η further, that for every set $U = \mathfrak{Z}(c_1, c_2, 1) + \boldsymbol{z}$ with $\boldsymbol{z} \in \{0\} \times \mathcal{B}_C^{d-1}$ and $c_1 < c_2$ satisfying $\xi_1 \leq c_1 < c_2 \leq \xi_2$ and $c_2 - c_1 \leq \eta$, we have

$$\int_U \nu_{\boldsymbol{y}}\Big(\Big\{M \in X_q(\boldsymbol{y}) : \#\Big(U \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M\Big) \geq 2\Big\}\Big)\, d\boldsymbol{y} \leq \varepsilon(c_2 - c_1). \tag{9.51}$$

We fix a splitting $\xi_1 = \theta_1 < \theta_2 < \cdots < \theta_n = \xi_2$ of the interval $[\xi_1, \xi_2)$ such that $\theta_{j+1} - \theta_j < \eta$ for each $j = 1, 2, \dots, n-1$. Note that (9.38) can be expressed as

$$\sum_{j=1}^{n-1} \lambda\big(\{\boldsymbol{x} \in D : \rho^{d-1}\tau_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{x}), \boldsymbol{v}(\boldsymbol{x}); \rho) \in [\theta_j, \theta_{j+1}), \\ \boldsymbol{w}_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{x}), \boldsymbol{v}(\boldsymbol{x}); \rho) \in -\mathfrak{U}E_1(\boldsymbol{x})^{-1}\}\big). \tag{9.52}$$

We now apply (9.46) and (9.48) for the lim inf and lim sup of each term in this sum. We get that the lim inf of the total expression is

$$\begin{aligned}(9.53)\quad &\geq \sum_{j=1}^{n-1} \int_{\boldsymbol{x}\in D} \mu_q\Big(\Big\{M \in X_q \,:\, \#\big((\mathfrak{Z}(0,\theta_j,1)+\boldsymbol{z_x})\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M\big)=0,\\ &\qquad \#\big((\mathfrak{Z}^{(\mathfrak{U})}(\theta_j,\theta_{j+1},1)+\boldsymbol{z_x})\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M\big)\geq 1,\\ &\qquad\qquad \#\big((\mathfrak{Z}(\theta_j,\theta_{j+1},1)+\boldsymbol{z_x})\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M\big)\leq 1\Big\}\Big)\,d\lambda(\boldsymbol{x}),\end{aligned}$$

where $\mathfrak{Z}^{(\mathfrak{U})}(c_1,c_2,\sigma)$ is defined as in (9.5).

We will next apply Proposition 7.3 to bound each term from below. Let us fix $j\in\{1,\dots,n-1\}$ and $\boldsymbol{x}\in D$ for the moment, set

$$\begin{aligned}(9.54)\qquad S=\Big\{M\in X_q\,:&\#\big((\mathfrak{Z}(0,\theta_j,1)+\boldsymbol{z_x})\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M\big)=0,\\ &\#\big((\mathfrak{Z}^{(\mathfrak{U})}(\theta_j,\theta_{j+1},1)+\boldsymbol{z_x})\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M\big)\geq 1\Big\}\end{aligned}$$

and denote by S' the subset of S which appears in (9.53) for our fixed $j,\boldsymbol{x}$. Set $U=\mathfrak{Z}^{(\mathfrak{U})}(\theta_j,\theta_{j+1},1)+\boldsymbol{z_x}$; then $S\subset\bigcup_{\boldsymbol{y}\in U}X_q(\boldsymbol{y})$ and also $\forall\,\boldsymbol{y}_1\neq\boldsymbol{y}_2\in U:$ $X_q(\boldsymbol{y}_1)\cap X_q(\boldsymbol{y}_2)\cap S'=\varnothing$, since $U\subset\mathfrak{Z}(\theta_j,\theta_{j+1},1)+\boldsymbol{z_x}$. Hence Proposition 7.3 applies, yielding

$$\begin{aligned}(9.55)\quad &\mu_q(S')=\int_U \nu_{\boldsymbol{y}}(S'\cap X_q(\boldsymbol{y}))\,d\boldsymbol{y}\geq\int_U\nu_{\boldsymbol{y}}(S\cap X_q(\boldsymbol{y}))\,d\boldsymbol{y}\\ &-\int_U\nu_{\boldsymbol{y}}\Big(\Big\{M\in X_q(\boldsymbol{y})\,:\,\#\Big(\big(\mathfrak{Z}(\theta_j,\theta_{j+1},1)+\boldsymbol{z_x}\big)\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M\Big)\geq 2\Big\}\Big)\,d\boldsymbol{y}.\end{aligned}$$

Here the first integral in the right-hand side equals

$$\int_{\theta_j}^{\theta_{j+1}}\int_{\mathfrak{U}_\perp} f_0(0,\theta_j,1,\boldsymbol{z_x},\xi\boldsymbol{e}_1+\boldsymbol{w}+\boldsymbol{z_x})\,d\boldsymbol{w}\,d\xi$$

(recall (9.49)), since each $M\in X_q(\boldsymbol{y})$ with $\boldsymbol{y}\in U$ automatically fulfills

$$\#\big((\mathfrak{Z}^{(\mathfrak{U})}(\theta_j,\theta_{j+1},1)+\boldsymbol{z_x})\cap(\mathbb{Z}^d+\boldsymbol{\alpha})M\big)\geq 1;$$

and the second integral is bounded from above by $\varepsilon(\theta_{j+1}-\theta_j)$, by (9.51). Adding this over all j and $\boldsymbol{x}$ we have now proved that the total expression in (9.53) is

$$\begin{aligned}(9.56)\quad &\geq -\varepsilon(\xi_2-\xi_1)\\ &\quad+\int_{\boldsymbol{x}\in D}\sum_{j=1}^{n-1}\int_{\theta_j}^{\theta_{j+1}}\int_{\mathfrak{U}_\perp} f_0(0,\theta_j,1,\boldsymbol{z_x},\xi\boldsymbol{e}_1+\boldsymbol{w}+\boldsymbol{z_x})\,d\boldsymbol{w}\,d\xi\,d\lambda(\boldsymbol{x}).\end{aligned}$$

Now for each $\langle\boldsymbol{x},j,\xi,\boldsymbol{w}\rangle$ which appears in the above integral, and which satisfies $\|\boldsymbol{w}+\boldsymbol{z_x}\|\geq C^{-1}$ if $d=2$, Lemma 8.4 applies, and yields

$$(9.57)\qquad \big|f_0(0,\theta_j,1,\boldsymbol{z_x},\xi\boldsymbol{e}_1+\boldsymbol{w}+\boldsymbol{z_x})-F_0(\xi,\boldsymbol{w},\boldsymbol{z_x})\big|\leq\varepsilon.$$

If $d = 2$ we note that the set $\{\boldsymbol{w} : \|\boldsymbol{w} + \boldsymbol{z}_{\boldsymbol{x}}\| < C^{-1}\}$ has measure $\le 2C^{-1} \le \varepsilon$ (w.r.t. the 1-dimensional Lebesgue measure $d\boldsymbol{w}$), and for these $\boldsymbol{w}$'s the difference in (9.57) is certainly ≤ 1, since $0 \le f_0, F_0 \le 1$ everywhere. Hence (9.56) is

$$\ge -2\varepsilon(\xi_2 - \xi_1) + \int_{\boldsymbol{x}\in D} \int_{\xi_1}^{\xi_2} \int_{\mathfrak{U}_\perp} \Big(-\varepsilon + F_0(\xi, \boldsymbol{w}, \boldsymbol{z}_{\boldsymbol{x}})\Big)\, d\boldsymbol{w}\, d\xi\, d\lambda(\boldsymbol{x}). \tag{9.58}$$

In conclusion, we have proved that this last expression is a lower bound for the lim inf of (9.38). But this is true for any $\varepsilon > 0$; hence the lim inf is in fact

$$\ge \int_{\boldsymbol{x}\in D} \int_{\xi_1}^{\xi_2} \int_{\mathfrak{U}_\perp} F_0(\xi, \boldsymbol{w}, \boldsymbol{z}_{\boldsymbol{x}})\, d\boldsymbol{w}\, d\xi\, d\lambda(\boldsymbol{x}). \tag{9.59}$$

The treatment of the lim sup is similar but a bit easier: With S and U as before we need only notice that by the upper bound in Proposition 7.3 we have

$$\begin{aligned} \mu_q(S) &\le \int_U \nu_{\boldsymbol{y}}(S \cap X_q(\boldsymbol{y}))\, d\boldsymbol{y} \\ &= \int_{\boldsymbol{x}\in D} \sum_{j=1}^{n-1} \int_{\theta_j}^{\theta_{j+1}} \int_{\mathfrak{U}_\perp} f_0(0, \theta_j, 1, \boldsymbol{z}_{\boldsymbol{x}}, \xi\boldsymbol{e}_1 + \boldsymbol{w} + \boldsymbol{z}_{\boldsymbol{x}})\, d\boldsymbol{w}\, d\xi\, d\lambda(\boldsymbol{x}). \end{aligned} \tag{9.60}$$

Now Lemma 8.4 is applied as before, and we obtain that the lim sup of (9.38) is

$$\le \varepsilon(\xi_2 - \xi_1) + \int_{\boldsymbol{x}\in D} \int_{\xi_1}^{\xi_2} \int_{\mathfrak{U}_\perp} \Big(\varepsilon + F_0(\xi, \boldsymbol{w}, \boldsymbol{z}_{\boldsymbol{x}})\Big)\, d\boldsymbol{w}\, d\xi\, d\lambda(\boldsymbol{x}). \tag{9.61}$$

Hence, by letting $\varepsilon \to 0$ and combining with our result for lim inf, we have finally proved our claim that (9.38) tends to (9.50) as $\rho \to 0$. This completes the proof of Theorem 4.4. □

Proof of Theorem 4.8. Let λ and f be given as in the statement of the theorem. By (4.19), the left-hand side of (4.21) equals

$$\lim_{\rho\to 0} \int_{\mathrm{S}_1^{d-1}} g\big(\boldsymbol{v}_0, \rho^{d-1}\tau_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{v}_0), \boldsymbol{v}_0; \rho), \boldsymbol{w}_1(\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{v}_0), \boldsymbol{v}_0; \rho)\big) d\lambda(\boldsymbol{v}_0), \tag{9.62}$$

where $g(\boldsymbol{v}_0, \xi, \boldsymbol{w}_1) = f\big(\boldsymbol{v}_0, \xi, \boldsymbol{v}_0 - 2(\boldsymbol{v}_0 \cdot \boldsymbol{w}_1)\boldsymbol{w}_1\big)$. Using Corollary 4.7 we obtain

$$\begin{aligned} &= \int_{\mathrm{S}_1'^{\,d-1}} \int_{\mathbb{R}_{>0}} \int_{\mathrm{S}_1^{d-1}} f\big(\boldsymbol{v}_0, \xi, \boldsymbol{v}_0 - 2(\boldsymbol{v}_0 \cdot (\boldsymbol{\omega} K(\boldsymbol{v}_0)^{-1}))\boldsymbol{\omega} K(\boldsymbol{v}_0)^{-1}\big) \\ &\qquad \times \Phi_{\boldsymbol{\alpha}}\big(\xi, \boldsymbol{\omega}_\perp, (\boldsymbol{\beta}(\boldsymbol{v}_0)K(\boldsymbol{v}_0))_\perp\big)\, \omega_1\, d\lambda(\boldsymbol{v}_0)\, d\xi\, d\mathrm{vol}_{\mathrm{S}_1^{d-1}}(\boldsymbol{\omega}). \end{aligned} \tag{9.63}$$

Now change the order of integration by moving the integral over $\boldsymbol{\omega} \in \mathrm{S}_1'^{\,d-1}$ to the innermost position, and then apply the variable substitution (4.23) in the innermost integral; note that this gives a diffeomorphism $\boldsymbol{\omega} \mapsto \boldsymbol{v}_1$ from $\mathrm{S}_1'^{\,d-1}$ onto $\mathrm{S}_1^{d-1} \setminus \{\boldsymbol{v}_0\}$ (the inverse map is given by $\boldsymbol{\omega} = \frac{\boldsymbol{e}_1 - \boldsymbol{v}_1 K(\boldsymbol{v}_0)}{\|\boldsymbol{v}_0 - \boldsymbol{v}_1\|}$). Recalling (4.22) we then see that (9.63) equals the right-hand side of (4.21), and we are done. □

9.4. *Proofs for Section* 2. Introduce $E_1 : D \to \mathrm{SO}(d)$ as in Section 9.1 and write λ also for the lift of λ to the variable $\boldsymbol{x} \in \mathbb{R}^{d-1}$, as before. Set

$$\mathfrak{C}_T(c,\sigma) = \big\{\boldsymbol{y} \in \mathbb{R}^d \setminus \{\boldsymbol{0}\} : cT \le \|\boldsymbol{y}\| < T,\ \|\boldsymbol{y}\|^{-1}\boldsymbol{y} \in \mathfrak{D}_T(\sigma)\big\} \times \begin{pmatrix} T^{-1} & \boldsymbol{0} \\ {}^{\mathrm{t}}\boldsymbol{0} & T^{1/(d-1)}1_{d-1} \end{pmatrix}. \tag{9.64}$$

Then

$$\mathcal{N}_{c,T}(\sigma, \boldsymbol{e}_1 E_1(\boldsymbol{x})^{-1}) = \#\big(\mathfrak{C}_T(c,\sigma)\Phi^{-t}(E_1(\boldsymbol{x})^{-1},\boldsymbol{0}) \cap (\mathbb{Z}^d + \boldsymbol{\alpha})M_0 \setminus \{\boldsymbol{0}\}\big), \tag{9.65}$$

with $T = \mathrm{e}^{(d-1)t}$. As before, Theorem 2.1 and Theorem 2.5 now follow from the theorems in Section 6 and the following lemma.

LEMMA 9.5. *Fix* $0 \le c < 1$. *Let* σ_T *be a continuous nonnegative function of* $T > 0$ *such that the limit* $\sigma_\infty = \lim_{T\to\infty} \sigma_T$ *exist. Then the union* $\cup_{T\ge 1}\mathfrak{C}_T(c,\sigma_T)$ *is bounded, and*

$$\lim(\inf \mathfrak{C}_T(c,\sigma_T))^\circ \supset \mathfrak{C}(c,\sigma_\infty)^\circ, \qquad \lim \overline{\sup \mathfrak{C}_T(c,\sigma_T)} \subset \overline{\mathfrak{C}(c,\sigma_\infty)}, \tag{9.66}$$

where $\mathfrak{C}(c,\sigma)$ *is as in* (2.14). *The boundary of* $\mathfrak{C}(c,\sigma)$ *has Lebesgue measure zero.*

Proof. From (2.9) we have $\mathfrak{D}_T(\sigma_T) = (\boldsymbol{e}_1 + \mathcal{B}^d_{r_T}) \cap \mathrm{S}^{d-1}_1$ where

$$T^{\frac{d}{d-1}} r_T \to \Big(\frac{d\sigma_\infty}{(1-c^d)\,\mathrm{vol}(\mathcal{B}^{d-1}_1)}\Big)^{\frac{1}{d-1}} = A(c,\sigma_\infty) \qquad \text{as } T \to \infty. \tag{9.67}$$

In particular, for T sufficiently large, if $\boldsymbol{y}$ is any point in $\mathfrak{C}_T(c,\sigma_T)$, and $\boldsymbol{y}' = Ty_1\boldsymbol{e}_1 + T^{-\frac{1}{d-1}}\boldsymbol{y}_\perp$, then $\|\boldsymbol{y}'\|^{-1}\boldsymbol{y}' \in \mathfrak{D}_T(\sigma_T)$ implies $y'_1 > 0$ and

$$\|\boldsymbol{y}'_\perp\| \le \frac{r_T\sqrt{4-r_T^2}}{2-r_T^2}\, y'_1,$$

thus $y_1 > 0$ and $\|\boldsymbol{y}_\perp\| \le (A(c,\sigma_\infty)+\eta)y_1$, where $\eta > 0$ can be made arbitrarily small. With these observations the proof of Lemma 9.5 is easily completed by mimicking the proof of Lemma 9.2. □

Index of notation

$\mathfrak{D}_T(\sigma,\boldsymbol{v})$	small disc on S_1^{d-1} of radius $\asymp T^{-d/(d-1)}$	1957
$\mathfrak{D}_T(\sigma)$	$=\mathfrak{D}_T(\sigma,\boldsymbol{e}_1)$	
$\mathfrak{D}(\epsilon,\boldsymbol{v})$	small disc on S_1^{d-1} of radius $\sim\epsilon$	1962
$\mathfrak{D}(\epsilon)$	$=\mathfrak{D}(\epsilon,\boldsymbol{e}_1)$	
$\boldsymbol{e}_1$	$=(1,0,\dots,0)$	
$\boldsymbol{e}_2$	$=(0,1,0,\dots,0)$	
$\mathscr{E}(\mathfrak{B},r)$	$=\{(\boldsymbol{x},g)\in\mathbb{R}^{d-1}\times X\,:\,\#(\mathfrak{B}\vert_{\boldsymbol{x}}\cap\mathbb{Z}^d g)\ge r\}$	1980
$\widehat{\mathscr{E}}_t$	the family of sets in (6.9) or (6.17)	1982, 1983
E_1	a map $D\to\mathrm{SO}(d)$	1974, 2015, 2022
$E_{c,\boldsymbol{\alpha}}(r,\sigma)$	limiting probability for $\mathscr{N}_{c,T}(\sigma,\boldsymbol{v})=r$	1958
$\widehat{E}_{c,\boldsymbol{\alpha}}(r,\sigma)$	limiting probability for $\widehat{\mathscr{N}}_{c,T}(\sigma,\boldsymbol{v})=r$	1960
$F_{c,\boldsymbol{\alpha}}(r,\sigma)$	limiting probability for $\mathscr{N}_{c,T}(\rho,\boldsymbol{v})=r$	1962
$F_c(r,\sigma)$	$=F_{c,\boldsymbol{\alpha}}(r,\sigma)$ for any $\boldsymbol{\alpha}\notin\mathbb{Q}^d$, universal limiting probability	1963
$F_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma)$	limiting probability for $\mathscr{N}_{c,T}(\rho,\boldsymbol{v},\boldsymbol{\beta}(\boldsymbol{v}))=r$	1963
$F^{(\mathfrak{U})}_{c,\boldsymbol{\alpha},\boldsymbol{\beta}}(r,\sigma)$	limiting probability for $\mathscr{N}^{(\mathfrak{U})}_{c,T}(\rho,\boldsymbol{x},\boldsymbol{\beta}(\boldsymbol{x}))=r$	2016
$F_{c,\boldsymbol{\alpha}}(r,\mathfrak{Q})$	limiting probability for $\mathscr{N}_{c,T}(\mathfrak{Q},\boldsymbol{v})=r$	1965
$f_r(c_1,c_2,\sigma,\boldsymbol{z},\boldsymbol{y})$	the volume function in (8.1) or (8.27)	2002, 2008
$F_r(\xi,\boldsymbol{w},\boldsymbol{z})$	$=f_r(0,\xi,1,\boldsymbol{z},\xi\boldsymbol{e}_1+\boldsymbol{w}+\boldsymbol{z})$	2002, 2008
$\mathfrak{F}_N$	$=\{\frac{h}{k}\,:\,h,k\in\mathbb{Z},\,0<h<k\le N\}$	2003
$g_r(c_1,c_2,\sigma,\boldsymbol{z},\boldsymbol{y})$	the volume function in (8.44) or (8.47)	2014
$G_r(\xi,\boldsymbol{w},\boldsymbol{z})$	$=g_r(0,\xi,1,\boldsymbol{z},\xi\boldsymbol{e}_1+\boldsymbol{w}+\boldsymbol{z})$	2014
H	$=\{g\in\mathrm{SL}(d,\mathbb{R})\,:\,\boldsymbol{e}_1 g=\boldsymbol{e}_1\}$	1989
I_q	$=[\Gamma(q):\Gamma(1)]$	1987
$\mathscr{K}_\rho$	complement of the set $\mathscr{B}_\rho^d+\mathscr{L}$ in $\mathbb{R}^d$ (the "billiard domain")	1952
$\lim(\inf\mathscr{E}_t)^\circ$	$=\cup_{t\ge t_0}\big(\cap_{s\ge t}\mathscr{E}_s\big)^\circ$	1977
$\lim\overline{\sup\mathscr{E}_t}$	$=\cap_{t\ge t_0}\overline{\cup_{s\ge t}\mathscr{E}_s}$	1977
$\mathscr{L}$	$=\mathbb{Z}^d M_{\mathbf{0}}$, euclidean lattice of covolume 1	1956
$\mathscr{L}_{\boldsymbol{\alpha}}$	$=(\mathbb{Z}^d+\boldsymbol{\alpha})M_{\mathbf{0}}$, affine lattice of covolume 1	1957
$\widehat{\mathscr{L}}_{\boldsymbol{\alpha}}$	set of visible lattice points	1959
$n_-(\boldsymbol{x})$	element in $\mathrm{ASL}(d,\mathbb{R})$	1971
$\mathscr{N}_{c,T}(\sigma,\boldsymbol{v})$	$=\#\{\boldsymbol{y}\in\mathscr{P}_T\,:\,\Vert\boldsymbol{y}\Vert^{-1}\boldsymbol{y}\in\mathfrak{D}_T(\sigma,\boldsymbol{v})\}$	1957
$\mathscr{N}_{c,T}(\sigma,K)$	$=\#(\mathscr{P}_T\cap\mathfrak{D}_T(\sigma)K)$	1959
$\mathscr{N}_{c,T}(\rho,\boldsymbol{v})$	number of spheres in direction $\boldsymbol{v}$	1962
$\mathscr{N}_{c,T}(\rho,\boldsymbol{v},\boldsymbol{w})$	as above, but includes shift by $\rho\boldsymbol{w}$	1963
$\mathscr{N}^{(\mathfrak{U})}_{c,T}(\rho,\boldsymbol{x},\boldsymbol{w})$	generalized version of $\mathscr{N}_{c,T}(\rho,\boldsymbol{v},\boldsymbol{w})$	2015
$\widehat{\mathscr{N}}_{c,T}(\sigma,\boldsymbol{v})$	analogue of $\mathscr{N}_{c,T}(\sigma,\boldsymbol{v})$ for visible lattice points	1960
$\mathscr{N}_{c,T}(\mathfrak{Q},\boldsymbol{v})$	number of $\mathfrak{Q}$'s in direction $\boldsymbol{v}$	1965
$p_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{v}_0,\xi,\boldsymbol{v}_1)$	joint limiting distribution for free path lengths and velocities	1971
$P_{\boldsymbol{\alpha}}(s)$	limiting gap distribution for directions of lattice points	1954
$\widehat{P}_{\boldsymbol{\alpha}}(s)$	limiting gap distribution for directions of visible lattice points	1961
$\mathscr{P}_T$	$=\mathscr{L}_{\boldsymbol{\alpha}}\cap\mathscr{B}_T^d(c)\setminus\{\mathbf{0}\}$	1957
$\widehat{\mathscr{P}}_T$	$=\widehat{\mathscr{L}}_{\boldsymbol{\alpha}}\cap\mathscr{B}_T^d(c)$	1959
$\mathrm{Proj}_{\{\boldsymbol{v}\}^\perp}$	orthogonal projection from $\mathbb{R}^d$ onto the orthogonal complement of $\boldsymbol{v}$	1964

$\boldsymbol{q}_1(\boldsymbol{q},\boldsymbol{v};\rho)$	location of first collision in $\mathbb{R}^d$	1968
$\boldsymbol{q}_{\rho,\boldsymbol{\beta}}(\boldsymbol{v})$	initial position $\boldsymbol{q}+\rho\boldsymbol{\beta}(\boldsymbol{v})$	1968
S_1^{d-1}	unit sphere in $\mathbb{R}^d$	
$\mathrm{S}_1'^{\,d-1}$	the hemisphere $\{\boldsymbol{v}=(v_1,\dots,v_d)\in\mathrm{S}_1^{d-1}\,:\,v_1>0\}$	1968
$s(\delta)$	see (8.8)	2003
$t_{\boldsymbol{k}}$	$=\gcd(qk_1,qk_2,\dots,qk_d)$	1993
$\boldsymbol{w}_1(\boldsymbol{q},\boldsymbol{v};\rho)$	location of first collision on S_1^{d-1}	1968
X	$=\mathrm{ASL}(d,\mathbb{Z})\backslash\mathrm{ASL}(d,\mathbb{R})$, space of affine lattices	1956
X_1	$=\mathrm{SL}(d,\mathbb{Z})\backslash\mathrm{SL}(d,\mathbb{R})$, space of lattices	1956
X_q	$=\Gamma(q)\backslash\mathrm{SL}(d,\mathbb{R})$	1956
$X(\boldsymbol{y})$	submanifold of X	1968
$X_q(\boldsymbol{y})$	submanifold of X_q	1968
$X_q(\boldsymbol{k},\boldsymbol{y})$	connected component of $X_q(\boldsymbol{y})$	1988
$X_q^{(t_0)}(\boldsymbol{y})$	subset of $X_q(\boldsymbol{y})$	1997
$\mathbb{Z}^d_*$	$=\mathbb{Z}^d\setminus\{\boldsymbol{0}\}$	1958
$\widehat{\mathbb{Z}}^d$	$=\widehat{\mathbb{Z}}^d_{\boldsymbol{0}}$, set of primitive lattice points	1959
$\widehat{\mathbb{Z}}^d_{\boldsymbol{\alpha}}$	$=\{\boldsymbol{x}\in(\mathbb{Z}^d+\boldsymbol{\alpha})\setminus\{\boldsymbol{0}\}\,:\,\gcd(q\boldsymbol{x})\le q\}$	1959
$\mathfrak{Z}(c,\sigma)$	the cylinder in (3.9)	1963
$\mathfrak{Z}_{\boldsymbol{v}}(c,\sigma)$	the cylinder in (3.14)	1964
$\mathfrak{Z}(c_1,c_2,\sigma)$	the cylinder in (4.15)	1969
$\widetilde{\mathfrak{Z}}(c_1,c_2,\sigma)$	modification of $\mathfrak{Z}(c_1,c_2,\sigma)$	2004
$\mathfrak{Z}^{(\mathfrak{U})}(c_1,c_2,\sigma)$	generalized cylinder in (9.5)	2016
$\mathfrak{Z}^{(\mathfrak{U})}(c_1,c_2,\sigma,\boldsymbol{\beta})$	generalized cylinder in (9.5)	2016
$\mathfrak{Z}_T^{(\mathfrak{U})}(c_1,c_2,\sigma,\boldsymbol{\beta})$	the set in (9.9)	2017
$\mathfrak{Z}(c,\mathfrak{D})$	the set in (9.28)	2020
$\mathfrak{Z}_T(c,\mathfrak{D})$	the set in (9.25)	2020
$\Gamma(q)$	principal congruence subgroup	1956
κ_q	the constant in (2.22)	1960
μ	Haar measure on $\mathrm{ASL}(d,\mathbb{R})$, probability measure on X	1956
μ_H	Haar measure on H	1989
μ_q	Haar measure on $\mathrm{SL}(d,\mathbb{R})$, probability measure on X_q	1956
ν	Liouville measure	1952
$\nu_{\boldsymbol{y}}$	volume measure on $X_q(\boldsymbol{y})$ or $X(\boldsymbol{y})$	1989, 1998
$\tau_1(\boldsymbol{q},\boldsymbol{v};\rho)$	$=\inf\{t>0:\boldsymbol{q}+t\boldsymbol{v}\notin\mathcal{K}_\rho\}$, free path length	1952
φ_t	the Lorentz flow	1952
$\Phi(\xi)$, $\Phi_{\mathscr{L},\boldsymbol{q}}(\xi)$	limiting distribution for the free path length	1952
$\Phi_{\boldsymbol{\alpha}}(\xi)$	alternative notation for $\Phi_{\mathscr{L},\boldsymbol{q}}(\xi)$	1966
$\Phi_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\xi)$	limiting distribution for the free path length	1967
$\Phi_{\boldsymbol{\alpha}}(\xi,\boldsymbol{w},\boldsymbol{z})$	joint limiting distribution for free path length and impact location	1969
Φ^t	element in $\mathrm{ASL}(d,\mathbb{R})$	1972
$\chi_{\mathscr{A}}$	characteristic function of a set $\mathscr{A}$	
Ω	domain of $f_r(c_1,c_2,\sigma,\boldsymbol{z},\boldsymbol{y})$	2002
Ω_C	Ω truncated	2003, 2014

References

[1] P. BÁLINT and I. P. TÓTH, Exponential decay of correlations in multi-dimensional dispersing billiards, *Ann. Henri Poincaré* **9** (2008), 1309–1369. MR 2010a:37069 Zbl 1162.37018

[2] P. M. BLEHER, Statistical properties of two-dimensional periodic Lorentz gas with infinite horizon, *J. Statist. Phys.* **66** (1992), 315–373. MR 93b:82052 Zbl 0925.82147

[3] F. P. BOCA, C. COBELI, and A. ZAHARESCU, Distribution of lattice points visible from the origin, *Comm. Math. Phys.* **213** (2000), 433–470. MR 2001j:11094 Zbl 0989.11049

[4] F. P. BOCA, R. N. GOLOGAN, and A. ZAHARESCU, The statistics of the trajectory of a certain billiard in a flat two-torus, *Comm. Math. Phys.* **240** (2003), 53–73. MR 2004f:37007 Zbl 1078.37006

[5] F. P. BOCA and A. ZAHARESCU, The correlations of Farey fractions, *J. London Math. Soc.* **72** (2005), 25–39. MR 2006h:11086 Zbl 1089.11037

[6] ———, On the correlations of directions in the Euclidean plane, *Trans. Amer. Math. Soc.* **358** (2006), 1797–1825. MR 2007e:11077 Zbl 1154.11022

[7] ———, The distribution of the free path lengths in the periodic two-dimensional Lorentz gas in the small-scatterer limit, *Comm. Math. Phys.* **269** (2007), 425–471. MR 2008f:37013 Zbl 1143.37002

[8] C. BOLDRIGHINI, L. A. BUNIMOVICH, and Y. G. SINAI, On the Boltzmann equation for the Lorentz gas, *J. Statist. Phys.* **32** (1983), 477–501. MR 85h:82025 Zbl 0583.76092

[9] J. BOURGAIN, F. GOLSE, and B. WENNBERG, On the distribution of free path lengths for the periodic Lorentz gas, *Comm. Math. Phys.* **190** (1998), 491–508. MR 98k:82146 Zbl 0910.60082

[10] L. A. BUNIMOVICH and Y. G. SINAI, Statistical properties of Lorentz gas with periodic configuration of scatterers, *Comm. Math. Phys.* **78** (1980/81), 479–497. MR 82m:82007 Zbl 0459.60099

[11] E. CAGLIOTI and F. GOLSE, On the distribution of free path lengths for the periodic Lorentz gas. III, *Comm. Math. Phys.* **236** (2003), 199–221. MR 2004h:82060 Zbl 1041.82016

[12] N. I. CHERNOV, Statistical properties of the periodic Lorentz gas. Multidimensional case, *J. Statist. Phys.* **74** (1994), 11–53. MR 95d:58083 Zbl 0946.37500

[13] D. DOLGOPYAT, D. SZÁSZ, and T. VARJÚ, Recurrence properties of planar Lorentz process, *Duke Math. J.* **142** (2008), 241–281. MR 2009h:37074 Zbl 1136.37022

[14] N. D. ELKIES and C. T. MCMULLEN, Gaps in $\sqrt{n}$ mod 1 and ergodic theory, *Duke Math. J.* **123** (2004), 95–139. MR 2005f:11143 Zbl 1063.11020

[15] A. ESKIN and C. MCMULLEN, Mixing, counting, and equidistribution in Lie groups, *Duke Math. J.* **71** (1993), 181–209. MR 95b:22025 Zbl 0798.11025

[16] G. GALLAVOTTI, Divergences and approach to equilibrium in the Lorentz and the wind-tree-models, *Phys. Rev.* **185** (1969), 308–322.

[17] F. GOLSE, The periodic Lorentz gas in the Boltzmann-Grad limit, in *International Congress of Mathematicians. Vol.* III, Eur. Math. Soc., Zürich, 2006, pp. 183–201. MR 2007k:82112 Zbl 1109.82020

[18] ———, On the periodic Lorentz gas and the Lorentz kinetic equation, *Ann. Fac. Sci. Toulouse Math.* **17** (2008), 735–749. MR 2010e:82080 Zbl 1166.82304

[19] F. GOLSE and B. WENNBERG, On the distribution of free path lengths for the periodic Lorentz gas. II, *M2AN Math. Model. Numer. Anal.* **34** (2000), 1151–1163. MR 2002a:82092 Zbl 1006.82025

[20] H. LORENTZ, Le mouvement des électrons dans les métaux, *Arch. Néerl* **10** (1905), 336–371.

[21] J. MARKLOF, The n-point correlations between values of a linear form, *Ergodic Theory Dynam. Systems* **20** (2000), 1127–1172. MR 2001m:11112 Zbl 0967.60021

[22] J. MARKLOF, Distribution modulo one and Ratner's theorem, in *Equidistribution in Number Theory, an Introduction*, *NATO Sci. Ser. II Math. Phys. Chem.* **237**, Springer-Verlag, New York, 2007, pp. 217–244. MR 2007k:11118 Zbl 1181.11048

[23] J. MARKLOF and A. STRÖMBERGSSON, The Boltzmann-Grad limit of the periodic Lorentz gas, *Ann. of Math.*, to appear. arXiv 0801.0612

[24] ———, The distribution of visible lattice points, in preparation.

[25] ———, The periodic Lorentz gas in the Boltzmann-Grad limit: Asymptotic estimates. arXiv 1008.3811

[26] ———, Kinetic transport in the two-dimensional periodic Lorentz gas, *Nonlinearity* **21** (2008), 1413–1422. MR 2009j:82029 Zbl 1153.82020

[27] I. MELBOURNE and M. NICOL, Almost sure invariance principle for nonuniformly hyperbolic systems, *Comm. Math. Phys.* **260** (2005), 131–146. MR 2006h:37047 Zbl 1084.37024

[28] ———, A vector-valued almost sure invariance principle for hyperbolic dynamical systems, *Ann. Probab.* **37** (2009), 478–505. MR 2010i:37016 Zbl 1176.37006

[29] G. PÓLYA, Zahlentheoretisches und Wahrscheinlichkeitstheoretisches über die Sichtweite im Walde, *Arch. Math. Phys.* **27** (1918), 135–142.

[30] M. RATNER, On Raghunathan's measure conjecture, *Ann. of Math.* **134** (1991), 545–607. MR 93a:22009 Zbl 0763.28012

[31] N. A. SHAH, Limit distributions of expanding translates of certain orbits on homogeneous spaces, *Proc. Indian Acad. Sci. Math. Sci.* **106** (1996), 105–125. MR 98b:22024 Zbl 0864.22004

[32] G. SHIMURA, *Introduction to the Arithmetic Theory of Automorphic Functions*, Publications of the Mathematical Society of Japan, No. 11, Iwanami Shoten, Publishers, Tokyo, 1971. MR 47 #3318 Zbl 0221.10029

[33] A. N. SHIRYAEV, *Probability*, second ed., *Grad. Texts in Math.* **95**, Springer-Verlag, New York, 1996, Translated from the first (1980) Russian edition by R. P. Boas. MR 97c:60003 Zbl 0870.00038

[34] C. L. SIEGEL, *Lectures on the Geometry of Numbers*, Springer-Verlag, New York, 1989. MR 91d:11070 Zbl 0691.10021

[35] H. SPOHN, The Lorentz process converges to a random flight process, *Comm. Math. Phys.* **60** (1978), 277–290. MR 58 #14861 Zbl 0381.60099

[36] A. STRÖMBERGSSON and A. VENKATESH, Small solutions to linear congruences and Hecke equidistribution, *Acta Arith.* **118** (2005), 41–78. MR 2006g:11204 Zbl 1082.11061

[37] D. SZÁSZ and T. VARJÚ, Limit laws and recurrence for the planar Lorentz process with infinite horizon, *J. Stat. Phys.* **129** (2007), 59–80. MR 2008k:37087 Zbl 1128.82011

(Received June 29, 2007)
(Revised January 14, 2008)

E-mail address: j.marklof@bristol.ac.uk
SCHOOL OF MATHEMATICS, UNIVERSITY OF BRISTOL, BRISTOL, BS8 1TW,
UNITED KINGDOM
http://www.maths.bris.ac.uk/∼majm/

E-mail address: astrombe@math.uu.se
DEPARTMENT OF MATHEMATICS, ROYAL INSTITUTE OF TECHNOLOGY,
SE-100 44 STOCKHOLM, SWEDEN
Current address: DEPARTMENT OF MATHEMATICS, BOX 480, UPPSALA UNIVERSITY,
SE-75106 UPPSALA, SWEDEN

http://www.math.uu.se/∼astrombe/