

# SUPERDIFFUSION IN THE PERIODIC LORENTZ GAS

JENS MARKLOF AND BÁLINT TÓTH

ABSTRACT. We prove a superdiffusive central limit theorem for the displacement of a test particle in the periodic Lorentz gas in the limit of large times  $t$  and low scatterer densities (Boltzmann-Grad limit). The normalization factor is  $\sqrt{t \log t}$ , where  $t$  is measured in units of the mean collision time. This result holds in any dimension and for a general class of finite-range scattering potentials.

## 1. INTRODUCTION

The periodic Lorentz gas is one of the iconic models of “chaotic” diffusion in deterministic systems. It describes the dynamics of a test-particle in an infinite periodic array of spherically symmetric scatterers. The main results characterizing the diffusive nature of the periodic Lorentz gas have to date been mainly restricted to the two-dimensional setting and hard-sphere scatterers. The first seminal result on this subject was the proof of a central limit theorem for the displacement of the test particle at large times  $t$  for the finite-horizon Lorentz gas by Bunimovich and Sinai [8]. For more general invariance principles see Melbourne and Nicol [21] and references therein. In the case of the infinite-horizon Lorentz gas, Bleher [5] pointed out that the mean-square displacement grows like  $t \log t$  when  $t \rightarrow \infty$ , as opposed to a linear growth in the finite-horizon case. The superdiffusive central limit theorem suggested in [5] was first proved by Szász and Varjú [27] for the discrete-time billiard map. Dolgopyat and Chernov [14] provided an alternative proof, and established the central limit theorem and invariance principle for the billiard flow. Analogous results hold for the stadium billiard (Bálint and Gouëzel [2]) and billiards with cusps (Bálint, Chernov and Dolgopyat [1]). The difficulty in extending the above findings to dimensions greater than two lies in the possibly exponential growth of the complexity of singularities (Bálint and Tóth [3, 4], Chernov [12]) and, in the case of infinite horizon, the subtle geometry of channels (Dettmann [13], Nándori, Szász and Varjú [22]).

In the present paper we prove an unconditional superdiffusive central limit theorem for the periodic Lorentz gas in any dimension  $d \geq 2$ , valid in the limit of low scatterer density (Boltzmann-Grad limit) and for a general class of finite-range scattering potentials. The precise setting is as follows. Let  $\mathcal{L} \subset \mathbb{R}^d$  be a fixed Euclidean lattice of covolume one (such as the cubic lattice  $\mathcal{L} = \mathbb{Z}^d$ ), and define the scaled lattice  $\mathcal{L}_r := r^{(d-1)/d} \mathcal{L}$ . At each point in  $\mathcal{L}_r$  we center a sphere of radius  $r$ . We consider a test particle that moves along straight lines with unit speed until it hits a sphere, where it is scattered elastically. The above scaling of scattering radius vs. lattice spacing ensures that the mean free path length (i.e., the average distance between consecutive collisions) has the limit  $\bar{\xi} = 1/v_{d-1}$  as  $r \rightarrow 0$ , where  $v_{d-1} = \pi^{\frac{d-1}{2}} / \Gamma(\frac{d+1}{2})$  denotes the volume of the unit ball in  $\mathbb{R}^{d-1}$ .

In the case of the classic Lorentz gas the scattering mechanism is given by specular reflection, but as in [19] we will here also allow more general spherically symmetric scattering maps. The precise conditions will be stated in Section 2.

---

*Date:* 24 March 2014.

2010 *Mathematics Subject Classification.* 37D50, 60F05, 82C40.

The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 291147. J.M. is furthermore supported by a Royal Society Wolfson Research Merit Award. The research of B.T. is partially supported by the Hungarian National Science Foundation (OTKA) through grant K100473.

The position of our test particle at time  $t$  is denoted by

$$(1.1) \quad \mathbf{x}_t = \mathbf{x}_t(\mathbf{x}_0, \mathbf{v}_0) \in \mathcal{K}_r := \mathbb{R}^d \setminus (\mathcal{L}_r + r\mathcal{B}_1^d),$$

where  $\mathbf{x}_0$  and  $\mathbf{v}_0$  are position and velocity at time  $t = 0$ , and  $\mathcal{B}_1^d$  is the open unit ball in  $\mathbb{R}^d$  centered at the origin. We use the convention that for any boundary point  $\mathbf{x}_0 \in \partial\mathcal{K}_r$  we choose the *outgoing* velocity  $\mathbf{v}_0$ , i.e. the velocity *after* the scattering. The corresponding phase space is denoted by  $\mathbb{T}^1(\mathcal{K}_r)$ . For notational reasons it is convenient to extend the dynamics to  $\mathbb{T}^1(\mathbb{R}^d) := \mathbb{R}^d \times \mathbb{S}_1^{d-1}$  by setting  $\mathbf{x}_t = \mathbf{x}_0$  for all initial conditions  $\mathbf{x}_0 \notin \mathcal{K}_r$ .

We consider the time evolution of a test particle with random initial data  $(\mathbf{x}_0, \mathbf{v}_0) \in \mathbb{T}^1(\mathbb{R}^d)$ , distributed according to a given Borel probability measure  $\Lambda$  on  $\mathbb{T}^1(\mathbb{R}^d)$ . The following superdiffusive central limit theorem, valid for small scattering radii and large times, is the main result of this paper.

**Theorem 1.1.** *Let  $d \geq 2$  and fix a Euclidean lattice  $\mathcal{L} \subset \mathbb{R}^d$  of covolume one. Assume  $(\mathbf{x}_0, \mathbf{v}_0)$  is distributed according to an absolutely continuous Borel probability measure  $\Lambda$  on  $\mathbb{T}^1(\mathbb{R}^d)$ . Then, for any bounded continuous  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ ,*

$$(1.2) \quad \lim_{t \rightarrow \infty} \lim_{r \rightarrow 0} \mathbf{E} f \left( \frac{\mathbf{x}_t - \mathbf{x}_0}{\Sigma_d \sqrt{t \log t}} \right) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(\mathbf{x}) e^{-\frac{1}{2}\|\mathbf{x}\|^2} d\mathbf{x},$$

with

$$(1.3) \quad \Sigma_d^2 := \frac{2^{1-d} v_{d-1}}{d^2(d+1)\zeta(d)}.$$

Here  $\zeta(d) := \sum_{n=1}^{\infty} n^{-d}$  denotes the Riemann zeta function. Theorem 1.1 will follow from its discrete-time analogue, Theorem 1.2 below. Let us denote by  $\mathbf{q}_n = \mathbf{q}_n(\mathbf{q}_0, \mathbf{v}_0) \in \partial\mathcal{K}_r$  ( $n = 1, 2, 3, \dots$ ) the location where the test particle with initial condition  $(\mathbf{q}_0, \mathbf{v}_0)$  leaves the  $n$ th scatterer. It is natural in this setting to assume  $\mathbf{q}_0 \in \partial\mathcal{K}_r$ . By the translational invariance of the lattice, we may in fact assume without loss of generality  $\mathbf{q}_0 \in r\mathbb{S}_1^{d-1}$ . For given exit velocity  $\mathbf{v}_0$ , we write

$$(1.4) \quad \mathbf{q}_0 = r(\mathbf{s}_0 + \mathbf{v}_0 \sqrt{1 - \|\mathbf{s}_0\|^2})$$

and stipulate in the following that the random variable  $\mathbf{s}_0$  is uniformly distributed in the unit disc orthogonal to  $\mathbf{v}_0$ .

**Theorem 1.2.** *Let  $d \geq 2$  and  $\mathcal{L}$  as above. Assume  $\mathbf{v}_0$  is distributed according to an absolutely continuous Borel probability measure  $\lambda$  on  $\mathbb{S}_1^{d-1}$ . Then, for any bounded continuous  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ ,*

$$(1.5) \quad \lim_{n \rightarrow \infty} \lim_{r \rightarrow 0} \mathbf{E} f \left( \frac{\mathbf{q}_n - \mathbf{q}_0}{\sigma_d \sqrt{n \log n}} \right) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(\mathbf{x}) e^{-\frac{1}{2}\|\mathbf{x}\|^2} d\mathbf{x},$$

with

$$(1.6) \quad \sigma_d^2 := \frac{2^{1-d}}{d^2(d+1)\zeta(d)} = \bar{\xi} \Sigma_d^2.$$

The starting point of our analysis is the paper [19], which proves that, for every fixed  $t > 0$ , the limit  $r \rightarrow 0$  in (1.2) (resp. (1.5)) exists and is given by a continuous-time (resp. discrete-time) Markov process. The main objective of the present study is therefore to prove a superdiffusive central limit theorem for each of these Markov processes. This is stated as Theorem 3.2 in Section 3 after a brief survey of the relevant results from [19]. The subsequent sections of the paper are devoted to the proof of Theorem 3.2.

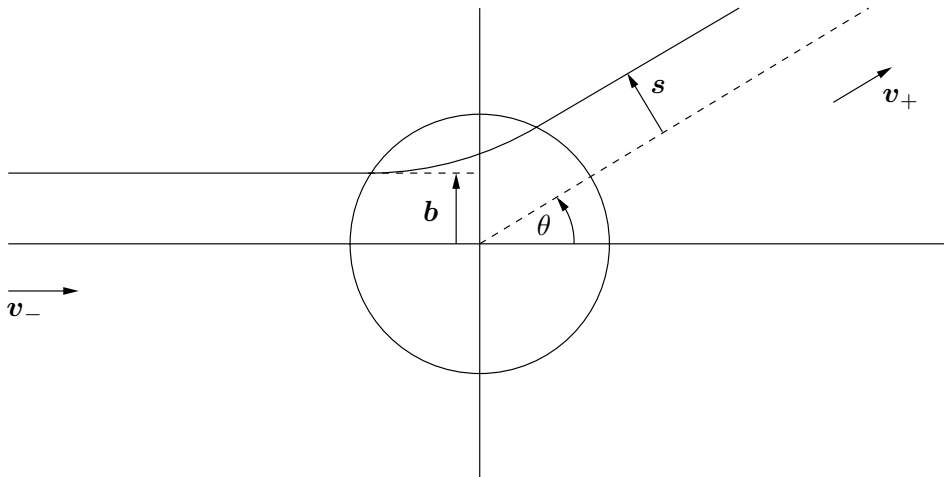


FIGURE 1. The scattering map.

## 2. THE SCATTERING MAP

We now specify the conditions on the scattering map that are assumed in Theorems 1.1 and 1.2. These are the same as in [19], with the additional simplifying assumption that the scattering map preserves angular momentum, cf. [19, Remark 2.3]. We describe the scattering map in units of  $r$ , i.e., the scatterer is represented as the open unit ball  $\mathcal{B}_1^d$ . Set

$$(2.1) \quad \mathcal{S} := \{(\mathbf{v}, \mathbf{b}) \in S_1^{d-1} \times \mathcal{B}_1^d \mid \mathbf{v} \cdot \mathbf{b} = 0\},$$

and consider the scattering map

$$(2.2) \quad \Theta : \mathcal{S} \rightarrow \mathcal{S}, \quad (\mathbf{v}_-, \mathbf{b}) \mapsto (\mathbf{v}_+, \mathbf{s}).$$

The incoming data is denoted by  $(\mathbf{v}_-, \mathbf{b}) \in \mathcal{S}$ , where  $\mathbf{v}_-$  is the velocity of the particle before the collision and  $\mathbf{b}$  the impact parameter, i.e., the point of impact on  $S_1^{d-1}$  projected onto the plane  $\{\mathbf{b} \in \mathbb{R}^d \mid \mathbf{v}_- \cdot \mathbf{b} = 0\}$ . The outgoing data is analogously defined as  $(\mathbf{v}_+, \mathbf{s}) \in \mathcal{S}$ , where  $\mathbf{v}_+$  is the velocity of the particle after the collision and  $\mathbf{s}$  the exit parameter, cf. Figure 1. Since we assume the scattering map is spherically symmetric, it is sufficient to define  $\Theta$  for  $(\mathbf{v}_-, \mathbf{b}) = (\mathbf{e}_1, w\mathbf{e}_2)$  for  $w \in [0, 1)$ , where  $\mathbf{e}_j$  denotes the unit vector in the  $j$ th coordinate direction. Any spherically symmetric scattering map (2.2) which preserves angular momentum is thus uniquely determined by

$$(2.3) \quad \Theta(\mathbf{e}_1, w\mathbf{e}_2) = (\mathbf{e}_1 \cos \theta(w) + \mathbf{e}_2 \sin \theta(w), -\mathbf{e}_1 w \sin \theta(w) + \mathbf{e}_2 w \cos \theta(w))$$

where  $\theta(w)$  is called the *scattering angle*.

To satisfy the conditions of [19], we assume in the statements of Theorems 1.1 and 1.2 that one of the following hypotheses is true (cf. Fig. 2):

- (A)  $\theta \in C^1([0, 1))$  is strictly decreasing with  $\theta(0) = \pi$  and  $\theta(w) > 0$ ;
- (B)  $\theta \in C^1([0, 1))$  is strictly increasing with  $\theta(0) = -\pi$  and  $\theta(w) < 0$ .

This assumption holds for a large class of scattering potentials, including muffin-tin Coulomb potentials, cf. [19]. In the case of hard-sphere scatterers we have  $\theta(w) = \pi - 2 \arcsin(w)$  and hence Hypothesis (A) holds.

Note that for more general impact parameters of the form

$$(2.4) \quad \mathbf{b} = \begin{pmatrix} 0 \\ \mathbf{w} \end{pmatrix}, \quad \mathbf{w} \in \mathcal{B}_1^{d-1} \setminus \{\mathbf{0}\},$$

we have (by spherical symmetry)

$$(2.5) \quad \Theta \left( \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} 0 \\ \mathbf{w} \end{pmatrix} \right) = \left( S(\mathbf{w}) \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}, S(\mathbf{w}) \begin{pmatrix} 0 \\ \mathbf{w} \end{pmatrix} \right)$$

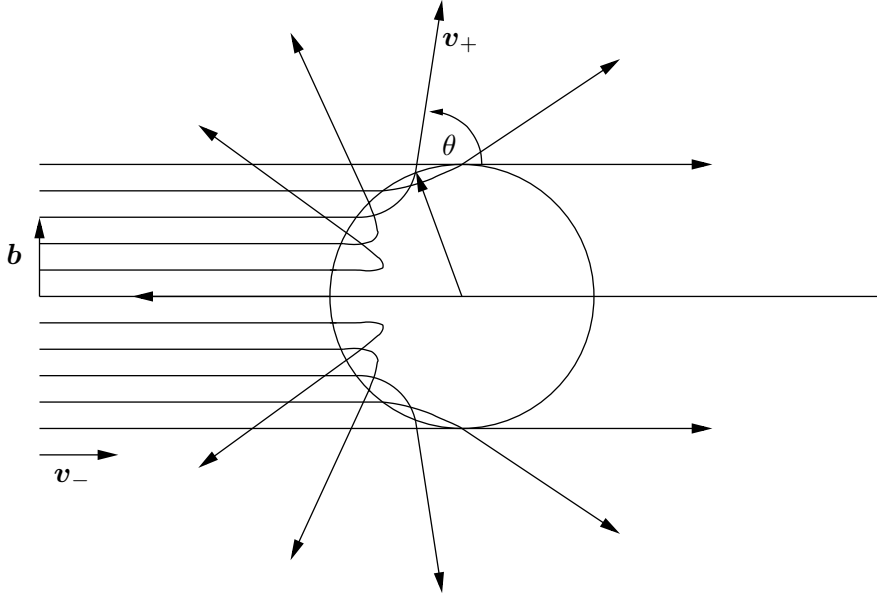


FIGURE 2. Illustration of a scattering map satisfying Hypothesis (A).

with the matrix

$$(2.6) \quad S(\mathbf{w}) = E(\theta(w)\widehat{\mathbf{w}}),$$

where

$$(2.7) \quad w := \|\mathbf{w}\| > 0, \quad \widehat{\mathbf{w}} := w^{-1}\mathbf{w} \in \mathbb{S}_1^{d-1}, \quad E(\mathbf{x}) := \exp \begin{pmatrix} 0 & -\mathbf{x} \\ \mathbf{x} & 0_{d-1} \end{pmatrix} \in \text{SO}(d).$$

More explicitly,

$$(2.8) \quad S(\mathbf{w}) = \begin{pmatrix} \cos \theta(w) & -\widehat{\mathbf{w}} \sin \theta(w) \\ \widehat{\mathbf{w}} \sin \theta(w) & 1_{d-1} - \widehat{\mathbf{w}} \otimes \widehat{\mathbf{w}} (1 - \cos \theta(w)) \end{pmatrix}.$$

We extend the definition of  $S(\mathbf{w})$  to  $\mathbf{w} = \mathbf{0}$  by setting  $S(\mathbf{0}) := -I_d \in \text{SO}(d)$  for  $d$  even and  $S(\mathbf{0}) := \begin{pmatrix} -I_{d-1} & \\ & 1 \end{pmatrix} \in \text{SO}(d)$  for  $d$  odd. This choice ensures that  $S(\mathbf{0})\mathbf{e}_1 = -\mathbf{e}_1$ .

For the case of general initial data  $(\mathbf{v}_-, \mathbf{b}) \in \mathcal{S}$ , assume  $R(\mathbf{v}_-) \in \text{SO}(d)$  and  $\mathbf{w} \in \mathcal{B}_1^{d-1}$  are chosen so that

$$(2.9) \quad \mathbf{v}_- = R(\mathbf{v}_-) \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}, \quad \mathbf{b} = R(\mathbf{v}_-) \begin{pmatrix} 0 \\ \mathbf{w} \end{pmatrix}.$$

Then

$$(2.10) \quad \Theta(\mathbf{v}_-, \mathbf{b}) = \left( R(\mathbf{v}_-)S(\mathbf{w}) \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}, R(\mathbf{v}_-)S(\mathbf{w}) \begin{pmatrix} 0 \\ \mathbf{w} \end{pmatrix} \right).$$

We use an inductive argument to work out the velocity  $\mathbf{v}_n$  after the  $n$ th collision, as well as the impact and exit parameters  $\mathbf{b}_n$  and  $\mathbf{s}_n$  of the  $n$ th collision.

**Lemma 2.1.** *Fix  $\mathbf{v}_0$  and  $R_0 \in \text{SO}(d)$  so that  $\mathbf{v}_0 = R_0\mathbf{e}_1$ , and denote by  $(\mathbf{v}_n)_{n \in \mathbb{N}}$ ,  $(\mathbf{b}_n)_{n \in \mathbb{N}}$ ,  $(\mathbf{s}_n)_{n \in \mathbb{N}}$  the sequence of velocities, impact and exit parameters of a given particle trajectory. Then there is a unique sequence  $(\mathbf{w}_n)_{n \in \mathbb{N}}$  in  $\mathcal{B}_1^{d-1}$  such that for all  $n \in \mathbb{N}$*

$$(2.11) \quad \mathbf{v}_n = R_n \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}, \quad \mathbf{b}_n = R_{n-1} \begin{pmatrix} 0 \\ \mathbf{w}_n \end{pmatrix}, \quad \mathbf{s}_n = R_n \begin{pmatrix} 0 \\ \mathbf{w}_n \end{pmatrix},$$

where

$$(2.12) \quad R_n := R_0 S(\mathbf{w}_1) \cdots S(\mathbf{w}_n).$$

*Proof.* We proceed by induction. We have  $\mathbf{v}_0 \cdot \mathbf{b}_1 = 0$  and thus  $\mathbf{e}_1 \cdot R_0^{-1} \mathbf{b}_1 = 0$ . We define  $\mathbf{w}_1 \in \mathcal{B}_1^{d-1}$  by

$$(2.13) \quad \begin{pmatrix} 0 \\ \mathbf{w}_1 \end{pmatrix} = R_0^{-1} \mathbf{b}_1.$$

Then the assumption (2.9) is satisfied and (2.10) yields

$$(2.14) \quad (\mathbf{v}_1, \mathbf{s}_1) = \Theta(\mathbf{v}_0, \mathbf{b}_1) = \left( R_0 S(\mathbf{w}_1) \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}, R_0 S(\mathbf{w}_1) \begin{pmatrix} 0 \\ \mathbf{w}_1 \end{pmatrix} \right).$$

which proves the case  $n = 1$ . Let us therefore assume the statement is true for  $n = k - 1$ . By the induction hypothesis, we have  $\mathbf{v}_{k-1} = R_{k-1} \mathbf{e}_1$ . Note that  $\mathbf{v}_{k-1} \cdot \mathbf{b}_k = 0$  implies  $\mathbf{e}_1 \cdot R_{k-1}^{-1} \mathbf{b}_k = 0$ , and define  $\mathbf{w}_k \in \mathcal{B}_1^{d-1}$  by

$$(2.15) \quad \begin{pmatrix} 0 \\ \mathbf{w}_k \end{pmatrix} = R_{k-1}^{-1} \mathbf{b}_k.$$

Therefore (2.9) holds with  $\mathbf{v}_- = \mathbf{v}_{k-1}$ ,  $\mathbf{b} = \mathbf{b}_k$ , and we can apply (2.10):

$$(2.16) \quad \begin{aligned} (\mathbf{v}_k, \mathbf{s}_k) &= \Theta(\mathbf{v}_{k-1}, \mathbf{b}_k) \\ &= \left( R_{k-1} S(\mathbf{w}_k) \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}, R_{k-1} S(\mathbf{w}_k) \begin{pmatrix} 0 \\ \mathbf{w}_k \end{pmatrix} \right) \\ &= \left( R_k \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}, R_k \begin{pmatrix} 0 \\ \mathbf{w}_k \end{pmatrix} \right), \end{aligned}$$

where  $R_k := R_{k-1} S(\mathbf{w}_k) = R_0 S(\mathbf{w}_1) \cdots S(\mathbf{w}_k)$ . This completes the proof.  $\square$

### 3. THE BOLTZMANN-GRAD LIMIT

We now recall the results of [18, 19] that are relevant to our investigation. Define the Markov chain

$$(3.1) \quad n \mapsto (\xi_n, \boldsymbol{\eta}_n)$$

on the state space  $\mathbb{R}_{>0} \times \mathcal{B}_1^{d-1}$  with transition probability

$$(3.2) \quad \mathbf{P}((\xi_n, \boldsymbol{\eta}_n) \in \mathcal{A} \mid \xi_{n-1}, \boldsymbol{\eta}_{n-1}) = \int_{\mathcal{A}} \Psi_0(\boldsymbol{\eta}_{n-1}, x, \mathbf{z}) dx dz.$$

We will discuss the transition kernel  $\Psi_0(\mathbf{w}, x, \mathbf{z})$  in detail in Section 5. At this point, it suffices to note that it is independent of  $\xi_{n-1}$  and symmetric, i.e.  $\Psi_0(\mathbf{w}, x, \mathbf{z}) = \Psi_0(\mathbf{z}, x, \mathbf{w})$ . It is also independent of the choice of  $\theta$ ,  $\mathcal{L}$  and  $\Lambda$  [18]. (Note that  $\Psi_0$  is related to the kernel  $\Phi_0$  studied in [18, 19, 20] by  $\Psi_0(\mathbf{w}, x, \mathbf{z}) = \Phi_0(x, \mathbf{w}, -\mathbf{z})$ .) Let

$$(3.3) \quad \Psi_0(x, \mathbf{z}) := \frac{1}{v_{d-1}} \int_{\mathcal{B}_1^{d-1}} \Psi_0(\mathbf{w}, x, \mathbf{z}) d\mathbf{w},$$

$$(3.4) \quad \Psi(x, \mathbf{z}) := \frac{1}{\bar{\xi}} \int_x^\infty \Psi_0(x', \mathbf{z}) dx',$$

with the mean free path length  $\bar{\xi} = 1/v_{d-1}$ . Both  $\Psi_0(\mathbf{z}, \mathbf{z})$  and  $\Psi(x, \mathbf{z})$  define probability densities on  $\mathbb{R}_{>0} \times \mathcal{B}_1^{d-1}$  with respect to  $dx dz$ . The first fact follows from the symmetry of the transition kernel, and the second from the relation

$$(3.5) \quad \int_{\mathcal{B}_1^{d-1} \times \mathbb{R}_{>0}} \Psi(x, \mathbf{z}) dx dz = \frac{1}{\bar{\xi}} \int_{\mathcal{B}_1^{d-1} \times \mathbb{R}_{>0}} x \Psi_0(x, \mathbf{z}) dx dz = 1.$$

Suppose in the following that the sequence of random variables

$$(3.6) \quad ((\xi_n, \boldsymbol{\eta}_n))_{n=1}^\infty$$

is given by the Markov chain (3.1), where  $(\xi_1, \boldsymbol{\eta}_1)$  has density either  $\Psi(x, \mathbf{z})$  (for the continuous time setting) or  $\Psi_0(x, \mathbf{z})$  (for the discrete time setting). The relation (3.4) between the two

reflects the fact that the continuous time Markov process is a suspension flow over the discrete time process, where the particle moves with unit speed between consecutive collisions; see [19, Sect. 6] for more details.

We assume in the following that  $R$  is a function  $S_1^{d-1} \rightarrow \text{SO}(d)$  which satisfies  $\mathbf{v} = R(\mathbf{v})\mathbf{e}_1$  and which is smooth when restricted to  $S_1^{d-1} \setminus \{-\mathbf{e}_1\}$ . An example is

$$(3.7) \quad R(\mathbf{v}) = E\left(\frac{2 \arcsin(\|\mathbf{v} - \mathbf{e}_1\|/2)}{\|\mathbf{v}_\perp\|} \mathbf{v}_\perp\right) \quad \text{for } \mathbf{v} \in S_1^{d-1} \setminus \{\mathbf{e}_1, -\mathbf{e}_1\},$$

where  $\mathbf{v}_\perp := (v_2, \dots, v_d) \in \mathbb{R}^{d-1}$ , and  $R(\mathbf{e}_1) = I$ ,  $R(-\mathbf{e}_1) = -I$ .

For  $n \in \mathbb{N}$ , define the following random variables:

$$(3.8) \quad \tau_n := \sum_{j=1}^n \xi_j, \quad \tau_0 := 0, \quad (\text{time to the } n\text{th collision});$$

$$(3.9) \quad \nu_t := \max\{n \in \mathbb{Z}_{\geq 0} : \tau_n \leq t\} \quad (\text{number of collisions within time } t);$$

$$(3.10) \quad \mathbf{V}_n := R(\mathbf{v}_0)S(\boldsymbol{\eta}_1) \cdots S(\boldsymbol{\eta}_n)\mathbf{e}_1, \quad \mathbf{V}_0 := \mathbf{v}_0, \quad (\text{velocity after the } n\text{th collision});$$

$$(3.11) \quad \mathbf{Q}_n := \sum_{j=1}^n \xi_j \mathbf{V}_{j-1} \quad (\text{discrete time displacement});$$

$$(3.12) \quad \mathbf{X}_t := \mathbf{Q}_{\nu_t} + (t - \tau_{\nu_t})\mathbf{V}_{\nu_t} \quad (\text{continuous time displacement}).$$

**Theorem 3.1** ([19]). (i) Under the hypotheses of Theorem 1.1, for any  $t > 0$ ,

$$(3.13) \quad \mathbf{x}_t - \mathbf{x}_0 \Rightarrow \mathbf{X}_t$$

as  $r \rightarrow 0$ , where the random variable  $(\xi_1, \boldsymbol{\eta}_1)$  has density  $\Psi(x, \mathbf{z})$ .

(ii) Under the hypotheses of Theorem 1.2, for any  $n \in \mathbb{N}$ ,

$$(3.14) \quad \mathbf{q}_n - \mathbf{q}_0 \Rightarrow \mathbf{Q}_n$$

as  $r \rightarrow 0$ , where the random variable  $(\xi_1, \boldsymbol{\eta}_1)$  has density  $\Psi_0(x, \mathbf{z})$ .

The main part of this paper is devoted to the proof of the following superdiffusive central limit theorem for the processes  $\mathbf{X}_t$  and  $\mathbf{Q}_n$ , which in turn implies Theorems 1.1 and 1.2. We will only assume that the random variable  $(\xi_1, \boldsymbol{\eta}_1)$  is such that the marginal distribution of  $\boldsymbol{\eta}_1$  is absolutely continuous on  $\mathcal{B}_1^{d-1}$  with respect to Lebesgue measure; there is no further assumption on the distribution of  $\xi_1$ . This hypothesis is satisfied for  $(\xi_1, \boldsymbol{\eta}_1)$  with density  $\Psi_0(x, \mathbf{z})$ , since

$$(3.15) \quad \bar{\Psi}_0(\mathbf{z}) := \int_0^\infty \Psi_0(x, \mathbf{z}) dx = \frac{1}{v_{d-1}} \int_{\mathbb{R}_{>0} \times \mathcal{B}_1^{d-1}} \Psi_0(\mathbf{z}, x, \mathbf{w}) dx d\mathbf{w} = \frac{1}{v_{d-1}}.$$

That is, the marginal distribution of  $\boldsymbol{\eta}_1$  is uniform on  $\mathcal{B}_1^{d-1}$ . We will later see that  $(\xi_1, \boldsymbol{\eta}_1)$  with density  $\Psi(x, \mathbf{z})$  also complies with the above hypothesis (cf. Proposition 10.1). The processes  $\mathbf{X}_t$  and  $\mathbf{Q}_n$  are independent of  $\mathbf{x}_0$  and  $\mathbf{q}_0$ , respectively, and we will in the following fix  $\mathbf{v}_0 \in S_1^{d-1}$ . Also, the required assumptions on the scattering angle  $\theta$  are significantly weaker than in the previous theorems.

**Theorem 3.2.** Let  $d \geq 2$ ,  $\mathbf{v}_0 \in S_1^{d-1}$  and assume that the marginal distribution of  $\boldsymbol{\eta}_1$  is absolutely continuous. Assume  $\theta : [0, 1] \rightarrow [-\pi, \pi]$  is measurable, so that

$$(3.16) \quad \text{meas}\{w \in [0, 1) : \theta(w) \notin \mathbb{Q}\} > 0.$$

Then (i)

$$(3.17) \quad \frac{\mathbf{X}_t}{\Sigma_d \sqrt{t \log t}} \Rightarrow \mathcal{N}(0, I_d),$$

and (ii)

$$(3.18) \quad \frac{\mathbf{Q}_n}{\sigma_d \sqrt{n \log n}} \Rightarrow \mathcal{N}(0, I_d),$$

where  $\mathcal{N}(0, I_d)$  is a centered normal random variable in  $\mathbb{R}^d$  with identity covariance matrix.

In view of Theorem 3.1, Theorem 3.2 implies Theorems 1.1 and 1.2.

It is interesting to compare the above results with the case of a random, rather than periodic, scatterer configuration, where the scatterers are placed at the points of a fixed realisation of a Poisson process in  $\mathbb{R}^d$ . In the case of fixed scattering radius there is, to the best of our knowledge, no proof of a central limit theorem even in dimension  $d = 2$ . In the Boltzmann-Grad limit, however, the work of Gallavotti [15], Spohn [24] and Boldrighini, Bunimovich and Sinai [6] shows that we have an analogue of Theorem 3.1, where the limit random flight process  $\mathbf{X}_t$  is governed by the linear Boltzmann equation. In this setting, (3.6) is a sequence of independent random variables, where  $\xi_n$  has density  $\Psi_0(x) = v_{d-1} \exp(-v_{d-1}x)$  and  $\boldsymbol{\eta}_n$  is uniformly distributed in  $\mathcal{B}_1^{d-1}$ . Routine techniques [23] show that in this case the central limit theorem holds for  $\mathbf{X}_t$  with a standard  $\sqrt{t}$  normalisation, and for  $\mathbf{Q}_n$  with a  $\sqrt{n}$  normalisation.

#### 4. OUTLINE OF THE PROOF OF THEOREM 3.2

We will now outline the central arguments in the proof of Theorem 3.2 (ii) for discrete time by reducing the statement to four main lemmas, whose proof is given in Section 9. The continuous-time case (i) follows from (ii) via technical estimates supplied in Section 11. We will assume from now on that  $(\xi_1, \boldsymbol{\eta}_1)$  has density  $\Psi_0(x, \mathbf{z})$ , and discuss the generalisation to more general distributions in Section 10. We note that for  $\boldsymbol{\eta}_0$  uniformly distributed in  $\mathcal{B}_1^{d-1}$ ,

$$(4.1) \quad \Psi_0(x, \mathbf{z}) = \mathbf{E} \Psi_0(\boldsymbol{\eta}_0, x, \mathbf{z}),$$

and it is therefore equivalent to consider instead of (3.6) the Markov chain

$$(4.2) \quad ((\xi_n, \boldsymbol{\eta}_n))_{n=0}^\infty$$

with the same transition probability (3.2),  $\boldsymbol{\eta}_0$  uniformly distributed in  $\mathcal{B}_1^{d-1}$  and  $\xi_0 = 0$ . The sequence

$$(4.3) \quad \underline{\boldsymbol{\eta}} = (\boldsymbol{\eta}_n)_{n=0}^\infty,$$

with  $\boldsymbol{\eta}_0$  as defined above, is itself generated by a Markov chain on the state space  $\mathcal{B}_1^{d-1}$  with transition probability

$$(4.4) \quad \mathbf{P}(\boldsymbol{\eta}_n \in \mathcal{A} \mid \boldsymbol{\eta}_{n-1}) = \int_{\mathcal{A}} K_0(\boldsymbol{\eta}_{n-1}, \mathbf{z}) d\mathbf{z}$$

where

$$(4.5) \quad K_0(\mathbf{w}, \mathbf{z}) := \int_0^\infty \Psi_0(\mathbf{w}, x, \mathbf{z}) dx.$$

The objective is to prove a central limit theorem of sums of the random variables  $\xi_n \mathbf{V}_{n-1}$ . The first observation is that these are of course not independent. If we, however, condition on the sequence  $\underline{\boldsymbol{\eta}}$ , then the  $\mathbf{V}_n$  are deterministic, and  $(\xi_n)_{n=1}^\infty$  is a sequence of independent (but not identically distributed) random variables,

$$(4.6) \quad \mathbf{P}(\xi_n \in (x, x + dx) \mid \underline{\boldsymbol{\eta}}) = \frac{\Psi_0(\boldsymbol{\eta}_{n-1}, x, \boldsymbol{\eta}_n) dx}{K_0(\boldsymbol{\eta}_{n-1}, \boldsymbol{\eta}_n)}.$$

The plan is now to apply the Lindeberg central limit theorem to the sum of independent random variables,  $\mathbf{Q}_n = \sum_{j=1}^n \xi_j \mathbf{V}_{j-1}$ , conditioned on  $\underline{\boldsymbol{\eta}}$ .

To this end we first truncate  $\mathbf{Q}_n$  by defining the random variable

$$(4.7) \quad \mathbf{Q}'_n := \sum_{j=1}^n \xi'_j \mathbf{V}_{j-1}$$

with

$$(4.8) \quad \xi'_j := \xi_j \mathbb{1}_{\{\xi_j^2 \leq j(\log j)^\gamma\}}$$

for some fixed  $\gamma \in (1, 2)$ . The following lemma tells us that it is sufficient to prove Theorem 3.2 (ii) for  $\mathbf{Q}'_n$  instead of  $\mathbf{Q}_n$ .

**Lemma 4.1.** *We have*

$$(4.9) \quad \sup_{n \in \mathbb{N}} \|\mathbf{Q}_n - \mathbf{Q}'_n\| < \infty$$

*almost surely.*

To prove the central limit theorem for  $\mathbf{Q}'_n$ , we center  $\xi'_j$  by setting

$$(4.10) \quad \tilde{\xi}_j = \xi'_j - m_j,$$

with the conditional expectation

$$(4.11) \quad m_j := \mathbf{E}(\xi'_j \mid \underline{\eta}) = \frac{K_{1,r_j}(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}{K_0(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}$$

where  $r_j := \sqrt{j(\log j)^\gamma}$  and

$$(4.12) \quad K_{1,r}(\mathbf{w}, \mathbf{z}) := \int_0^r x \Psi_0(\mathbf{w}, x, \mathbf{z}) dx.$$

Let

$$(4.13) \quad \tilde{\mathbf{Q}}_n := \sum_{j=1}^n \tilde{\xi}_j \mathbf{V}_{j-1}.$$

The following lemma shows that  $\mathbf{Q}'_n$  and  $\tilde{\mathbf{Q}}_n$  are close relative to  $\sqrt{n \log n}$ .

**Lemma 4.2.** *The sequence of random variables*

$$(4.14) \quad \frac{\mathbf{Q}'_n - \tilde{\mathbf{Q}}_n}{\sqrt{n \log \log n}}$$

*is tight if  $d = 2$ , and*

$$(4.15) \quad \frac{\mathbf{Q}'_n - \tilde{\mathbf{Q}}_n}{\sqrt{n}}$$

*is tight if  $d \geq 3$ .*

It is therefore sufficient to prove Theorem 3.2 (ii) for  $\tilde{\mathbf{Q}}_n$  in place of  $\mathbf{Q}_n$ . This will be achieved by applying the Lindeberg central limit theorem to the conditional sum as alluded to above. We begin by estimating the conditional variance. Set

$$(4.16) \quad a_j^2 := \mathbf{Var}(\tilde{\xi}_j \mid \underline{\eta}) = \frac{K_{2,r_j}(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}{K_0(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)} - m_j^2,$$

with

$$(4.17) \quad K_{2,r}(\mathbf{w}, \mathbf{z}) := \int_0^r x^2 \Psi_0(\mathbf{w}, x, \mathbf{z}) dx.$$

**Lemma 4.3.** *There is a constant  $\sigma_d > 0$  such that, for  $n \rightarrow \infty$ ,*

$$(4.18) \quad \frac{\mathbf{E}(\tilde{\mathbf{Q}}_n \otimes \tilde{\mathbf{Q}}_n \mid \underline{\eta})}{n \log n} = \frac{\sum_{j=1}^n a_j^2 \mathbf{V}_{j-1} \otimes \mathbf{V}_{j-1}}{n \log n} \xrightarrow{\mathbf{P}} \sigma_d^2 I_d.$$



By taking the trace in (4.18), we have in particular

$$(4.19) \quad \frac{A_n^2}{n \log n} \xrightarrow{\mathbf{P}} d \sigma_d^2$$

for

$$(4.20) \quad A_n^2 := \sum_{j=1}^n a_j^2 = \mathbf{E}(\|\tilde{\mathbf{Q}}_n\|^2 \mid \underline{\eta}).$$

The next lemma verifies the Lindeberg conditions for random  $\underline{\eta}$ .

**Lemma 4.4.** *For any fixed  $\varepsilon > 0$ ,*

$$(4.21) \quad A_n^{-2} \sum_{j=1}^n \mathbf{E}(\tilde{\xi}_j^2 \mathbb{1}_{\{\tilde{\xi}_j^2 > \varepsilon^2 A_n^2\}} \mid \underline{\eta}) \xrightarrow{\mathbf{P}} 0$$

as  $n \rightarrow \infty$ .

Given these lemmas, let us now conclude the proof of the fact that

$$(4.22) \quad \mathbf{Y}_n := \frac{\tilde{\mathbf{Q}}_n}{\sigma_d \sqrt{n \log n}} \Rightarrow \mathcal{N}(0, I_d).$$

By Chebyshev's inequality we have, for any  $K > 0$ ,

$$(4.23) \quad \mathbf{P}(\|\mathbf{Y}_n\| > K \mid \underline{\eta}) \leq \frac{1}{K^2} \mathbf{E}(\|\mathbf{Y}_n\|^2 \mid \underline{\eta}),$$

and thus, for any  $\kappa > 0$ ,

$$(4.24) \quad \begin{aligned} \mathbf{P}(\|\mathbf{Y}_n\| > K) &\leq \frac{\kappa^2}{K^2} + \mathbf{P}(\mathbf{E}(\|\mathbf{Y}_n\|^2 \mid \underline{\eta}) > \kappa^2) \\ &= \frac{\kappa^2}{K^2} + \mathbf{P}(A_n^2 > \kappa^2 \sigma_d^2 n \log n). \end{aligned}$$

By (4.19), the second term on the right hand side of (4.24) converges to 0 as  $n \rightarrow \infty$ , if we choose  $\kappa = d$ , say. So (4.24) implies that the sequence of random variables  $\mathbf{Y}_n$  is tight. By the Helly-Prokhorov theorem, there is an infinite subset  $S_1 \subset \mathbb{N}$  so that  $\mathbf{Y}_n$  converges in distribution along  $n \in S_1$  to some limit  $\mathbf{Y}$ . Assume for a contradiction that  $\mathbf{Y}$  is *not* distributed according to  $\mathcal{N}(0, I_d)$ . The Borel-Cantelli lemma implies that there is an infinite subset  $S_2 \subset S_1$ , so that in the statements of Lemmas 4.3 and 4.4 we have almost-sure convergence along  $n \in S_2$ :

$$(4.25) \quad \frac{\mathbf{E}(\tilde{\mathbf{Q}}_n \otimes \tilde{\mathbf{Q}}_n \mid \underline{\eta})}{n \log n} \xrightarrow{\text{a.s.}} \sigma_d^2 I_d,$$

$$(4.26) \quad \frac{A_n^2}{n \log n} \xrightarrow{\text{a.s.}} d \sigma_d^2,$$

and

$$(4.27) \quad A_n^{-2} \sum_{j=1}^n \mathbf{E}(\tilde{\xi}_j^2 \mathbb{1}_{\{\tilde{\xi}_j^2 > \varepsilon^2 A_n^2\}} \mid \underline{\eta}) \xrightarrow{\text{a.s.}} 0.$$

The hypotheses of the Lindeberg central limit theorem are met, and we infer that  $\mathbf{Y}_n \Rightarrow \mathcal{N}(0, I_d)$  for  $n \in S_2$ . (We use the Lindeberg theorem for *triangular arrays* of independent random variables, since we have verified the Lindeberg conditions only along a subsequence.) This, however, contradicts our assumption that  $\mathbf{Y}$  is not normal, and hence  $\mathcal{N}(0, I_d)$  is indeed the unique limit point of any converging subsequence. This in turn implies that every sequence converges, and therefore completes the proof of (4.22). In view of Lemmas 4.1 and 4.2, this implies Theorem 3.2 (ii) (still under the assumption that  $(\xi_1, \boldsymbol{\eta}_1)$  has density  $\Psi_0(x, \mathbf{z})$ ).

Let us briefly describe the further contents of this paper. In Section 5 we recall the basic properties of the transition kernel  $\Psi_0(\mathbf{w}, x, \mathbf{z})$  from [20]. Section 6 establishes key estimates for

the moments  $K_{p,r}(\mathbf{w}, \mathbf{z})$ ,  $m_j$  and  $a_j$  introduced above. In Sections 7 and 8 we prove spectral gap estimates and exponential mixing for the discrete time Markov process defined in (4.4). The estimates from Sections 6–8 are the main input in the proof of Lemmas 4.1–4.4, which is given in Section 9. In Section 10 we show that the discrete-time statement in Theorem 3.2 (ii) holds for more general initial distributions than  $\Psi_0(x, \mathbf{z})$ . It holds in particular for  $\Psi(x, \mathbf{z})$ , which appears in the continuous-time variant. Section 11 explains how to pass from discrete to continuous time, thus completing the proof of Theorem 3.2 (i).

## 5. THE TRANSITION KERNEL

In dimension  $d = 2$  we have the following explicit formula for the transition kernel. For  $w, z \in (-1, 1)$ ,

$$(5.1) \quad \Psi_0(w, x, z) = \frac{6}{\pi^2} \begin{cases} \mathbb{1}_{\{0 \leq x < \frac{1}{1+w}\}} + \mathbb{1}_{\{\frac{1}{1+w} \leq x < \frac{1}{1+z}\}} \frac{x^{-1} - 1 - z}{w - z} & \text{if } 0 < w, -w < z < w, \\ \mathbb{1}_{\{0 \leq x < \frac{1}{1+z}\}} + \mathbb{1}_{\{\frac{1}{1+z} \leq x < \frac{1}{1+w}\}} \frac{x^{-1} - 1 - w}{z - w} & \text{if } 0 < z, -z < w < z, \\ \mathbb{1}_{\{0 \leq x < \frac{1}{1-w}\}} + \mathbb{1}_{\{\frac{1}{1-w} \leq x < \frac{1}{1-z}\}} \frac{x^{-1} - 1 + z}{z - w} & \text{if } w < 0, w < z < -w, \\ \mathbb{1}_{\{0 \leq x < \frac{1}{1-z}\}} + \mathbb{1}_{\{\frac{1}{1-z} \leq x < \frac{1}{1-w}\}} \frac{x^{-1} - 1 + w}{w - z} & \text{if } z < 0, z < w < -z. \end{cases}$$

This formula has been derived, independently and with different methods, by Marklof and Strömbergsson [17], Caglioti and Golse [10, 11] and by Bykovskii and Ustinov [9].

In dimension  $d \geq 3$  we have no such explicit formulas for the transition kernel. We recall from [18, 19, 20] the following properties. If  $d \geq 3$ , the function

$$(5.2) \quad \Psi_0 : \mathcal{B}_1^{d-1} \times \mathbb{R}_{>0} \times \mathcal{B}_1^{d-1} \rightarrow [0, 1]$$

is continuous.  $\Psi_0(\mathbf{w}, x, \mathbf{z})$  depends only on  $x$ ,  $w := \|\mathbf{w}\|$ ,  $z := \|\mathbf{z}\|$  and the angle  $\varphi := \varphi(\mathbf{w}, \mathbf{z}) \in [0, \pi]$  between the vectors  $\mathbf{w}, \mathbf{z} \in \mathcal{B}_1^{d-1}$ . Note that in dimension  $d = 2$  the angle  $\varphi$  can only take the values 0 and  $\pi$ . For statements that are specific to dimension  $d = 2$ , we will often use  $w \in (-1, 1)$  instead of  $\mathbf{w}$ , and  $|w|$  instead of  $w = \|\mathbf{w}\|$ . We recall once more that  $\Psi_0(\mathbf{w}, x, \mathbf{z}) = \Phi_0(x, \mathbf{w}, -\mathbf{z})$  in the notation of [18, 19, 20], and so in particular the angle  $\varphi$  between  $\mathbf{w}, \mathbf{z}$  becomes  $\pi - \varphi$ .

Our proofs will exploit the following estimates on the transition kernel [20, 26]. All bounds are uniform in  $x > 0$  and  $\mathbf{w}, \mathbf{z} \in \mathcal{B}_1^{d-1}$ . We have by [20, Thm. 1.1],

$$(5.3) \quad \frac{1 - 2^{d-1} v_{d-1} x}{\zeta(d)} \leq \Psi_0(\mathbf{w}, x, \mathbf{z}) \leq \frac{1}{\zeta(d)}.$$

Furthermore, by [20, Thm. 1.7], there exists a continuous and uniformly bounded function  $F_{0,d} : \mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  such that

$$(5.4) \quad \Psi_0(\mathbf{w}, x, \mathbf{z}) = x^{-2+\frac{2}{d}} F_{0,d} \left( x^{\frac{2}{d}}(1-z), x^{\frac{2}{d}}(1-w), x^{\frac{1}{d}}(\pi - \varphi) \right) + O(E),$$

where the error term is

$$(5.5) \quad E = \begin{cases} x^{-2} & \text{if } d = 2, \\ x^{-2} \log(2 + \min(x, (\pi - \varphi)^{-1})) & \text{if } d = 3, \\ \min(x^{-2}, x^{-3+\frac{2}{d-1}} (\pi - \varphi)^{2-d+\frac{2}{d-1}}) & \text{if } d \geq 4. \end{cases}$$

It is noted in [20] that  $F_{0,d}$  is uniformly bounded from below for  $t_1, t_2, \alpha$  near zero. That is, there is a small constant  $c > 0$  which only depends on  $d$  such that

$$(5.6) \quad \max(t_1, t_2, \alpha) < c \implies F_{0,d}(t_2, t_1, \alpha) > c.$$

Furthermore, the support of  $F_{0,d}$  is contained in  $(0, c'] \times (0, c'] \times \mathbb{R}_{\geq 0}$  for some  $c' > 0$ , and for any fixed  $t_1, t_2 > 0$ , the function  $F_{0,d}(t_1, t_2, \cdot)$  has compact support.

In dimension  $d \geq 3$ , the following upper bound will prove useful [26, Thm. 1.8]:

$$(5.7) \quad \Psi_0(\mathbf{w}, x, \mathbf{z}) \ll \begin{cases} x^{-2} \min\left\{1, (x\varphi^{d-2})^{-1+\frac{2}{d-1}}\right\} & \text{if } \varphi \in [0, \frac{\pi}{2}] \\ x^{-2+\frac{2}{d}} \min\left\{1, (x(\pi-\varphi)^d)^{-1+\frac{2}{d(d-1)}}\right\} & \text{if } \varphi \in [\frac{\pi}{2}, \pi]. \end{cases}$$

The notation  $f \ll g$  is here defined as  $f = O(g)$ , i.e., there exists a constant  $C > 0$  such that  $|f| \leq C|g|$ .

The support of  $\Psi_0(\mathbf{w}, x, \mathbf{z})$  is described by a continuous function  $x_0 : \mathcal{B}_1^{d-1} \times \mathcal{B}_1^{d-1} \rightarrow \mathbb{R}_{>0}$ . We have  $\Psi_0(\mathbf{w}, x, \mathbf{z}) > 0$  holds if and only if  $x < x_0(\mathbf{w}, \mathbf{z})$ . Set

$$(5.8) \quad t := t(\mathbf{w}, \mathbf{z}) := \max(1-w, 1-z) \in (0, 1].$$

If  $d \geq 3$ , then [26, Prop. 1.9] tell us that

$$(5.9) \quad x_0(\mathbf{w}, \mathbf{z}) \asymp \begin{cases} \max(t^{-\frac{d-2}{2}}, t^{-\frac{d-1}{2}} \varphi) & \text{if } \varphi \in [0, \frac{\pi}{2}] \\ \min(t^{-\frac{d}{2}}, t^{-\frac{d-1}{2}} (\pi-\varphi)^{-1}) & \text{if } \varphi \in [\frac{\pi}{2}, \pi]. \end{cases}$$

(If  $\varphi = \pi$  then the right hand side of (5.9) should be interpreted as  $t^{-\frac{d}{2}}$ .)

For the distribution of free path length between consecutive collisions,

$$(5.10) \quad \Psi_0(x) = \frac{1}{v_{d-1}} \int_{\mathcal{B}_1^{d-1}} \int_{\mathcal{B}_1^{d-1}} \Psi_0(\mathbf{w}, x, \mathbf{z}) d\mathbf{w} d\mathbf{z},$$

we have the following tail estimate ( $\Psi_0$  is denoted  $\bar{\Phi}_0$  in [20]): For  $x \rightarrow \infty$ ,

$$(5.11) \quad \Psi_0(x) = \Theta_d x^{-3} + O(x^{-3-\frac{2}{d}}) \times \begin{cases} 1 & \text{if } d = 2 \\ \log x & \text{if } d = 3 \\ 1 & \text{if } d \geq 4 \end{cases}$$

with

$$(5.12) \quad \Theta_d := \frac{2^{2-d}}{d(d+1)\zeta(d)}.$$

This asymptotic estimate sharpens earlier upper and lower bounds by Bourgain, Golse and Wennberg [7, 16]. Note that the variances in Theorems 1.1 and 1.2 are related to the above tail via

$$(5.13) \quad \Sigma_d^2 = \frac{\Theta_d}{2d\bar{\xi}}, \quad \sigma_d^2 = \frac{\Theta_d}{2d}.$$

## 6. MOMENT ESTIMATES

We now provide key estimates of the random variables introduced in the previous section. For  $p = 0, 1, 2$  and  $r > 0$ , set

$$(6.1) \quad K_{p,r}(\mathbf{w}, \mathbf{z}) := \int_0^r x^p \Psi_0(\mathbf{w}, x, \mathbf{z}) dx$$

and

$$(6.2) \quad K_p(\mathbf{w}, \mathbf{z}) := \int_0^\infty x^p \Psi_0(\mathbf{w}, x, \mathbf{z}) dx.$$

We furthermore define the random variables (recall Section 4)

$$(6.3) \quad m_j := \mathbf{E}(\xi'_j | \underline{\boldsymbol{\eta}}) = \frac{K_{1,r_j}(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}{K_0(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}, \quad \mu_j := \mathbf{E}(\xi_j | \underline{\boldsymbol{\eta}}) = \frac{K_1(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}{K_0(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)},$$

$$(6.4) \quad b_j^2 := \mathbf{E}(\xi'_j | \underline{\boldsymbol{\eta}})^2 = \frac{K_{2,r_j}(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}{K_0(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}, \quad \beta_j^2 := \mathbf{E}(\xi_j | \underline{\boldsymbol{\eta}})^2 = \frac{K_2(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}{K_0(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)},$$

$$(6.5) \quad a_j^2 := \mathbf{Var}(\xi'_j \mid \underline{\eta}) = b_j^2 - m_j^2, \quad \alpha_j^2 := \mathbf{Var}(\xi_j \mid \underline{\eta}) = \beta_j^2 - \mu_j^2,$$

and

$$(6.6) \quad A_n^2 := \sum_{j=1}^n a_j^2,$$

with  $r_j = \sqrt{j(\log j)^\gamma}$  for some fixed  $\gamma \in (1, 2)$ .

**Lemma 6.1.** *Let  $d = 2$ . Then, for  $w, z \in (-1, 1)$ ,*

$$(6.7) \quad K_0(w, z) = \frac{6}{\pi^2} \begin{cases} \frac{1}{w-z} \ln \frac{1+w}{1+z} & \text{if } w+z \geq 0, \\ \frac{1}{z-w} \ln \frac{1-w}{1-z} & \text{if } w+z \leq 0, \end{cases}$$

$$(6.8) \quad K_1(w, z) = \frac{3}{\pi^2} \begin{cases} \frac{1}{(1+w)(1+z)} & \text{if } w+z \geq 0, \\ \frac{1}{(1-w)(1-z)} & \text{if } w+z \leq 0, \end{cases}$$

$$(6.9) \quad K_2(w, z) = \frac{1}{\pi^2} \begin{cases} \frac{2+w+z}{(1+w)^2(1+z)^2} & \text{if } w+z \geq 0, \\ \frac{2-w-z}{(1-w)^2(1-z)^2} & \text{if } w+z \leq 0. \end{cases}$$

*Proof.* These follow from the explicit formula (5.1) by direct computation.  $\square$

**Lemma 6.2.** *Let  $d \geq 3$ . Then*

$$(6.10) \quad \inf_{\mathbf{w}, \mathbf{z} \in \mathcal{B}_1^{d-1}} K_0(\mathbf{w}, \mathbf{z}) \geq \frac{1}{2^d v_{d-1} \zeta(d)} > 0,$$

$$(6.11) \quad \sup_{\mathbf{w}, \mathbf{z} \in \mathcal{B}_1^{d-1}} K_0(\mathbf{w}, \mathbf{z}) < \infty.$$

*Proof.* We have

$$(6.12) \quad K_0(\mathbf{w}, \mathbf{z}) \geq \int_0^y \Psi_0(\mathbf{w}, x, \mathbf{z}) dx$$

for any  $y \geq 0$ . Theorem 1.1 in [20] states that for  $x > 0$  and  $\mathbf{w}, \mathbf{z} \in \mathcal{B}_1^{d-1}$ ,

$$(6.13) \quad \Psi_0(\mathbf{w}, x, \mathbf{z}) \geq \frac{1 - 2^{d-1} v_{d-1} x}{\zeta(d)},$$

and the lower bound follows with the choice  $y = (2^{d-1} v_{d-1})^{-1}$ . The upper bound follows from (5.7) which tells us that

$$(6.14) \quad \Psi_0(\mathbf{w}, x, \mathbf{z}) = O(x^{-2+\frac{2}{d}}),$$

where the implied constant is independent of  $x, \mathbf{w}, \mathbf{z}$ .  $\square$

**Lemma 6.3.** *Let  $d \geq 3$ . For  $\mathbf{w}, \mathbf{z} \in \mathcal{B}_1^{d-1}$ ,*

$$(6.15) \quad K_1(\mathbf{w}, \mathbf{z}) \gg \min(t^{-1}, (\pi - \varphi)^{-2}),$$

and

$$(6.16) \quad K_1(\mathbf{w}, \mathbf{z}) \ll \begin{cases} 1 + \log \max(t^{-\frac{d-2}{2}}, t^{-\frac{d-1}{2}} \varphi) & \text{if } \varphi \in [0, \frac{\pi}{2}], \\ \min(t^{-1}, t^{-1+\frac{1}{d}} (\pi - \varphi)^{-\frac{2}{d}}) & \text{if } \varphi \in [\frac{\pi}{2}, \pi]. \end{cases}$$

*Proof.* As to the lower bound (6.15), we note that by (5.4)

$$(6.17) \quad K_1(\mathbf{w}, \mathbf{z}) \geq \int_{\substack{x>0 \\ x^{\frac{2}{d}}t(\mathbf{w}, \mathbf{z})<c \\ x^{\frac{1}{d}}(\pi-\varphi)<c}} x \left\{ x^{-2+\frac{2}{d}} F_{0,d} \left( x^{\frac{2}{d}}(1-z), x^{\frac{2}{d}}(1-w), x^{\frac{1}{d}}(\pi-\varphi) \right) - |O(E)| \right\} dx,$$

which, in view of (5.5) and (5.6), implies

$$(6.18) \quad \begin{aligned} K_1(\mathbf{w}, \mathbf{z}) &> c \int_{\substack{x>0 \\ x^{\frac{2}{d}}t(\mathbf{w}, \mathbf{z})<c \\ x^{\frac{1}{d}}(\pi-\varphi)<c}} \left\{ x^{-1+\frac{2}{d}} - |O(x^{-1} \log(2+x))| \right\} dx \\ &\gg \min(t(\mathbf{w}, \mathbf{z})^{-1}, (\pi-\varphi)^{-2}). \end{aligned}$$

The upper bound (6.16) follows from (5.7) and (5.9): for  $\varphi \in [0, \frac{\pi}{2}]$ ,

$$(6.19) \quad K_1(\mathbf{w}, \mathbf{z}) \ll 1 + \int_1^{x_0(\mathbf{w}, \mathbf{z})} x^{-1} dx \ll 1 + \log \max(t^{-\frac{d-2}{2}}, t^{-\frac{d-1}{2}} \varphi),$$

and for  $\varphi \in [\frac{\pi}{2}, \pi]$ , we have

$$(6.20) \quad K_1(\mathbf{w}, \mathbf{z}) \ll \int_0^{x_0(\mathbf{w}, \mathbf{z})} x^{-1+\frac{2}{d}} dx \ll \min(t^{-1}, t^{-1+\frac{1}{d}}(\pi-\varphi)^{-\frac{2}{d}}).$$

□

**Lemma 6.4.** *Let  $d \geq 3$ . Then, for  $\mathbf{w}, \mathbf{z} \in \mathcal{B}_1^{d-1}$ ,*

$$(6.21) \quad K_2(\mathbf{w}, \mathbf{z}) \gg \min(t^{-(1+\frac{d}{2})}, (\pi-\varphi)^{-(d+2)}),$$

and

$$(6.22) \quad K_2(\mathbf{w}, \mathbf{z}) \ll \begin{cases} \max(t^{-\frac{d-2}{2}}, t^{-\frac{d-1}{2}} \varphi) & \text{if } \varphi \in [0, \frac{\pi}{2}], \\ \min(t^{-(1+\frac{d}{2})}, t^{-\frac{d+1}{2}+\frac{1}{d}}(\pi-\varphi)^{-(1+\frac{2}{d})}) & \text{if } \varphi \in [\frac{\pi}{2}, \pi]. \end{cases}$$

*Proof.* The lower bound (6.21) follows from

$$(6.23) \quad K_2(\mathbf{w}, \mathbf{z}) \geq \int_{\substack{x>0 \\ x^{\frac{2}{d}}t(\mathbf{w}, \mathbf{z})<c \\ x^{\frac{1}{d}}(\pi-\varphi)<c}} x^2 \left\{ x^{-2+\frac{2}{d}} F_{0,d} \left( x^{\frac{2}{d}}(1-z), x^{\frac{2}{d}}(1-w), x^{\frac{1}{d}}(\pi-\varphi) \right) - |O(E)| \right\} dx,$$

and (5.5) and (5.6). Hence

$$(6.24) \quad \begin{aligned} K_2(\mathbf{w}, \mathbf{z}) &> c \int_{\substack{x>0 \\ x^{\frac{2}{d}}t(\mathbf{w}, \mathbf{z})<c \\ x^{\frac{1}{d}}(\pi-\varphi)<c}} \left\{ x^{\frac{2}{d}} - |O(\log(2+x))| \right\} dx \\ &\gg \min(t(\mathbf{w}, \mathbf{z})^{-(1+\frac{d}{2})}, (\pi-\varphi)^{-(d+2)}). \end{aligned}$$

The upper bound (6.22) follows from (5.7) and (5.9): for  $\varphi \in [0, \frac{\pi}{2}]$ ,

$$(6.25) \quad K_2(\mathbf{w}, \mathbf{z}) \ll \int_0^{x_0(\mathbf{w}, \mathbf{z})} dx \ll \max(t^{-\frac{d-2}{2}}, t^{-\frac{d-1}{2}} \varphi),$$

and for  $\varphi \in [\frac{\pi}{2}, \pi]$ , we have

$$(6.26) \quad K_2(\mathbf{w}, \mathbf{z}) \ll \int_0^{x_0(\mathbf{w}, \mathbf{z})} x^{\frac{2}{d}} dx \ll \min(t^{-(1+\frac{d}{2})}, t^{-\frac{d+1}{2}+\frac{1}{d}}(\pi-\varphi)^{-(1+\frac{2}{d})}).$$

□

**Proposition 6.5.** *Let  $d = 2$ . For  $u \rightarrow \infty$ ,*

$$(6.27) \quad \mathbf{P}(\mu_j > u) \sim \frac{3}{4\pi^2} \frac{1}{u^2 \log u}.$$

*Proof.* By the invariance of the integrand under  $(w, z) \mapsto -(w, z)$  and  $(w, z) \mapsto (z, w)$  we have

$$(6.28) \quad \begin{aligned} \mathbf{P}(\mu_j > u) &= \frac{1}{2} \int_{(-1,1)^2} \mathbb{1}_{\{K_1(w,z) > u K_0(w,z)\}} K_0(w, z) \, dw \, dz \\ &= 2 \int_{|z| < w < 1} \mathbb{1}_{\{K_1(w,z) > u K_0(w,z)\}} K_0(w, z) \, dw \, dz. \end{aligned}$$

In this range of integration, we have explicitly

$$(6.29) \quad K_0(w, z) = \frac{6}{\pi^2} \frac{1}{w-z} \ln \frac{1+w}{1+z}$$

and

$$(6.30) \quad \frac{K_0(w, z)}{K_1(w, z)} = 2 \frac{(1+w)(1+z)}{w-z} \ln \frac{1+w}{1+z}.$$

Using the variable substitution

$$(6.31) \quad x = w - z, \quad y = \frac{1+z}{w-z},$$

we have, with the shorthand  $f(y) = 2y(1+y) \ln(1+y^{-1})$ ,

$$(6.32) \quad \begin{aligned} \mathbf{P}(\mu_j > u) &= \frac{12}{\pi^2} \int_0^\infty \int_{\frac{2}{1+2y}}^{\frac{2}{1+y}} \mathbb{1}_{\{xf(y) < u^{-1}\}} \ln(1+y^{-1}) \, dx \, dy \\ &= \frac{12}{\pi^2} \int_0^\infty \mathbb{1}_{\{\frac{2f(y)}{1+y} < u^{-1}\}} \left( \frac{2}{1+y} - \frac{2}{1+2y} \right) \ln(1+y^{-1}) \, dy \\ &\quad + \frac{12}{\pi^2} \int_0^\infty \mathbb{1}_{\{\frac{2f(y)}{1+2y} < u^{-1} < \frac{2f(y)}{1+y}\}} \left( \frac{1}{uf(y)} - \frac{2}{1+2y} \right) \ln(1+y^{-1}) \, dy. \end{aligned}$$

The first term equals

$$(6.33) \quad \begin{aligned} \frac{12}{\pi^2} \int_0^\infty \mathbb{1}_{\{4y \ln(1+y^{-1}) < u^{-1}\}} (2y + O(y^2)) \ln(1+y^{-1}) \, dy \\ = \frac{12}{\pi^2} (y^2 \ln(1+y^{-1}) + O(y^2)) \Big|_{0+}^{4y \ln(1+y^{-1})=u^{-1}} = \frac{3}{4\pi^2} \frac{1}{u^2 \ln u} (1 + o(1)). \end{aligned}$$

The second term is bounded above by

$$(6.34) \quad \begin{aligned} \frac{12}{\pi^2} \int_0^\infty \mathbb{1}_{\{\frac{2f(y)}{1+2y} < u^{-1} < \frac{2f(y)}{1+y}\}} \left( \frac{2}{1+y} - \frac{2}{1+2y} \right) \ln(1+y^{-1}) \, dy \\ = \frac{12}{\pi^2} (y^2 \ln(1+y^{-1}) + O(y^2)) \Big|_{\frac{2f(y)}{1+y}=u^{-1}}^{\frac{2f(y)}{1+2y}=u^{-1}} = o\left(\frac{1}{u^2 \ln u}\right). \end{aligned}$$

This proves (6.27).  $\square$

**Proposition 6.6.** *Let  $d \geq 3$ . There are constants  $c_2 > c_1 > 0$ , such that for  $u \geq 1$ ,*

$$(6.35) \quad c_1 u^{-(1+\frac{d}{2})} \leq \mathbf{P}(\mu_j > u) \leq c_2 u^{-(1+\frac{d}{2})}.$$

*Proof.* The upper bounds in (6.16) for  $K_1$ , and the upper and lower bounds for  $K_0$  in Lemma 6.2 imply

$$(6.36) \quad \int_{K_1 > u K_0} K_0 \, d\mathbf{w} \, d\mathbf{z} \ll \int_{\substack{\log \max(t^{-\frac{d-2}{2}}, t^{-\frac{d-1}{2}} \varphi) \gg u \\ 0 \leq \varphi \leq \frac{\pi}{2}}} d\mathbf{w} \, d\mathbf{z} + \int_{\substack{\min(t^{-1}, t^{-1+\frac{1}{d}}(\pi-\varphi)^{-\frac{2}{d}}) \gg u \\ \frac{\pi}{2} \leq \varphi \leq \pi}} d\mathbf{w} \, d\mathbf{z}.$$

Using spherical coordinates, we see that the second term is, up to a multiplicative constant, equal to

$$(6.37) \quad \int_{\substack{0 < z < w < 1, \frac{\pi}{2} \leq \varphi \leq \pi \\ \min((1-z)^{-1}, (1-z)^{-1+\frac{1}{d}}(\pi-\varphi)^{-\frac{2}{d}}) \gg u}} w^{d-2} dw z^{d-2} dz (\sin \varphi)^{d-3} d\varphi.$$

We now substitute  $x = 1 - w$ ,  $y = 1 - z$ ,  $\phi = \pi - \varphi$ , and integrate  $x$  over  $(0, y)$ , and then over  $y$ . This yields (again up to a multiplicative constant)

$$(6.38) \quad \begin{aligned} & \int_0^{\pi/2} \min(u^{-1}, u^{-\frac{d}{d-1}} \phi^{-\frac{2}{d-1}})^2 (\sin \phi)^{d-3} d\phi \\ &= u^{-2} \int_0^{u^{-1/2}} (\sin \phi)^{d-3} d\phi + u^{-\frac{2d}{d-1}} \int_{u^{-1/2}}^{\pi/2} \phi^{-\frac{4}{d-1}} (\sin \phi)^{d-3} d\phi \\ &\ll u^{-(1+\frac{d}{2})}. \end{aligned}$$

The first term in (6.36) can be dealt with similarly, and yields a lower order contribution. This proves the upper bound

$$(6.39) \quad \mathbf{P}(\mu_j > u) \ll u^{-(1+\frac{d}{2})}.$$

As to the lower bound in (6.35), we use (6.15) (and the same variable substitutions as above):

$$(6.40) \quad \begin{aligned} \int_{K_1 > u K_0} K_0 d\mathbf{w} d\mathbf{z} &\gg \int_{\substack{\min(t^{-1}, (\pi-\varphi)^{-2}) \gg u \\ \frac{\pi}{2} \leq \varphi \leq \pi}} d\mathbf{w} d\mathbf{z} \\ &\gg u^{-2} \int_0^{u^{-1/2}} (\sin \phi)^{d-3} d\phi \\ &\gg u^{-(1+\frac{d}{2})}. \end{aligned}$$

□

**Proposition 6.7.**

$$(6.41) \quad \mathbf{P}(m_j > u) = \begin{cases} O(u^{-2}(\log u)^{-1}) \mathbb{1}_{\{u \leq r_j\}} & (d = 2) \\ O(u^{-(1+\frac{d}{2})}) \mathbb{1}_{\{u \leq r_j\}} & (d \geq 3). \end{cases}$$

*Proof.* This follows from Propositions 6.5 and 6.6, since  $m_j \leq \mu_j$  and  $m_j \leq r_j$ . □

**Proposition 6.8.**

$$(6.42) \quad \mathbf{E}(m_j) = \bar{\xi} + O(r_j^{-1}).$$

*Proof.* We have

$$(6.43) \quad \begin{aligned} \mathbf{E}(m_j) &= \frac{1}{v_{d-1}} \int_{\mathcal{B}_1^{d-1}} \int_{\mathcal{B}_1^{d-1}} K_{1,r_j}(\mathbf{w}, \mathbf{z}) d\mathbf{w} d\mathbf{z} \\ &= \int_0^{r_j} x \Psi_0(x) dx \\ &= \bar{\xi} - \int_{r_j}^{\infty} x \Psi_0(x) dx, \end{aligned}$$

and (6.42) follows from the asymptotics (5.11). □

**Proposition 6.9.**

$$(6.44) \quad \mathbf{E}(m_j^2) = \begin{cases} O(\log \log j) & (d = 2) \\ O(1) & (d \geq 3). \end{cases}$$

*Proof.* This is a direct corollary of Proposition 6.7.  $\square$

**Proposition 6.10.** For  $d \geq 2$  and  $j \rightarrow \infty$ ,

$$(6.45) \quad \mathbf{E}(a_j^2) = \frac{\Theta_d}{2} \log j + O(\log \log j),$$

$$(6.46) \quad \mathbf{E}(b_j^2) = \frac{\Theta_d}{2} \log j + O(\log \log j).$$

*Proof.* We have

$$(6.47) \quad \mathbf{E}(b_j^2) = \frac{1}{v_{d-1}} \int_{\mathcal{B}_1^{d-1}} \int_{\mathcal{B}_1^{d-1}} K_{2,r_j}(\mathbf{w}, \mathbf{z}) d\mathbf{w} d\mathbf{z} = \int_0^{r_j} x^2 \Psi_0(x) dx,$$

and hence (6.46) follows from the asymptotics (5.11). Relation (6.45) is a consequence of Proposition 6.9.  $\square$

**Proposition 6.11.** For  $d \geq 2$ ,

$$(6.48) \quad \mathbf{E}(a_j^4) \leq \mathbf{E}(b_j^4) = O(r_j^2).$$

*Proof.* We have

$$(6.49) \quad \begin{aligned} \mathbf{E}(b_j^4) &= \frac{1}{v_{d-1}} \int_{\mathcal{B}_1^{d-1}} \int_{\mathcal{B}_1^{d-1}} \frac{K_{2,r_j}(\mathbf{w}, \mathbf{z})^2}{K_0(\mathbf{w}, \mathbf{z})} d\mathbf{w} d\mathbf{z} \\ &\leq 2 \int_0^{r_j} x^4 \Psi_0(x) dx, \end{aligned}$$

and the claim follows from (5.11).  $\square$

**Proposition 6.12.** Let  $d = 2$ . For  $u \rightarrow \infty$ ,

$$(6.50) \quad \mathbf{P}(\alpha_n > u) \sim \frac{1}{2\pi^2 u^2}.$$

$$(6.51) \quad \mathbf{P}(\beta_n > u) \sim \frac{1}{2\pi^2 u^2}.$$

*Proof.* We follow the same strategy as in the proof of Proposition 6.5. For

$$(6.52) \quad f(y) = \frac{6y^2(1+y)^2}{1+2y} \ln(1+y^{-1}),$$

the variable substitution (6.31) yields

$$(6.53) \quad \begin{aligned} \mathbf{P}(\beta_n > u) &= \frac{12}{\pi^2} \int_0^\infty \int_{\frac{2}{1+2y}}^{\frac{2}{1+y}} \mathbb{1}_{\{x^2 f(y) < u^{-2}\}} \ln(1+y^{-1}) dx dy \\ &= \frac{12}{\pi^2} \int_0^\infty \mathbb{1}_{\{\frac{4f(y)}{(1+y)^2} < u^{-2}\}} \left( \frac{2}{1+y} - \frac{2}{1+2y} \right) \ln(1+y^{-1}) dy \\ &\quad + \frac{12}{\pi^2} \int_0^\infty \mathbb{1}_{\{\frac{4f(y)}{(1+2y)^2} < u^{-2} < \frac{4f(y)}{(1+y)^2}\}} \left( \frac{1}{uf(y)} - \frac{2}{1+2y} \right) \ln(1+y^{-1}) dy. \end{aligned}$$

The leading order contribution comes from the first term, which evaluates to

$$(6.54) \quad \begin{aligned} \frac{12}{\pi^2} \int_0^\infty \mathbb{1}_{\{\frac{24y^2}{1+2y} \ln(1+y^{-1}) < u^{-2}\}} (2y + O(y^2)) \ln(1+y^{-1}) dy \\ = \frac{12}{\pi^2} (y^2 \ln(1+y^{-1}) + O(y^2)) \Big|_{0+}^{\frac{24y^2}{1+2y} \ln(1+y^{-1}) = u^{-2}} = \frac{1}{2\pi^2 u^2} (1 + o(1)). \end{aligned}$$

This proves (6.51). To see that (6.50) has the same asymptotics, recall that  $\beta_n^2 - \alpha_n^2 = \mu_n^2$  and (6.27).  $\square$



**Proposition 6.13.** *Let  $d \geq 3$ . There are constants  $c_2 > c_1 > 0$ , such that for  $u \geq 1$ ,*

$$(6.55) \quad c_1 u^{-2} \leq \mathbf{P}(\alpha_n > u) \leq \mathbf{P}(\beta_n > u) \leq c_2 u^{-2}.$$

*Proof.* We exploit the bounds in Lemma 6.4. For the upper bound,

$$(6.56) \quad \int_{K_2 > u^2 K_0} K_0 d\mathbf{w} d\mathbf{z} \ll \int_{\substack{\max(t^{-\frac{d-2}{2}}, t^{-\frac{d-1}{2}} \varphi) \gg u^2 \\ 0 \leq \varphi \leq \frac{\pi}{2}}} d\mathbf{w} d\mathbf{z} + \int_{\substack{\min(t^{-(1+\frac{d}{2})}, t^{-\frac{d+1}{2} + \frac{1}{d}(\pi-\varphi)^{-(1+\frac{2}{d})}) \gg u^2 \\ \frac{\pi}{2} \leq \varphi \leq \pi}} d\mathbf{w} d\mathbf{z}.$$

Set

$$(6.57) \quad \alpha := \frac{4}{d+2}, \quad \beta := \frac{4d}{d(d+1)-2}, \quad \gamma := \frac{2d+4}{d(d+1)-2},$$

and note that

$$(6.58) \quad \frac{\alpha - \beta}{\gamma} = -\frac{2}{d+2}.$$

Using polar co-ordinates as before, the second term in (6.56) evaluates to

$$(6.59) \quad \begin{aligned} & \int_0^{\pi/2} \min(u^{-\alpha}, u^{-\beta} \phi^{-\gamma})^2 (\sin \phi)^{d-3} d\phi \\ &= u^{-2\alpha} \int_0^{u^{\frac{\alpha-\beta}{\gamma}}} (\sin \phi)^{d-3} d\phi + u^{-2\beta} \int_{u^{\frac{\alpha-\beta}{\gamma}}}^{\pi/2} \phi^{-2\gamma} (\sin \phi)^{d-3} d\phi \\ &\ll u^{-2\alpha} u^{(d-2)\frac{\alpha-\beta}{\gamma}} + u^{-2\beta} u^{(d-2-2\gamma)\frac{\alpha-\beta}{\gamma}} = 2u^{-2}. \end{aligned}$$

A similar calculation shows that the first term in (6.56) produces a lower order contribution. This establishes the upper bound in (6.55). For the lower bound for  $\mathbf{P}(\beta_n > u)$ ,

$$(6.60) \quad \begin{aligned} \int_{K_2 > u^2 K_0} K_0 d\mathbf{w} d\mathbf{z} &\gg \int_{\substack{\min(t^{-(1+\frac{d}{2})}, (\pi-\varphi)^{-(d+2)}) \gg u^2 \\ \frac{\pi}{2} \leq \varphi \leq \pi}} d\mathbf{w} d\mathbf{z} \\ &\gg u^{-2\alpha} \int_0^{u^{-\frac{2}{d+2}}} (\sin \phi)^{d-3} d\phi \\ &\gg u^{-2\alpha} u^{-2\frac{d-2}{d+2}} = u^{-2}. \end{aligned}$$

The lower bound for  $\mathbf{P}(\alpha_n > u)$  follows by combining the lower bound for  $\mathbf{P}(\beta_n > u)$  with (6.35).  $\square$

## 7. SPECTRAL GAPS

Let  $V$  be a finite-dimensional real vector space with inner product  $\langle \cdot, \cdot \rangle_V$  and norm  $\|x\| := \langle x, x \rangle^{1/2}$ . Denote by  $\mathcal{H} = L^2(\mathcal{B}_1^{d-1}, V, v_{d-1}^{-1} d\mathbf{w})$  the Hilbert space of square-integrable functions  $\mathcal{B}_1^{d-1} \rightarrow V$  with inner product

$$(7.1) \quad \langle f, g \rangle := \frac{1}{v_{d-1}} \int_{\mathcal{B}_1^{d-1}} \langle f(\mathbf{w}), g(\mathbf{w}) \rangle_V d\mathbf{w},$$

and norm  $\|f\| := \langle f, f \rangle^{1/2}$ . We will also denote by  $\|\cdot\|$  the corresponding operator norm on  $\mathcal{H} \rightarrow \mathcal{H}$ . In the following,  $(\rho, V)$  will denote a representation of  $\text{SO}(d)$  with group homomorphism  $\rho : \text{SO}(d) \rightarrow \text{O}(V)$ .

Define the following operators on  $\mathcal{H}$ :

$$(7.2) \quad Pf(\mathbf{w}) := \int_{\mathcal{B}_1^{d-1}} K_0(\mathbf{w}, \mathbf{z}) f(\mathbf{z}) d\mathbf{z},$$

$$(7.3) \quad \Pi f(\mathbf{w}) := \frac{1}{v_{d-1}} \int_{\mathcal{B}_1^{d-1}} f(\mathbf{z}) d\mathbf{z},$$

$$(7.4) \quad Uf(\mathbf{w}) := \rho(S(\mathbf{w}))f(\mathbf{w}).$$

We have

$$(7.5) \quad \Pi P = P\Pi = \Pi.$$

Denote by  $\mathcal{H}_0 := \Pi\mathcal{H}$  the subspace of constant functions, and by  $\mathcal{H}_1 = (I - \Pi)\mathcal{H}$  its orthogonal complement. (This means that all components of  $f \in \mathcal{H}_1$  have zero mean.) Note that for  $f \in \mathcal{H}_0$  we have  $Pf = f$ , and for  $f \in \mathcal{H}_1$  we have  $Pf \in \mathcal{H}_1$ .

**Proposition 7.1.** *The operator  $P$  has the spectral gap  $1 - \omega_0$  with*

$$(7.6) \quad \omega_0 := \|P - \Pi\| \leq 1 - \frac{1}{2^d \zeta(d)}.$$

*Proof.* This follows from the standard Doeblin argument. Note that, since  $K_0(\mathbf{w}, \mathbf{z})$  is the kernel of a stochastic transition operator with respect to  $d\mathbf{w}$  on  $\mathcal{B}_1^{d-1}$ , we have

$$(7.7) \quad J := v_{d-1} \inf K_0(\mathbf{w}, \mathbf{z}) \leq 1.$$

If  $J = 1$ , we have  $P = \Pi$  and thus  $\omega_0 = 0$ . Assume therefore  $0 \leq J < 1$ . Then

$$(7.8) \quad Q := (1 - J)^{-1}(P - J\Pi)$$

is itself a stochastic transition operator, with the same stationary measure. Using (7.5), we can write

$$(7.9) \quad P = \Pi + (1 - J)(I - \Pi)Q(I - \Pi),$$

and so

$$(7.10) \quad \|P - \Pi\| \leq (1 - J)\|Q\| = 1 - J.$$

The claim of the proposition now follows from (6.10).  $\square$

**Lemma 7.2.** *Let  $\theta : [0, 1) \rightarrow \mathbb{R}$  be measurable, so that*

$$(7.11) \quad \text{meas}\{w \in [0, 1) : \theta(w) \notin \mathbb{Q}\} > 0,$$

*and let  $(\rho, V)$  be a non-trivial irreducible representation of  $\text{SO}(d)$ . Then*

$$(7.12) \quad \delta_\rho := \left\| \frac{1}{v_{d-1}} \int_{\mathcal{B}_1^{d-1}} \rho(S(\mathbf{w})) d\mathbf{w} \right\| < 1.$$

*Proof.* For any fixed  $\mathbf{e} \in \mathcal{S}_1^{d-1}$ , let  $\Lambda_{\mathbf{e}}$  be the push-forward of Lebesgue measure on  $[0, 1)$  under the map

$$(7.13) \quad [0, 1) \rightarrow \text{SO}(d), \quad w \mapsto S(w\mathbf{e}).$$

The group generated by the support of  $\Lambda_{\mathbf{e}}$  is, by assumption (7.11), dense in the subgroup

$$(7.14) \quad \{E(\phi\mathbf{e}) : \phi \in [0, 2\pi)\} \simeq \text{SO}(2),$$

with  $E(\mathbf{x})$  as in (2.6). Next, let  $\Lambda$  be the push-forward of Lebesgue measure on  $\mathcal{B}_1^{d-1}$  under the map

$$(7.15) \quad \mathcal{B}_1^{d-1} \rightarrow \text{SO}(d), \quad \mathbf{w} \mapsto S(\mathbf{w}).$$

The above observation, together with the fact that

$$(7.16) \quad \{E(\mathbf{x}) : \mathbf{x} \in \mathcal{B}_{2\pi}^{d-1}\}$$

generates  $\text{SO}(d)$ , implies that the group generated by the support of  $\Lambda$  is dense in  $\text{SO}(d)$ . The claim now follows from well known arguments [25].  $\square$

**Proposition 7.3.** *Let  $\theta$  and  $(\rho, V)$  be as in Lemma 7.2. Then the operator  $PU$  has spectral radius*

$$(7.17) \quad \omega_\rho := \lim_{n \rightarrow \infty} \|(PU)^n\|^{1/n} < 1.$$

*Proof.* It is sufficient to prove  $\|PUP\| < 1$ , that is

$$(7.18) \quad \sup_{f \in \mathcal{H}: f \neq 0} \frac{\|PUPf\|^2}{\|f\|^2} < 1.$$

We may restrict to functions of the form  $f = \alpha f_0 + f_1$ , where  $\alpha > 0$  and  $f_0 \in \mathcal{H}_0$ ,  $f_1 \in \mathcal{H}_1$  with  $\|f_0\| = \|f_1\| = 1$ . Note that in this case  $\|f\|^2 = \alpha^2 + 1$ , and hence the supremum (7.18) equals

$$(7.19) \quad \sup_{\alpha > 0} \frac{1}{\alpha^2 + 1} \sup_{f_0, f_1} \|PUP(\alpha f_0 + f_1)\|^2.$$

Now,

$$(7.20) \quad \begin{aligned} \sup_{f_1} \|PUP(\alpha f_0 + f_1)\|^2 &= \sup_{f_1} \|PU(\alpha f_0 + Pf_1)\|^2 \\ &\leq \sup_{f_1} \|PU(\alpha f_0 + \omega_0 f_1)\|^2 \end{aligned}$$

since

$$(7.21) \quad \{Pf_1 : f_1 \in \mathcal{H}_1, \|f_1\| = 1\} \subseteq \{f_1 \in \mathcal{H}_1, \|f_1\| \leq \omega_0\}.$$

We have

$$(7.22) \quad U(\alpha f_0 + \omega_0 f_1) = \alpha y_0 + \omega_0 y_1 + \alpha \tilde{y}_0 + \omega_0 \tilde{y}_1,$$

where

$$(7.23) \quad y_0 := \Pi U f_0 \in \mathcal{H}_0, \quad \tilde{y}_0 := (I - \Pi) U f_0 \in \mathcal{H}_1,$$

$$(7.24) \quad y_1 := \Pi U f_1 \in \mathcal{H}_0, \quad \tilde{y}_1 := (I - \Pi) U f_1 \in \mathcal{H}_1.$$

Therefore,

$$(7.25) \quad \begin{aligned} &\|PU(\alpha f_0 + \omega_0 f_1)\|^2 \\ &= \|\alpha y_0 + \omega_0 y_1\|^2 + \|P(\alpha \tilde{y}_0 + \omega_0 \tilde{y}_1)\|^2 \\ &\leq \|\alpha y_0 + \omega_0 y_1\|^2 + \omega_0^2 \|\alpha \tilde{y}_0 + \omega_0 \tilde{y}_1\|^2 \\ &= \alpha^2 (\omega_0^2 + (1 - \omega_0^2) \|y_0\|^2) + 2\alpha \omega_0 (1 - \omega_0^2) \langle y_0, y_1 \rangle + \omega_0^2 (\omega_0^2 + (1 - \omega_0^2) \|y_1\|^2). \end{aligned}$$

In the last equality we have used the relations

$$(7.26) \quad \langle \tilde{y}_0, \tilde{y}_1 \rangle = -\langle y_0, y_1 \rangle,$$

which follows from  $\langle U f_0, U f_1 \rangle = \langle f_0, f_1 \rangle = 0$ , and

$$(7.27) \quad \|\tilde{y}_0\|^2 = 1 - \|y_0\|^2, \quad \|\tilde{y}_1\|^2 = 1 - \|y_1\|^2.$$

Since  $f_0$  is a constant function with  $\|f_0\| = 1$ , we have by Lemma 7.2  $\|y_0\| \leq \delta_\rho < 1$ . Furthermore  $\|y_1\| \leq 1$ . This shows

$$(7.28) \quad \sup_{f \in \mathcal{H}: f \neq 0} \frac{\|PUPf\|^2}{\|f\|^2} \leq \sup_{\alpha > 0} \sup_{\substack{\|y_0\| \leq \delta_\rho \\ \|y_1\| \leq 1}} E(\alpha),$$

with

$$(7.29) \quad E(\alpha) := \frac{\alpha^2 (\omega_0^2 + (1 - \omega_0^2) \|y_0\|^2) + 2\alpha \omega_0 (1 - \omega_0^2) \langle y_0, y_1 \rangle + \omega_0^2 (\omega_0^2 + (1 - \omega_0^2) \|y_1\|^2)}{\alpha^2 + 1}.$$

The final step in the proof of Proposition 7.3 is now to show that

$$(7.30) \quad \sup_{\alpha > 0} \sup_{\substack{\|y_0\| \leq \delta_\rho \\ \|y_1\| \leq 1}} E(\alpha) < 1.$$

To achieve this, first note that

$$(7.31) \quad \sup_{\substack{\|y_0\| \leq \delta_\rho \\ \|y_1\| \leq 1}} E(0) = \sup_{\|y_1\| \leq 1} \omega_0^2(\omega_0^2 + (1 - \omega_0^2)\|y_1\|^2) \leq \omega_0^2 < 1$$

and

$$(7.32) \quad \lim_{\alpha \rightarrow \infty} \sup_{\substack{\|y_0\| \leq \delta_\rho \\ \|y_1\| \leq 1}} E(\alpha) = \sup_{\|y_0\| \leq \delta_\rho} \omega_0^2 + (1 - \omega_0^2)\|y_0\|^2 < 1.$$

To prove, (7.30), it is therefore sufficient that the quadratic equation

$$(7.33) \quad E(\alpha) = 1$$

has no positive real solution for all  $\|y_0\| \leq \delta_\rho$ ,  $\|y_1\| \leq 1$ . This in turn holds, if the discriminant of Eq. (7.33) is strictly negative, i.e.

$$(7.34) \quad \sup_{\substack{\|y_0\| \leq \delta_\rho \\ \|y_1\| \leq 1}} [-4(1 - \omega_0^2)^2 \{1 - \|y_0\|^2 + \omega_0^2[(1 - \|y_0\|^2)(1 - \|y_1\|^2) - \langle y_0, y_1 \rangle^2]\}] < 0.$$

Because  $1 - \omega_0^2 > 0$  and  $1 - \|y_0\|^2 \geq 1 - \delta_\rho^2 > 0$ , it remains to be shown that

$$(7.35) \quad \langle y_0, y_1 \rangle^2 \leq (1 - \|y_0\|^2)(1 - \|y_1\|^2).$$

To this end, apply eqs. (7.26), (7.27) and the Cauchy-Schwarz inequality,

$$(7.36) \quad |\langle y_0, y_1 \rangle| = |\langle \tilde{y}_0, \tilde{y}_1 \rangle| \leq \|\tilde{y}_0\| \|\tilde{y}_1\| = \sqrt{(1 - \|y_0\|^2)} \sqrt{(1 - \|y_1\|^2)},$$

which completes the proof.  $\square$

## 8. EXPONENTIAL MIXING

We will now apply the spectral estimates of the previous section to obtain exponential mixing rates.

Denote by  $(\rho_1, V_1) = (\text{id}, \mathbb{R}^d)$  the natural representation of  $\text{SO}(d)$ , and by  $(\rho_2, V_2)$  the adjoint representation of  $\text{SO}(d)$  on the vector space  $V_2$  of real symmetric traceless  $d \times d$  matrices defined by

$$(8.1) \quad \rho_2(R) : M \mapsto RM {}^tR.$$

The inner product on  $V_1$  is the standard Euclidean inner product

$$(8.2) \quad \langle \mathbf{x}_1, \mathbf{x}_2 \rangle_{V_1} := \mathbf{x}_1 \cdot \mathbf{x}_2,$$

and on  $V_2$  the Hilbert-Schmidt inner product

$$(8.3) \quad \langle M_1, M_2 \rangle_{V_2} := \text{tr}(M_1 M_2).$$

**Proposition 8.1.** *Fix any  $\omega \in [\omega_0, 1) \cap (\omega_{\rho_1}, 1) \cap (\omega_{\rho_2}, 1)$ ,  $m \in \mathbb{N}$  and  $p = 0, 1, 2$ . Then there is a constant  $C_m > 0$  such that, for all  $n_1, n_2 \in \mathbb{N}$ ,  $\mathbf{v}_0, \mathbf{e} \in \mathbb{S}_1^{d-1}$  and all measurable  $f, g : (\mathcal{B}_1^{d-1})^{m+1} \rightarrow \mathbb{R}$  with*

$$(8.4) \quad \mathbf{E}(f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m)^2) < \infty, \quad \mathbf{E}(g(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m)^2) < \infty,$$

we have

$$(8.5) \quad \left| \text{Cov}((\mathbf{e} \cdot \mathbf{V}_{n_1})^p f(\boldsymbol{\eta}_{n_1}, \dots, \boldsymbol{\eta}_{n_1+m}), (\mathbf{e} \cdot \mathbf{V}_{n_2})^p g(\boldsymbol{\eta}_{n_2}, \dots, \boldsymbol{\eta}_{n_2+m})) \right| \\ \leq C_m \omega^{|n_1 - n_2| - m} \sqrt{\mathbf{E}(f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m)^2)} \sqrt{\mathbf{E}(g(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m)^2)}.$$

*Proof.* Assume without loss of generality that  $n_1 \leq n_2$ . We have

$$(8.6) \quad \mathbf{Cov} \left( (e \cdot \mathbf{V}_{n_1})^p f(\boldsymbol{\eta}_{n_1}, \dots, \boldsymbol{\eta}_{n_1+m}), (e \cdot \mathbf{V}_{n_2})^p g(\boldsymbol{\eta}_{n_2}, \dots, \boldsymbol{\eta}_{n_2+m}) \right) \\ = \mathbf{E} \left[ \mathbf{Cov} \left( (e \cdot \mathbf{V}_{n_1})^p f(\boldsymbol{\eta}_{n_1}, \dots, \boldsymbol{\eta}_{n_1+m}), (e \cdot \mathbf{V}_{n_2})^p g(\boldsymbol{\eta}_{n_2}, \dots, \boldsymbol{\eta}_{n_2+m}) \mid \mathbf{V}_{n_1} \right) \right].$$

It is therefore sufficient to prove (8.5) conditioned on  $\mathbf{V}_{n_1}$  with a constant  $C_m$  independent on  $\mathbf{V}_{n_1}$ , or equivalently, to show that for any  $n \in \mathbb{Z}_{\geq 0}$ ,  $\mathbf{v}_0, \mathbf{e} \in \mathbb{S}_1^{d-1}$ ,

$$(8.7) \quad \left| \mathbf{Cov} \left( f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m), (e \cdot \mathbf{V}_n)^p g(\boldsymbol{\eta}_n, \dots, \boldsymbol{\eta}_{n+m}) \right) \right| \\ \leq C_m \omega^{n-m} \sqrt{\mathbf{E} \left( f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m)^2 \right)} \sqrt{\mathbf{E} \left( g(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m)^2 \right)}.$$

The case  $n \leq m$  follows from the Cauchy-Schwarz inequality. We assume therefore in the following that  $n > m$ .

*Case A:*  $p = 0$ . Define the functions  $\mathcal{B}_1^{d-1} \rightarrow \mathbb{R}$

$$(8.8) \quad \tilde{f}(\mathbf{w}) := \mathbf{E} \left( f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m) \mid \boldsymbol{\eta}_m = \mathbf{w} \right),$$

$$(8.9) \quad \tilde{g}(\mathbf{w}) := \mathbf{E} \left( g(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m) \mid \boldsymbol{\eta}_0 = \mathbf{w} \right).$$

Then

$$(8.10) \quad \mathbf{Cov} \left( f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m), g(\boldsymbol{\eta}_n, \dots, \boldsymbol{\eta}_{n+m}) \right) = \mathbf{Cov} \left( \tilde{f}(\boldsymbol{\eta}_m), \tilde{g}(\boldsymbol{\eta}_n) \right) \\ = \langle (I - \Pi) \tilde{f}, (P^{n-m} - \Pi) \tilde{g} \rangle \\ = \langle (I - \Pi) \tilde{f}, (P - \Pi)^{n-m} \tilde{g} \rangle,$$

and so

$$(8.11) \quad \left| \mathbf{Cov} \left( f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m), g(\boldsymbol{\eta}_n, \dots, \boldsymbol{\eta}_{n+m}) \right) \right| \leq \|\tilde{f}\| \|(P - \Pi)^{n-m} \tilde{g}\| \\ \leq \omega_0^{n-m} \|\tilde{f}\| \|\tilde{g}\|$$

in view of Proposition 7.1. Finally,

$$(8.12) \quad \|\tilde{f}\|^2 \leq \mathbf{E} \left( f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m)^2 \right),$$

since  $\tilde{f}$  is obtained from  $f$  via orthogonal projection, thus decreasing the L<sup>2</sup>-norm (likewise for  $\tilde{g}$ ). This proves (8.7) for  $p = 0$ .

*Case B:*  $p = 1$ . Set  $\tilde{\mathbf{e}} = R(\mathbf{v}_0)^{-1} \mathbf{e}$ , and

$$(8.13) \quad \bar{f} := \mathbf{E} f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m).$$

Now

$$(8.14) \quad \mathbf{Cov} \left( f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m), (e \cdot \mathbf{V}_n) g(\boldsymbol{\eta}_n, \dots, \boldsymbol{\eta}_{n+m}) \right) \\ = \mathbf{Cov} \left( f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m), (\tilde{\mathbf{e}} \cdot S(\boldsymbol{\eta}_1) \cdots S(\boldsymbol{\eta}_n) \mathbf{e}_1) g(\boldsymbol{\eta}_n, \dots, \boldsymbol{\eta}_{n+m}) \right) \\ = \mathbf{E} \left( [f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m) - \bar{f}] (\tilde{\mathbf{e}} \cdot S(\boldsymbol{\eta}_1) \cdots S(\boldsymbol{\eta}_n) \mathbf{e}_1) g(\boldsymbol{\eta}_n, \dots, \boldsymbol{\eta}_{n+m}) \right).$$

By using the vector-valued functions  $\mathcal{B}_1^{d-1} \rightarrow \mathbb{R}^d$

$$(8.15) \quad \tilde{f}(\mathbf{w}) := \mathbf{E} \left( [f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m) - \bar{f}] {}^t S(\boldsymbol{\eta}_m) \cdots {}^t S(\boldsymbol{\eta}_1) \mid \boldsymbol{\eta}_m = \mathbf{w} \right) \tilde{\mathbf{e}},$$

$$(8.16) \quad \tilde{g}(\mathbf{w}) := \mathbf{E} \left( g(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m) \mid \boldsymbol{\eta}_0 = \mathbf{w} \right) \mathbf{e}_1,$$

we find

$$(8.17) \quad \mathbf{Cov} \left( f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m), (e \cdot \mathbf{V}_n) g(\boldsymbol{\eta}_n, \dots, \boldsymbol{\eta}_{n+m}) \right) \\ = \mathbf{E} \left( \tilde{f}(\boldsymbol{\eta}_m) \cdot S(\boldsymbol{\eta}_{m+1}) \cdots S(\boldsymbol{\eta}_n) \tilde{g}(\boldsymbol{\eta}_n) \right) \\ = \langle \tilde{f}, (PU)^{n-m} \tilde{g} \rangle.$$

We conclude from Proposition 7.3 applied to the natural representation  $(\rho_1, V_1)$ :

$$(8.18) \quad \left| \mathbf{Cov} \left( f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m), (e \cdot \mathbf{V}_n) g(\boldsymbol{\eta}_n, \dots, \boldsymbol{\eta}_{n+m}) \right) \right| \leq c \omega^{n-m} \|\tilde{f}\| \|\tilde{g}\|$$

for some  $c > 0$ . Finally, because orthogonal projection decreases the  $L^2$ -norm,

$$\begin{aligned}
\|\tilde{f}\|^2 &\leq \mathbf{E} (\| [f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m) - \bar{f}] {}^tS(\boldsymbol{\eta}_m) \cdots {}^tS(\boldsymbol{\eta}_1) \tilde{\mathbf{e}} \|_{V_1}^2) \\
(8.19) \quad &= \mathbf{E} ([f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m) - \bar{f}]^2) \\
&\leq \mathbf{E} (f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m)^2).
\end{aligned}$$

*Case C:*  $p = 2$ . The vector  $\tilde{\mathbf{e}}$  and the expectation  $\bar{f}$  are defined as in Case B. We have

$$\begin{aligned}
(8.20) \quad &\mathbf{Cov} (f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m), (\mathbf{e} \cdot \mathbf{V}_n)^2 g(\boldsymbol{\eta}_n, \dots, \boldsymbol{\eta}_{n+m})) \\
&= \mathbf{Cov} (f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m), (\tilde{\mathbf{e}} \cdot S(\boldsymbol{\eta}_1) \cdots S(\boldsymbol{\eta}_n) \mathbf{e}_1)^2 g(\boldsymbol{\eta}_n, \dots, \boldsymbol{\eta}_{n+m})) \\
&= \mathbf{E} ([f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m) - \bar{f}] (\tilde{\mathbf{e}} \cdot S(\boldsymbol{\eta}_1) \cdots S(\boldsymbol{\eta}_n) \mathbf{e}_1)^2 g(\boldsymbol{\eta}_n, \dots, \boldsymbol{\eta}_{n+m})).
\end{aligned}$$

We write

$$\begin{aligned}
(8.21) \quad &(\tilde{\mathbf{e}} \cdot S(\boldsymbol{\eta}_1) \cdots S(\boldsymbol{\eta}_n) \mathbf{e}_1)^2 = \frac{1}{d} + \text{tr} [\tilde{E} S(\boldsymbol{\eta}_1) \cdots S(\boldsymbol{\eta}_n) E_1 {}^t(S(\boldsymbol{\eta}_1) \cdots S(\boldsymbol{\eta}_n))] \\
&= \frac{1}{d} + \langle \tilde{E}, \rho(S(\boldsymbol{\eta}_1) \cdots S(\boldsymbol{\eta}_n)) E_1 \rangle_{V_2},
\end{aligned}$$

where  $V_2$  is the vector space of symmetric traceless  $d \times d$  matrices, and

$$(8.22) \quad E_1 := \mathbf{e}_1 \otimes \mathbf{e}_1 - \frac{1}{d} I_d \in V_2, \quad \tilde{E} := \tilde{\mathbf{e}} \otimes \tilde{\mathbf{e}} - \frac{1}{d} I_d \in V_2.$$

The constant term  $\frac{1}{d}$  in (8.21) contributes to (8.20) the term

$$\begin{aligned}
(8.23) \quad &\frac{1}{d} |\mathbf{E} ([f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m) - \bar{f}] g(\boldsymbol{\eta}_n, \dots, \boldsymbol{\eta}_{n+m}))| \\
&\leq \frac{1}{d} \omega_0^{n-m} \sqrt{\mathbf{E} (f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m)^2)} \sqrt{\mathbf{E} (g(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m)^2)}.
\end{aligned}$$

This follows from our discussion in Case A ( $p = 0$ ). The non-constant term in (8.21) is handled in analogy with Case B. Define functions  $\mathcal{B}_1^{d-1} \rightarrow V_2$

$$(8.24) \quad \tilde{f}(\mathbf{w}) := \mathbf{E} ([f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m) - \bar{f}] \rho({}^tS(\boldsymbol{\eta}_m) \cdots {}^tS(\boldsymbol{\eta}_1)) \mid \boldsymbol{\eta}_m = \mathbf{w}) \tilde{E},$$

$$(8.25) \quad \tilde{g}(\mathbf{w}) := \mathbf{E} (g(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m) \mid \boldsymbol{\eta}_0 = \mathbf{w}) E_1,$$

so that the non-constant contribution to (8.20) becomes

$$\begin{aligned}
(8.26) \quad &\mathbf{E} ([f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m) - \bar{f}] (\tilde{\mathbf{e}} \cdot S(\boldsymbol{\eta}_1) \cdots S(\boldsymbol{\eta}_n) \mathbf{e}_1)^2 g(\boldsymbol{\eta}_n, \dots, \boldsymbol{\eta}_{n+m})) \\
&= \mathbf{E} ([f(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m) - \bar{f}] \langle \tilde{E}, \rho(S(\boldsymbol{\eta}_1) \cdots S(\boldsymbol{\eta}_n)) E_1 \rangle_{V_2} g(\boldsymbol{\eta}_n, \dots, \boldsymbol{\eta}_{n+m})) \\
&= \langle \tilde{f}, (PU)^{n-m} \tilde{g} \rangle.
\end{aligned}$$

We now apply Proposition 7.3 with the adjoint representation  $(\rho_2, V_2)$ . This yields the desired bound.  $\square$

We will also require the following estimate.

**Proposition 8.2.** *Fix  $\omega$  and  $m$  as in Proposition 8.1. Then there is a constant  $\tilde{C} > 0$  such that, for all  $n \in \mathbb{N}$ ,  $\mathbf{v}_0, \mathbf{e} \in \mathcal{S}_1^{d-1}$  and all measurable  $g : (\mathcal{B}_1^{d-1})^{m+1} \rightarrow \mathbb{R}$  with*

$$(8.27) \quad \mathbf{E} (g(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m)^2) < \infty,$$

we have

$$(8.28) \quad |\mathbf{E} ((\mathbf{e} \cdot \mathbf{V}_n) g(\boldsymbol{\eta}_n, \dots, \boldsymbol{\eta}_{n+m}))| \leq \tilde{C} \omega^n \sqrt{\mathbf{E} (g(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m)^2)}.$$

$$(8.29) \quad |\mathbf{E} ((\mathbf{e} \cdot \mathbf{V}_n)^2 - d^{-1}) g(\boldsymbol{\eta}_n, \dots, \boldsymbol{\eta}_{n+m})| \leq \tilde{C} \omega^n \sqrt{\mathbf{E} (g(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_m)^2)}.$$

*Proof.* With  $\tilde{g}$  defined as in the previous proof, Case B, we have

$$(8.30) \quad \mathbf{E}((\mathbf{e} \cdot \mathbf{V}_n)g(\boldsymbol{\eta}_n, \dots, \boldsymbol{\eta}_{n+m})) = \langle 1, (PU)^n \tilde{g} \rangle.$$

The bound (8.28) now follows from Proposition 7.3. Relation (8.29) follows similarly from Case C of the previous proof.  $\square$

## 9. PROOF OF THE MAIN LEMMAS

We now turn to the proofs of the four main lemmas in Section 4.

*Proof of Lemma 4.1.* We have

$$(9.1) \quad \|\mathbf{Q}_n - \mathbf{Q}'_n\| \leq \sum_{j=1}^n \zeta_j$$

with

$$(9.2) \quad \zeta_j := \xi_j \mathbb{1}_{\{\xi_j^2 > j(\log j)^\gamma\}}.$$

By (5.11), we have

$$(9.3) \quad \mathbf{P}(\zeta_j \neq 0) = O(j^{-1}(\log j)^{-\gamma})$$

This is summable (since  $\gamma > 1$ ) and so, by the Borel-Cantelli lemma,  $\zeta_j \neq 0$  only for finitely many  $j$ . This proves Lemma 4.1.  $\square$

*Proof of Lemma 4.2.* Set  $\zeta_j := (\mathbf{e} \cdot \mathbf{V}_{j-1})m_j$ . We need to show that, for every  $\mathbf{e} \in S_1^{d-1}$ , the sequence of random variables

$$(9.4) \quad \begin{cases} \frac{\sum_{j=1}^n \zeta_j}{\sqrt{n \log \log n}} & (d = 2) \\ \frac{\sum_{j=1}^n \zeta_j}{\sqrt{n}} & (d \geq 3) \end{cases}$$

is tight. Now choose in Proposition 8.2  $m = 1$  and  $g(\mathbf{w}, \mathbf{z}) = K_{1,r_j}(\mathbf{w}, \mathbf{z})/K_0(\mathbf{w}, \mathbf{z})$ , and use Proposition 6.9 to bound  $\mathbf{E}(g(\boldsymbol{\eta}_0, \boldsymbol{\eta}_1)^2) = \mathbf{E}(m_j^2) = O(\log \log j)$  for  $d = 2$ , and  $= O(1)$  for  $d \geq 3$ . This proves  $\mathbf{E}(\zeta_j) = O(\omega^j)$ . Therefore

$$(9.5) \quad \mathbf{Var}(\zeta_j^2) = \mathbf{E}(\zeta_j^2) + O(\omega^{2j}).$$

Proposition 6.9 yields

$$(9.6) \quad \mathbf{E}(\zeta_j^2) \leq \mathbf{E}(m_j^2) = \begin{cases} O(\log \log j) & (d = 2) \\ O(1) & (d \geq 3). \end{cases}$$

Due to Proposition 8.1, we also have

$$(9.7) \quad \mathbf{Cov}(\zeta_i, \zeta_j) \leq \sqrt{\mathbf{E}(\zeta_i^2)} \sqrt{\mathbf{E}(\zeta_j^2)} \omega^{|i-j|}.$$

Hence

$$(9.8) \quad \mathbf{E} \left[ \left( \sum_{j=1}^n \zeta_j \right)^2 \right] = \begin{cases} O(n \log \log n) & (d = 2) \\ O(n) & (d \geq 3), \end{cases}$$

which establishes the tightness of (9.4) and thus Lemma 4.2.  $\square$

*Proof of Lemma 4.3.* Let  $\zeta_j := (\mathbf{e} \cdot \mathbf{V}_{j-1})^2 a_j^2$ . It is sufficient to prove that for any unit vector  $\mathbf{e} \in S_1^{d-1}$

$$(9.9) \quad \frac{\sum_{j=1}^n \zeta_j}{n \log n} \xrightarrow{\mathbf{P}} \sigma_d^2.$$

Using (8.29) in Proposition 8.2, and Proposition 6.10,

$$(9.10) \quad \mathbf{E}(\zeta_j) = \frac{\Theta_d}{2d} \log j + O(\log \log j),$$

and, by Proposition 6.11,

$$(9.11) \quad \mathbf{E}(a_j^4) = O(j(\log j)^\gamma).$$

Furthermore, due to Proposition 8.1 ( $p = 2$ ), we have

$$(9.12) \quad \mathbf{Cov}(\zeta_i, \zeta_j) \leq \sqrt{\mathbf{E}(a_i^4)} \sqrt{\mathbf{E}(a_j^4)} \omega^{|i-j|}.$$

Hence,

$$(9.13) \quad \mathbf{E}\left(\sum_{j=1}^n \zeta_j\right) = \frac{\Theta_d}{2d} n \log n + O(n \log \log n)$$

and

$$(9.14) \quad \mathbf{Var}\left(\sum_{j=1}^n \zeta_j\right) = O(n^2(\log n)^\gamma) = o((n \log n)^2)$$

since  $\gamma < 2$ . This proves Lemma 4.3.  $\square$

*Proof of Lemma 4.4.* In view of the asymptotic relation for  $A_n$  in (4.19) we have to prove that for any  $\varepsilon > 0$

$$(9.15) \quad \frac{\sum_{j=1}^n \mathbf{E}(\tilde{\xi}_j^2 \mathbb{1}_{\{\tilde{\xi}_j^2 > \varepsilon^2 n \log n\}} \mid \underline{\eta})}{n \log n} \xrightarrow{\mathbf{P}} 0.$$

The lower tail  $\tilde{\xi}_j < -\varepsilon\sqrt{n \log n}$  is estimated by

$$(9.16) \quad \begin{aligned} \mathbf{E}((\xi'_j - m_j)^2 \mathbb{1}_{\{\xi'_j - m_j < -\varepsilon\sqrt{n \log n}\}} \mid \underline{\eta}) &\leq m_j^2 \mathbb{1}_{\{m_j > \varepsilon\sqrt{n \log n}\}} \\ &\leq m_j^2 \mathbb{1}_{\{m_j > \varepsilon\sqrt{j \log j}\}}. \end{aligned}$$

Proposition 6.7 yields

$$(9.17) \quad \mathbf{P}(m_j > \varepsilon\sqrt{j \log j}) = \begin{cases} O((\varepsilon^2 j (\log j)^2)^{-1}) & (d = 2) \\ O((\varepsilon^2 j \log j)^{-(\frac{1}{2} + \frac{d}{4})}) & (d \geq 3). \end{cases}$$

Since this is summable, we have, by the Borel-Cantelli lemma,

$$(9.18) \quad \frac{\sum_{j=1}^n \mathbf{E}(\tilde{\xi}_j^2 \mathbb{1}_{\{\tilde{\xi}_j < -\varepsilon\sqrt{n \log n}\}} \mid \underline{\eta})}{n \log n} \xrightarrow{\text{a.s.}} 0.$$

For the upper tail  $\tilde{\xi}_j > \varepsilon\sqrt{n \log n}$  we have

$$(9.19) \quad \begin{aligned} \zeta_{n,j} &:= \mathbf{E}((\xi'_j - m_j)^2 \mathbb{1}_{\{\xi'_j - m_j > \varepsilon\sqrt{n \log n}\}} \mid \underline{\eta}) \\ &\leq \mathbf{E}(\xi_j'^2 \mathbb{1}_{\{\xi'_j > \varepsilon\sqrt{n \log n}\}} \mid \underline{\eta}) \\ &= \mathbf{E}(\xi_j^2 \mathbb{1}_{\{\varepsilon^2 n \log n < \xi_j^2 \leq j(\log j)^\gamma\}} \mid \underline{\eta}). \end{aligned}$$

On the other hand, in view of (5.11), we have for  $n \rightarrow \infty$ ,

$$(9.20) \quad \mathbf{E}(\zeta_{n,j}) \ll \log \sqrt{\frac{n(\log n)^\gamma}{\varepsilon^2 n \log n}} \sim \frac{\gamma - 1}{2} \log \log n,$$

and therefore

$$(9.21) \quad \frac{\mathbf{E}(\sum_{j=1}^n \zeta_{n,j})}{n \log n} \rightarrow 0.$$

From (9.18) and (9.21), the assertion of Lemma 4.4 follows.  $\square$



10. GENERAL INITIAL DATA

Up to now we have assumed that  $(\xi_1, \boldsymbol{\eta}_1)$  has density  $\Psi_0(x, \mathbf{z})$ . We now extend the above results to more general initial data  $(\xi_1, \boldsymbol{\eta}_1)$ , where the only assumption is that the marginal distribution of  $\boldsymbol{\eta}_1$  is absolutely continuous with respect to Lebesgue measure on  $\mathcal{B}_1^{d-1}$ .

*Proof of Theorem 3.2 (ii) for general initial data.* Since

$$(10.1) \quad \frac{\xi_1 \mathbf{V}_0}{\sqrt{n \log n}} \xrightarrow{\mathbf{P}} \mathbf{0},$$

it is sufficient to show that

$$(10.2) \quad \frac{\sum_{j=2}^n \xi_j \mathbf{V}_{j-1}}{\sigma_d \sqrt{n \log n}} \Rightarrow \mathcal{N}(0, I_d)$$

where  $\boldsymbol{\eta}_1$  has (by assumption) an absolutely continuous distribution and  $\xi_1 = 0$ . By an obvious re-labelling, this is equivalent to showing that

$$(10.3) \quad \frac{\sum_{j=1}^n \xi_j \mathbf{V}_{j-1}}{\sigma_d \sqrt{n \log n}} \Rightarrow \mathcal{N}(0, I_d)$$

where  $\boldsymbol{\eta}_0$  has an absolutely continuous distribution. In view of the remarks following Eq. (4.1), the only difference from the proof of Theorem 3.2 is now that  $\boldsymbol{\eta}_0$  is distributed according to an absolutely continuous probability measure, rather than Lebesgue measure. Because *tightness*, *almost sure convergence* and *convergence in probability* continue to hold when passing from a measure to a measure which is absolutely continuous with respect to the first, the Lemmas in Section 4 remain valid also in the present setting. The proof of Theorem 3.2 for general initial data therefore follows from these lemmas in the same way as for the density  $\Psi_0(x, \mathbf{z})$ , as described at the end of Section 4.  $\square$

The following proposition shows that, if  $(\xi_1, \boldsymbol{\eta}_1)$  has density  $\Psi(x, \mathbf{z})$  (which appears in the continuous-time setting of the Boltzmann-Grad limit, Theorem 3.1 (i)), then the marginal distribution of  $\boldsymbol{\eta}_1$  is absolutely continuous. Let

$$(10.4) \quad \bar{\Psi}(\mathbf{z}) := \int_0^\infty \Psi(x, \mathbf{z}) dx$$

**Proposition 10.1.**

$$(10.5) \quad \bar{\Psi} \in L^1(\mathcal{B}_1^{d-1}, d\mathbf{z}).$$

*Proof.* The function  $\Psi(x, \mathbf{z})$  is continuous, and in view of Corollary 1.2 and Theorem 1.11 in [20], uniformly bounded. The latter theorem produces a precise asymptotics of  $\Psi(x, \mathbf{z})$ , which implies

$$(10.6) \quad \Psi(x, \mathbf{z}) = O(x^{-2+\frac{2}{d}}),$$

uniformly for all  $x > 0$ ,  $\mathbf{z} \in \mathcal{B}_1^{d-1}$ . Thus, for  $d \geq 3$ ,  $\bar{\Psi}(\mathbf{z})$  is uniformly bounded, and hence  $\bar{\Psi} \in L^1(\mathcal{B}_1^{d-1}, d\mathbf{z})$  as required.

In dimension  $d = 2$ , there is an explicit formula for  $\Psi(x, z)$ , cf. [17, Eq. (30)], which yields (see the last displayed equation of that paper) for  $x \rightarrow \infty$  and  $z \in (-1, 1)$ ,

$$(10.7) \quad \Psi(x, z) = \frac{3}{2\pi^2} (1-u)^2 x^{-1} + O(x^{-2})$$

if  $u := x(1 - |z|) \in [0, 1)$ , and

$$(10.8) \quad \Psi(x, z) = 0$$

if  $x(1 - |z|) \notin [0, 1)$ . The implied constant in (10.7) is independent of  $x$  and  $u$ . The above asymptotics (and the fact that  $\Psi(x, z)$  is uniformly bounded) imply

$$(10.9) \quad \bar{\Psi}(z) := \int_0^\infty \Psi(x, z) dx = \log \frac{1}{1 - |z|} + O(1),$$

which holds uniformly for all  $z \in (-1, 1)$ . We conclude that  $\bar{\Psi} \in L^1((-1, 1), dz)$ .  $\square$

## 11. FROM DISCRETE TO CONTINUOUS TIME

The following proposition, together with Theorem 3.2 (ii), immediately implies Theorem 3.2 (i). Let us denote by

$$(11.1) \quad n_t := \lfloor \bar{\xi}^{-1} t \rfloor$$

the (integer part of the) expected number of collisions within time  $t$ .

**Proposition 11.1.** *For any  $\varepsilon > 0$*

$$(11.2) \quad \frac{\|\mathbf{X}_t - \mathbf{Q}_{n_t}\|}{t^{5/12+\varepsilon}} \xrightarrow{\mathbf{P}} 0,$$

as  $t \rightarrow \infty$ .

By the same argument as in the proof of Theorem 3.2 (see Section 10), it is sufficient to prove Proposition 11.1 in the case when  $(\xi_1, \boldsymbol{\eta}_1)$  has density  $\Psi_0(x, \boldsymbol{z})$ . We will assume this from now on. Furthermore, note that the left hand side of (11.2) is independent of the choice of  $\boldsymbol{v}_0$ . We may therefore assume without loss of generality that  $\boldsymbol{v}_0$  is a random variable uniformly distributed in  $S_1^{d-1}$  (this will only be used in the justification of rel. (11.31) below). The proof of Proposition 11.1, which is given at the end of this section, exploits the following three lemmas.

**Lemma 11.2.**

$$(11.3) \quad \frac{\tau_n - n\bar{\xi}}{\sigma_d \sqrt{d n \log n}} \Rightarrow \mathcal{N}(0, 1),$$

as  $n \rightarrow \infty$ .

*Proof.* This is a simple variant of the proof of Theorem 3.2 (ii).  $\square$

**Lemma 11.3.** *For all  $n \in \mathbb{N}$  and  $u \geq 2$ ,*

$$(11.4) \quad \mathbf{P}(\|\mathbf{Q}_n\| > u) = O\left(\frac{n \log u}{u^2}\right).$$

*Proof.* We begin by observing that

$$(11.5) \quad \mathbf{P}(\|\mathbf{Q}_n\| > u) \leq \mathbf{P}\left(\left\|\sum_{j=1}^n \xi_j^* \mathbf{V}_{j-1}\right\| > \frac{u}{2}\right) + \mathbf{P}\left(\left\|\sum_{j=1}^n \mu_j \mathbf{V}_{j-1}\right\| > \frac{u}{2}\right),$$

where  $\xi_j^* = \xi_j - \mu_j$  and  $\mu_j$  is the conditional expectation of the (untruncated)  $\xi_j$  defined in (6.3). Recall also the definition of the corresponding conditional variance  $\alpha_j$  in (6.5). Now,

$$(11.6) \quad \begin{aligned} \mathbf{P}\left(\left\|\sum_{j=1}^n \xi_j^* \mathbf{V}_{j-1}\right\| > \frac{u}{2}\right) &\leq \mathbf{P}\left(\left\|\sum_{j=1}^n \xi_j^* \mathbf{V}_{j-1} \mathbb{1}_{\{\alpha_j \leq u\}}\right\| > \frac{u}{4}\right) + \mathbf{P}\left(\left\|\sum_{j=1}^n \xi_j^* \mathbf{V}_{j-1} \mathbb{1}_{\{\alpha_j > u\}}\right\| > \frac{u}{4}\right) \\ &\leq \frac{16}{u^2} \sum_{j=1}^n \mathbf{E}(\alpha_j^2 \mathbb{1}_{\{\alpha_j \leq u\}}) + \sum_{j=1}^n \mathbf{P}(\alpha_j > u) \\ &= O\left(\frac{n \log u}{u^2}\right) + O\left(\frac{n}{u^2}\right), \end{aligned}$$

where we have used Chebyshev's inequality and, in the last bound, Propositions 6.12 ( $d = 2$ ) and 6.13 ( $d \geq 3$ ).

The second term in (11.5) is bounded similarly: We have

$$(11.7) \quad \begin{aligned} \mathbf{P}\left(\left\|\sum_{j=1}^n \mu_j \mathbf{V}_{j-1}\right\| > \frac{u}{2}\right) &\leq \mathbf{P}\left(\left\|\sum_{j=1}^n \mu_j \mathbf{V}_{j-1} \mathbb{1}_{\{\mu_j \leq u\}}\right\| > \frac{u}{4}\right) + \mathbf{P}\left(\left\|\sum_{j=1}^n \mu_j \mathbf{V}_{j-1} \mathbb{1}_{\{\mu_j > u\}}\right\| > \frac{u}{4}\right) \\ &\leq \frac{16}{u^2} \mathbf{E}\left(\left\|\sum_{j=1}^n \mu_j \mathbf{V}_{j-1} \mathbb{1}_{\{\mu_j \leq u\}}\right\|^2\right) + \sum_{j=1}^n \mathbf{P}(\mu_j > u). \end{aligned}$$

To control the first term on the right hand side of (11.7), it is sufficient to bound

$$(11.8) \quad \mathbf{E}\left[\left(\sum_{j=1}^n \zeta_j\right)^2\right], \quad \zeta_j := (\mathbf{e} \cdot \mathbf{V}_{j-1}) \mu_j \mathbb{1}_{\{\mu_j \leq u\}},$$

for arbitrary  $\mathbf{e} \in \mathbb{S}_1^{d-1}$ . We follow the same steps as in the proof of Lemma (4.2). Let us first show that

$$(11.9) \quad \mathbf{E}(\zeta_j) = \begin{cases} O(\omega^j \sqrt{\log \log u}) & (d = 2) \\ O(\omega^j) & (d \geq 3). \end{cases}$$

To this end choose in Proposition 8.2  $m = 1$  and

$$(11.10) \quad g(\mathbf{w}, \mathbf{z}) = \frac{K_1(\mathbf{w}, \mathbf{z})}{K_0(\mathbf{w}, \mathbf{z})} \mathbb{1}_{\{K_1(\mathbf{w}, \mathbf{z}) \leq u K_0(\mathbf{w}, \mathbf{z})\}}.$$

Propositions 6.5 and 6.6 yield

$$(11.11) \quad \mathbf{E}(g(\boldsymbol{\eta}_0, \boldsymbol{\eta}_1)^2) = \mathbf{E}(\mu_j^2 \mathbb{1}_{\{\mu_j \leq u\}}) = \begin{cases} O(\log \log u) & (d = 2) \\ O(1) & (d \geq 3), \end{cases}$$

and hence (11.9).

The above implies

$$(11.12) \quad \mathbf{Var}(\zeta_j^2) = \mathbf{E}(\zeta_j^2) + \begin{cases} O(\omega^{2j} \log \log u) & (d = 2) \\ O(\omega^{2j}) & (d \geq 3), \end{cases}$$

where

$$(11.13) \quad \mathbf{E}(\zeta_j^2) \leq \mathbf{E}(\mu_j^2 \mathbb{1}_{\{\mu_j \leq u\}}) = \begin{cases} O(\log \log u) & (d = 2) \\ O(1) & (d \geq 3). \end{cases}$$

Due to Proposition 8.1, we also have

$$(11.14) \quad \mathbf{Cov}(\zeta_i, \zeta_j) \leq \sqrt{\mathbf{E}(\zeta_i^2)} \sqrt{\mathbf{E}(\zeta_j^2)} \omega^{|i-j|}.$$

We conclude

$$(11.15) \quad \mathbf{E}\left[\left(\sum_{j=1}^n \zeta_j\right)^2\right] = \begin{cases} O(n \log \log u) & (d = 2) \\ O(n) & (d \geq 3). \end{cases}$$

The second term on the right hand side of (11.7) is controlled by the tail estimates in Propositions 6.5 and 6.6. The overall result is

$$(11.16) \quad \mathbf{P}\left(\left\|\sum_{j=1}^n \mu_j \mathbf{V}_{j-1}\right\| > \frac{u}{2}\right) = \begin{cases} O\left(\frac{n \log \log u}{u^2}\right) + O\left(\frac{n}{u^2 \log u}\right) & (d = 2) \\ O\left(\frac{n}{u^2}\right) + O\left(\frac{n}{u^{1+\frac{d}{2}}}\right) & (d \geq 3), \end{cases}$$

which completes the proof of the lemma.  $\square$

**Lemma 11.4.** For any  $\varepsilon > 0$  and  $\delta > 0$ ,

$$(11.17) \quad \lim_{t \rightarrow \infty} \mathbf{P}(|\nu_t - n_t| > \delta t^{1/2+\varepsilon}) = 0,$$

$$(11.18) \quad \lim_{n \rightarrow \infty} \mathbf{P}\left(\max_{1 \leq j \leq n} \xi_j > \delta n^{1/2+\varepsilon}\right) = 0,$$

$$(11.19) \quad \lim_{n \rightarrow \infty} \mathbf{P}\left(\max_{1 \leq m \leq n} \|\mathbf{Q}_m\| > \delta n^{5/6+\varepsilon}\right) = 0.$$

*Proof of (11.17).* Note that, for any  $N \in \mathbb{Z}_{\geq 0}$ ,  $t \geq 0$ ,

$$(11.20) \quad \nu_t \geq N \Leftrightarrow \tau_N \leq t,$$

and therefore, with  $N(t) := \lfloor n_t + \delta t^{1/2+\varepsilon} \rfloor$ ,

$$(11.21) \quad \mathbf{P}(\nu_t - n_t > \delta t^{1/2+\varepsilon}) \leq \mathbf{P}(\nu_t \geq N(t)) = \mathbf{P}(\tau_{N(t)} \leq t).$$

On the other hand, it follows from Lemma 11.2 that

$$(11.22) \quad \lim_{t \rightarrow \infty} \mathbf{P}(\tau_{N(t)} < t) = 0.$$

Similarly, for  $M(t) := \lfloor n_t - \delta t^{1/2+\varepsilon} \rfloor$ , we have

$$(11.23) \quad \mathbf{P}(\nu_t - n_t < -\delta t^{1/2+\varepsilon}) = \mathbf{P}(\tau_{M(t)} > t),$$

and Lemma 11.2 implies

$$(11.24) \quad \lim_{t \rightarrow \infty} \mathbf{P}(\tau_{M(t)} > t) = 0.$$

□

*Proof of (11.18).* We use the simplest union bound and Markov's inequality:

$$(11.25) \quad \mathbf{P}\left(\max_{1 \leq j \leq n} \xi_j > \delta n^{1/2+\varepsilon}\right) \leq n \mathbf{P}(\xi_1 > \delta n^{1/2+\varepsilon}) \leq n \frac{\mathbf{E}(\xi_1^{2-\varepsilon})}{\delta^{2-\varepsilon} n^{(1/2+\varepsilon)(2-\varepsilon)}}.$$

This sequence converges to 0 for  $\varepsilon < \frac{3}{2}$ .

□

*Proof of (11.19).* Note first that

$$(11.26) \quad \max_{1 \leq m \leq n} \|\mathbf{Q}_m\| \leq \max_{1 \leq m \leq n^{2/3}} \|\mathbf{Q}_{m \lfloor n^{1/3} \rfloor}\| + n^{1/3} \max_{1 \leq m \leq n} \xi_m.$$

Hence

$$(11.27) \quad \mathbf{P}\left(\max_{1 \leq m \leq n} \|\mathbf{Q}_m\| > \delta n^{5/6+\varepsilon}\right) \\ \leq \mathbf{P}\left(\max_{1 \leq m \leq n^{2/3}} \|\mathbf{Q}_{m \lfloor n^{1/3} \rfloor}\| > \frac{\delta}{2} n^{5/6+\varepsilon}\right) + \mathbf{P}\left(\max_{1 \leq m \leq n} \xi_m > \frac{\delta}{2} n^{1/2+\varepsilon}\right).$$

The second term on the right hand side of (11.27) converges to zero, due to (11.18). From (11.4) it follows that

$$(11.28) \quad \mathbf{P}\left(\max_{1 \leq m \leq n^{2/3}} \|\mathbf{Q}_{m \lfloor n^{1/3} \rfloor}\| > \delta n^{5/6+\varepsilon}\right) \ll \frac{\log n}{\delta^2 n^{5/3+2\varepsilon}} \sum_{m=1}^{\lfloor n^{2/3} \rfloor} m n^{1/3} \ll \frac{\log n}{\delta^2 n^{2\varepsilon}} \rightarrow 0.$$

This completes the proof of Lemma 11.4.

□

*Proof of Proposition 11.1.* Since

$$(11.29) \quad \|\mathbf{X}_t - \mathbf{Q}_{n_t}\| \leq \|\mathbf{Q}_{\nu_t} - \mathbf{Q}_{n_t}\| + \xi_{\nu_t+1},$$

we have

$$(11.30) \quad \mathbf{P}(\|\mathbf{X}_t - \mathbf{Q}_{n_t}\| > \delta t^{5/12+\varepsilon}) \leq \mathbf{P}(|\nu_t - n_t| > t^{1/2+\varepsilon}) \\ + \mathbf{P}\left(\max_{|m| \leq t^{1/2+\varepsilon}} \xi_{n_t+m} > \frac{\delta}{2} t^{5/12+\varepsilon}\right) + \mathbf{P}\left(\max_{|m| \leq t^{1/2+\varepsilon}} \|\mathbf{Q}_{n_t+m} - \mathbf{Q}_{n_t}\| > \frac{\delta}{2} t^{5/12+\varepsilon}\right),$$

and therefore, by stationarity of the Markov process (3.2) (recall that here we may assume without loss of generality that  $\mathbf{v}_0$  is uniformly distributed in  $S_1^{d-1}$ ),

$$(11.31) \quad \mathbf{P}(\|\mathbf{X}_t - \mathbf{Q}_{n_t}\| > \delta t^{5/12+\varepsilon}) \leq \mathbf{P}(|\nu_t - n_t| > t^{1/2+\varepsilon}) \\ + \mathbf{P}\left(\max_{1 \leq m \leq 2t^{1/2+\varepsilon}} \xi_m > \frac{\delta}{2} t^{5/12+\varepsilon}\right) + \mathbf{P}\left(\max_{1 \leq m \leq 2t^{1/2+\varepsilon}} \|\mathbf{Q}_m\| > \frac{\delta}{4} t^{5/12+\varepsilon}\right).$$

The three terms on the right hand side of (11.31) are controlled by Lemma 11.4. This completes the proof of Proposition 11.1.  $\square$

## REFERENCES

1. P Bálint, N Chernov, and D Dolgopyat, *Limit theorems for dispersing billiards with cusps*, Communications in Mathematical Physics **308** (2011), no. 2, 479–510.
2. Péter Bálint and Sébastien Gouëzel, *Limit theorems in the stadium billiard*, Communications in Mathematical Physics **263** (2006), no. 2, 461–512.
3. Péter Bálint and Imre Péter Tóth, *Exponential decay of correlations in multi-dimensional dispersing billiards*, Annales Henri Poincaré. A Journal of Theoretical and Mathematical Physics **9** (2008), no. 7, 1309–1369.
4. ———, *Example for exponential growth of complexity in a finite horizon multi-dimensional dispersing billiard*, Nonlinearity **25** (2012), no. 5, 1275–1297.
5. P M Bleher, *Statistical properties of two-dimensional periodic Lorentz gas with infinite horizon*, Journal of Statistical Physics **66** (1992), no. 1-2, 315–373.
6. C Boldrighini, L A Bunimovich, and Ya G Sinai, *On the Boltzmann equation for the Lorentz gas*, Journal of Statistical Physics **32** (1983), no. 3, 477–501.
7. J Bourgain, F Golse, and B Wennberg, *On the distribution of free path lengths for the periodic Lorentz gas*, Communications in Mathematical Physics **190** (1998), no. 3, 491–508.
8. L A Bunimovich and Ya. G. Sinai, *Statistical properties of Lorentz gas with periodic configuration of scatterers*, Communications in Mathematical Physics **78** (1980), no. 4, 479–497.
9. V A Bykovskii and A V Ustinov, *The statistics of particle trajectories in the nonhomogeneous Sinai problem for a two-dimensional lattice*, Rossiiskaya Akademiya Nauk. Izvestiya. Seriya Matematicheskaya **73** (2009), no. 4, 17–36.
10. Emanuele Caglioti and Francois Golse, *The Boltzmann-Grad limit of the periodic Lorentz gas in two space dimensions*, Comptes Rendus Mathématique. Académie des Sciences. Paris **346** (2008), no. 7-8, 477–482.
11. ———, *On the Boltzmann-Grad limit for the two dimensional periodic Lorentz gas*, Journal of Statistical Physics **141** (2010), no. 2, 264–317.
12. N I Chernov, *Statistical properties of the periodic Lorentz gas. Multidimensional case*, Journal of Statistical Physics **74** (1994), no. 1-2, 11–53.
13. Carl P Dettmann, *New horizons in multidimensional diffusion: the Lorentz gas and the Riemann hypothesis*, Journal of Statistical Physics **146** (2012), no. 1, 181–204.
14. D I Dolgopyat and N I Chernov, *Anomalous current in periodic Lorentz gases with an infinite horizon*, Rossiiskaya Akademiya Nauk. Moskovskoe Matematicheskoe Obshchestvo. Uspekhi Matematicheskikh Nauk **64** (2009), no. 4(388), 73–124.
15. G Gallavotti, *Divergences and approach to equilibrium in the Lorentz and the wind-tree-models*, Physical Review **185** (1969), 308–322.
16. Francois Golse and Bernt Wennberg, *On the distribution of free path lengths for the periodic Lorentz gas. II*, M2AN. Mathematical Modelling and Numerical Analysis **34** (2000), no. 6, 1151–1163.
17. Jens Marklof and Andreas Strömbergsson, *Kinetic transport in the two-dimensional periodic Lorentz gas*, Nonlinearity **21** (2008), no. 7, 1413–1422.
18. ———, *The distribution of free path lengths in the periodic Lorentz gas and related lattice point problems*, Annals of Mathematics. Second Series **172** (2010), no. 3, 1949–2033.
19. ———, *The Boltzmann-Grad limit of the periodic Lorentz gas*, Annals of Mathematics. Second Series **174** (2011), no. 1, 225–298.
20. ———, *The periodic Lorentz gas in the Boltzmann-Grad limit: asymptotic estimates*, Geometric and Functional Analysis **21** (2011), no. 3, 560–647.
21. Ian Melbourne and Matthew Nicol, *A vector-valued almost sure invariance principle for hyperbolic dynamical systems*, The Annals of Probability **37** (2009), no. 2, 478–505.
22. P. Nandori, D. Szasz, and T. Varju, *Tail asymptotics of free path lengths for the periodic Lorentz process. On Dettmann's geometric conjectures*, e-print arXiv:1210.2231 (2012).
23. George C Papanicolaou, *Asymptotic analysis of transport processes*, Bulletin of the American Mathematical Society **81** (1975), 330–392.

24. H Spohn, *The Lorentz process converges to a random flight process*, Communications in Mathematical Physics **60** (1978), 277–290.
25. Karl Stromberg, *Probabilities on a compact group*, Transactions of the American Mathematical Society **94** (1960), 295–309.
26. Andreas Strömbergsson, *On the probability of a random lattice avoiding a large convex set*, Proceedings of the London Mathematical Society. Third Series **103** (2011), no. 6, 950–1006.
27. Domokos Szász and Tamás Varjú, *Limit laws and recurrence for the planar Lorentz process with infinite horizon*, J. Stat. Phys. **129** (2007), no. 1, 59–80.

JENS MARKLOF, SCHOOL OF MATHEMATICS, UNIVERSITY OF BRISTOL, BRISTOL BS8 1TW, U.K.  
[j.marklof@bristol.ac.uk](mailto:j.marklof@bristol.ac.uk)

BÁLINT TÓTH, SCHOOL OF MATHEMATICS, UNIVERSITY OF BRISTOL, BRISTOL BS8 1TW, U.K., AND  
INSTITUTE OF MATHEMATICS, BUDAPEST UNIVERSITY OF TECHNOLOGY, BUDAPEST, HUNGARY  
[balint.toth@bristol.ac.uk](mailto:balint.toth@bristol.ac.uk), [balint@math.bme.hu](mailto:balint@math.bme.hu)