

# THE BOLTZMANN-GRAD LIMIT OF THE PERIODIC LORENTZ GAS

JENS MARKLOF AND ANDREAS STRÖMBERGSSON

ABSTRACT. We study the dynamics of a point particle in a periodic array of spherical scatterers, and construct a stochastic process that governs the time evolution for random initial data in the limit of low scatterer density (Boltzmann-Grad limit). A generic path of the limiting process is a piecewise linear curve whose consecutive segments are generated by a Markov process with memory two.

## CONTENTS

|  |    |
|--|----|
| 1. Introduction  | 2  |
| 1.1. The joint distribution of path segments             | 2  |
| 1.2. A limiting stochastic process for the billiard flow | 4  |
| 1.3. Macroscopic initial conditions                      | 5  |
| 1.4. Outline of the paper                                | 7  |
| 2. First collision                                       | 7  |
| 2.1. Location of the first collision                     | 7  |
| 2.2. Scattering maps                                     | 9  |
| 2.3. Velocity after the first impact                     | 10 |
| 2.4. A uniform version of Theorem 2.2                    | 11 |
| 3. Iterated scattering maps                              | 17 |
| 3.1. Two lemmas  | 17 |
| 3.2. The main proposition                                | 20 |
| 4. Loss of memory  | 23 |
| 4.1. Statement of the main theorem                       | 23 |
| 4.2. Sets of good initial velocities                     | 24 |
| 4.3. Proof of Theorem 4.1                                | 27 |
| 4.4. Proof of Theorem 1.1                                | 33 |
| 4.5. Proof of Theorem 1.3                                | 34 |
| 5. Convergence to the stochastic process $\Xi(t)$        | 36 |
| 5.1. Four lemmas   | 36 |
| 5.2. Proof of Theorem 5.1                                | 39 |
| 5.3. A counterexample                                    | 40 |
| 5.4. Macroscopic initial conditions                      | 40 |
| 6. A continuous-time Markov process                      | 41 |
| 6.1. Transcription of Theorem 1.4                        | 41 |
| 6.2. A semigroup of propagators                          | 43 |
| 6.3. The Fokker-Planck-Kolmogorov equation               | 45 |
| References   | 50 |

---

*Date:* 3 January 2008; to appear in the Annals of Mathematics.

J.M. has been supported by EPSRC Research Grants GR/T28058/01 and GR/S87461/01, and a Philip Leverhulme Prize. A.S. is a Royal Swedish Academy of Sciences Research Fellow supported by a grant from the Knut and Alice Wallenberg Foundation.

## 1. INTRODUCTION

The Lorentz gas describes an ensemble of non-interacting point particles in an infinite array of spherical scatterers. It was originally developed by Lorentz [18] in 1905 to model, in the limit of low scatterer density (Boltzmann-Grad limit), the stochastic properties of the motion of electrons in a metal. In the present paper we consider the case of a periodic array of scatterers, and construct a stochastic process that indeed governs the macroscopic dynamics of a particle cloud in the Boltzmann-Grad limit. The corresponding result has been known for some time in the case of a Poisson-distributed (rather than periodic) configuration of scatterers. Here the limiting process corresponds to a solution of the linear Boltzmann equation, see Gallavotti [13], Spohn [27], and Boldrighini, Bunimovich and Sinai [4]. It already follows from the estimates in [5, 16] that the linear Boltzmann equation does not hold in the periodic set-up; this was pointed out recently by Golse [14, 15].

Our results complement classical studies in ergodic theory that characterize the stochastic properties of the periodic Lorentz gas in the limit of long times, see [6, 2, 9, 22, 23, 1, 26, 12] for details.

To state our main results, consider an ensemble of non-interacting point particles moving in an array of spherical scatterers which are placed at the vertices of a euclidean lattice  $\mathcal{L} \subset \mathbb{R}^d$  of covolume one (Figure 1). The dynamics of each particle is governed by the billiard flow

$$(1.1) \quad \varphi_t : T^1(\mathcal{K}_\rho) \rightarrow T^1(\mathcal{K}_\rho), \quad (\mathbf{q}_0, \mathbf{v}_0) \mapsto (\mathbf{q}(t), \mathbf{v}(t))$$

where  $\mathcal{K}_\rho \subset \mathbb{R}^d$  is the complement of the set  $\mathcal{B}_\rho^d + \mathcal{L}$  (the “billiard domain”), and  $T^1(\mathcal{K}_\rho) = \mathcal{K}_\rho \times S_1^{d-1}$  is its unit tangent bundle (the “phase space”).  $\mathcal{B}_\rho^d$  denotes the open ball of radius  $\rho$ , centered at the origin. A point in  $T^1(\mathcal{K}_\rho)$  is parametrized by  $(\mathbf{q}, \mathbf{v})$ , with  $\mathbf{q} \in \mathcal{K}_\rho$  denoting the position and  $\mathbf{v} \in S_1^{d-1}$  the velocity of the particle. The Liouville measure on  $T^1(\mathcal{K}_\rho)$  is

$$(1.2) \quad d\nu(\mathbf{q}, \mathbf{v}) = d\text{vol}_{\mathbb{R}^d}(\mathbf{q}) d\text{vol}_{S_1^{d-1}}(\mathbf{v})$$

where  $\text{vol}_{\mathbb{R}^d}$  and  $\text{vol}_{S_1^{d-1}}$  refer to the Lebesgue measures on  $\mathbb{R}^d$  and  $S_1^{d-1}$ , respectively. For the purpose of this introduction we will restrict our attention to Lorentz’ classical set-up, where the scatterers are assumed to be hard spheres. Our results in fact also hold for scattering processes described by smooth potentials, see Sec. 2.2 for details.

If the initial condition  $(\mathbf{q}_0, \mathbf{v}_0)$  is random, the billiard flow gives rise to the stochastic process

$$(1.3) \quad \{(\mathbf{q}(t), \mathbf{v}(t)) : t \in \mathbb{R}_{>0}\}$$

which we will refer to as the Lorentz process. The central result of this paper is the existence of a limiting stochastic process  $\{\Xi(t) : t \in \mathbb{R}_{>0}\}$  of the Lorentz process in the Boltzmann-Grad limit  $\rho \rightarrow 0$ . We begin with a study of the distribution of path segments of the billiard flow between collisions.

**1.1. The joint distribution of path segments.** The billiard flow  $\varphi_t$  induces a billiard map on the boundary  $\partial T^1(\mathcal{K}_\rho)$ ,

$$(1.4) \quad (\mathbf{q}_{n-1}, \mathbf{v}_{n-1}) \mapsto (\mathbf{q}_n, \mathbf{v}_n),$$

where  $\mathbf{q}_n, \mathbf{v}_n$  denote position and velocity at the  $n$ th collision in the *outgoing* configuration, i.e.,

$$(1.5) \quad (\mathbf{q}_n, \mathbf{v}_n) = \lim_{\epsilon \rightarrow 0^+} \varphi_{\tau_1(\mathbf{q}_{n-1}, \mathbf{v}_{n-1}; \rho) + \epsilon}(\mathbf{q}_{n-1}, \mathbf{v}_{n-1}).$$

Here  $\tau_1$  denotes the free path length, defined by

$$(1.6) \quad \tau_1(\mathbf{q}, \mathbf{v}; \rho) = \inf\{t > 0 : \mathbf{q} + t\mathbf{v} \notin \mathcal{K}_\rho\}.$$

We will later also use the parametrization  $\partial T^1(\mathcal{K}_\rho) = (S_\rho^{d-1} + \mathcal{L}) \times S_1^{d-1}$ , so that  $\mathbf{q}_n = \mathbf{m}_n + \rho \mathbf{w}_n$ , where  $\mathbf{w}_n \in S_1^{d-1}$  and  $\mathbf{m}_n \in \mathcal{L}$  are the position on the ball and ball label at the  $n$ th collision.

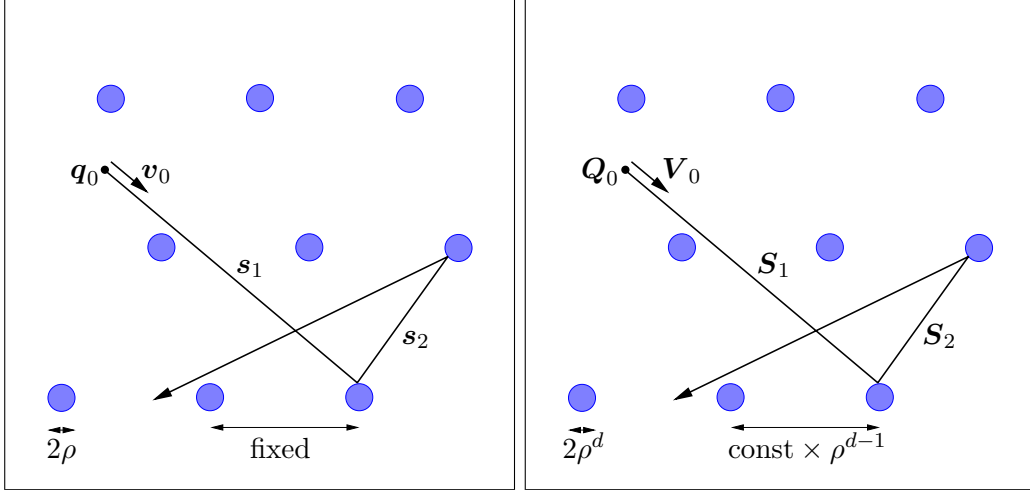


FIGURE 1. Left: The periodic Lorentz gas in “microscopic” coordinates—the lattice  $\mathcal{L}$  remains fixed as the radius  $\rho$  of the scatterer tends to zero. Right: The periodic Lorentz gas in “macroscopic” coordinates—both the lattice constant and the radius of each scatterer tend to zero, in such a way that the mean free path length remains finite. The vectors  $\mathbf{s}_1, \mathbf{s}_2, \dots$  (resp.  $\mathbf{S}_1, \mathbf{S}_2, \dots$ ) represent the segments of the billiard path between collisions.

The time elapsed between the  $(n-1)$ th and  $n$ th hit is defined as the  $n$ th collision time

$$(1.7) \quad \tau_n(\mathbf{q}_0, \mathbf{v}_0; \rho) = \tau_1(\mathbf{q}_{n-1}, \mathbf{v}_{n-1}; \rho).$$

We express the  $n$ th path segment by the vector

$$(1.8) \quad \mathbf{s}_n(\mathbf{q}_0, \mathbf{v}_0; \rho) := \tau_n(\mathbf{q}_0, \mathbf{v}_0; \rho) \mathbf{v}_{n-1}(\mathbf{q}_0, \mathbf{v}_0; \rho).$$

The central result of [19] is the proof of a limiting distribution for the first collision time  $\tau_1$ , and further refined versions that also take into account the particle’s direction after the reflection. We will here extend these results to find a joint limiting distribution for the first  $n$  segments of the billiard path with initial coordinates  $(\mathbf{q}_0, \mathbf{v}_0)$ , where the position  $\mathbf{q}_0$  is fixed and the velocity  $\mathbf{v}_0$  random. The precise statement is the following.

Here and in the remainder of this paper we will use the standard representation  $\mathcal{L} = \mathbb{Z}^d M_0$ , where  $M_0 \in \text{SL}(d, \mathbb{R})$ . We will also use the notation  $\widehat{\mathbf{S}} := \|\mathbf{S}\|^{-1} \mathbf{S}$ . We set

$$(1.9) \quad \mathcal{B}_n := \{(\mathbf{S}_1, \dots, \mathbf{S}_n) \in (\mathbb{R}^d \setminus \{\mathbf{0}\})^n : \widehat{\mathbf{S}}_{j+1} \neq \widehat{\mathbf{S}}_j \ (j = 1, \dots, n-1)\}.$$

**Theorem 1.1.** *Fix a lattice  $\mathcal{L} = \mathbb{Z}^d M_0$  and a point  $\mathbf{q}_0 \in \mathbb{R}^d \setminus \mathcal{L}$ , and write  $\boldsymbol{\alpha} = -\mathbf{q}_0 M_0^{-1}$ . Then for each  $n \in \mathbb{Z}_{>0}$  there exists a function  $P_{\boldsymbol{\alpha}}^{(n)} : \mathcal{B}_n \rightarrow \mathbb{R}_{\geq 0}$  such that, for any Borel probability measure  $\lambda$  on  $\mathbb{S}_1^{d-1}$  which is absolutely continuous with respect to  $\text{vol}_{\mathbb{S}_1^{d-1}}$ , and for any set  $\mathcal{A} \subset \mathbb{R}^{nd}$  with boundary of Lebesgue measure zero,*

$$(1.10) \quad \lim_{\rho \rightarrow 0} \lambda(\{\mathbf{v}_0 \in \mathbb{S}_1^{d-1} : (\mathbf{s}_1(\mathbf{q}_0, \mathbf{v}_0; \rho), \dots, \mathbf{s}_n(\mathbf{q}_0, \mathbf{v}_0; \rho)) \in \rho^{-(d-1)} \mathcal{A}\}) \\ = \int_{\mathcal{A}} P_{\boldsymbol{\alpha}}^{(n)}(\mathbf{S}_1, \dots, \mathbf{S}_n) \lambda'(\widehat{\mathbf{S}}_1) d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_1) \cdots d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_n),$$

where  $\lambda' \in L^1(\mathbb{S}_1^{d-1})$  is the Radon-Nikodym derivative of  $\lambda$  with respect to  $\text{vol}_{\mathbb{S}_1^{d-1}}$ . Furthermore, there is a function  $\Psi : \mathcal{B}_3 \rightarrow \mathbb{R}_{\geq 0}$  such that

$$(1.11) \quad P_{\boldsymbol{\alpha}}^{(n)}(\mathbf{S}_1, \dots, \mathbf{S}_n) = P_{\boldsymbol{\alpha}}^{(2)}(\mathbf{S}_1, \mathbf{S}_2) \prod_{j=3}^n \Psi(\mathbf{S}_{j-2}, \mathbf{S}_{j-1}, \mathbf{S}_j)$$

for all  $n \geq 3$  and all  $(\mathbf{S}_1, \dots, \mathbf{S}_n) \in \mathcal{B}_n$ .

The above condition  $\mathbf{q}_0 \in \mathbb{R}^d \setminus \mathcal{L}$  ensures that  $\tau_1$  is defined for  $\rho$  sufficiently small. In Sec. 4.4 we also consider variants of Theorem 1.1 where the initial position is near  $\mathcal{L}$ , e.g.,  $\mathbf{q}_0 \in \partial\mathcal{K}_\rho$ .

We define the probability measure corresponding to (1.10) by

$$(1.12) \quad \mu_{\alpha,\lambda}^{(n)}(\mathcal{A}) := \int_{\mathcal{A}} P_{\alpha}^{(n)}(\mathbf{S}_1, \dots, \mathbf{S}_n) \lambda'(\widehat{\mathbf{S}}_1) d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_1) \cdots d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_n).$$

Note in particular that  $\mu_{\alpha,\lambda}^{(n+1)}(\mathcal{A} \times \mathbb{R}^d) = \mu_{\alpha,\lambda}^{(n)}(\mathcal{A})$ .

*Remark 1.1.* In probabilistic terms, Theorem 1.1 states that the discrete-time stochastic process  $\{\rho^{d-1} \mathbf{s}_n(\mathbf{q}_0, \mathbf{v}_0; \rho) : n \in \mathbb{Z}_{>0}\}$  converges in the limit  $\rho \rightarrow 0$  to

$$(1.13) \quad \{\mathbf{S}_n : n \in \mathbb{Z}_{>0}\},$$

a Markov process with memory two. As we shall see,  $\Psi(\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3)$  is in fact independent of  $\|\mathbf{S}_1\|$ .

*Remark 1.2.* If  $d \geq 3$  then  $P_{\alpha}^{(n)}$  is continuous on all of  $\mathcal{B}_n$ . If  $d = 2$  then  $P_{\alpha}^{(n)}$  is continuous except possibly at points  $(\mathbf{S}_1, \dots, \mathbf{S}_n) \in \mathcal{B}_n$  with  $\widehat{\mathbf{S}}_2 = -\widehat{\mathbf{S}}_1$  or  $\widehat{\mathbf{S}}_{j+2} = \widehat{\mathbf{S}}_j R_{\mathbf{S}_{j+1}}$  for some  $1 \leq j \leq n-2$ , where  $R_{\mathbf{S}} \in O(2)$  denotes reflection in the line  $\mathbb{R}\mathbf{S}$ . Cf. Remark 4.7 below.

*Remark 1.3.* Note that  $\Psi$  is independent of  $\mathcal{L}$  and  $\mathbf{q}_0$ , and  $P_{\alpha}^{(n)}$  depends only on the choice of  $\alpha$ . This means in particular that  $\Psi$  and  $P_{\alpha}^{(n)}$  are rotation-invariant, i.e., for any  $K \in O(d)$  we have

$$(1.14) \quad \Psi(\mathbf{S}_1 K, \mathbf{S}_2 K, \mathbf{S}_3 K) = \Psi(\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3)$$

and

$$(1.15) \quad P_{\alpha}^{(n)}(\mathbf{S}_1 K, \dots, \mathbf{S}_n K) = P_{\alpha}^{(n)}(\mathbf{S}_1, \dots, \mathbf{S}_n).$$

For  $\alpha \in \mathbb{R}^d \setminus \mathbb{Q}^d$  also  $P_{\alpha}^{(n)} =: P^{(n)}$  is in fact independent of  $\alpha$ , cf. Remark 4.6 below. Explicit formulas and asymptotic properties of the limiting distributions will be presented in [20] and [21], respectively.

*Remark 1.4.* The case  $n = 1$  of course leads to the distribution of the free path length discussed in [19]; cf. also [10, 5, 16, 7, 3] for earlier results.

**1.2. A limiting stochastic process for the billiard flow.** In Theorem 1.1 we have identified a Markov process with memory two that describes the limiting distribution of billiard paths with random initial data  $(\mathbf{q}_0, \mathbf{v}_0)$ . Let us denote by

$$(1.16) \quad \{\Xi(t) : t \in \mathbb{R}_{>0}\},$$

the continuous-time stochastic process that is obtained by moving with unit speed along the random paths of the Markov process (1.13). The process is fully specified by the probability

$$(1.17) \quad \mathbb{P}_{\alpha,\lambda}(\Xi(t_1) \in \mathcal{D}_1, \dots, \Xi(t_M) \in \mathcal{D}_M)$$

that  $\Xi(t)$  visits the sets  $\mathcal{D}_1, \dots, \mathcal{D}_M \subset \mathbb{T}^1(\mathbb{R}^d)$  at times  $t = t_1, \dots, t_M$ , with  $M$  arbitrarily large. To give a precise definition of (1.17) set  $T_0 := 0$ ,  $T_n := \sum_{j=1}^n \|\mathbf{S}_j\|$ , and define the probability that  $\Xi(t)$  is in the set  $\mathcal{D}_1$  at time  $t_1$  after exactly  $n_1$  hits, in the set  $\mathcal{D}_2$  at time  $t_2$  after exactly  $n_2$  hits, etc., by

$$(1.18) \quad \mathbb{P}_{\alpha,\lambda}^{(n)}(\Xi(t_1) \in \mathcal{D}_1, \dots, \Xi(t_M) \in \mathcal{D}_M \text{ and } T_{n_1} \leq t_1 < T_{n_1+1}, \dots, T_{n_M} \leq t_M < T_{n_M+1}) \\ := \mu_{\alpha,\lambda}^{(n+1)}(\{(\mathbf{S}_1, \dots, \mathbf{S}_{n+1}) : \Xi_{n_j}(t_j) \in \mathcal{D}_j, T_{n_j} \leq t_j < T_{n_j+1} (j = 1, \dots, M)\})$$

with  $\mathbf{n} := (n_1, \dots, n_M)$ ,  $n := \max(n_1, \dots, n_M)$ , and

$$(1.19) \quad \Xi_n(t) := \left( \sum_{j=1}^n \mathbf{S}_j + (t - T_n) \widehat{\mathbf{S}}_{n+1}, \widehat{\mathbf{S}}_{n+1} \right).$$

Note that the choice  $T_n \leq t < T_{n+1}$  of semi-open intervals is determined by the use of the outgoing configuration, recall (1.5). The formal definition of (1.17) is thus

$$(1.20) \quad \mathbb{P}_{\alpha, \lambda}(\Xi(t_1) \in \mathcal{D}_1, \dots, \Xi(t_M) \in \mathcal{D}_M) \\ := \sum_{\mathbf{n} \in \mathbb{Z}_{\geq 0}^M} \mathbb{P}_{\alpha, \lambda}^{(\mathbf{n})}(\Xi(t_1) \in \mathcal{D}_1, \dots, \Xi(t_M) \in \mathcal{D}_M \text{ and } T_{n_1} \leq t_1 < T_{n_1+1}, \dots, T_{n_M} \leq t_M < T_{n_M+1}).$$

The following theorem shows that the Lorentz process (1.3), suitably rescaled, converges to the stochastic process (1.16) as  $\rho \rightarrow 0$ . Given any set  $\mathcal{D} \subset \mathbb{T}^1(\mathbb{R}^d)$  we say that  $t \geq 0$  is  $\mathcal{D}$ -admissible if

$$(1.21) \quad \text{vol}_{\mathbb{S}_1^{d-1}}(\{\widehat{\mathbf{S}}_1 \in \mathbb{S}_1^{d-1} : (t\widehat{\mathbf{S}}_1, \widehat{\mathbf{S}}_1) \in \partial\mathcal{D}\}) = 0.$$

We write  $\text{adm}(\mathcal{D})$  for the set of all  $\mathcal{D}$ -admissible numbers  $t \geq 0$ .

**Theorem 1.2.** *Fix a lattice  $\mathcal{L} = \mathbb{Z}^d M_0$  and a point  $\mathbf{q}_0 \in \mathbb{R}^d \setminus \mathcal{L}$ , set  $\alpha = -\mathbf{q}_0 M_0^{-1}$ , and let  $\lambda$  be a Borel probability measure on  $\mathbb{S}_1^{d-1}$  which is absolutely continuous with respect to  $\text{vol}_{\mathbb{S}_1^{d-1}}$ . Then, for any subsets  $\mathcal{D}_1, \dots, \mathcal{D}_M \subset \mathbb{T}^1(\mathbb{R}^d)$  with boundary of Lebesgue measure zero, and any numbers  $t_j \in \text{adm}(\mathcal{D}_j)$  ( $j = 1, \dots, M$ ),*

$$(1.22) \quad \lim_{\rho \rightarrow 0} \lambda(\{\mathbf{v}_0 \in \mathbb{S}_1^{d-1} : (\rho^{d-1} \mathbf{q}(\rho^{-(d-1)} t_j), \mathbf{v}(\rho^{-(d-1)} t_j)) \in \mathcal{D}_j, j = 1, \dots, M\}) \\ = \mathbb{P}_{\alpha, \lambda}(\Xi(t_1) \in \mathcal{D}_1, \dots, \Xi(t_M) \in \mathcal{D}_M).$$

The convergence is uniform for  $(t_1, \dots, t_M)$  in compact subsets of  $\text{adm}(\mathcal{D}_1) \times \dots \times \text{adm}(\mathcal{D}_M)$ .

*Remark 1.5.* The condition  $t_j \in \text{adm}(\mathcal{D}_j)$  cannot be disposed with. For example, (1.22) is in general false in the case  $M = 1$ ,  $\mathcal{D}_1 = \mathcal{B}_{t_1}^d \times \mathbb{S}_1^{d-1}$ . We prove this in Section 5.3. Note however that no admissibility condition is required in the macroscopic analogue of Theorem 1.2, see Theorem 1.4 below.

**1.3. Macroscopic initial conditions.** In view of the rescaling applied in the previous section it is natural to consider the “macroscopic” billiard flow

$$(1.23) \quad F_t : \mathbb{T}^1(\rho^{d-1} \mathcal{K}_\rho) \rightarrow \mathbb{T}^1(\rho^{d-1} \mathcal{K}_\rho) \\ (\mathbf{Q}_0, \mathbf{V}_0) \mapsto (\mathbf{Q}(t), \mathbf{V}(t)) = (\rho^{d-1} \mathbf{q}(\rho^{-(d-1)} t), \mathbf{v}(\rho^{-(d-1)} t)),$$

and take random initial conditions  $(\mathbf{Q}_0, \mathbf{V}_0)$  with respect to some fixed probability measure  $\Lambda$ . We will establish the analogous limit laws as in the previous sections. Although the macroscopic versions are less general (they are obtained by averaging over  $\mathbf{q}_0$ ), they appear more natural from a physical viewpoint, where one is interested in the time evolution of a macroscopic particle cloud; cf. the discussion at end of this section.

The  $n$ th path segment in these macroscopic coordinates is

$$(1.24) \quad \mathbf{S}_n(\mathbf{Q}_0, \mathbf{V}_0; \rho) := \rho^{d-1} \mathbf{s}_n(\rho^{-(d-1)} \mathbf{Q}_0, \mathbf{V}_0; \rho).$$

**Theorem 1.3.** *Fix a lattice  $\mathcal{L}$  and let  $\Lambda$  be a Borel probability measure on  $\mathbb{T}^1(\mathbb{R}^d)$  which is absolutely continuous with respect to Lebesgue measure. Then, for each  $n \in \mathbb{Z}_{>0}$ , and for any set  $\mathcal{A} \subset \mathbb{R}^d \times \mathbb{R}^{nd}$  with boundary of Lebesgue measure zero,*

$$(1.25) \quad \lim_{\rho \rightarrow 0} \Lambda(\{(\mathbf{Q}_0, \mathbf{V}_0) \in \mathbb{T}^1(\rho^{d-1} \mathcal{K}_\rho) : (\mathbf{Q}_0, \mathbf{S}_1(\mathbf{Q}_0, \mathbf{V}_0; \rho), \dots, \mathbf{S}_n(\mathbf{Q}_0, \mathbf{V}_0; \rho)) \in \mathcal{A}\}) \\ = \int_{\mathcal{A}} P^{(n)}(\mathbf{S}_1, \dots, \mathbf{S}_n) \Lambda'(\mathbf{Q}_0, \widehat{\mathbf{S}}_1) d\text{vol}_{\mathbb{R}^d}(\mathbf{Q}_0) d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_1) \cdots d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_n),$$

with  $P^{(n)}$  as in Remark 1.3, and where  $\Lambda'$  is the Radon-Nikodym derivative of  $\Lambda$  with respect to  $\text{vol}_{\mathbb{R}^d} \times \text{vol}_{\mathbb{S}_1^{d-1}}$ .

The probability measure corresponding to the above limiting distribution is defined by

$$(1.26) \quad \mu_\Lambda^{(n)}(\mathcal{A}) := \int_{\mathcal{A}} P^{(n)}(\mathbf{S}_1, \dots, \mathbf{S}_n) \Lambda'(\mathbf{Q}_0, \widehat{\mathbf{S}}_1) d\text{vol}_{\mathbb{R}^d}(\mathbf{Q}_0) d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_1) \cdots d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_n).$$

We redefine the stochastic process (1.16) by specifying the probability

$$(1.27) \quad \mathbb{P}_\Lambda(\Xi(t_1) \in \mathcal{D}_1, \dots, \Xi(t_M) \in \mathcal{D}_M)$$

via the measure  $\mu_\Lambda^{(n)}$  by the same construction as in Section 1.2. The only essential difference is that we need to replace (1.19) by

$$(1.28) \quad \Xi_n(t) := \left( \mathbf{Q}_0 + \sum_{j=1}^n \mathbf{S}_j + (t - T_n) \widehat{\mathbf{S}}_{n+1}, \widehat{\mathbf{S}}_{n+1} \right).$$

Note that formally  $\mathbb{P}_\Lambda = \mathbb{P}_{\alpha, \lambda}$  if  $\Lambda'(\mathbf{Q}, \mathbf{V}) = \delta(\mathbf{Q})\lambda'(\mathbf{V})$ , and  $\alpha \in \mathbb{R}^d \setminus \mathbb{Q}^d$ .

**Theorem 1.4.** *Fix a lattice  $\mathcal{L}$  and let  $\Lambda$  be a Borel probability measure on  $\mathbb{T}^1(\mathbb{R}^d)$  which is absolutely continuous with respect to Lebesgue measure. Then, for any  $t_1, \dots, t_M \in \mathbb{R}_{\geq 0}$ , and any subsets  $\mathcal{D}_1, \dots, \mathcal{D}_M \subset \mathbb{T}^1(\mathbb{R}^d)$  with boundary of Lebesgue measure zero,*

$$(1.29) \quad \lim_{\rho \rightarrow 0} \Lambda(\{(\mathbf{Q}_0, \mathbf{V}_0) \in \mathbb{T}^1(\rho^{d-1}\mathcal{K}_\rho) : (\mathbf{Q}(t_1), \mathbf{V}(t_1)) \in \mathcal{D}_1, \dots, (\mathbf{Q}(t_M), \mathbf{V}(t_M)) \in \mathcal{D}_M\}) \\ = \mathbb{P}_\Lambda(\Xi(t_1) \in \mathcal{D}_1, \dots, \Xi(t_M) \in \mathcal{D}_M).$$

The convergence is uniform for  $t_1, \dots, t_M$  in compact subsets of  $\mathbb{R}_{\geq 0}$ .

The time evolution of an initial particle cloud  $f \in L^1(\mathbb{T}^1(\rho^{d-1}\mathcal{K}_\rho))$  in the periodic Lorentz gas is described by the operator  $L_{t, \rho}$  defined by

$$(1.30) \quad [L_{t, \rho} f](\mathbf{Q}, \mathbf{V}) = f(F_t^{-1}(\mathbf{Q}, \mathbf{V})).$$

To allow a  $\rho$ -independent choice of the initial density  $f$ , it is convenient to extend the action of  $F_t$  from  $\mathbb{T}^1(\rho^{d-1}\mathcal{K}_\rho)$  to  $\mathbb{T}^1(\mathbb{R}^d)$  by setting  $F_t = \text{id}$  on  $\mathbb{T}^1(\mathbb{R}^d) \setminus \mathbb{T}^1(\rho^{d-1}\mathcal{K}_\rho)$ . We fix the Liouville measure on  $\mathbb{T}^1(\mathbb{R}^d)$  to be the standard Lebesgue measure

$$(1.31) \quad d\nu(\mathbf{Q}, \mathbf{V}) = d\text{vol}_{\mathbb{R}^d}(\mathbf{Q}) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{V}).$$

Theorem 1.4 now implies the existence of a limiting operator  $L_t$  that describes the evolution of the particle cloud in the Boltzmann-Grad limit. More precisely, for every set  $\mathcal{D}$  with boundary of Lebesgue measure zero, we have

$$(1.32) \quad \lim_{\rho \rightarrow 0} \int_{\mathcal{D}} [L_{t, \rho} f](\mathbf{Q}, \mathbf{V}) d\nu(\mathbf{Q}, \mathbf{V}) = \int_{\mathcal{D}} [L_t f](\mathbf{Q}, \mathbf{V}) d\nu(\mathbf{Q}, \mathbf{V}),$$

uniformly for  $t$  on compacta in  $\mathbb{R}_{\geq 0}$ , and  $L_t$  is defined by the relation

$$(1.33) \quad \int_{\mathcal{D}} [L_t f](\mathbf{Q}, \mathbf{V}) d\nu(\mathbf{Q}, \mathbf{V}) = \mathbb{P}_\Lambda(\Xi(t) \in \mathcal{D}),$$

for any absolutely continuous  $\Lambda$ , any Borel subset  $\mathcal{D} \subset \mathbb{T}^1(\mathbb{R}^d)$ , and  $f = \Lambda'$ . We note that  $L_t$  commutes with the translation operators  $\{T_{\mathbf{R}} : \mathbf{R} \in \mathbb{R}^d\}$ ,

$$(1.34) \quad [T_{\mathbf{R}} f](\mathbf{Q}, \mathbf{V}) := f(\mathbf{Q} - \mathbf{R}, \mathbf{V}),$$

and, in view of Remark 1.3, with the rotation operators  $\{R_K : K \in \text{O}(d)\}$ ,

$$(1.35) \quad [R_K f](\mathbf{Q}, \mathbf{V}) := f(\mathbf{Q}K, \mathbf{V}K).$$

It was already pointed out by Golse [15], that the weak-\* limit of any converging subsequence  $L_{t, \rho_i} f$  ( $\rho_i \rightarrow 0$ ) does not satisfy the linear Boltzmann equation. His arguments use the a priori estimates in [5, 16], and do not require knowledge of the existence of the limit (1.32). The fundamental reason behind the failure of the linear Boltzmann equation is that, perhaps surprisingly,  $\{L_t : t \geq 0\}$  is not a semigroup. We will show in Section 6 how to overcome this problem by considering an extended stochastic process that keeps track not only of position  $\mathbf{Q}$  and current velocity  $\mathbf{V}$ , but also the free path length  $\mathcal{T}$  until the next collision, and

the velocity  $\mathbf{V}_+$  thereafter. We will establish that the extended process is Markovian, and derive the corresponding Fokker-Planck-Kolmogorov equation describing the evolution of the particle density in the extended phase space. A similar approach has recently been explored by Caglioti and Golse in the two-dimensional case [8]. Their result is however conditional on an independence hypothesis, which is equivalent to the Markov property established by our Theorem 1.3 above.

**1.4. Outline of the paper.** The key ingredient in the present work is Theorem 4.8 of [19] (restated as Theorem 2.2 below for general scattering maps), which yields the joint limiting distribution for the free path length and velocity after the next collision, given that the initial position and velocity are taken at random with respect to a fixed probability measure. The proofs of Theorems 1.1 and 1.3 are based on a uniform version of Theorem 2.2, where the fixed probability measures are replaced by certain equismooth families, see Section 2 for details. Section 3 provides technical information on the  $n$ th iterate of the scattering maps, which in conjunction with the uniform version of Theorem 2.2 yields the proof of Theorems 1.1 and 1.3 (Section 4). In Section 5 we prove that the dynamics in the periodic Lorentz gas converges in the Boltzmann-Grad limit to a stochastic process  $\Xi(t)$ , and thus establish Theorems 1.2 and 1.4. We finally derive the substitute for the linear Boltzmann equation in Section 6, by extending  $\Xi(t)$  to a Markov process and calculating its Fokker-Planck-Kolmogorov equation.

## 2. FIRST COLLISION

We begin by reviewing the central result of [19].

**2.1. Location of the first collision.** We fix a lattice  $\mathcal{L} = \mathbb{Z}^d M_0$  with  $M_0 \in \text{SL}(d, \mathbb{R})$ , once and for all. Recall that  $\mathcal{K}_\rho \subset \mathbb{R}^d$  is the complement of the set  $\mathcal{B}_\rho^d + \mathcal{L}$  and that the free path length for the initial condition  $(\mathbf{q}, \mathbf{v}) \in \mathbb{T}^1(\mathcal{K}_\rho)$  is defined as

$$(2.1) \quad \tau_1(\mathbf{q}, \mathbf{v}; \rho) = \inf\{t > 0 : \mathbf{q} + t\mathbf{v} \notin \mathcal{K}_\rho\}.$$

Note that  $\tau_1(\mathbf{q}, \mathbf{v}; \rho) = \infty$  can only happen for a set of  $\mathbf{v}$ 's of measure zero with respect to  $\text{vol}_{\mathbb{S}_1^{d-1}}$ . In fact we have  $\tau_1(\mathbf{q}, \mathbf{v}; \rho) < \infty$  whenever the  $d$  coordinates of  $\mathbf{v}M_0^{-1} \in \mathbb{R}^d$  are linearly independent over  $\mathbb{Q}$  (note that this condition is independent of  $\mathbf{q}$ ), since then each orbit of the linear flow  $\mathbf{x} \mapsto \mathbf{x} + t\mathbf{v}$  is dense on  $\mathbb{R}^d/\mathcal{L}$ .

The position of the particle when hitting the first scatterer is

$$(2.2) \quad \mathbf{q}_1(\mathbf{q}, \mathbf{v}; \rho) := \mathbf{q} + \tau_1(\mathbf{q}, \mathbf{v}; \rho)\mathbf{v}.$$

As in Sec. 1.1 we write  $\mathbf{q}_1(\mathbf{q}, \mathbf{v}; \rho) = \mathbf{m}_1 + \rho\mathbf{w}_1$  with  $\mathbf{m}_1 \in \mathcal{L}$  and  $\mathbf{w}_1 = \mathbf{w}_1(\mathbf{q}, \mathbf{v}; \rho) \in \mathbb{S}_1^{d-1}$ .

Let us fix a map  $K : \mathbb{S}_1^{d-1} \rightarrow \text{SO}(d)$  such that  $\mathbf{v}K(\mathbf{v}) = \mathbf{e}_1$  for all  $\mathbf{v} \in \mathbb{S}_1^{d-1}$ ; we assume that  $K$  is smooth when restricted to  $\mathbb{S}_1^{d-1}$  minus one point. For example, we may choose  $K$  as  $K(\mathbf{e}_1) = I$ ,  $K(-\mathbf{e}_1) = -I$  and

$$(2.3) \quad K(\mathbf{v}) = E\left(-\frac{2 \arcsin(\|\mathbf{v} - \mathbf{e}_1\|/2)}{\|\mathbf{v}_\perp\|} \mathbf{v}_\perp\right)$$

for  $\mathbf{v} \in \mathbb{S}_1^{d-1} \setminus \{\mathbf{e}_1, -\mathbf{e}_1\}$ , where

$$(2.4) \quad \mathbf{v}_\perp := (v_2, \dots, v_d) \in \mathbb{R}^{d-1}, \quad E(\mathbf{w}) = \exp\begin{pmatrix} 0 & \mathbf{w} \\ -\mathbf{w} & 0 \end{pmatrix} \in \text{SO}(d).$$

Then  $K$  is smooth when restricted to  $\mathbb{S}_1^{d-1} \setminus \{-\mathbf{e}_1\}$ .

It is evident that  $-\mathbf{w}_1 K(\mathbf{v}) \in \mathbb{S}_1^{d-1}$ , with the hemisphere

$$(2.5) \quad \mathbb{S}_1^{d-1} = \{\mathbf{v} = (v_1, \dots, v_d) \in \mathbb{S}_1^{d-1} : v_1 > 0\}.$$

Let  $\beta$  be a continuous function  $\mathbb{S}_1^{d-1} \rightarrow \mathbb{R}^d$ . If  $\mathbf{q} \in \mathcal{L}$  we assume that  $(\beta(\mathbf{v}) + \mathbb{R}_{>0}\mathbf{v}) \cap \mathcal{B}_1^d = \emptyset$  for all  $\mathbf{v} \in \mathbb{S}_1^{d-1}$ . We will consider initial conditions of the form  $(\mathbf{q}_{\rho, \beta}(\mathbf{v}), \mathbf{v}) \in \mathbb{T}^1(\mathcal{K}_\rho)$ , where

$\mathbf{q}_{\rho,\beta}(\mathbf{v}) = \mathbf{q} + \rho\beta(\mathbf{v})$  and where  $\mathbf{v}$  is picked at random in  $S_1^{d-1}$ . Note that for fixed  $\mathbf{q}$  and  $\beta$  we indeed have  $\mathbf{q}_{\rho,\beta}(\mathbf{v}) \in \mathcal{K}_\rho$  for all  $\mathbf{v} \in S_1^{d-1}$ , so long as  $\rho$  is sufficiently small.

For the statement of the theorem below, we recall the definition of the manifolds  $X_q(\mathbf{y})$  and  $X(\mathbf{y})$  from [19, Sec. 7]: If  $q \in \mathbb{Z}_{>0}$  and  $\alpha \in q^{-1}\mathbb{Z}^d$ , then we set  $X_q = \Gamma(q) \backslash \mathrm{SL}(d, \mathbb{R})$ , and define, for each  $\mathbf{y} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ ,

$$(2.6) \quad X_q(\mathbf{y}) := \{M \in X_q : \mathbf{y} \in (\mathbb{Z}^d + \alpha)M\}.$$

We also set  $X = \mathrm{ASL}(d, \mathbb{Z}) \backslash \mathrm{ASL}(d, \mathbb{R})$  where  $\mathrm{ASL}(d, \mathbb{R}) = \mathrm{SL}(d, \mathbb{R}) \ltimes \mathbb{R}^d$  is the semidirect product group with multiplication law  $(M, \xi)(M', \xi') = (MM', \xi M' + \xi')$ ; we let  $\mathrm{ASL}(d, \mathbb{R})$  act on  $\mathbb{R}^d$  through  $\mathbf{y} \mapsto \mathbf{y}(M, \xi) := \mathbf{y}M + \xi$ . Now for each  $\mathbf{y} \in \mathbb{R}^d$  we define

$$(2.7) \quad X(\mathbf{y}) := \{g \in X : \mathbf{y} \in \mathbb{Z}^d g\}.$$

The spaces  $X_q(\mathbf{y})$  and  $X(\mathbf{y})$  carry natural probability measures  $\nu_{\mathbf{y}}$  whose properties are discussed in [19, Sec. 7].

We will also use the notation

$$(2.8) \quad \mathbf{x}_\perp = \mathbf{x} - (\mathbf{x} \cdot \mathbf{e}_1)\mathbf{e}_1, \quad \text{for } \mathbf{x} \in \mathbb{R}^d.$$

The following is a restatement of [19, Theorem 4.4].

**Theorem 2.1.** *Fix a lattice  $\mathcal{L} = \mathbb{Z}^d M_0$ . Let  $\mathbf{q} \in \mathbb{R}^d$  and  $\alpha = -\mathbf{q}M_0^{-1}$ . There exists a function  $\Phi_\alpha : \mathbb{R}_{>0} \times (\{0\} \times \mathcal{B}_1^{d-1}) \times (\{0\} \times \mathbb{R}^{d-1}) \rightarrow \mathbb{R}_{\geq 0}$  such that for any Borel probability measure  $\lambda$  on  $S_1^{d-1}$  absolutely continuous with respect to  $\mathrm{vol}_{S_1^{d-1}}$ , any subset  $\mathcal{U} \subset S_1^{d-1}$  with  $\mathrm{vol}_{S_1^{d-1}}(\partial\mathcal{U}) = 0$ , and  $0 \leq \xi_1 < \xi_2$ , we have*

$$(2.9) \quad \lim_{\rho \rightarrow 0} \lambda(\{\mathbf{v} \in S_1^{d-1} : \rho^{d-1} \tau_1(\mathbf{q}_{\rho,\beta}(\mathbf{v}), \mathbf{v}; \rho) \in [\xi_1, \xi_2], -\mathbf{w}_1(\mathbf{q}_{\rho,\beta}(\mathbf{v}), \mathbf{v}; \rho)K(\mathbf{v}) \in \mathcal{U}\}) \\ = \int_{\xi_1}^{\xi_2} \int_{\mathcal{U}_\perp} \int_{S_1^{d-1}} \Phi_\alpha(\xi, \mathbf{w}, (\beta(\mathbf{v})K(\mathbf{v}))_\perp) d\lambda(\mathbf{v}) d\mathbf{w} d\xi,$$

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

$$(2.10) \quad \Phi_\alpha(\xi, \mathbf{w}, \mathbf{z}) = \begin{cases} \nu_{\mathbf{y}}(\{M \in X_q(\mathbf{y}) : (\mathbb{Z}^d + \alpha)M \cap (\mathfrak{Z}(0, \xi, 1) + \mathbf{z}) = \emptyset\}) & \text{if } \alpha \in q^{-1}\mathbb{Z}^d \\ \nu_{\mathbf{y}}(\{g \in X(\mathbf{y}) : \mathbb{Z}^d g \cap (\mathfrak{Z}(0, \xi, 1) + \mathbf{z}) = \emptyset\}) & \text{if } \alpha \notin \mathbb{Q}^d, \end{cases}$$

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

$$(2.11) \quad \mathfrak{Z}(c_1, c_2, \sigma) = \{(x_1, \dots, x_d) \in \mathbb{R}^d : c_1 < x_1 < c_2, \|(x_2, \dots, x_d)\| < \sigma\}.$$

*Remark 2.1.* For  $\alpha \in \mathbb{Q}^d$  the function  $\Phi_\alpha(\xi, \mathbf{w}, \mathbf{z})$  is Borel measurable, and in fact only depends on  $\alpha$  and the four real numbers  $\xi, \|\mathbf{w}\|, \|\mathbf{z}\|, \mathbf{z} \cdot \mathbf{w}$ . Also for  $\alpha \in \mathbb{Q}^d$ , if we restrict to  $\|\mathbf{z}\| \leq 1$  [and if  $d=2$ :  $\mathbf{z} + \mathbf{w} \neq \mathbf{0}$ ], then  $\Phi_\alpha(\xi, \mathbf{w}, \mathbf{z})$  is jointly continuous in the three variables  $\xi, \mathbf{w}, \mathbf{z}$ . If  $\alpha \notin \mathbb{Q}^d$  then  $\Phi_\alpha(\xi, \mathbf{w}, \mathbf{z})$  is everywhere continuous in the three variables, and it is independent of both  $\alpha$  and  $\mathbf{z}$ ; in fact it only depends on  $\xi$  and  $\|\mathbf{w}\|$ . We will therefore set  $\Phi(\xi, \mathbf{w}) := \Phi_\alpha(\xi, \mathbf{w}, \mathbf{z})$  for  $\alpha \notin \mathbb{Q}$ .

We have, for all  $\alpha \in \mathbb{R}^d$  and all  $\mathbf{z} \in \{0\} \times \mathbb{R}^{d-1}$ ,

$$(2.12) \quad \int_0^\infty \int_{\{0\} \times \mathcal{B}_1^{d-1}} \Phi_\alpha(\xi, \mathbf{w}, \mathbf{z}) d\mathbf{w} d\xi = 1.$$

The convergence in this integral is uniform, i.e.

$$(2.13) \quad \int_T^\infty \int_{\{0\} \times \mathcal{B}_1^{d-1}} \Phi_\alpha(\xi, \mathbf{w}, \mathbf{z}) d\mathbf{w} d\xi \rightarrow 0$$

uniformly with respect to  $\alpha$  and  $\mathbf{z}$  as  $T \rightarrow \infty$ . Cf. [19, Remark 4.5, (8.37) and Lemma 8.15].



*Remark 2.2.* We may extend the function  $\Phi(\xi, \mathbf{w})$  to the larger set  $\mathbb{R}_{\geq 0} \times (\{0\} \times \mathcal{B}_1^{d-1})$  by letting  $\Phi(0, \mathbf{w}) := 1$  for all  $\mathbf{w}$ . This definition is natural, since it makes  $\Phi(\xi, \mathbf{w})$  continuous (jointly in both variables) at each point with  $\xi = 0$ . The proof of this is an immediate extension of the discussion in [19, Sec. 8.2]. In the case  $\alpha \in \mathbb{Q}$ , the function  $\Phi_\alpha(\xi, \mathbf{w}, \mathbf{z})$  can be extended to a continuous function on  $\mathbb{R}_{\geq 0} \times (\{0\} \times \mathcal{B}_1^{d-1}) \times (\{0\} \times \mathcal{B}_1^{d-1})$  by setting  $\Phi_\alpha(0, \mathbf{w}, \mathbf{z}) := \kappa_q$ , where  $\kappa_q$  is the density of lattice points visible from  $\alpha$ , as defined in [19, Sec. 2.4]. The proof of this fact follows from the volume computations in [19, Sec. 7.1] and the argument used in [19, Sec. 8.1].

**2.2. Scattering maps.** As indicated above, the results of this paper extend to the case of a Lorentz gas, where the scattering process at a hard sphere is replaced by a smooth radial potential. To obtain the correct scaling in the Boltzmann-Grad limit, we assume the scattering potential is of the form  $V(\mathbf{q}/\rho)$ , where  $V(\mathbf{q})$  has compact support in the unit ball  $\mathcal{B}_1^d$ . We will refer to the rescaled ball  $\mathcal{B}_\rho^d = \rho\mathcal{B}_1^d$  as the *interaction region*.

It is most convenient to phrase the required assumptions in terms of a scattering map that describes the dynamics at each scatterer. Let

$$(2.14) \quad \mathcal{S}_- := \{(\mathbf{v}, \mathbf{w}) \in \mathbb{S}_1^{d-1} \times \mathbb{S}_1^{d-1} : \mathbf{v} \cdot \mathbf{w} < 0\}$$

be the set of *incoming data*  $(\mathbf{v}_-, \mathbf{w}_-)$ , i.e., the velocity and relative position with which the particle enters the interaction region. The corresponding *outgoing data* is parametrized by the set

$$(2.15) \quad \mathcal{S}_+ := \{(\mathbf{v}, \mathbf{w}) \in \mathbb{S}_1^{d-1} \times \mathbb{S}_1^{d-1} : \mathbf{v} \cdot \mathbf{w} > 0\}.$$

We define the scattering map by

$$(2.16) \quad \Theta : \mathcal{S}_- \rightarrow \mathcal{S}_+, \quad (\mathbf{v}_-, \mathbf{w}_-) \mapsto (\mathbf{v}_+, \mathbf{w}_+).$$

In the case of the original Lorentz gas the scattering map is given by specular reflection,

$$(2.17) \quad \Theta(\mathbf{v}, \mathbf{w}) = (\mathbf{v} - 2(\mathbf{v} \cdot \mathbf{w})\mathbf{w}, \mathbf{w});$$

scattering maps corresponding to smooth potentials can be readily obtained from classical results, see e.g. [24, Chapter 5].

In the following we will treat the scattering process as instantaneous. In the case of potentials the particle will of course spend a non-zero amount of time in the interaction region, but—under standard assumptions on the potential—this time will tend to zero when  $\rho \rightarrow 0$ .

Let  $\Theta_1(\mathbf{v}, \mathbf{w}) \in \mathbb{S}_1^{d-1}$  and  $\Theta_2(\mathbf{v}, \mathbf{w}) \in \mathbb{S}_1^{d-1}$  be the projection of  $\Theta(\mathbf{v}, \mathbf{w}) \in \mathbb{S}_1^{d-1}$  onto the first and second component, respectively. We assume throughout this paper that

- (i) the scattering map  $\Theta$  is *spherically symmetric*, i.e., if  $(\mathbf{v}_+, \mathbf{w}_+) = \Theta(\mathbf{v}, \mathbf{w})$  then  $(\mathbf{v}_+K, \mathbf{w}_+K) = \Theta(\mathbf{v}K, \mathbf{w}K)$  for all  $K \in O(d)$ ;
- (ii)  $\mathbf{v}_+$  and  $\mathbf{w}_+$  are contained in the subspace spanned by  $\mathbf{v}$  and  $\mathbf{w}$ ;
- (iii) if  $\mathbf{w} = -\mathbf{v}$  then  $\mathbf{v}_+ = -\mathbf{v}$ ;
- (iv)  $\Theta : \mathcal{S}_- \rightarrow \mathcal{S}_+$  is  $C^1$  and for each fixed  $\mathbf{v} \in \mathbb{S}_1^{d-1}$  the map  $\mathbf{w} \mapsto \Theta_1(\mathbf{v}, \mathbf{w})$  is a  $C^1$  diffeomorphism from  $\{\mathbf{w} \in \mathbb{S}_1^{d-1} : \mathbf{v} \cdot \mathbf{w} < 0\}$  onto some open subset of  $\mathbb{S}_1^{d-1}$ .

The above conditions are for example satisfied for the scattering map of a “muffin-tin” Coulomb potential,  $V(\mathbf{q}) = \alpha \max(\|\mathbf{q}\|^{-1} - 1, 0)$  with  $\alpha \notin \{-2E, 0\}$ , where  $E$  denotes the total energy. Conditions (iii) and (iv) help to simplify the presentation of the proofs, but are not essential. It is for instance not necessary that in (iv) the map  $\mathbf{w} \mapsto \Theta_1(\mathbf{v}, \mathbf{w})$  is invertible as long as it has finitely many pre-images, which allows a larger class of scattering potentials; condition (iii) can be dropped entirely.

We will write  $\varphi(\mathbf{v}, \mathbf{u}) \in [0, \pi]$  for the angle between any two vectors  $\mathbf{v}, \mathbf{u} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ . Using the spherical symmetry and  $\Theta_1(\mathbf{v}, -\mathbf{v}) = -\mathbf{v}$  one sees that there exists a constant  $0 \leq B_\Theta < \pi$  such that for each  $\mathbf{v} \in \mathbb{S}_1^{d-1}$ , the image of the diffeomorphism  $\mathbf{w} \mapsto \Theta_1(\mathbf{v}, \mathbf{w})$  is

$$(2.18) \quad \mathcal{V}_\mathbf{v} := \{\mathbf{u} \in \mathbb{S}_1^{d-1} : \varphi(\mathbf{v}, \mathbf{u}) > B_\Theta\}.$$

Let us write  $\beta_{\mathbf{v}}^- : \mathcal{V}_{\mathbf{v}} \rightarrow \{\mathbf{w} \in \mathbb{S}_1^{d-1} : \mathbf{v} \cdot \mathbf{w} < 0\}$  for the inverse map. Then  $\beta_{\mathbf{v}}^-$  is spherically symmetric in the sense that  $\beta_{\mathbf{v}K}^-(\mathbf{u}K) = \beta_{\mathbf{v}}^-(\mathbf{u})K$  for all  $\mathbf{v} \in \mathbb{S}_1^{d-1}$ ,  $\mathbf{u} \in \mathcal{V}_{\mathbf{v}}$ ,  $K \in O(d)$ , and in particular  $\beta_{\mathbf{v}}^-(\mathbf{u})$  is jointly  $C^1$  in  $\mathbf{v}, \mathbf{u}$ .

We also define

$$(2.19) \quad \beta_{\mathbf{v}}^+(\mathbf{u}) = \Theta_2(\mathbf{v}, \beta_{\mathbf{v}}^-(\mathbf{u})) \quad (\mathbf{v} \in \mathbb{S}_1^{d-1}, \mathbf{u} \in \mathcal{V}_{\mathbf{v}}).$$

The map  $\beta^+$  is also spherically symmetric and jointly  $C^1$  in  $\mathbf{v}, \mathbf{u}$ . In terms of the original scattering situation, the point of our notation is the following: Given any  $\mathbf{v}_-, \mathbf{v}_+ \in \mathbb{S}_1^{d-1}$ , there exist  $\mathbf{w}_-, \mathbf{w}_+ \in \mathbb{S}_1^{d-1}$  such that  $\Theta(\mathbf{v}_-, \mathbf{w}_-) = (\mathbf{v}_+, \mathbf{w}_+)$  if and only if  $\varphi(\mathbf{v}_-, \mathbf{v}_+) > B_{\Theta}$ , and in this case  $\mathbf{w}_-$  and  $\mathbf{w}_+$  are uniquely determined, as  $\mathbf{w}_{\pm} = \beta_{\mathbf{v}_{\pm}}^{\pm}(\mathbf{v}_{\pm})$ .

For example, in the case of specular reflection (2.17) we have  $B_{\Theta} = 0$  and

$$(2.20) \quad \mathbf{w}_+ = \mathbf{w}_- = \frac{\mathbf{v}_+ - \mathbf{v}_-}{\|\mathbf{v}_+ - \mathbf{v}_-\|}.$$

*Remark 2.3.* In the case of specular reflection, the flow  $F_t$  preserves the Liouville measure  $\nu$ , but this does not hold in the case of a general scattering map satisfying (i)–(iv). Indeed, a necessary and sufficient condition for  $F_t$  to preserve  $\nu$  is that the scattering map  $\Theta$  is a diffeomorphism from  $\mathcal{S}_-$  onto  $\mathcal{S}_+$  which carries the volume measure  $|\mathbf{v} \cdot \mathbf{w}| d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{w})$  on  $\mathcal{S}_-$  to  $(\mathbf{v} \cdot \mathbf{w}) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{w})$  on  $\mathcal{S}_+$ . Maps with this property can be classified explicitly: Define the functions  $\vartheta_1, \vartheta_2 : (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow \mathbb{R}$  through

$$(2.21) \quad \Theta_j(\mathbf{e}_1, -(\cos \varphi)\mathbf{e}_1 + (\sin \varphi)\mathbf{e}_2) = -(\cos \vartheta_j(\varphi))\mathbf{e}_1 + (\sin \vartheta_j(\varphi))\mathbf{e}_2.$$

In view of (iii) we may then assume  $\vartheta_j(0) = 0$ , and take  $\vartheta_1, \vartheta_2$  to be continuous. (Then  $\vartheta_1, \vartheta_2$  are both odd and  $C^1$ , and  $\vartheta_1$  is a  $C^1$  diffeomorphism from  $(-\frac{\pi}{2}, \frac{\pi}{2})$  onto  $(B_{\Theta} - \pi, \pi - B_{\Theta})$ .) In this notation, one checks by a computation that  $\Theta$  carries  $|\mathbf{v} \cdot \mathbf{w}| d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{w})$  to  $(\mathbf{v} \cdot \mathbf{w}) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{w})$  if and only if, for all  $\varphi \in (-\frac{\pi}{2}, \frac{\pi}{2}) \setminus \{0\}$ ,

$$(2.22) \quad \left| \frac{\sin(\vartheta_2(\varphi) - \vartheta_1(\varphi))}{\sin \varphi} \right|^{d-2} \cdot \left| \frac{\cos(\vartheta_2(\varphi) - \vartheta_1(\varphi))}{\cos \varphi} \right| \cdot |\vartheta_2'(\varphi) - \vartheta_1'(\varphi)| = 1.$$

This is seen to hold if and only if  $\vartheta_2(\varphi) = \vartheta_1(\varphi) - \varphi$  for all  $\varphi \in (-\frac{\pi}{2}, \frac{\pi}{2})$  or  $\vartheta_2(\varphi) = \vartheta_1(\varphi) + \varphi$  for all  $\varphi \in (-\frac{\pi}{2}, \frac{\pi}{2})$ . (In physical terms, this reflects the preservation of the angular momentum  $\mathbf{w} \wedge \mathbf{v}$ , or its reversal, respectively.) Translating this condition in terms of  $\beta^{\pm}$ , we conclude that  $F_t$  preserves the Liouville measure if and only if

$$(2.23) \quad \beta_{\mathbf{v}_1}^+(\mathbf{v}_2) \equiv \beta_{-\mathbf{v}_2}^-(\mathbf{v}_1) \quad \text{or} \quad \beta_{\mathbf{v}_1}^+(\mathbf{v}_2) \equiv \beta_{\mathbf{v}_2}^-(\mathbf{v}_1)R_{\{\mathbf{v}_2\}^{\perp}},$$

where  $R_{\{\mathbf{v}_2\}^{\perp}} \in O(d)$  denotes orthogonal reflection in the hyperplane  $\{\mathbf{v}_2\}^{\perp} \subset \mathbb{R}^d$ .

### 2.3. Velocity after the first impact.

**Theorem 2.2.** *Let  $\lambda$  be a Borel probability measure on  $\mathbb{S}_1^{d-1}$  absolutely continuous with respect to  $\text{vol}_{\mathbb{S}_1^{d-1}}$ . For any bounded continuous function  $f : \mathbb{S}_1^{d-1} \times \mathbb{R}_{>0} \times \mathbb{S}_1^{d-1} \rightarrow \mathbb{R}$ ,*

$$(2.24) \quad \lim_{\rho \rightarrow 0} \int_{\mathbb{S}_1^{d-1}} f(\mathbf{v}_0, \rho^{d-1} \tau_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho), \mathbf{v}_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda(\mathbf{v}_0) \\ = \int_{\mathbb{S}_1^{d-1}} \int_{\mathbb{R}_{>0}} \int_{\mathbb{S}_1^{d-1}} f(\mathbf{v}_0, \xi, \mathbf{v}_1) p_{\alpha, \beta}(\mathbf{v}_0, \xi, \mathbf{v}_1) d\lambda(\mathbf{v}_0) d\xi d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_1),$$

where the probability density  $p_{\alpha, \beta}$  is defined by

$$(2.25) \quad p_{\alpha, \beta}(\mathbf{v}_0, \xi, \mathbf{v}_1) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_1) = \begin{cases} \Phi_{\alpha}(\xi, \mathbf{w}, (\beta(\mathbf{v}_0)K(\mathbf{v}_0))_{\perp}) d\mathbf{w} & \text{if } \mathbf{v}_1 \in \mathcal{V}_{\mathbf{v}_0} \\ 0 & \text{if } \mathbf{v}_1 \notin \mathcal{V}_{\mathbf{v}_0}, \end{cases}$$

with

$$(2.26) \quad \mathbf{w} = -\beta_{\mathbf{e}_1}^-(\mathbf{v}_1 K(\mathbf{v}_0))_{\perp} \in \{0\} \times \mathbb{B}_1^{d-1}.$$

*Remark 2.4.* The density  $p_{\alpha,\beta}(\mathbf{v}_0, \xi, \mathbf{v}_1)$  is independent of the choice of the function  $K(\mathbf{v}_0)$ , since  $\Phi_\alpha(\xi, \mathbf{w}, \mathbf{z})$  only depends on the four real numbers  $\xi, \|\mathbf{w}\|, \|\mathbf{z}\|, \mathbf{w} \cdot \mathbf{z}$ , cf. Remark 2.1. It also follows from Remark 2.1 and Remark 2.2 that if  $\alpha \notin \mathbb{Q}^d$  then  $p_{\alpha,\beta}(\mathbf{v}_0, \xi, \mathbf{v}_1) =: p(\mathbf{v}_0, \xi, \mathbf{v}_1)$  is independent of  $\alpha, \beta$ , and is continuous at each point  $(\mathbf{v}_0, \xi, \mathbf{v}_1) \in S_1^{d-1} \times \mathbb{R}_{\geq 0} \times S_1^{d-1}$  with  $\mathbf{v}_1 \in \mathcal{V}_{\mathbf{v}_0}$ . The same continuity statement also holds for  $p_{\alpha,\beta}(\mathbf{v}_0, \xi, \mathbf{v}_1)$  if  $\alpha \in \mathbb{Q}^d$  and  $\sup \|\beta\| \leq 1$ , except possibly when  $d = 2, \xi > 0$  and  $\beta_{e_1}^-(\mathbf{v}_1 K(\mathbf{v}_0))_\perp = (\beta(\mathbf{v}_0) K(\mathbf{v}_0))_\perp$ .

*Remark 2.5.* The relationship between  $p_{\alpha,\beta}(\mathbf{v}_0, \xi, \mathbf{v}_1)$  and  $\Phi_\alpha(\xi, \mathbf{w}, \mathbf{z})$  when  $\mathbf{v}_1 \in \mathcal{V}_{\mathbf{v}_0}$  can be expressed more explicitly as

$$(2.27) \quad p_{\alpha,\beta}(\mathbf{v}_0, \xi, \mathbf{v}_1) = J(\mathbf{v}_0, \mathbf{v}_1) \Phi_\alpha\left(\xi, -\beta_{e_1}^-(\mathbf{v}_1 K(\mathbf{v}_0))_\perp, (\beta(\mathbf{v}_0) K(\mathbf{v}_0))_\perp\right),$$

where the factor  $J(\mathbf{v}_0, \mathbf{v}_1) > 0$  is a function of  $\varphi = \varphi(\mathbf{v}_0, -\mathbf{v}_1)$  given by

$$(2.28) \quad J(\mathbf{v}_0, \mathbf{v}_1) = \begin{cases} \left| \frac{\sin \omega(\varphi)}{\sin \varphi} \right|^{d-2} |\omega'(\varphi)| \cos \omega(\varphi) & \text{if } \varphi > 0 \\ |\omega'(0)|^{d-1} & \text{if } \varphi = 0, \end{cases}$$

with  $\omega = \vartheta_1^{-1} : (B_\Theta - \pi, \pi - B_\Theta) \rightarrow (-\frac{\pi}{2}, \frac{\pi}{2})$  being the inverse of the map  $\vartheta_1$  defined in Remark 2.3. In particular in the case of specular reflection (2.17) we have  $\vartheta_1^{-1}(\varphi) = \varphi/2$ , and we recover the formula [19, (4.24)].

*Proof of Theorem 2.2.* The proof follows exactly the same steps as the proof of Theorem 4.8 in [19, Sec. 9.3]: The left hand side of (2.24) equals

$$(2.29) \quad \lim_{\rho \rightarrow 0} \int_{S_1^{d-1}} g(\mathbf{v}_0, \rho^{d-1} \tau_1(\mathbf{q}_{\rho,\beta}(\mathbf{v}_0), \mathbf{v}_0; \rho), \mathbf{w}_1(\mathbf{q}_{\rho,\beta}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda(\mathbf{v}_0),$$

where  $g(\mathbf{v}_0, \xi, \mathbf{w}_1) = f(\mathbf{v}_0, \xi, \Theta_1(\mathbf{v}_0, \mathbf{w}_1))$ . Using [19, Cor. 4.7] (which in fact is a corollary of Theorem 2.1 in the present paper) we obtain

$$(2.30) \quad \begin{aligned} &= \int_{S_1^{d-1}} \int_{\mathbb{R}_{>0}} \int_{S_1^{d-1}} f(\mathbf{v}_0, \xi, \Theta_1(\mathbf{v}_0, -\omega K(\mathbf{v}_0)^{-1})) \\ &\quad \times \Phi_\alpha(\xi, \omega_\perp, (\beta(\mathbf{v}_0) K(\mathbf{v}_0))_\perp) \omega_1 d\lambda(\mathbf{v}_0) d\xi d\text{vol}_{S_1^{d-1}}(\omega). \end{aligned}$$

Now change the order of integration by moving the integral over  $\omega \in S_1^{d-1}$  to the innermost position, and then apply the variable substitution  $\mathbf{v}_1 = \Theta_1(e_1, -\omega) K(\mathbf{v}_0)^{-1} = \Theta_1(\mathbf{v}_0, -\omega K(\mathbf{v}_0)^{-1})$  in the innermost integral; note that this gives a diffeomorphism  $\omega \mapsto \mathbf{v}_1$  from  $S_1^{d-1}$  onto  $\mathcal{V}_{\mathbf{v}_0}$ , with the inverse map given by  $\omega = -\beta_{e_1}^-(\mathbf{v}_1 K(\mathbf{v}_0)) = -\beta_{\mathbf{v}_0}^-(\mathbf{v}_1) K(\mathbf{v}_0)$ . Recalling (2.26) we then see that (2.30) equals the right hand side of (2.24), and we are done.  $\square$

**2.4. A uniform version of Theorem 2.2.** In order to prove Theorem 4.1 we need a version of Theorem 2.2 which is uniform over certain families of  $\beta$ 's,  $f$ 's and  $\lambda$ 's.

Given any subset  $\mathcal{W} \subset S_1^{d-1}$  we let  $\partial_\varepsilon \mathcal{W} \subset S_1^{d-1}$  denote the  $\varepsilon$ -neighborhood of its boundary, i.e.

$$(2.31) \quad \partial_\varepsilon \mathcal{W} := \{\mathbf{v} \in S_1^{d-1} : \exists \mathbf{w} \in \partial \mathcal{W} : \varphi(\mathbf{v}, \mathbf{w}) < \varepsilon\}.$$

**Definition 2.1.** A family  $F$  of Borel subsets of  $S_1^{d-1}$  is called *equismooth* if for every  $\delta > 0$  there is some  $\varepsilon > 0$  such that  $\text{vol}_{S_1^{d-1}}(\partial_\varepsilon \mathcal{W}) < \delta$  for all  $\mathcal{W} \in F$ .

**Definition 2.2.** A family  $F$  of measures on  $S_1^{d-1}$  is called *equismooth* if there exist an equicontinuous and uniformly bounded family  $F'$  of functions from  $S_1^{d-1}$  to  $\mathbb{R}_{\geq 0}$  and an equismooth family  $F''$  of Borel subsets of  $S_1^{d-1}$ , such that each  $\mu \in F$  can be expressed as  $\mu = (g \cdot \text{vol}_{S_1^{d-1}})|_{\mathcal{W}}$  for some  $g \in F', \mathcal{W} \in F''$ .

**Theorem 2.3.** Fix a lattice  $\mathcal{L} = \mathbb{Z}^d M_0$  and a point  $\mathbf{q} \in \mathbb{R}^d$ , and write  $\boldsymbol{\alpha} = -\mathbf{q}M_0^{-1}$ . Let  $F_1$  be an equismooth family of probability measures on  $S_1^{d-1}$ , let  $F_2$  be a uniformly bounded and equicontinuous family of functions  $f : S_1^{d-1} \times \mathbb{R}_{>0} \times S_1^{d-1} \rightarrow \mathbb{R}$ , and let  $F_3$  be a uniformly bounded and equicontinuous family of functions  $\boldsymbol{\beta} : S_1^{d-1} \rightarrow \mathbb{R}^d$  such that if  $\mathbf{q} \in \mathcal{L}$  then  $(\boldsymbol{\beta}(\mathbf{v}) + \mathbb{R}_{>0}\mathbf{v}) \cap \mathcal{B}_1^d = \emptyset$  for all  $\boldsymbol{\beta} \in F_3$ ,  $\mathbf{v} \in S_1^{d-1}$ . Then the limit relation

$$(2.32) \quad \lim_{\rho \rightarrow 0} \int_{S_1^{d-1}} f(\mathbf{v}_0, \rho^{d-1} \tau_1(\mathbf{q}_{\rho, \boldsymbol{\beta}}(\mathbf{v}_0), \mathbf{v}_0; \rho), \mathbf{v}_1(\mathbf{q}_{\rho, \boldsymbol{\beta}}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda(\mathbf{v}_0) \\ = \int_{S_1^{d-1}} \int_{\mathbb{R}_{>0}} \int_{S_1^{d-1}} f(\mathbf{v}_0, \xi, \mathbf{v}_1) p_{\boldsymbol{\alpha}, \boldsymbol{\beta}}(\mathbf{v}_0, \xi, \mathbf{v}_1) d\lambda(\mathbf{v}_0) d\xi d\text{vol}_{S_1^{d-1}}(\mathbf{v}_1)$$

holds uniformly with respect to all  $\lambda \in F_1$ ,  $f \in F_2$ ,  $\boldsymbol{\beta} \in F_3$ .

This uniform version of Theorem 2.2 will actually be derived as a corollary of Theorem 2.2. We first need a lemma. Let  $\mathfrak{U}_\eta$  denote a small neighborhood of the boundary in  $S_1^{d-1}$ :

$$(2.33) \quad \mathfrak{U}_\eta := \{\mathbf{w} \in S_1^{d-1} : \varphi(\mathbf{w}, \mathbf{e}_1) > \frac{\pi}{2} - \eta\} \quad (0 < \eta < 1).$$

**Lemma 2.4.** Given  $B > 0$  there exists some number  $\rho_0(B, \mathcal{L}, \mathbf{q}) > 0$  such that for every  $0 < \rho < \rho_0(B, \mathcal{L}, \mathbf{q})$ ,  $\mathbf{v} \in S_1^{d-1}$ ,  $K \in \text{SO}(d)$ ,  $\boldsymbol{\beta}, \boldsymbol{\beta}' \in \mathbb{R}^d$  and  $0 < \eta < \frac{1}{10}$  such that

$$(2.34) \quad \mathbf{v}K = \mathbf{e}_1, \quad \|\boldsymbol{\beta}\| \leq B, \quad \|\boldsymbol{\beta} - \boldsymbol{\beta}'\| \leq \eta, \quad \boldsymbol{\beta} \cdot \mathbf{v} \geq 2, \quad \boldsymbol{\beta}' \cdot \mathbf{v} \geq 2,$$

one of the following holds:

- (i)  $\tau_1(\mathbf{q} + \rho\boldsymbol{\beta}, \mathbf{v}, \rho) = \tau_1(\mathbf{q} + \rho\boldsymbol{\beta}', \mathbf{v}, \rho) = \infty$ ,
- (ii)  $-\mathbf{w}_1(\mathbf{q} + \rho\boldsymbol{\beta}, \mathbf{v}, (1 + \eta)\rho)K \in \mathfrak{U}_{3\sqrt{\eta}}$ ,
- (iii)  $|\tau_1(\mathbf{q} + \rho\boldsymbol{\beta}', \mathbf{v}; \rho) - \tau_1(\mathbf{q} + \rho\boldsymbol{\beta}, \mathbf{v}; \rho)| < 3\rho\sqrt{\eta}$   
and  $\|\mathbf{w}_1(\mathbf{q} + \rho\boldsymbol{\beta}', \mathbf{v}; \rho) - \mathbf{w}_1(\mathbf{q} + \rho\boldsymbol{\beta}, \mathbf{v}; \rho)\| < \sqrt{2\eta}$ .

*Remark 2.6.* The conditions  $\boldsymbol{\beta} \cdot \mathbf{v} \geq 2$ ,  $\boldsymbol{\beta}' \cdot \mathbf{v} \geq 2$  are only needed in the case  $\mathbf{q} \in \mathcal{L}$ , so as to guarantee that the rays  $\mathbf{q} + \rho\boldsymbol{\beta} + \mathbb{R}_{>0}\mathbf{v}$  and  $\mathbf{q} + \rho\boldsymbol{\beta}' + \mathbb{R}_{>0}\mathbf{v}$  lie outside the ball  $\mathbf{q} + \mathcal{B}_\rho^d$ , and also outside the ball  $\mathbf{q} + \mathcal{B}_{(1+\eta)\rho}^d$ .

*Proof of Lemma 2.4.* We choose  $\rho_0 = \rho_0(B, \mathcal{L}, \mathbf{q}) > 0$  so small that all the balls  $\mathbf{m} + \mathcal{B}_{2\rho_0}^d$  ( $\mathbf{m} \in \mathcal{L} \setminus \{\mathbf{q}\}$ ) have disjoint closures, and are each disjoint from  $\mathbf{q} + \mathcal{B}_{(B+1)\rho_0}^d$ . Now fix any  $\rho, \mathbf{v}, K, \boldsymbol{\beta}, \boldsymbol{\beta}', \eta$  as in the statement of the lemma.

After an auxiliary rotation we may assume  $\mathbf{v} = \mathbf{e}_1$  and  $K = 1_d$ . Write  $\tau = \tau_1(\mathbf{q} + \rho\boldsymbol{\beta}, \mathbf{e}_1; \rho)$  and  $\mathbf{w} = \mathbf{w}_1(\mathbf{q} + \rho\boldsymbol{\beta}, \mathbf{e}_1; \rho)$ , let  $\tau'$  and  $\mathbf{w}'$  be the corresponding data for  $\boldsymbol{\beta}'$ , and write  $\tilde{\rho} = (1 + \eta)\rho$ ,  $\tilde{\tau} = \tau_1(\mathbf{q} + \rho\boldsymbol{\beta}, \mathbf{e}_1, \tilde{\rho})$  and  $\tilde{\mathbf{w}} = \mathbf{w}_1(\mathbf{q} + \rho\boldsymbol{\beta}, \mathbf{e}_1, \tilde{\rho})$ . By our assumption on  $\rho_0$  we see that  $\tau, \tau', \tilde{\tau}$  are well-defined numbers in  $\mathbb{R}_{>0} \cup \{\infty\}$ .

For any ray  $\gamma \subset \mathbb{R}^d$  and any point  $\mathbf{m} \in \mathbb{R}^d$  we write  $\delta(\gamma, \mathbf{m}) = \inf_{\mathbf{p} \in \gamma} \|\mathbf{p} - \mathbf{m}\|$  for the distance between  $\gamma$  and  $\mathbf{m}$ . If  $\tilde{\tau} = \infty$  then  $\delta(\mathbf{q} + \rho\boldsymbol{\beta} + \mathbb{R}_{>0}\mathbf{e}_1, \mathbf{m}) \geq \tilde{\rho}$  for all  $\mathbf{m} \in \mathcal{L}$ , and thus also  $\delta(\mathbf{q} + \rho\boldsymbol{\beta}' + \mathbb{R}_{>0}\mathbf{e}_1, \mathbf{m}) \geq \tilde{\rho} - \rho\|\boldsymbol{\beta} - \boldsymbol{\beta}'\| \geq \rho$  for all  $\mathbf{m} \in \mathcal{L}$ , so that  $\tau = \tau' = \infty$ . Hence from now on we may assume  $\tilde{\tau} < \infty$ .

By the definition of  $\tilde{\tau}$ ,  $\tilde{\mathbf{w}}$ , there is a unique  $\tilde{\mathbf{m}} \in \mathcal{L} \setminus \{\mathbf{q}\}$  such that

$$(2.35) \quad \mathbf{q} + \rho\boldsymbol{\beta} + \tilde{\tau}\mathbf{e}_1 = \tilde{\mathbf{m}} + \tilde{\rho}\tilde{\mathbf{w}},$$

and the ray  $\mathbf{q} + \rho\boldsymbol{\beta} + \mathbb{R}_{>0}\mathbf{e}_1$  does not intersect any previous ball, i.e.

$$(2.36) \quad \forall \mathbf{m} \in \mathcal{L} : \quad \mathbf{m} \cdot \mathbf{e}_1 < \tilde{\mathbf{m}} \cdot \mathbf{e}_1 \implies \delta(\mathbf{q} + \rho\boldsymbol{\beta} + \mathbb{R}_{>0}\mathbf{e}_1, \mathbf{m}) \geq \tilde{\rho}.$$

Now assume  $-\tilde{\mathbf{w}} \notin \mathfrak{U}_{3\sqrt{\eta}}$ . We will prove that (iii) holds. From (2.35) we get

$$(2.37) \quad \delta(\mathbf{q} + \rho\boldsymbol{\beta} + \mathbb{R}_{>0}\mathbf{e}_1, \tilde{\mathbf{m}}) = \tilde{\rho}\|\tilde{\mathbf{w}}_\perp\| \leq (1 + \eta) \cos(3\sqrt{\eta})\rho < (1 - \eta)\rho$$

and thus also

$$(2.38) \quad \delta(\mathbf{q} + \rho\boldsymbol{\beta}' + \mathbb{R}_{>0}\mathbf{e}_1, \tilde{\mathbf{m}}) < \rho.$$

Hence both the rays  $\mathbf{q} + \rho\boldsymbol{\beta} + \mathbb{R}_{>0}\mathbf{e}_1$  and  $\mathbf{q} + \rho\boldsymbol{\beta}' + \mathbb{R}_{>0}\mathbf{e}_1$  intersect the ball  $\tilde{\mathbf{m}} + \mathcal{B}_\rho^d$ . Note also that  $\delta(\mathbf{q} + \rho\boldsymbol{\beta} + \mathbb{R}_{>0}\mathbf{e}_1, \mathbf{m}) > \tilde{\rho}$  and thus  $\delta(\mathbf{q} + \rho\boldsymbol{\beta}' + \mathbb{R}_{>0}\mathbf{e}_1, \mathbf{m}) > \rho$  for all  $\mathbf{m} \in \mathcal{L}$  with  $\mathbf{m} \cdot \mathbf{e}_1 < \tilde{\mathbf{m}} \cdot \mathbf{e}_1$ , by (2.36). Hence neither  $\tau$  nor  $\tau'$  can arise from any previous intersection, i.e. we must have  $\mathbf{q} + \rho\boldsymbol{\beta} + \tau\mathbf{e}_1 = \tilde{\mathbf{m}} + \rho\mathbf{w}$  and  $\mathbf{q} + \rho\boldsymbol{\beta}' + \tau'\mathbf{e}_1 = \tilde{\mathbf{m}} + \rho\mathbf{w}'$ . This implies

$$(2.39) \quad \|(\mathbf{w}' - \mathbf{w})_\perp\| = \|(\boldsymbol{\beta}' - \boldsymbol{\beta})_\perp\| \leq \eta$$

and thus, using  $-\mathbf{w}', -\mathbf{w} \in S_1^{d-1}$ , we conclude  $\|\mathbf{w}' - \mathbf{w}\| < \sqrt{2}\eta$ . Finally from  $(\tau - \tau')\mathbf{e}_1 = \rho(\mathbf{w} - \mathbf{w}') - \rho(\boldsymbol{\beta} - \boldsymbol{\beta}')$  we conclude

$$(2.40) \quad |\tau - \tau'| < \rho(\sqrt{2}\eta + \eta) < 3\rho\sqrt{\eta}.$$

Hence (iii) holds.  $\square$

*Proof of Theorem 2.3. Step 1: Proof in the case that  $F_1$  and  $F_2$  are singleton sets,  $F_1 = \{\lambda\}$  and  $F_2 = \{f\}$ , and furthermore  $f$  is uniformly continuous.* Let  $\varepsilon > 0$  be given. Let  $R = 1 + \sup_{S_1 \times \mathbb{R}_{>0} \times S_1} |f|$ . Set  $B = 3 + \sup_{\boldsymbol{\beta} \in F_3} \sup_{S_1^{d-1}} \|\boldsymbol{\beta}\|$  and

$$(2.41) \quad F'_3 = \{\mathbf{v} \mapsto \boldsymbol{\beta}(\mathbf{v}) + B\mathbf{v} : \boldsymbol{\beta} \in F_3\}.$$

Then  $\boldsymbol{\beta}_1(\mathbf{v}) \cdot \mathbf{v} \geq 3$  and  $\|\boldsymbol{\beta}_1(\mathbf{v})\| < 2B$  hold for all  $\boldsymbol{\beta}_1 \in F'_3$  and  $\mathbf{v} \in S_1^{d-1}$ .

Because of Remark 2.1 and  $0 \leq \Phi_\alpha(\xi, \mathbf{w}, \mathbf{z}) \leq 1$ , the following limit holds uniformly with respect to all  $\mathbf{z} \in \{0\} \times \mathbb{R}^{d-1}$ :

$$(2.42) \quad \lim_{\delta \rightarrow 0} \left( \int_0^\delta + \int_{\delta^{-1}}^\infty \right) \int_{\{0\} \times \mathcal{B}_1^{d-1}} \Phi_\alpha(\xi, \mathbf{w}, \mathbf{z}) d\mathbf{w} d\xi = 0.$$

Furthermore, by [19, Lemma 8.8], for any given  $\delta > 0$  there is some  $0 < \eta < \frac{1}{10}$  so small that

$$(2.43) \quad \int_{\{0\} \times \mathcal{B}_1^{d-1}} |\Phi_\alpha(\xi, \mathbf{w}, \mathbf{z}_1) - \Phi_\alpha(\xi, \mathbf{w}, \mathbf{z}_2)| d\mathbf{w} < \frac{\delta\varepsilon}{R}$$

holds for all  $\xi \in [T^{-1}, T]$  and all  $\mathbf{z}_1, \mathbf{z}_2 \in \{0\} \times \mathcal{B}_{2B}^{d-1}$  with  $\|\mathbf{z}_1 - \mathbf{z}_2\| \leq \eta$ . As a consequence of (2.42) and (2.43), we can now fix  $\eta > 0$  so small that, for any two (measurable) functions  $\mathbf{z}_1, \mathbf{z}_2 : S_1^{d-1} \rightarrow \{0\} \times \mathcal{B}_{2B}^{d-1}$  satisfying  $\|\mathbf{z}_1(\mathbf{v}) - \mathbf{z}_2(\mathbf{v})\| \leq \eta$  ( $\forall \mathbf{v} \in S_1^{d-1}$ ), we have

$$(2.44) \quad \int_0^\infty \int_{\{0\} \times \mathcal{B}_1^{d-1}} \int_{S_1^{d-1}} |\Phi_\alpha(\xi, \mathbf{w}, \mathbf{z}_1(\mathbf{v})) - \Phi_\alpha(\xi, \mathbf{w}, \mathbf{z}_2(\mathbf{v}))| d\lambda(\mathbf{v}) d\mathbf{w} d\xi < \frac{2\varepsilon}{R}.$$

Again using (2.42) (and  $\lim_{\eta \rightarrow 0} \text{vol}_{\mathbb{R}^{d-1}}((\mathcal{U}_{3\sqrt{\eta}})_\perp) = 0$ ), we see that by possibly further shrinking  $\eta$ , we may also assume that

$$(2.45) \quad \int_0^\infty \int_{(\mathcal{U}_{3\sqrt{\eta}})_\perp} \int_{S_1^{d-1}} \Phi_\alpha(\xi, \mathbf{w}, (\boldsymbol{\beta}(\mathbf{v})K(\mathbf{v}))_\perp) d\lambda(\mathbf{v}) d\mathbf{w} d\xi < \frac{2\varepsilon}{R}$$

holds for all (measurable) functions  $\boldsymbol{\beta} : S_1^{d-1} \rightarrow \mathbb{R}^d$ .

Since  $f$  is uniformly continuous there is some  $\delta > 0$  such that  $|f(\mathbf{v}_0, \xi, \mathbf{v}_1) - f(\mathbf{v}_0, \xi', \mathbf{v}'_1)| < \varepsilon$  holds whenever  $|\xi - \xi'| < \delta$  and  $\|\mathbf{v}_1 - \mathbf{v}'_1\| < \delta$ . Furthermore, our assumptions on the scattering function  $\Theta$  imply that  $\Theta_1(\mathbf{e}_1, \mathbf{w})$  is uniformly continuous in  $\mathbf{w} \in -S_1^{d-1}$ ; hence by possibly further shrinking  $\eta$  we may assume that

$$(2.46) \quad \forall \mathbf{w}, \mathbf{w}' \in -S_1^{d-1} : \|\mathbf{w} - \mathbf{w}'\| < \sqrt{2}\eta \implies \|\Theta_1(\mathbf{e}_1, \mathbf{w}) - \Theta_1(\mathbf{e}_1, \mathbf{w}')\| < \delta.$$

Since the family  $F'_3$  is uniformly bounded and equicontinuous, there is a finite set  $F''_3$  of continuous functions  $\boldsymbol{\beta}_2 : S_1^{d-1} \rightarrow \mathcal{B}_{2B}^d$  such that for each  $\boldsymbol{\beta}_1 \in F'_3$  there is some  $\boldsymbol{\beta}_2 \in F''_3$  with  $\sup_{S_1^{d-1}} \|\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2\| < \eta$ .

Given  $\beta_2 \in F_3''$ , let us write  $\tilde{\beta}_2$  for the function  $\tilde{\beta}_2(\mathbf{v}) := (1 + \eta)^{-1}\beta_2(\mathbf{v})$ . Now since  $F_3''$  is finite, by Theorem 2.2 and Theorem 2.1 combined with (2.45) there is some  $\rho_0 > 0$  such that the following two inequalities hold for all  $\rho \in (0, 2\rho_0)$  and all  $\beta_2 \in F_3''$ :

$$(2.47) \quad \left| \int_{S_1^{d-1}} f(\mathbf{v}_0, \rho^{d-1}\tau_1(\mathbf{q}_{\rho, \beta_2}(\mathbf{v}_0), \mathbf{v}_0; \rho), \mathbf{v}_1(\mathbf{q}_{\rho, \beta_2}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda(\mathbf{v}_0) \right. \\ \left. - \int_{S_1^{d-1}} \int_{\mathbb{R}_{>0}} \int_{S_1^{d-1}} f(\mathbf{v}_0, \xi, \mathbf{v}_1) p_{\alpha, \beta_2}(\mathbf{v}_0, \xi, \mathbf{v}_1) d\lambda(\mathbf{v}_0) d\xi d\text{vol}_{S_1^{d-1}}(\mathbf{v}_1) \right| < \varepsilon;$$

$$(2.48) \quad \lambda(\{\mathbf{v}_0 \in S_1^{d-1} : -\mathbf{w}_1(\mathbf{q}_{\rho, \tilde{\beta}_2}(\mathbf{v}_0), \mathbf{v}_0; \rho)K(\mathbf{v}_0) \in \mathfrak{U}_{3\sqrt{\eta}}\}) < \frac{3\varepsilon}{R}.$$

By possibly further shrinking  $\rho_0$  we may assume that  $(3\sqrt{\eta} + B)\rho_0^d < \delta$ , and  $\rho_0 < \rho_0(2B, \mathcal{L}, \mathbf{q})$  (the constant from Lemma 2.4), and that every ball  $\mathbf{m} + \mathcal{B}_{\rho_0}^d$  with  $\mathbf{m} \in \mathcal{L} \setminus \{\mathbf{q}\}$  lies outside  $\mathbf{q} + \mathcal{B}_{2B\rho_0}^d$ .

We now claim that for all  $\rho \in (0, \rho_0)$  and all  $\beta \in F_3$  we have

$$(2.49) \quad \left| \int_{S_1^{d-1}} f(\mathbf{v}_0, \rho^{d-1}\tau_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho), \mathbf{v}_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda(\mathbf{v}_0) \right. \\ \left. - \int_{S_1^{d-1}} \int_{\mathbb{R}_{>0}} \int_{S_1^{d-1}} f(\mathbf{v}_0, \xi, \mathbf{v}_1) p_{\alpha, \beta}(\mathbf{v}_0, \xi, \mathbf{v}_1) d\lambda(\mathbf{v}_0) d\xi d\text{vol}_{S_1^{d-1}}(\mathbf{v}_1) \right| < 10\varepsilon.$$

This will complete the proof of Step 1, since  $\varepsilon > 0$  was arbitrary.

To prove (2.49), let  $\rho \in (0, \rho_0)$  and  $\beta \in F_3$  be given. Set  $\beta_1(\mathbf{v}) \equiv \beta(\mathbf{v}) + B\mathbf{v}$ ; then  $\beta_1 \in F_3'$ , and hence there is some  $\beta_2 \in F_3''$  such that  $\sup_{S_1^{d-1}} \|\beta_1 - \beta_2\| < \eta$ . Now for each  $\mathbf{v}_0 \in S_1^{d-1}$ , Lemma 2.4 applies to  $\rho$ ,  $\mathbf{v}_0$ ,  $K(\mathbf{v}_0)$ ,  $\beta_2(\mathbf{v}_0)$ ,  $\beta_1(\mathbf{v}_0)$ ,  $\eta$ , and hence either

- (i)  $\tau_1(\mathbf{q}_{\rho, \beta_2}(\mathbf{v}_0), \mathbf{v}_0, \rho) = \tau_1(\mathbf{q}_{\rho, \beta_1}(\mathbf{v}_0), \mathbf{v}_0, \rho) = \infty$ , or
- (ii)  $-\mathbf{w}_1(\mathbf{q}_{\tilde{\rho}, \tilde{\beta}_2}(\mathbf{v}_0), \mathbf{v}_0, \tilde{\rho})K(\mathbf{v}_0) \in \mathfrak{U}_{3\sqrt{\eta}}$  with  $\tilde{\rho} = \rho(1 + \eta)$ , or
- (iii)  $|\tau_1(\mathbf{q}_{\rho, \beta_1}(\mathbf{v}_0), \mathbf{v}_0; \rho) - \tau_1(\mathbf{q}_{\rho, \beta_2}(\mathbf{v}_0), \mathbf{v}_0; \rho)| < 3\rho\sqrt{\eta}$   
and  $\|\mathbf{w}_1(\mathbf{q}_{\rho, \beta_1}(\mathbf{v}_0), \mathbf{v}_0; \rho) - \mathbf{w}_1(\mathbf{q}_{\rho, \beta_2}(\mathbf{v}_0), \mathbf{v}_0; \rho)\| < \sqrt{2\eta}$ .

But

$$(2.50) \quad \tau_1(\mathbf{q}_{\rho, \beta_1}(\mathbf{v}_0), \mathbf{v}_0; \rho) = \tau_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho) - B\rho$$

and

$$(2.51) \quad \mathbf{w}_1(\mathbf{q}_{\rho, \beta_1}(\mathbf{v}_0), \mathbf{v}_0; \rho) = \mathbf{w}_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho),$$

since every ball  $\mathbf{m} + \mathcal{B}_{\rho}^d$  ( $\mathbf{m} \in \mathcal{L} \setminus \{\mathbf{q}\}$ ) lies outside  $\mathbf{q} + \mathcal{B}_{2B\rho}^d$  and  $(\beta(\mathbf{v}_0) + \mathbb{R}_{>0}\mathbf{v}_0) \cap \mathcal{B}_1^d = \emptyset$  if  $\mathbf{q} \in \mathcal{L}$ . Hence if (iii) holds then we have (using our assumptions on  $\rho_0$ )

$$(2.52) \quad |\rho^{d-1}\tau_1(\mathbf{q}_{\rho, \beta_2}(\mathbf{v}_0), \mathbf{v}_0; \rho) - \rho^{d-1}\tau_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho)| < (3\sqrt{\eta} + B)\rho^d < \delta$$

and also, via (2.46),

$$(2.53) \quad \|\mathbf{v}_1(\mathbf{q}_{\rho, \beta_2}(\mathbf{v}_0), \mathbf{v}_0; \rho) - \mathbf{v}_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho)\| < \delta.$$

Therefore, by our choice of  $\delta$ ,

$$(2.54) \quad \left| f(\mathbf{v}_0, \rho^{d-1}\tau_1(\mathbf{q}_{\rho, \beta_2}(\mathbf{v}_0), \mathbf{v}_0; \rho), \mathbf{v}_1(\mathbf{q}_{\rho, \beta_2}(\mathbf{v}_0), \mathbf{v}_0; \rho)) \right. \\ \left. - f(\mathbf{v}_0, \rho^{d-1}\tau_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho), \mathbf{v}_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho)) \right| < \varepsilon.$$

Regarding the other possibilities (i) and (ii), the set  $\{\mathbf{v}_0 \in S_1^{d-1} : \tau_1(\mathbf{q}_{\rho, \beta_1}(\mathbf{v}_0), \mathbf{v}_0, \rho) = \infty\}$  has measure zero with respect to  $\text{vol}_{S_1^{d-1}}$  (see Sec. 2.1), and hence also with respect to  $\lambda$ . Furthermore we have

$$(2.55) \quad \lambda(\{\mathbf{v}_0 \in S_1^{d-1} : -\mathbf{w}_1(\mathbf{q}_{\tilde{\rho}, \tilde{\beta}_2}(\mathbf{v}_0), \mathbf{v}_0, \tilde{\rho})K(\mathbf{v}_0) \in \mathfrak{U}_{3\sqrt{\eta}}\}) < \frac{3\varepsilon}{R}$$

using (2.48) and  $\bar{\rho} < 2\rho_0$ . Hence, since the difference in (2.54) is *always*  $< 2R$ , we conclude

$$(2.56) \quad \left| \int_{S_1^{d-1}} f(\mathbf{v}_0, \rho^{d-1} \tau_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho), \mathbf{v}_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda(\mathbf{v}_0) \right. \\ \left. - \int_{S_1^{d-1}} f(\mathbf{v}_0, \rho^{d-1} \tau_1(\mathbf{q}_{\rho, \beta_2}(\mathbf{v}_0), \mathbf{v}_0; \rho), \mathbf{v}_1(\mathbf{q}_{\rho, \beta_2}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda(\mathbf{v}_0) \right| < 7\varepsilon.$$

Using (2.44) and the definition of  $p_{\alpha, \beta}$  in Theorem 2.2 we also get

$$(2.57) \quad \int_{S_1^{d-1}} \int_{\mathbb{R}_{>0}} \int_{S_1^{d-1}} \left| p_{\alpha, \beta_2}(\mathbf{v}_0, \xi, \mathbf{v}_1) - p_{\alpha, \beta}(\mathbf{v}_0, \xi, \mathbf{v}_1) \right| d\lambda(\mathbf{v}_0) d\xi d\text{vol}_{S_1^{d-1}}(\mathbf{v}_1) < \frac{2\varepsilon}{R}.$$

Combining (2.47), (2.56) and (2.57), we conclude that (2.49) holds, thus completing the proof of Step 1.

*Step 2: Proof of Theorem 2.3 in the case that  $F_1$  is a singleton set,  $F_1 = \{\lambda\}$ , and  $F_2$  has the property that  $\cup_{f \in F_2} \text{supp}(f)$  is compact.* In this case, since the family  $F_2$  is also uniformly bounded and equicontinuous, for any given  $\varepsilon > 0$  there exists a finite subfamily  $F_2' \subset F_2$  such that for every  $f \in F_2$  there is some  $f_0 \in F_2'$  with  $\sup_{S_1 \times \mathbb{R}_{>0} \times S_1} |f - f_0| < \varepsilon$ . Hence the desired result follows from Step 1 by a standard approximation argument.

*Step 3: Proof of Theorem 2.3 in the case that  $F_1$  is a singleton set,  $F_1 = \{\lambda\}$ .* We first prove that the following limit holds uniformly with respect to all  $\beta \in F_3$ :

$$(2.58) \quad \lim_{(\rho, \delta) \rightarrow (0,0)} \lambda(\{\mathbf{v}_0 \in S_1^{d-1} : \rho^{d-1} \tau_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho) \notin [\delta, \delta^{-1}]\}) = 0.$$

To prove this, for each  $0 < \delta < \frac{1}{2}$  we fix  $c_\delta : \mathbb{R}_{>0} \rightarrow [0, 1]$  to be some continuous function with  $\chi_{[2\delta, (2\delta)^{-1}]} \leq c_\delta \leq \chi_{[\delta, \delta^{-1}]}$ , and view  $c_\delta$  as a function on  $S_1^{d-1} \times \mathbb{R}_{>0} \times S_1^{d-1}$  via projection onto the second component. Applying Step 1 for the families  $F_1, F_2' = \{c_\delta\}, F_3$ , one proves that for any  $\varepsilon, \delta > 0$  there exists some  $\rho_0 > 0$  such that

$$(2.59) \quad \lambda(\{\mathbf{v}_0 \in S_1^{d-1} : \rho^{d-1} \tau_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho) \notin [\delta, \delta^{-1}]\}) \\ < \varepsilon + \left( \int_0^{2\delta} + \int_{(2\delta)^{-1}}^\infty \right) \int_{\{0\} \times \mathcal{B}_1^{d-1}} \int_{S_1^{d-1}} \Phi_\alpha(\xi, \mathbf{w}, (\beta(\mathbf{v})K(\mathbf{v}))_\perp) d\lambda(\mathbf{v}) d\mathbf{w} d\xi$$

for all  $\rho \in (0, \rho_0)$  and all  $\beta \in F_3$ . Thus (2.58) follows using (2.42) and monotonicity in  $\delta$ .

Now note that for each fixed  $\delta$ , the family  $F_2'' := \{c_\delta \cdot f : f \in F_2\}$  is uniformly bounded and equicontinuous, and all functions in  $F_2''$  have support contained in the compact set  $S_1^{d-1} \times [\delta, \delta^{-1}] \times S_1^{d-1}$ . Hence Step 2 applies to the families  $F_1, F_2'', F_3$ . Now the desired claim follows upon letting  $\delta \rightarrow 0$  and using (2.42) and (2.58) to control the error caused by replacing  $f$  by  $c_\delta \cdot f$  in both sides of (2.32).

*Step 4: Proof of Theorem 2.3 in the general case.* Since  $F_1$  is equismooth there exist an equicontinuous and uniformly bounded family  $F'$  of functions from  $S_1^{d-1}$  to  $\mathbb{R}_{\geq 0}$  and an equismooth family  $F''$  of Borel subsets of  $S_1^{d-1}$  such that each  $\lambda \in F_1$  can be expressed as  $\lambda = (g \cdot \text{vol}_{S_1^{d-1}})|_{\mathcal{W}}$  with  $g \in F', \mathcal{W} \in F''$ . Set  $R = 1 + \sup_{f \in F_2} \sup_{S_1 \times \mathbb{R}_{>0} \times S_1} |f|$  and  $S = 1 + \sup_{g \in F'} \sup_{S_1^{d-1}} g$ .

Let  $\varepsilon > 0$  be given. Fix  $\eta > 0$  so small that

$$(2.60) \quad \forall g \in F' : \forall \mathbf{v}_0, \mathbf{v}'_0 \in S_1^{d-1} : \varphi(\mathbf{v}_0, \mathbf{v}'_0) < \eta \implies \|g(\mathbf{v}_0) - g(\mathbf{v}'_0)\| < \frac{\varepsilon}{R \text{vol}(S_1^{d-1})}$$

and

$$(2.61) \quad \forall \mathcal{W} \in F'' : \text{vol}_{S_1^{d-1}}(\partial_\eta \mathcal{W}) < \frac{\varepsilon}{RS}.$$

Choose a partition  $S_1^{d-1} = \bigsqcup_{j=1}^n D_j$  of  $S_1^{d-1}$  into Borel subsets  $D_1, \dots, D_n$  which are pathwise connected, of positive volume and each has diameter  $< \eta$  (with respect the metric  $\varphi$  on  $S_1^{d-1}$ ). Let  $\lambda_j$  be the probability measure  $\lambda_j = \text{vol}_{S_1^{d-1}}(D_j)^{-1}(\text{vol}_{S_1^{d-1}})_{|D_j}$  for  $j = 1, \dots, n$ . By Step 3 applied  $n$  times, there exists some  $\rho_0 > 0$  such that for all  $\rho \in (0, \rho_0)$ ,  $j \in \{1, \dots, n\}$ ,  $f \in F_2$  and  $\beta \in F_3$ , we have

$$(2.62) \quad \left| \int_{S_1^{d-1}} f(\mathbf{v}_0, \rho^{d-1} \tau_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho), \mathbf{v}_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda_j(\mathbf{v}_0) \right. \\ \left. - \int_{S_1^{d-1}} \int_{\mathbb{R}_{>0}} \int_{S_1^{d-1}} f(\mathbf{v}_0, \xi, \mathbf{v}_1) p_{\alpha, \beta}(\mathbf{v}_0, \xi, \mathbf{v}_1) d\lambda_j(\mathbf{v}_0) d\xi d\text{vol}_{S_1^{d-1}}(\mathbf{v}_1) \right| < \frac{\varepsilon}{S \text{vol}(S_1^{d-1})}.$$

We now claim that for all  $\rho \in (0, \rho_0)$ ,  $\lambda \in F_1$ ,  $f \in F_2$  and  $\beta \in F_3$ , we have

$$(2.63) \quad \left| \int_{S_1^{d-1}} f(\mathbf{v}_0, \rho^{d-1} \tau_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho), \mathbf{v}_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda(\mathbf{v}_0) \right. \\ \left. - \int_{S_1^{d-1}} \int_{\mathbb{R}_{>0}} \int_{S_1^{d-1}} f(\mathbf{v}_0, \xi, \mathbf{v}_1) p_{\alpha, \beta}(\mathbf{v}_0, \xi, \mathbf{v}_1) d\lambda(\mathbf{v}_0) d\xi d\text{vol}_{S_1^{d-1}}(\mathbf{v}_1) \right| < 5\varepsilon.$$

This will complete the proof of Theorem 2.3, since  $\varepsilon > 0$  was arbitrary.

To prove (2.63), let  $\rho \in (0, \rho_0)$ ,  $\lambda \in F_1$ ,  $f \in F_2$  and  $\beta \in F_3$  be given. Take  $g \in F'$  and  $\mathcal{W} \in F''$  such that  $\lambda = (g \cdot \text{vol}_{S_1^{d-1}})_{|\mathcal{W}}$ . Set

$$(2.64) \quad M := \{j : D_j \subset \mathcal{W}\}, \quad M' := \{j : D_j \not\subset \mathcal{W} \text{ and } D_j \not\subset S_1^{d-1} \setminus \mathcal{W}\}.$$

Since  $D_j$  is pathwise connected, for each  $j \in M'$  there is some point  $\mathbf{p} \in D_j \cap \partial\mathcal{W}$ . Therefore, and because  $D_j$  has diameter  $< \eta$ , we have  $D_j \subset \partial_\eta \mathcal{W}$ . Hence  $\sum_{j \in M'} \text{vol}_{S_1^{d-1}}(D_j) \leq \text{vol}_{S_1^{d-1}}(\partial_\eta \mathcal{W}) < \frac{\varepsilon}{RS}$ .

For each  $j \in \{1, \dots, n\}$  we fix a point  $\mathbf{p}_j \in D_j$  and set  $g_j := g(\mathbf{p}_j)$ ; then note that  $|g(\mathbf{v}_0) - g_j| < \frac{\varepsilon}{R \text{vol}(S_1^{d-1})}$  for all  $\mathbf{v}_0 \in D_j$ , by our choice of  $\eta$  in (2.60). Hence

$$(2.65) \quad \left| \int_{S_1^{d-1}} f(\mathbf{v}_0, \rho^{d-1} \tau_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho), \mathbf{v}_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda(\mathbf{v}_0) \right. \\ \left. - \sum_{j \in M} g_j \text{vol}_{S_1^{d-1}}(D_j) \int_{S_1^{d-1}} f(\mathbf{v}_0, \rho^{d-1} \tau_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho), \mathbf{v}_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda_j(\mathbf{v}_0) \right| \\ \leq \sum_{j \in M} \int_{D_j} R \frac{\varepsilon}{R \text{vol}(S_1^{d-1})} d\text{vol}_{S_1^{d-1}}(\mathbf{v}_0) + \sum_{j \in M'} \int_{D_j} RS d\text{vol}_{S_1^{d-1}}(\mathbf{v}_0) < 2\varepsilon.$$

Combining this with (2.62), applied for each  $j \in M$ , we get

$$(2.66) \quad \left| \int_{S_1^{d-1}} f(\mathbf{v}_0, \rho^{d-1} \tau_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho), \mathbf{v}_1(\mathbf{q}_{\rho, \beta}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda(\mathbf{v}_0) \right. \\ \left. - \sum_{j \in M} g_j \text{vol}_{S_1^{d-1}}(D_j) \int_{S_1^{d-1}} \int_{\mathbb{R}_{>0}} \int_{S_1^{d-1}} f(\mathbf{v}_0, \xi, \mathbf{v}_1) p_{\alpha, \beta}(\mathbf{v}_0, \xi, \mathbf{v}_1) d\lambda_j(\mathbf{v}_0) d\xi d\text{vol}_{S_1^{d-1}}(\mathbf{v}_1) \right| \\ < 2\varepsilon + \sum_{j \in M} S \text{vol}_{S_1^{d-1}}(D_j) \frac{\varepsilon}{S \text{vol}(S_1^{d-1})} \leq 3\varepsilon.$$



On the other hand, using  $\int_{S_1^{d-1}} \int_{\mathbb{R}_{>0}} p_{\alpha,\beta}(\mathbf{v}_0, \xi, \mathbf{v}_1) d\xi d\text{vol}_{S_1^{d-1}}(\mathbf{v}_1) = 1$  (true for all  $\mathbf{v}_0 \in S_1^{d-1}$ , cf. (2.25) and (2.12)) and the same bounds as in (2.65), we get

$$(2.67) \quad \left| \int_{S_1^{d-1}} \int_{\mathbb{R}_{>0}} \int_{S_1^{d-1}} f(\mathbf{v}_0, \xi, \mathbf{v}_1) p_{\alpha,\beta}(\mathbf{v}_0, \xi, \mathbf{v}_1) d\lambda(\mathbf{v}_0) d\xi d\text{vol}_{S_1^{d-1}}(\mathbf{v}_1) \right. \\ \left. - \sum_{j \in M} g_j \text{vol}_{S_1^{d-1}}(D_j) \int_{S_1^{d-1}} \int_{\mathbb{R}_{>0}} \int_{S_1^{d-1}} f(\mathbf{v}_0, \xi, \mathbf{v}_1) p_{\alpha,\beta}(\mathbf{v}_0, \xi, \mathbf{v}_1) d\lambda_j(\mathbf{v}_0) d\xi d\text{vol}_{S_1^{d-1}}(\mathbf{v}_1) \right| \\ \leq \sum_{j \in M} \int_{D_j} R \frac{\varepsilon}{R \text{vol}(S_1^{d-1})} d\text{vol}_{S_1^{d-1}}(\mathbf{v}_0) + \sum_{j \in M'} \int_{D_j} RS d\text{vol}_{S_1^{d-1}}(\mathbf{v}_0) < 2\varepsilon.$$

Combining (2.66) and (2.67) we obtain (2.63), and we are done.  $\square$

### 3. ITERATED SCATTERING MAPS

To prepare for the proof of the main result of this paper, Theorem 4.1, this section provides a detailed study of the map obtained by iterated scattering in a given sequence of balls  $\mathbf{m}_j + \mathcal{B}_\rho^d$ , when  $\rho$  is small. The central result of this section is Proposition 3.3 below.

**3.1. Two lemmas.** Given  $\mathbf{v} \in S_1^{d-1}$ , the tangent space  $T_{\mathbf{v}}(S_1^{d-1})$  is naturally identified with  $\{\mathbf{v}\}^\perp \subset \mathbb{R}^d$ . For  $\mathbf{h} \in T_{\mathbf{v}}(S_1^{d-1})$  we write  $D_{\mathbf{h}}$  for the corresponding derivative. We use the standard Riemannian metric for  $S_1^{d-1}$ , and denote by  $T_{\mathbf{v}}^1(S_1^{d-1})$  the set of unit vectors in  $T_{\mathbf{v}}(S_1^{d-1})$ . For any open subset  $\mathcal{V} \subset S_1^{d-1}$  we write  $T^1(\mathcal{V}) = \bigsqcup_{\mathbf{v} \in \mathcal{V}} T_{\mathbf{v}}^1(S_1^{d-1})$  for the unit tangent bundle of  $\mathcal{V}$ .

For  $\mathbf{w} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$  and  $\eta \geq 0$  we define  $\mathcal{V}_{\mathbf{w}} := \mathcal{V}_{\tilde{\mathbf{w}}}$  (cf. (2.18)) and

$$(3.1) \quad \mathcal{V}_{\mathbf{w}}^\eta := \left\{ \mathbf{u} \in S_1^{d-1} : \varphi(\mathbf{u}, \mathbf{w}) > B_\Theta + \eta \right\} \subseteq \mathcal{V}_{\mathbf{w}};$$

thus in particular  $\mathcal{V}_{\mathbf{w}}^0 = \mathcal{V}_{\mathbf{w}}$ . Set

$$(3.2) \quad C_\eta := 1 + \max \left( \sup_{\mathbf{h} \in T^1(\mathcal{V}_{\mathbf{v}}^\eta)} \|D_{\mathbf{h}}\beta_{\mathbf{v}}^+\|, \sup_{\mathbf{h} \in T^1(\mathcal{V}_{\mathbf{v}}^\eta)} \|D_{\mathbf{h}}\beta_{\mathbf{v}}^-\| \right).$$

Then  $C_\eta$  is independent of  $\mathbf{v}$ , depends continuously on  $\eta \in (0, \pi - B_\Theta)$ , and may approach infinity as  $\eta \rightarrow 0$ .

For any  $\mathbf{s} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$  we let  $\lambda_{\mathbf{s}}$  be the probability measure on  $S_1^{d-1}$  which gives the direction of a ray after it has been scattered in the ball  $\mathcal{B}_1^d$ , given that the incoming ray has direction  $\mathbf{s}$  and is part of the line  $\mathbf{x} + \mathbb{R}\mathbf{s}$  with  $\mathbf{x}$  picked at random in the  $(d-1)$ -dimensional unit ball  $\{\mathbf{s}\}^\perp \cap \mathcal{B}_1^d$ , with respect to the  $(d-1)$ -dimensional Lebesgue measure. In particular for  $\mathbf{s} = \mathbf{e}_1$  we have

$$(3.3) \quad \lambda_{\mathbf{e}_1} = \text{vol}(\mathcal{B}_1^{d-1})^{-1} \cdot (V_0)_*(\text{vol}_{\mathbb{R}^{d-1}}),$$

where  $V_0 : \mathcal{B}_1^{d-1} \rightarrow S_1^{d-1}$  is the map

$$(3.4) \quad V_0(\mathbf{x}) = \Theta_1(\mathbf{e}_1, (-\sqrt{1 - \|\mathbf{x}\|^2}, \mathbf{x})).$$

Thus  $V_0$  is a diffeomorphism from  $\mathcal{B}_1^{d-1}$  onto  $\mathcal{V}_{\mathbf{e}_1}$ . For general  $\mathbf{s} \neq \mathbf{0}$  we have  $\lambda_{\mathbf{s}} = K_*(\lambda_{\mathbf{e}_1})$  where  $K \in \text{SO}(d)$  is any rotation such that  $\hat{\mathbf{s}} = \mathbf{e}_1 K$ . The following lemma shows that for any subset  $M \subset \mathcal{V}_{\mathbf{s}}^\eta$ , the renormalized  $S_1^{d-1}$ -volume measure on  $M$  and the renormalized  $\lambda_{\mathbf{s}}$ -measure on  $M$  are comparable with a controllable distortion factor.

**Lemma 3.1.** *Given  $\eta, \varepsilon > 0$ , there exist constants  $K_\eta > 1$  and  $c_{\eta, \varepsilon} > 0$  such that for every Borel subset  $M \subset \mathcal{V}_{\mathbf{s}}^\eta$  with  $\text{vol}_{S_1^{d-1}}(M) > 0$ , we have*

$$(3.5) \quad \lambda_{\mathbf{s}}(M)^{-1} \lambda_{\mathbf{s}|M} = g \cdot (\text{vol}_{S_1^{d-1}}(M))^{-1} \cdot \text{vol}_{S_1^{d-1}|M}$$

for some continuous function  $g : M \rightarrow \mathbb{R}_{>0}$  with  $g(\mathbf{u}) \in [K_\eta^{-1}, K_\eta]$  for all  $\mathbf{u} \in M$ .

If furthermore  $M$  has diameter  $\leq c_{\eta,\varepsilon}$ , i.e. if  $\varphi(\mathbf{u}_1, \mathbf{u}_2) \leq c_{\eta,\varepsilon}$  for all  $\mathbf{u}_1, \mathbf{u}_2 \in M$ , then  $g(\mathbf{u}) \in [1 - \varepsilon, 1 + \varepsilon]$  for all  $\mathbf{u} \in M$ .

*Proof.* Without loss of generality we may assume  $\mathbf{s} = \mathbf{e}_1$ . Set  $\nu = (\text{vol}_{\mathbb{S}_1^{d-1}}(M))^{-1} \cdot \text{vol}_{\mathbb{S}_1^{d-1}|_M}$  and  $\lambda = \lambda_{\mathbf{e}_1}(M)^{-1} \lambda_{\mathbf{e}_1|_M}$ . By (3.3) we have  $\lambda = C \cdot f|_M \cdot \nu$ , where  $f : \mathcal{V}_{\mathbf{e}_1} \rightarrow \mathbb{R}_{>0}$  is the Jacobian determinant of the inverse map  $V_0^{-1}$  from  $\mathcal{V}_{\mathbf{e}_1}$  to  $\mathcal{B}_1^{d-1}$ , and  $C > 0$  is a constant determined by  $C \int_M f d\nu = 1$  (in particular if the ratio of  $\sup_M f$  and  $\inf_M f$  is close to 1 then the function  $C \cdot f|_M$  is uniformly close to 1). Now the lemma follows from the fact that both  $f$  and  $f^{-1}$  are bounded and uniformly continuous on  $\mathcal{V}_{\mathbf{e}_1}^\eta$ , since  $\mathcal{V}_{\mathbf{e}_1}^\eta$  has compact closure in  $\mathcal{V}_{\mathbf{e}_1}$ .  $\square$

We define  $\lambda_{\mathbf{s}}^\eta$  to be the probability measure which is obtained by restricting  $\lambda_{\mathbf{s}}$  to  $\mathcal{V}_{\mathbf{s}}^\eta$ , and normalizing appropriately.

Given  $\mathbf{r}, \mathbf{s} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ , a number  $\rho > 0$  and a continuous function  $\beta : \mathcal{V}_{\mathbf{r}}^\eta \rightarrow \mathbb{R}^d$ , we set

$$(3.6) \quad \Omega = \left\{ \mathbf{v} \in \mathcal{V}_{\mathbf{r}}^\eta : (\rho\beta(\mathbf{v}) + \mathbb{R}_{>0}\mathbf{v}) \cap (\mathbf{s} + \mathcal{B}_\rho^d) \neq \emptyset \right\}.$$

For  $\mathbf{v} \in \Omega$  we set

$$(3.7) \quad \tau(\mathbf{v}) = \tau_{\rho,\mathbf{s},\beta}(\mathbf{v}) := \inf\{t > 0 : \rho\beta(\mathbf{v}) + t\mathbf{v} \in \mathbf{s} + \mathcal{B}_\rho^d\},$$

let  $W(\mathbf{v}) = W_{\rho,\mathbf{s},\beta}(\mathbf{v})$  be the impact location on  $\mathbb{S}_1^{d-1}$ , i.e., the point for which  $\rho\beta(\mathbf{v}) + \tau(\mathbf{v})\mathbf{v} = \mathbf{s} + \rho W(\mathbf{v})$ , and let

$$(3.8) \quad V(\mathbf{v}) = V_{\rho,\mathbf{s},\beta}(\mathbf{v}) := \Theta_1(\mathbf{v}, W(\mathbf{v})) \in \mathbb{S}_1^{d-1},$$

the outgoing direction after the ray  $\rho\beta(\mathbf{v}) + \mathbb{R}_{>0}\mathbf{v}$  is scattered in the sphere  $\mathbf{s} + \mathbb{S}_\rho^{d-1}$ .

**Lemma 3.2.** *Given any  $0 < \eta < \frac{\pi - B_\Theta}{100}$ ,  $C \geq 10$  and  $\varepsilon > 0$ , there exists a constant  $\tilde{\rho}_0 = \tilde{\rho}_0(\eta, C, \varepsilon) > 0$  such that all the following statements hold for any  $\rho \in (0, \tilde{\rho}_0)$ , any  $\mathbf{r}, \mathbf{s} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$  with  $\|\mathbf{s}\| \geq C^{-1}$  and  $\varphi(\mathbf{r}, \mathbf{s}) > B_\Theta + 2\eta$ , and for any  $C^1$ -function  $\beta : \mathcal{V}_{\mathbf{r}}^\eta \rightarrow \mathbb{R}^d$  with  $\sup_{\mathcal{V}_{\mathbf{r}}^\eta} \|\beta\| \leq C$  and  $\sup_{\mathbf{h} \in \mathbb{T}^1(\mathcal{V}_{\mathbf{r}}^\eta)} \|D_{\mathbf{h}}\beta\| \leq C$ :*

- (i) *Let  $\bar{V} = \bar{V}_{\rho,\mathbf{s},\beta}$  be the restriction of  $V = V_{\rho,\mathbf{s},\beta}$  to  $V^{-1}(\mathcal{V}_{\mathbf{s}}^\eta)$ ; then  $\bar{V}$  is a  $C^1$  diffeomorphism onto  $\mathcal{V}_{\mathbf{s}}^\eta$ .*
- (ii) *If  $M \subset \mathcal{V}_{\mathbf{s}}^\eta$  is any Borel subset with  $\lambda_{\mathbf{s}}(M) > 0$  and if  $\mu$  denotes the measure  $\text{vol}_{\mathbb{S}_1^{d-1}}$  restricted to  $\bar{V}^{-1}(M)$  and rescaled to be a probability measure, then  $\bar{V}_* \mu = g \cdot \lambda_{\mathbf{s}}(M)^{-1} \lambda_{\mathbf{s}}|_M$  for some continuous function  $g : M \rightarrow [1 - \varepsilon, 1 + \varepsilon]$ .*
- (iii) *Define the  $C^1$  maps  $B^\pm = B_{\rho,\mathbf{s},\beta}^\pm : \mathcal{V}_{\mathbf{s}}^\eta \rightarrow \mathbb{S}_1^{d-1}$  through  $B^\pm(\mathbf{u}) = \beta_{\bar{V}^{-1}(\mathbf{u})}^\pm(\mathbf{u})$ . Then  $\|B^\pm(\mathbf{u}) - \beta_{\mathbf{s}}^\pm(\mathbf{u})\| < \varepsilon$  for all  $\mathbf{u} \in \mathcal{V}_{\mathbf{s}}^\eta$  and  $\|D_{\mathbf{h}}B^\pm\| < C_\eta$  for all  $\mathbf{h} \in \mathbb{T}^1(\mathcal{V}_{\mathbf{s}}^\eta)$ .*

*Proof.* Since  $\Omega$ ,  $W_{\rho,\mathbf{s},\beta}$  and  $V_{\rho,\mathbf{s},\beta}$  are invariant under  $\langle \rho, \mathbf{s} \rangle \mapsto \langle c\rho, c\mathbf{s} \rangle$  for any  $c > 0$ , it suffices to treat the case  $\|\mathbf{s}\| = 1$ . After an auxiliary rotation we may then assume  $\mathbf{s} = \mathbf{e}_1$ . From now on we will always keep  $\rho < \frac{\eta}{20C}$ . Then  $\rho\beta(\mathbf{v})$  certainly lies outside the ball  $\mathbf{e}_1 + \mathcal{B}_\rho^d$  for all  $\mathbf{v} \in \mathcal{V}_{\mathbf{r}}^\eta$ . Let us write  $\mathbf{a} = \mathbf{a}(\mathbf{v}) = \rho^{-1}\mathbf{e}_1 - \beta(\mathbf{v})$  and note that  $\|\mathbf{a}\| > \frac{1}{\rho} > 10^3$  for all  $\mathbf{v}$ . Now for each  $\mathbf{v} \in \Omega$  the ray  $\mathbb{R}_{>0}\mathbf{v}$  hits  $\mathbf{a} + \mathcal{B}_1^d$ , and thus  $\varphi(\mathbf{v}, \mathbf{a}) \leq \arcsin(\|\mathbf{a}\|^{-1}) \leq \frac{\pi}{2\|\mathbf{a}\|} < \pi\rho$ . Also  $\varphi(\mathbf{a}, \mathbf{e}_1) \leq \arcsin(\rho\|\beta(\mathbf{v})\|) \leq \frac{\pi}{2}C\rho$ . Hence by the triangle inequality for  $\varphi$ , using  $\rho < \frac{\eta}{20C}$  we obtain

$$(3.9) \quad \varphi(\mathbf{v}, \mathbf{e}_1) < \left(\pi + \frac{\pi}{2}C\right)\rho < 2C\rho < \frac{\eta}{10}, \quad \forall \mathbf{v} \in \Omega.$$

We now compute

$$(3.10) \quad \Omega = \left\{ \mathbf{v} \in \mathbb{S}_1^{d-1} : 1 - \|\mathbf{a}\|^2 + (\mathbf{a} \cdot \mathbf{v})^2 > 0 \right\},$$

which is automatically a subset of  $\mathcal{V}_{\mathbf{r}}^\eta$ , because of the assumption  $\varphi(\mathbf{r}, \mathbf{e}_1) > B_\Theta + 2\eta$ , and furthermore, for all  $\mathbf{v} \in \Omega$ ,  $\mathbf{h} \in \mathbb{T}_{\mathbf{v}}(\mathbb{S}_1^{d-1})$ ,

$$(3.11) \quad W(\mathbf{v}) = -\mathbf{a} + \left( (\mathbf{a} \cdot \mathbf{v}) - \sqrt{1 - \|\mathbf{a}\|^2 + (\mathbf{a} \cdot \mathbf{v})^2} \right) \mathbf{v}$$

and

$$(3.12) \quad D_{\mathbf{h}}W = D_{\mathbf{h}}\beta + \left( \mathbf{a} \cdot \mathbf{h} - \mathbf{v} \cdot D_{\mathbf{h}}\beta - \frac{\mathbf{a} \cdot D_{\mathbf{h}}\beta + (\mathbf{a} \cdot \mathbf{v})(\mathbf{a} \cdot \mathbf{h} - \mathbf{v} \cdot D_{\mathbf{h}}\beta)}{\sqrt{1 - \|\mathbf{a}\|^2 + (\mathbf{a} \cdot \mathbf{v})^2}} \right) \mathbf{v} \\ + \left( \mathbf{a} \cdot \mathbf{v} - \sqrt{1 - \|\mathbf{a}\|^2 + (\mathbf{a} \cdot \mathbf{v})^2} \right) \mathbf{h},$$

using the standard embedding  $\mathbf{h} \in T_{\mathbf{v}}(\mathbb{S}_1^{d-1}) \subset \mathbb{R}^d$ .

Since  $\Omega \subset \mathbb{S}_1^{d-1}$  is contained in a small neighborhood of  $\mathbf{e}_1$  (see (3.9)), the map  $\mathbf{v} \mapsto \beta(\mathbf{e}_1)_{\perp} + \rho^{-1}\mathbf{v}_{\perp}$  gives a diffeomorphism from  $\Omega$  onto

$$(3.13) \quad \Omega_I := \beta(\mathbf{e}_1)_{\perp} + \rho^{-1}\Omega_{\perp} \subset \{0\} \times \mathbb{R}^{d-1},$$

which transforms the  $\mathbb{S}_1^{d-1}$  volume measure into the standard  $\mathbb{R}^{d-1}$  Lebesgue volume measure scaled by  $\rho^{d-1}$  times a distortion factor which is uniformly close to 1 when  $\rho$  is small. The inverse map is (using the identification  $\{0\} \times \mathbb{R}^{d-1} = \mathbb{R}^{d-1}$ )

$$(3.14) \quad I : \Omega_I \rightarrow \Omega, \quad \mathbf{x} \mapsto Q(\rho(\mathbf{x} - \beta(\mathbf{e}_1)_{\perp})),$$

where  $Q$  is the map

$$(3.15) \quad Q : \mathcal{B}_1^{d-1} \rightarrow \mathbb{S}_1^{d-1}, \quad \mathbf{y} \mapsto (\sqrt{1 - \|\mathbf{y}\|^2}, \mathbf{y}).$$

Now, for (i) and (ii), it suffices to prove that (i) and (ii) hold with the map  $V$  replaced throughout by  $V_I := V \circ I : \Omega_I \rightarrow \mathbb{S}_1^{d-1}$ ,  $\text{vol}_{\mathbb{S}_1^{d-1}}$  replaced by  $\text{vol}_{\mathbb{R}^{d-1}}$ , and  $\varepsilon$  replaced by  $\varepsilon/2$ .

Recall the definition (3.4) of the diffeomorphism  $V_0$  from  $\mathcal{B}_1^{d-1}$  onto  $\mathcal{V}_{\mathbf{e}_1}$ . Using the spherical invariance of  $\Theta_1$  it follows that  $V_0^{-1}(\mathcal{V}_{\mathbf{e}_1}^{\eta/2}) = \mathcal{B}_r^{d-1}$  for some  $r = r(\Theta, \eta) < 1$ .

From (3.9) we get  $\Omega_I \subset \{0\} \times \mathcal{B}_{3C}^{d-1}$ . Now for all  $\mathbf{x} \in \mathcal{B}_{3C}^{d-1}$  we compute, with  $\mathbf{v} = Q(\rho(\mathbf{x} - \beta(\mathbf{e}_1)_{\perp}))$  and  $\mathbf{a} = \rho^{-1}\mathbf{e}_1 - \beta(\mathbf{v})$  as before,

$$(3.16) \quad \mathbf{a} \cdot \mathbf{v} = \rho^{-1} - \beta_1(\mathbf{v}) + \frac{1}{2}(\|\beta(\mathbf{e}_1)_{\perp}\|^2 - \|\mathbf{x}\|^2)\rho + O(\rho^2);$$

$$(3.17) \quad \|\mathbf{a}\|^2 = \rho^{-2} - 2\beta_1(\mathbf{v})\rho^{-1} + \|\beta(\mathbf{e}_1)_{\perp}\|^2 + O(\rho),$$

where  $\beta_1(\mathbf{v}) := \beta(\mathbf{v}) \cdot \mathbf{e}_1$ . Here and in the rest of the proof, the implied constant in any big  $O$  depends only on  $\Theta, C, \eta, \varepsilon$ . It follows that

$$(3.18) \quad 1 - \|\mathbf{a}\|^2 + (\mathbf{a} \cdot \mathbf{v})^2 = 1 - \|\mathbf{x}\|^2 + O(\rho),$$

and hence by (3.10), if  $\rho$  is sufficiently small,

$$(3.19) \quad V_0^{-1}(\mathcal{V}_{\mathbf{e}_1}^{\eta/2}) = \mathcal{B}_r^{d-1} \subset \Omega_I.$$

Let  $W_I := W \circ I : \Omega_I \rightarrow \mathbb{S}_1^{d-1}$ . By a computation using (3.11) and (3.16)–(3.18) we find that for each  $\mathbf{x} \in \mathcal{B}_r^{d-1}$ ,

$$(3.20) \quad W_I(\mathbf{x}) = W(\mathbf{v}) = W_0(\mathbf{x}) + O(\rho), \quad \text{with } W_0(\mathbf{x}) = \left( -\sqrt{1 - \|\mathbf{x}\|^2}, \mathbf{x} \right).$$

By a similar computation using (3.12) we also obtain, for  $\rho$  sufficiently small,

$$(3.21) \quad (D_{\mathbf{h}}W)_{\perp} = \rho^{-1}\mathbf{h} + O(\|\mathbf{h}\|)$$

for all  $\mathbf{v} \in I(\mathcal{B}_r^{d-1})$  and  $\mathbf{h} \in T_{\mathbf{v}}(\mathbb{S}_1^{d-1})$ . Therefore, for all  $\mathbf{x} \in \mathcal{B}_r^{d-1}$  and  $\mathbf{u} \in T_{\mathbf{x}}(\Omega_I) \cong \{0\} \times \mathbb{R}^{d-1}$ ,

$$(3.22) \quad (D_{\mathbf{u}}W_I)_{\perp} = \rho^{-1}D_{\mathbf{u}}I + O(\|D_{\mathbf{u}}I\|) = \mathbf{u} + O(\rho\|\mathbf{u}\|),$$

since  $D_{\mathbf{u}}I = \rho\mathbf{u} + O(\rho^2\|\mathbf{u}\|)$ . We also have, trivially,  $(D_{\mathbf{u}}W_0)_{\perp} = \mathbf{u}$ . But we know

$$(3.23) \quad D_{\mathbf{u}}W_I \in T_{W_I(\mathbf{x})}(\mathbb{S}_1^{d-1}) = \{W_I(\mathbf{x})\}^{\perp} \subset \mathbb{R}^d, \quad D_{\mathbf{u}}W_0 \in \{W_0(\mathbf{x})\}^{\perp} \subset \mathbb{R}^d.$$

For  $\rho$  sufficiently small, (3.20) implies that  $W_I(\mathbf{x})$  and  $W_0(\mathbf{x})$  are close on  $\mathbb{S}_1^{d-1}$  and both have  $\mathbf{e}_1$ -component  $< -\frac{1}{2}\sqrt{1 - r^2}$ . Hence (3.22),  $(D_{\mathbf{u}}W_0)_{\perp} = \mathbf{u}$  and (3.23) imply

$$(3.24) \quad D_{\mathbf{u}}W_I = D_{\mathbf{u}}W_0 + O(\rho\|\mathbf{u}\|).$$

It follows from (3.20) and (3.24) that the function  $(W_I)|_{\mathcal{B}_r^{d-1}}$  converges to  $(W_0)|_{\mathcal{B}_r^{d-1}}$  in  $C^1$ -norm as  $\rho \rightarrow 0$ .

To continue, note that  $V_I(\mathbf{x}) = \Theta_1(I(\mathbf{x}), W_I(\mathbf{x}))$ . The function  $I : \Omega_I \rightarrow \Omega$  tends to the constant function  $\mathbf{x} \mapsto \mathbf{e}_1$  in  $C^1$ -norm, as  $\rho \rightarrow 0$ . Using that  $\Theta_1 : \mathcal{S}_- \rightarrow S_1^{d-1}$  is  $C^1$  and that  $I(\mathcal{B}_r^{d-1}) \times W_I(\mathcal{B}_r^{d-1})$  is a subset with compact closure in  $\mathcal{S}_-$  for  $\rho$  sufficiently small, we conclude that the map  $(V_I)|_{\mathcal{B}_r^{d-1}} : \mathcal{B}_r^{d-1} \rightarrow S_1^{d-1}$  tends to  $(V_0)|_{\mathcal{B}_r^{d-1}}$  in  $C^1$ -norm as  $\rho \rightarrow 0$ . Since the latter map is a  $C^1$  diffeomorphism from  $\mathcal{B}_r^{d-1}$  onto  $\mathcal{V}_{\mathbf{e}_1}^{\eta/2}$  which is a restriction of a diffeomorphism from all of  $\mathcal{B}_1^{d-1}$ , it follows by a standard argument in differential geometry (cf., e.g., [17, Sec. 2.1]) that, for  $\rho$  small enough,  $(V_I)|_{\mathcal{B}_r^{d-1}}$  must be a diffeomorphism as well, mapping  $\mathcal{B}_r^{d-1}$  onto some open subset of  $S_1^{d-1}$  which contains  $\mathcal{V}_{\mathbf{e}_1}^\eta$ , and that (ii) holds with  $\bar{V}$  replaced by  $(V_I)|_{\mathcal{B}_r^{d-1}}$  (and  $\text{vol}_{S_1^{d-1}}$  replaced by  $\text{vol}_{\mathbb{R}^{d-1}}$ , and  $\varepsilon$  replaced by  $\varepsilon/2$ ).

Hence to conclude the proof of (i) and (ii), it only remains to prove that the map  $\bar{V}$  is injective, or equivalently that  $V_I(\mathbf{x}) \notin \mathcal{V}_{\mathbf{e}_1}^\eta$  for all  $\mathbf{x} \in \Omega_I \setminus \mathcal{B}_r^{d-1}$ . A computation shows that if  $\rho$  is small and  $\mathbf{x} \in \Omega_I \setminus \mathcal{B}_r^{d-1}$ , then  $W_I(\mathbf{x}) \cdot \mathbf{e}_1 \geq -\sqrt{1-r^2} - O(\rho)$ , and thus

$$(3.25) \quad \varphi(I(\mathbf{x}), W_I(\mathbf{x})) \leq \varphi(\mathbf{e}_1, W_I(\mathbf{x})) + O(\rho) \leq \frac{\pi}{2} + \arccos(r) + O(\rho).$$

It follows from our assumptions on  $\Theta_1$  that  $\varphi(I(\mathbf{x}), V_I(\mathbf{x}))$  is an increasing  $C^1$  function of  $\varphi(I(\mathbf{x}), W_I(\mathbf{x})) \in (\frac{\pi}{2}, \pi]$ , and  $\varphi(I(\mathbf{x}), V_I(\mathbf{x})) = B_\Theta + \frac{\eta}{2}$  when  $\varphi(I(\mathbf{x}), W_I(\mathbf{x})) = \frac{\pi}{2} + \arccos(r)$ . Hence

$$(3.26) \quad \varphi(\mathbf{e}_1, V_I(\mathbf{x})) \leq \varphi(I(\mathbf{x}), V_I(\mathbf{x})) + O(\rho) \leq B_\Theta + \frac{\eta}{2} + O(\rho), \quad \forall \mathbf{x} \in \Omega_I \setminus \mathcal{B}_r^{d-1},$$

and in particular if  $\rho$  is sufficiently small then  $V_I(\mathbf{x}) \notin \mathcal{V}_{\mathbf{e}_1}^\eta$  for all  $\mathbf{x} \in \Omega_I \setminus \mathcal{B}_r^{d-1}$ , as desired.

Finally we turn to (iii). Set  $\bar{V}_I = (V_I)|_{V_I^{-1}(\mathcal{V}_{\mathbf{e}_1}^\eta)}$ . Then, since  $(V_I)|_{\mathcal{B}_r^{d-1}}$  tends to  $(V_0)|_{\mathcal{B}_r^{d-1}}$  in  $C^1$ -norm, we have  $D_{\mathbf{u}}(\bar{V}_I^{-1}) = O(\|\mathbf{u}\|)$  for all  $\mathbf{x} \in \mathcal{V}_{\mathbf{e}_1}^\eta$  and  $\mathbf{u} \in T_{\mathbf{x}}(\mathcal{V}_{\mathbf{e}_1}^\eta)$ , so long as  $\rho$  is sufficiently small. But

$$(3.27) \quad \bar{V}^{-1} = I \circ \bar{V}_I^{-1} : \mathcal{V}_{\mathbf{e}_1}^\eta \rightarrow S_1^{d-1}$$

and  $I$  tends to the constant function  $\mathbf{x} \rightarrow \mathbf{e}_1$  in  $C^1$ -norm. Hence also  $\bar{V}^{-1}$  tends to the constant function  $\mathbf{x} \rightarrow \mathbf{e}_1$  in  $C^1$ -norm as  $\rho \rightarrow 0$ . Therefore, by continuity and the definition of  $C_\eta$ , (3.2), claim (iii) is established.  $\square$

**3.2. The main proposition.** Let us consider a particle trajectory that follows the ray emerging at  $\mathbf{m}_0 + \rho\boldsymbol{\beta}(\mathbf{v}_0)$  in direction  $\mathbf{v}_0$ , is scattered at the ball  $\mathbf{m}_1 + \mathcal{B}_\rho^d$ , exits with velocity  $\mathbf{v}_1$  and moves with constant speed until scattered at  $\mathbf{m}_2 + \mathcal{B}_\rho^d$ , and so on; after the final scattering at  $\mathbf{m}_n + \mathcal{B}_\rho^d$  the particle exists with velocity  $\mathbf{v}_n$ . We call  $\bar{\mathbf{m}} = (\mathbf{m}_0, \mathbf{m}_1, \dots, \mathbf{m}_n)$  the *scattering sequence* associated with this trajectory. We denote by  $\bar{\boldsymbol{\beta}}^-(\mathbf{v}_n)$  and  $\bar{\boldsymbol{\beta}}^+(\mathbf{v}_n) \in S_1^{d-1}$  the position of impact and exit on the last scatterer; thus  $\bar{\boldsymbol{\beta}}^-(\mathbf{v}_n) = \boldsymbol{\beta}_{\mathbf{v}_{n-1}}^-(\mathbf{v}_n)$  and  $\bar{\boldsymbol{\beta}}^+(\mathbf{v}_n) = \boldsymbol{\beta}_{\mathbf{v}_{n-1}}^+(\mathbf{v}_n)$ . We call  $\mathbf{v}_0$  and  $\mathbf{v}_n$  the *initial* and *final velocity*, respectively. Set furthermore

$$(3.28) \quad \mathbf{s}_k := \mathbf{m}_k - \mathbf{m}_{k-1}.$$

Note that  $\mathbf{s}_k$  differs from the path segment  $\tau_k \mathbf{v}_{k-1}$  defined in (1.8) by  $\ll \rho$ .

Given positive constants  $\eta, C$ , we say a sequence  $\bar{\mathbf{m}} = (\mathbf{m}_0, \mathbf{m}_1, \dots, \mathbf{m}_n)$  of scatterer positions is  $(\eta, C)$ -*admissible*, if

$$(3.29) \quad \|\mathbf{s}_k\| \geq C^{-1}, \quad (k = 1, \dots, n)$$

and

$$(3.30) \quad \varphi(\mathbf{s}_{k+1}, \mathbf{s}_k) > B_\Theta + 2\eta, \quad (k = 1, \dots, n-1).$$

**Proposition 3.3.** *Given any  $N \in \mathbb{Z}_{>0}$ ,  $0 < \eta < \frac{\pi - B_\Theta}{100}$ ,  $C \geq 10$  and  $\varepsilon > 0$ , there exists a constant  $\rho_0 = \rho_0(N, \eta, C, \varepsilon) > 0$  such that for any  $\rho \in (0, \rho_0)$ , any  $(\eta, C)$ -admissible sequence*

$\vec{\mathbf{m}} = (\mathbf{m}_0, \mathbf{m}_1, \dots, \mathbf{m}_n)$  with  $n \leq N$ , and any  $C^1$  function  $\beta : S_1^{d-1} \rightarrow \mathbb{R}^d$  with  $\sup_{S_1^{d-1}} \|\beta\| \leq C$  and  $\sup_{\mathbf{h} \in T^1(S_1^{d-1})} \|D_{\mathbf{h}}\beta\| \leq C$ , the following holds:

- (i) There exists a  $C^1$  diffeomorphism  $\Phi = \Phi_{\rho, \vec{\mathbf{m}}, \beta}$  from an open subset  $\Delta = \Delta_{\rho, \vec{\mathbf{m}}, \beta} \subset S_1^{d-1}$  onto  $\mathcal{V}_{s_n}^\eta$  such that, for every  $\mathbf{v} \in S_1^{d-1}$ ,  $\mathbf{u} \in \mathcal{V}_{s_n}^\eta$ , the following statements are equivalent:
  - (a) There is a particle trajectory with scattering sequence  $\vec{\mathbf{m}}$ , initial velocity  $\mathbf{v}_0 = \mathbf{v}$  and final velocity  $\mathbf{v}_n = \mathbf{u}$ .
  - (b)  $\mathbf{v} \in \Delta$  and  $\Phi(\mathbf{v}) = \mathbf{u}$ .
- (ii) The position of impact and exit on the last scatterer,  $\bar{\beta}^- = \bar{\beta}_{\rho, \vec{\mathbf{m}}, \beta}^-$  and  $\bar{\beta}^+ = \bar{\beta}_{\rho, \vec{\mathbf{m}}, \beta}^+$ , are  $C^1$  maps from  $\mathcal{V}_{s_n}^\eta$  to  $S_1^{d-1}$ , satisfying  $\|\bar{\beta}^\pm(\mathbf{u}) - \beta_{s_n}^\pm(\mathbf{u})\| < \varepsilon$  for all  $\mathbf{u} \in \mathcal{V}_{s_n}^\eta$  and  $\sup_{\mathbf{h} \in T^1(\mathcal{V}_{s_n}^\eta)} \|D_{\mathbf{h}}\bar{\beta}^\pm\| < C_\eta$ .
- (iii) If  $\mu$  is the  $S_1^{d-1}$  volume measure restricted to  $\Delta$  and rescaled to be a probability measure, then  $\Phi_*\mu = g \cdot \lambda_{s_n}^\eta$  for some continuous function  $g : \mathcal{V}_{s_n}^\eta \rightarrow [1 - \varepsilon, 1 + \varepsilon]$ .

*Proof.* We assume that  $C \geq C_\eta$ . Fix  $\varepsilon_1 \in (0, \varepsilon)$  so small that  $(1 + \varepsilon_1)^{2N-1} < 1 + \varepsilon$  and  $(1 - \varepsilon_1)^{2N-1} > 1 - \varepsilon$ , and then fix

$$(3.31) \quad \rho_0 = \rho_0(N, \eta, C, \varepsilon) := \min(\tilde{\rho}_0(\eta, C, \varepsilon_1), \frac{c_{\eta, \varepsilon_1}}{2\pi C}, \frac{\eta}{20C^2}),$$

where  $\tilde{\rho}_0(\eta, C, \varepsilon_1)$  is as in Lemma 3.2, and  $c_{\eta, \varepsilon_1}$  is as in Lemma 3.1.

Now take arbitrary  $\rho \in (0, \rho_0)$  and  $\mathbf{m}_0, \dots, \mathbf{m}_n$  and  $\beta$  satisfying the assumptions of the proposition. Using the notation introduced in the context of Lemma 3.2 (setting  $\mathbf{r} = -\mathbf{s}_1$ , say, and considering the restriction of  $\beta$  to  $\mathcal{V}_{\mathbf{r}}^\eta$ ), we set

$$(3.32) \quad \Omega^{(1)} = V_{\rho, \mathbf{s}_1, \beta}^{-1}(\mathcal{V}_{\mathbf{s}_1}^\eta),$$

$$(3.33) \quad V^{(1)} = \bar{V}_{\rho, \mathbf{s}_1, \beta} : \Omega^{(1)} \rightarrow S_1^{d-1},$$

$$(3.34) \quad \beta^{(1)} = B_{\rho, \mathbf{s}_1, \beta}^+ : \mathcal{V}_{\mathbf{s}_1}^\eta \rightarrow S_1^{d-1},$$

and recursively for  $k = 1, 2, \dots, n-1$  ( $\mathbf{s} = \mathbf{s}_{k+1} = \mathbf{m}_{k+1} - \mathbf{m}_k$  and  $\mathbf{r} = \mathbf{s}_k = \mathbf{m}_k - \mathbf{m}_{k-1}$ , in the the context of Lemma 3.2),

$$(3.35) \quad \Omega^{(k+1)} = (V^{(k)})^{-1}(V_{\rho, \mathbf{s}_{k+1}, \beta^{(k)}}^{-1}(\mathcal{V}_{\mathbf{s}_{k+1}}^\eta)),$$

$$(3.36) \quad V^{(k+1)} = \bar{V}_{\rho, \mathbf{s}_{k+1}, \beta^{(k)}} \circ V^{(k)} : \Omega^{(k+1)} \rightarrow S_1^{d-1},$$

$$(3.37) \quad \beta^{(k+1)} = B_{\rho, \mathbf{s}_{k+1}, \beta^{(k)}}^+ : \mathcal{V}_{\mathbf{s}_{k+1}}^\eta \rightarrow S_1^{d-1}.$$

Then  $\Omega^{(n)} \subset \Omega^{(n-1)} \subset \dots \subset \Omega^{(1)}$  by definition.

Lemma 3.2 implies, by induction over  $k$ , that  $V^{(k)}$  is a  $C^1$  diffeomorphism from  $\Omega^{(k)}$  onto  $\mathcal{V}_{\mathbf{s}_k}^\eta$  and that  $\beta^{(k)}$  is a  $C^1$  map from  $\mathcal{V}_{\mathbf{s}_k}^\eta$  to  $S_1^{d-1}$  satisfying  $\sup_{\mathbf{h} \in T^1(\mathcal{V}_{\mathbf{s}_k}^\eta)} \|D_{\mathbf{h}}\beta^{(k)}\| < C_\eta \leq C$ , for each  $k = 1, \dots, n$ .

Now set  $\Delta = \Omega^{(n)}$  and  $\Phi = V^{(n)}$ . Then  $\Phi$  is a  $C^1$  diffeomorphism from  $\Delta$  onto  $\mathcal{V}_{s_n}^\eta$ .

*Proof of (i).* The implication “(b) $\Rightarrow$ (a)” in Proposition 3.3 follows directly by construction. Indeed, suppose  $\mathbf{v} \in \Delta$  and  $\mathbf{u} = \Phi(\mathbf{v})$ . Then since  $\mathbf{v} = \mathbf{v}_0 \in \Omega^{(1)}$  the ray  $\mathbf{m}_0 + \rho\beta(\mathbf{v}) + \mathbb{R}_{>0}\mathbf{v}$  hits the ball  $\mathbf{m}_1 + \mathcal{B}_\rho^d$  at the point  $\mathbf{m}_1 + \rho W_{\rho, \mathbf{s}_1, \beta}(\mathbf{v})$ , and after scattering we obtain the ray  $\mathbf{m}_1 + \rho\beta^{(1)}(V^{(1)}(\mathbf{v})) + \mathbb{R}_{>0}V^{(1)}(\mathbf{v})$ . Similarly it follows from our definitions that for each  $k \in \{1, \dots, n-1\}$ , the ray  $\mathbf{m}_k + \rho\beta^{(k)}(V^{(k)}(\mathbf{v})) + \mathbb{R}_{>0}V^{(k)}(\mathbf{v})$  hits the ball  $\mathbf{m}_{k+1} + \mathcal{B}_\rho^d$  and after scattering gives the ray  $\mathbf{m}_{k+1} + \rho\beta^{(k+1)}(V^{(k+1)}(\mathbf{v})) + \mathbb{R}_{>0}V^{(k+1)}(\mathbf{v})$ . Using this for  $k = 1, \dots, n-1$  we see that (a) holds for the given vectors  $\mathbf{v} = \mathbf{v}_0$ ,  $\mathbf{u} = \mathbf{v}_n$ .

Conversely, we now prove “(a) $\Rightarrow$ (b)”. Suppose that  $\mathbf{v} = \mathbf{v}_0 \in S_1^{d-1}$  and  $\mathbf{u} = \mathbf{v}_n \in \mathcal{V}_{s_n}^\eta$  satisfy the assumptions in (a), i.e. the ray  $\mathbf{m}_0 + \rho\beta(\mathbf{v}_0) + \mathbb{R}_{>0}\mathbf{v}_0$  hits  $\mathbf{m}_1 + \mathcal{B}_\rho^d$ ; after scattering in this ball we get a ray which hits  $\mathbf{m}_2 + \mathcal{B}_\rho^d$ , and so on for  $\mathbf{m}_3 + \mathcal{B}_\rho^d, \dots, \mathbf{m}_n + \mathcal{B}_\rho^d$ , and after the final scattering in  $\mathbf{m}_n + \mathcal{B}_\rho^d$  we get a ray with direction  $\mathbf{v}_n = \mathbf{u}$ . Let  $\mathbf{v}_k \in S_1^{d-1}$  be the

direction of the ray leaving the ball  $\mathbf{m}_k + \mathcal{B}_\rho^d$  in this scenario. For each  $k \in \{1, \dots, n-1\}$ , by assumption there exists some ray with direction  $\mathbf{v}_k$  and starting point on  $\mathbf{m}_k + \mathcal{S}_\rho^{d-1}$  which hits the ball  $\mathbf{m}_{k+1} + \mathcal{B}_\rho^d$ ; this implies  $\varphi(\mathbf{v}_k, \mathbf{s}_{k+1}) < \frac{\eta}{10}$  by the same argument that led to (3.9), using  $\rho < \frac{\eta}{20C^2}$ . Since we are also assuming  $\varphi(\mathbf{s}_k, \mathbf{s}_{k+1}) > B_\Theta + 2\eta$  we conclude  $\mathbf{v}_k \in \mathcal{V}_{\mathbf{s}_k}^\eta$ . This is also true for  $k = n$ , since  $\mathbf{v}_n = \mathbf{u}$  and  $\mathbf{u} \in \mathcal{V}_{\mathbf{s}_n}^\eta$  by assumption.

Now since the ray  $\mathbf{m}_0 + \rho\beta(\mathbf{v}_0) + \mathbb{R}_{>0}\mathbf{v}_0$  hits  $\mathbf{m}_1 + \mathcal{B}_\rho^d$  and gives a ray with direction  $\mathbf{v}_1 \in \mathcal{V}_{\mathbf{s}_1}^\eta$  after scattering, we must have, by our definitions,  $\mathbf{v}_0 \in V_{\rho, \mathbf{s}_1, \beta}^{-1}(\mathcal{V}_{\mathbf{s}_1}^\eta) = \Omega^{(1)}$  and  $V^{(1)}(\mathbf{v}_0) = \mathbf{v}_1$ ; in fact the ray obtained after the first scattering equals

$$(3.38) \quad \mathbf{m}_1 + \rho\beta^{(1)}(\mathbf{v}_1) + \mathbb{R}_{>0}\mathbf{v}_1, \quad \mathbf{v}_1 = V^{(1)}(\mathbf{v}_0).$$

Similarly one now proves by induction that for each  $k \in \{1, \dots, n\}$  we have  $\mathbf{v}_0 \in \Omega^{(k)}$ , and the ray obtained after the  $k$ th scattering in our scenario equals

$$(3.39) \quad \mathbf{m}_k + \rho\beta^{(k)}(\mathbf{v}_k) + \mathbb{R}_{>0}\mathbf{v}_k, \quad \mathbf{v}_k = V^{(k)}(\mathbf{v}_0).$$

In particular for  $k = n$  we obtain  $\mathbf{v} = \mathbf{v}_0 \in \Delta$  and  $\mathbf{v}_n = \Phi(\mathbf{v})$ , i.e. (b) holds for the given vectors  $\mathbf{v} = \mathbf{v}_0$ ,  $\mathbf{u} = \mathbf{v}_n$ . This completes the proof of the implication “(a) $\Rightarrow$ (b)”.

*Proof of (ii).* It follows from the above discussion that the functions  $\bar{\beta}^\pm$  defined in Proposition 3.3 are the same as

$$(3.40) \quad \beta^{(n)} = B_{\rho, \mathbf{s}_n, \beta^{(n-1)}}^+ : \mathcal{V}_{\mathbf{s}_n}^\eta \rightarrow \mathcal{S}_1^{d-1}$$

and

$$(3.41) \quad B_{\rho, \mathbf{s}_n, \beta^{(n-1)}}^- : \mathcal{V}_{\mathbf{s}_n}^\eta \rightarrow \mathcal{S}_1^{d-1},$$

respectively, where if  $n = 1$  we understand  $\beta^{(0)} = \beta$ . Hence the claims about  $\bar{\beta}^\pm$  are direct consequences of Lemma 3.2 (iii) (since  $\varepsilon_1 < \varepsilon$ ).

*Proof of (iii).* When  $n = 1$  the statement follows directly from Lemma 3.2 (ii); thus from now on we assume  $n \geq 2$ . Let us write  $\mu^{(k)} = V_*^{(k)}\mu$  for  $k = 1, \dots, n$ , so that  $\mu^{(n)} = \Phi_*\mu$ . We know that for each  $\mathbf{v} \in V^{(1)}(\Delta)$  the ray  $\mathbf{m}_1 + \rho\beta^{(1)}(\mathbf{v}) + \mathbb{R}_{>0}\mathbf{v}$  hits  $\mathbf{m}_2 + \mathcal{B}_\rho^d$ . Hence  $\varphi(\mathbf{v}, \mathbf{s}_2) < \pi C\rho$  by the same argument that led to (3.9). It follows that

$$(3.42) \quad \varphi(\mathbf{v}, \mathbf{v}') < 2\pi C\rho \leq c_{\eta, \varepsilon_1}$$

for all  $\mathbf{v}, \mathbf{v}' \in V^{(1)}(\Delta)$ , by our choice of  $\rho$ . Hence Lemma 3.2 (ii), using  $\rho < \tilde{\rho}_0(\eta, C, \varepsilon_1)$ , together with Lemma 3.1 imply that

$$(3.43) \quad \mu^{(1)} = (\bar{V}_{\rho, \mathbf{s}_1, \beta})_*(\mu) = g_1 \cdot \nu_1$$

where  $\nu_1$  is the  $\mathcal{S}_1^{d-1}$  volume measure restricted to  $V^{(1)}(\Delta)$  and rescaled to be a probability measure, and  $g_1$  is some continuous function from  $V^{(1)}(\Delta)$  to  $[(1 - \varepsilon_1)^2, (1 + \varepsilon_1)^2]$ .

Repeating the same argument, using

$$(3.44) \quad \mu^{(k+1)} = (\bar{V}_{\rho, \mathbf{s}_{k+1}, \beta^{(k)}})_*(\mu^{(k)}),$$

we obtain  $\mu^{(k)} = g_k \cdot \nu_k$  for each  $k = 2, \dots, n-1$ , where  $\nu_k$  is the  $\mathcal{S}_1^{d-1}$  volume measure restricted to  $V^{(k)}(\Delta)$  and rescaled to be a probability measure, and  $g_k$  is some continuous function from  $V^{(k)}(\Delta)$  to  $[(1 - \varepsilon_1)^{2k}, (1 + \varepsilon_1)^{2k}]$ . Using this fact for  $k = n-1$  together with one more application of Lemma 3.2 (ii), we finally obtain  $\Phi_*\mu = \mu^{(n)} = g \cdot \lambda_{\mathbf{s}_n}^\eta$ , where  $g$  is some continuous function from  $V^{(n)}(\Delta) = \mathcal{V}_{\mathbf{s}_n}^\eta$  to  $[(1 - \varepsilon_1)^{2n-1}, (1 + \varepsilon_1)^{2n-1}]$ , thus proving the desired claim.  $\square$

## 4. LOSS OF MEMORY

**4.1. Statement of the main theorem.** For  $\mathbf{v}_0 \in S_1^{d-1}$ , we define the probability density  $p_{\mathbf{0}, \beta_{\mathbf{v}_0}^+}(\mathbf{v}_1, \xi, \mathbf{v}_2)$  on  $S_1^{d-1} \times \mathbb{R}_{\geq 0} \times S_1^{d-1}$  by

$$(4.1) \quad p_{\mathbf{0}, \beta_{\mathbf{v}_0}^+}(\mathbf{v}_1, \xi, \mathbf{v}_2) d\text{vol}_{S_1^{d-1}}(\mathbf{v}_2) = \begin{cases} \Phi_{\mathbf{0}}(\xi, \mathbf{w}, (\beta_{\mathbf{v}_0}^+(\mathbf{v}_1)K(\mathbf{v}_1))_{\perp}) d\mathbf{w} & \text{if } \mathbf{v}_1 \in \mathcal{V}_{\mathbf{v}_0}, \mathbf{v}_2 \in \mathcal{V}_{\mathbf{v}_1} \\ 0 & \text{otherwise,} \end{cases}$$

with  $\mathbf{w} = -\beta_{\mathbf{e}_1}^-(\mathbf{v}_2 K(\mathbf{v}_1))_{\perp} \in \{0\} \times \mathcal{B}_1^{d-1}$ .

*Remark 4.1.* As in Remark 2.4, the density  $p_{\mathbf{0}, \beta_{\mathbf{v}_0}^+}(\mathbf{v}_1, \xi, \mathbf{v}_2)$  is independent of the choice of the function  $K(\mathbf{v}_1)$ ; also  $p_{\mathbf{0}, \beta_{\mathbf{v}_0}^+}(\mathbf{v}_1, \xi, \mathbf{v}_2)$  is continuous at each point  $(\mathbf{v}_1, \xi, \mathbf{v}_2) \in S_1^{d-1} \times \mathbb{R}_{\geq 0} \times S_1^{d-1}$  with  $\mathbf{v}_1 \in \mathcal{V}_{\mathbf{v}_0}$ ,  $\mathbf{v}_2 \in \mathcal{V}_{\mathbf{v}_1}$ , except possibly when  $d = 2$ ,  $\xi > 0$  and  $\beta_{\mathbf{v}_1}^-(\mathbf{v}_2) = -\beta_{\mathbf{v}_0}^+(\mathbf{v}_1)R_{\mathbf{v}_1}$  ( $\Leftrightarrow \mathbf{v}_2 = \Theta_1(\mathbf{v}_1, -\beta_{\mathbf{v}_0}^+(\mathbf{v}_1)R_{\mathbf{v}_1})$ ). Here  $R_{\mathbf{v}_1}$  denotes reflection in the line  $\mathbb{R}\mathbf{v}_1$ .

*Remark 4.2.* Due to the spherical symmetry of the scattering map  $\Theta$  (cf. Section 2.2) we have

$$(4.2) \quad p_{\mathbf{0}, \beta_{\mathbf{v}_0 K}^+}(\mathbf{v}_1 K, \xi, \mathbf{v}_2 K) = p_{\mathbf{0}, \beta_{\mathbf{v}_0}^+}(\mathbf{v}_1, \xi, \mathbf{v}_2)$$

for any  $K \in O(d)$ .

*Remark 4.3.* The explicit formula in Remark 2.5 carries over directly to the present case.

*Remark 4.4.* By (2.12) we have, whenever  $\mathbf{v}_1 \in \mathcal{V}_{\mathbf{v}_0}$ ,

$$(4.3) \quad \int_{S_1^{d-1} \times \mathbb{R}_{>0}} p_{\mathbf{0}, \beta_{\mathbf{v}_0}^+}(\mathbf{v}_1, \xi, \mathbf{v}_2) d\xi d\text{vol}_{S_1^{d-1}}(\mathbf{v}_2) = 1.$$

For the definition of  $p_{\alpha, \beta_0}$ , recall (2.25). The analogue of relation (4.3) of course also holds for  $p_{\alpha, \beta_0}$ , again by (2.12).

**Theorem 4.1.** Fix a lattice  $\mathcal{L} = \mathbb{Z}^d M_0$  and a point  $\mathbf{q} \in \mathbb{R}^d$ , set  $\alpha = -\mathbf{q}M_0^{-1}$ , and let  $\beta_0 : S_1^{d-1} \rightarrow \mathbb{R}^d$  be a  $C^1$  function. If  $\mathbf{q} \in \mathcal{L}$  we assume that  $(\beta_0(\mathbf{v}) + \mathbb{R}_{>0}\mathbf{v}) \cap \mathcal{B}_1^d = \emptyset$  for all  $\mathbf{v} \in S_1^{d-1}$ . Then for any Borel probability measure  $\lambda_0$  on  $S_1^{d-1}$  which is absolutely continuous with respect to  $\text{vol}_{S_1^{d-1}}$ , and for any bounded continuous function  $f_0 : S_1^{d-1} \times (\mathbb{R}_{>0} \times S_1^{d-1})^n \rightarrow \mathbb{R}$ , we have

$$(4.4) \quad \begin{aligned} & \lim_{\rho \rightarrow 0} \int_{S_1^{d-1}} f_0(\mathbf{v}_0, \rho^{d-1} \tau_1(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho), \mathbf{v}_1(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho), \dots, \mathbf{v}_n(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda_0(\mathbf{v}_0) \\ &= \int_{S_1^{d-1} \times (\mathbb{R}_{>0} \times S_1^{d-1})^n} f_0(\mathbf{v}_0, \xi_1, \mathbf{v}_1, \dots, \xi_n, \mathbf{v}_n) p_{\alpha, \beta_0}(\mathbf{v}_0, \xi_1, \mathbf{v}_1) p_{\mathbf{0}, \beta_{\mathbf{v}_0}^+}(\mathbf{v}_1, \xi_2, \mathbf{v}_2) \\ & \quad \cdots p_{\mathbf{0}, \beta_{\mathbf{v}_{n-2}}^+}(\mathbf{v}_{n-1}, \xi_n, \mathbf{v}_n) d\lambda_0(\mathbf{v}_0) \prod_{k=1}^n d\xi_k d\text{vol}_{S_1^{d-1}}(\mathbf{v}_k). \end{aligned}$$

Note that the case  $n = 1$  specializes to the statement of Theorem 2.2. Hence from now on we will assume  $n \geq 2$ .

The proof of Theorem 4.1 is given in Section 4.3, and is based on an iterative application of Theorem 2.2. The idea is that at the  $n$ th step, for a given, small  $\rho > 0$ , we apply Theorem 2.2 once for each possible sequence of balls  $\mathbf{m}_1 + \mathcal{B}_{\rho}^d$ ,  $\mathbf{m}_2 + \mathcal{B}_{\rho}^d$ ,  $\dots$ ,  $\mathbf{m}_{n-1} + \mathcal{B}_{\rho}^d$  ( $\mathbf{m}_k \in \mathcal{L}$ ) causing the first  $n - 1$  collisions in the orbit of the flow  $\varphi_t$ . For each such sequence  $\{\mathbf{m}_k\}_{k=1}^{n-1}$  we apply Theorem 2.2 with  $\lambda$  as the probability measure on  $S_1^{d-1}$  describing the random variable  $\mathbf{v}_{n-1}(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho)$  conditioned on  $\mathbf{v}_0$  leading to  $\{\mathbf{m}_k\}_{k=1}^{n-1}$ , and with  $\beta$  as the function “ $\mathbf{v}_{n-1}(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho) \mapsto \mathbf{w}_{n-1}(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho)$ ” (this makes sense once we restrict to  $\mathbf{v}_0$  leading to the fixed sequence  $\{\mathbf{m}_k\}_{k=1}^{n-1}$ , see Prop. 3.3). Since  $\rho$  is small,  $\beta$  is well approximated by  $\beta_{\mathbf{v}_{n-2}}^+$  (see Prop. 3.3), and the resulting limiting density is  $p_{\mathbf{0}, \beta_{\mathbf{v}_{n-2}}^+}(\mathbf{v}_{n-1}, \xi_n, \mathbf{v}_n)$ .

Note that in Theorem 2.2,  $p_{\alpha, \beta}$  does not depend on the choice of  $\lambda$ ; hence our limiting density  $p_{\mathbf{0}, \beta_{\mathbf{v}_{n-2}}^+}(\mathbf{v}_{n-1}, \xi_n, \mathbf{v}_n)$  is independent of the measure  $d\lambda(\mathbf{v}_{n-1})$ , and thus independent of  $\xi_1, \dots, \xi_{n-1}$  and  $\mathbf{v}_0, \dots, \mathbf{v}_{n-3}$ .

To make the above argument rigorous, we need to use the uniform version Theorem 2.3 in place of Theorem 2.2, for note that both  $\lambda$  and  $\beta$  depend on the hit sequence  $\{\mathbf{m}_k\}_{k=1}^{n-1}$ , as well as on  $\rho$ . For the application of Theorem 2.3 we need to prove that all our  $\lambda$ 's are contained in some equismooth family, and furthermore that  $\beta$  is indeed sufficiently well approximated by  $\beta_{\mathbf{v}_{n-2}}^+$  as stated above.

**4.2. Sets of good initial velocities.** In this section we will define sets of initial velocities  $\mathbf{v}_0$  with good properties, and prove two lemmas which will later allow us to see that the complement of these sets have small  $\lambda_0$ -measure.

Let  $n \geq 2$  and  $\beta_0$  be given as in Theorem 4.1, and set

$$(4.5) \quad C = 10 \max \left( 1, \sup_{S_1^{d-1}} \|\beta_0\|, \sup_{\mathbf{h} \in \mathbb{T}^1(S_1^{d-1})} \|D_{\mathbf{h}}\beta_0\|, \sup_{\mathbf{y} \in \mathcal{L} \setminus \{\mathbf{0}\}} \|\mathbf{y}\|^{-1}, \sup_{\mathbf{y} \in (\mathcal{L}-\mathbf{q}) \setminus \{\mathbf{0}\}} \|\mathbf{y}\|^{-1} \right).$$

Let  $\eta$  and  $\rho$  be arbitrary numbers subject to the conditions

$$(4.6) \quad 0 < \eta < \frac{\pi - B_{\Theta}}{100} \quad \text{and} \quad 0 < \rho < \min \left( \frac{\eta}{C^3 C_{\eta}}, \rho_0(n, \eta, C, \eta) \right),$$

where  $\rho_0(n, \eta, C, \eta)$  is as in Proposition 3.3, and  $C_{\eta}$  as in (3.2). These two numbers  $\eta$  and  $\rho$  will be kept fixed throughout the present section. Recall (2.33), and define  $\omega(\eta) \in (0, \frac{\pi}{2})$  to be the angle such that

$$(4.7) \quad \mathfrak{U}_{\omega(\eta)} = \mathfrak{U}_{2\sqrt{\eta}} \cup -\beta_{e_1}^- \left( \mathcal{V}_{e_1} \setminus \overline{\mathcal{V}_{e_1}^{20\eta}} \right).$$

We also fix a partition  $S_1^{d-1} = \bigsqcup_{j=1}^N D_j$  of the sphere  $S_1^{d-1}$  into Borel subsets  $D_1, \dots, D_N$ , each of positive volume and boundary of measure zero, and with diameter  $< \eta/C_{\eta}$  with respect to the metric  $\varphi$ .

For  $\mathbf{v} \in S_1^{d-1}$  and  $k \geq 1$  we let

$$(4.8) \quad \mathbf{w}_k(\mathbf{v}) = \mathbf{w}_k(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}), \mathbf{v}; \rho) \in S_1^{d-1}, \quad \mathbf{m}_k(\mathbf{v}) = \mathbf{m}_k(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}), \mathbf{v}; \rho) \in \mathcal{L}$$

be the impact position and the ball label at the  $k$ th collision for initial condition  $(\mathbf{q} + \rho\beta_0(\mathbf{v}), \mathbf{v}) \in \mathbb{T}^1(\mathcal{K}_{\rho})$ . Let  $\mathbf{v}_k(\mathbf{v}) = \mathbf{v}_k(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}), \mathbf{v}; \rho) \in S_1^{d-1}$  be the velocity directly after this collision, and  $\tau_k(\mathbf{v}) = \tau_k(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}), \mathbf{v}; \rho) \in \mathbb{R}_{>0}$  be the time elapsed between the  $(k-1)$ th and  $k$ th collision. (Cf. (1.4) and (1.7).) We also set  $\mathbf{m}_0 = \mathbf{q} \in \mathbb{R}^d$  and

$$(4.9) \quad \mathbf{s}_k(\mathbf{v}) = \mathbf{s}_k(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}), \mathbf{v}; \rho) = \mathbf{m}_k(\mathbf{v}) - \mathbf{m}_{k-1}(\mathbf{v})$$

for  $k \geq 1$ . Note that  $\|\mathbf{s}_k(\mathbf{v})\| \geq 10/C$  always holds, by our choice of  $C$ .

Given any  $\mathbf{a} \in S_1^{d-1}$  we let  $[\mathbf{a}]$  be the unique set  $D_j$  for which  $\mathbf{a} \in D_j$ . Set  $\mathcal{U}^{(0)} = S_1^{d-1}$ , and define the subsets  $\mathcal{U}^{(0)} \supset \mathcal{U}^{(1)} \supset \mathcal{U}^{(2)} \supset \dots$  recursively as follows: For each  $k \geq 1$ , we let  $\mathcal{U}^{(k)}$  be the set of all  $\mathbf{v} \in \mathcal{U}^{(k-1)}$  satisfying the following three conditions:

- (I)  $\tau_k(\mathbf{v}) < \infty$ ,
- (II)  $[\mathbf{v}_k(\mathbf{v})] \subset \mathcal{V}_{\mathbf{s}_k}^{10\eta}$ , and
- (III) for each  $\mathbf{u} \in [\mathbf{v}_k(\mathbf{v})]$  there is some  $\mathbf{v}' \in \mathcal{U}^{(k-1)}$  satisfying

$$(4.10) \quad \begin{cases} \mathbf{m}_{\ell}(\mathbf{v}') = \mathbf{m}_{\ell}(\mathbf{v}), & \ell = 1, \dots, k; \\ \mathbf{v}_k(\mathbf{v}') = \mathbf{u}. \end{cases}$$

If  $\mathbf{v} \in \mathcal{U}^{(k)}$  and  $\mathbf{u} \in [\mathbf{v}_k(\mathbf{v})]$ , then every  $\mathbf{v}' \in \mathcal{U}^{(k-1)}$  satisfying (4.10) will actually lie in  $\mathcal{U}^{(k)}$ .

Given a vector  $\mathbf{v} \in \mathcal{U}^{(k)}$  ( $1 \leq k \leq n$ ) we let  $M_{\mathbf{v}}^k$  be the set of those indices  $i$  for which (4.10) holds for some  $\mathbf{u} \in D_i$  and  $\mathbf{v}' \in \mathcal{U}^{(k)}$ . Then note that  $\cup_{i \in M_{\mathbf{v}}^k} D_i \subset \mathcal{V}_{\mathbf{s}_k}^{10\eta}$  and for each  $\mathbf{u} \in \cup_{i \in M_{\mathbf{v}}^k} D_i$  there is some  $\mathbf{v}' \in \mathcal{U}^{(k)}$  such that (4.10) holds. Write  $\vec{\mathbf{m}}_{\mathbf{v}}^k := (\mathbf{m}_0, \dots, \mathbf{m}_k)$  where  $\mathbf{m}_0 = \mathbf{q}$  and  $\mathbf{m}_{\ell} = \mathbf{m}_{\ell}(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}), \mathbf{v}; \rho)$  ( $\ell = 1, \dots, k$ ) as before. For each  $\ell \in \{1, \dots, k-1\}$  we



have  $\mathbf{v}_\ell \in \mathcal{V}_{\mathbf{s}_\ell}^{10\eta}$  since  $\mathbf{v} \in \mathcal{U}^{(\ell)}$ , and also, since the ray  $\mathbf{m}_\ell + \rho\boldsymbol{\beta}_{\mathbf{v}_{\ell-1}}^+(\mathbf{v}_\ell) + \mathbb{R}_{>0}\mathbf{v}_\ell$  hits  $\mathbf{m}_{\ell+1} + \mathcal{B}_\rho^d$  we have

$$(4.11) \quad \varphi(\mathbf{v}_\ell, \mathbf{s}_{\ell+1}) \leq \arcsin \frac{2\rho}{\|\mathbf{s}_{\ell+1}\|} \leq \arcsin \frac{C\rho}{5} < \frac{1}{2}C\rho < \eta;$$

hence  $\varphi(\mathbf{s}_\ell, \mathbf{s}_{\ell+1}) > B_\Theta + 9\eta$ . Also recall  $\rho < \rho_0(n, \eta, C, \eta)$ . Hence the data  $\vec{\mathbf{m}}_{\mathbf{v}}^k$ ,  $\rho$  and  $\boldsymbol{\beta}_0$  satisfy all the assumptions of Proposition 3.3. Now taking  $\Phi = \Phi_{\rho, \vec{\mathbf{m}}_{\mathbf{v}}^k, \boldsymbol{\beta}_0} : \Delta \xrightarrow{\sim} \mathcal{V}_{\mathbf{s}_k}^\eta$  to be the diffeomorphism in that proposition, it follows that for every  $\mathbf{u} \in \cup_{i \in M_{\mathbf{v}}^k} D_i$  there is a *unique*  $\mathbf{v}' \in \mathcal{U}^{(k)}$  for which (4.10) holds, namely  $\mathbf{v}' = \Phi^{-1}(\mathbf{u})$ . Proposition 3.3 also implies that for any such pair  $\mathbf{u}, \mathbf{v}'$ , we have regarding the starting point of the  $\mathbf{u}$ -ray leaving the  $\mathbf{m}_k$ -ball:  $\boldsymbol{\beta}_{\mathbf{v}_{k-1}(\mathbf{q}_{\rho, \boldsymbol{\beta}_0}(\mathbf{v}'), \mathbf{v}'; \rho)}^+(\mathbf{u}) = \bar{\boldsymbol{\beta}}^+(\mathbf{u})$  with  $\bar{\boldsymbol{\beta}}^+ = \bar{\boldsymbol{\beta}}_{\rho, \vec{\mathbf{m}}_{\mathbf{v}}^k, \boldsymbol{\beta}_0}^+ : \mathcal{V}_{\mathbf{s}_k}^\eta \rightarrow S_1^{d-1}$ .

**Lemma 4.2.** *Assume  $1 \leq k \leq n-1$ ,  $\mathbf{v} \in \mathcal{U}^{(k)}$  and  $\bar{\boldsymbol{\beta}}^+ = \bar{\boldsymbol{\beta}}_{\rho, \vec{\mathbf{m}}_{\mathbf{v}}^k, \boldsymbol{\beta}_0}^+$ . If  $\mathbf{v} \notin \mathcal{U}^{(k+1)}$  then one of the following statements holds:*

- (i)  $\tau_{k+1}(\mathbf{v}) = \infty$ ,
- (ii)  $\mathbf{v}_k(\mathbf{v}) \in \partial_{C\rho}(\cup_{i \in M_{\mathbf{v}}^k} D_i)$ ,
- (iii)  $-\mathbf{w}_1(\rho\bar{\boldsymbol{\beta}}^+(\mathbf{v}_k), \mathbf{v}_k; \rho)K(\mathbf{v}_k) \in \mathfrak{U}_{\omega(\eta)}$ ,
- (iv)  $-\mathbf{w}_1(\rho(\bar{\boldsymbol{\beta}}^+(\mathbf{v}_k) + 3\mathbf{v}_k), \mathbf{v}_k; (1+\eta)\rho)K(\mathbf{v}_k) \in \mathfrak{U}_{\omega(\eta)}$ .

*Proof.* Fix any vector  $\mathbf{v} \in \mathcal{U}^{(k)} \setminus \mathcal{U}^{(k+1)}$ , assume  $\tau_{k+1}(\mathbf{v}) < \infty$  so that  $\mathbf{w}_{k+1} = \mathbf{w}_{k+1}(\mathbf{v})$  and  $\mathbf{v}_{k+1} = \mathbf{v}_{k+1}(\mathbf{v})$  exist, and write  $D_j = [\mathbf{v}_{k+1}]$ . Now by the definition of  $\mathcal{U}^{(k+1)}$  we have either  $D_j \not\subset \mathcal{V}_{\mathbf{s}_{k+1}}^{10\eta}$ , or else there is some  $\mathbf{u} \in D_j$  such that there does *not* exist any  $\mathbf{v}' \in \mathcal{U}^{(k)}$  with

$$(4.12) \quad \begin{cases} \mathbf{m}_\ell(\mathbf{v}') = \mathbf{m}_\ell(\mathbf{v}), & \ell = 1, \dots, k+1; \\ \mathbf{v}_{k+1}(\mathbf{v}') = \mathbf{u}. \end{cases}$$

First assume that  $D_j \not\subset \mathcal{V}_{\mathbf{s}_{k+1}}^{10\eta}$ , i.e. there is some  $\mathbf{u} \in D_j$  with  $\mathbf{u} \notin \mathcal{V}_{\mathbf{s}_{k+1}}^{10\eta}$ . Then  $\varphi(\mathbf{u}, \mathbf{s}_{k+1}) \leq B_\Theta + 10\eta$ ; also  $\varphi(\mathbf{v}_{k+1}, \mathbf{u}) < \eta$  since  $D_j$  has diameter less than  $\eta$ , and  $\varphi(\mathbf{s}_{k+1}, \mathbf{v}_k) < \frac{1}{2}C\rho < \eta$  as in (4.11). Hence  $\varphi(\mathbf{v}_{k+1}, \mathbf{v}_k) < B_\Theta + 12\eta$ , and using  $\mathbf{w}_{k+1} = \boldsymbol{\beta}_{\mathbf{v}_k}^-(\mathbf{v}_{k+1})$  and the spherical invariance of  $\boldsymbol{\beta}^-$  this implies  $-\mathbf{w}_{k+1}K(\mathbf{v}_k) \in \mathfrak{U}_{\omega(\eta)}$  (cf. (4.7)). But

$$(4.13) \quad \mathbf{w}_{k+1} = \mathbf{w}_1(\mathbf{m}_k + \rho\bar{\boldsymbol{\beta}}^+(\mathbf{v}_k), \mathbf{v}_k; \rho) = \mathbf{w}_1(\rho\bar{\boldsymbol{\beta}}^+(\mathbf{v}_k), \mathbf{v}_k; \rho),$$

since  $\mathcal{L} - \mathbf{m}_k = \mathcal{L}$ . Hence (iii) in Lemma 4.2 holds.

It remains to consider the case when  $D_j \subset \mathcal{V}_{\mathbf{s}_{k+1}}^{10\eta}$  but there is some  $\mathbf{u} \in D_j$  (which we now consider as fixed) such that (4.12) does not hold for any  $\mathbf{v}' \in \mathcal{U}^{(k)}$ . As in the discussion preceding the lemma we have  $\varphi(\mathbf{s}_\ell, \mathbf{s}_{\ell+1}) > B_\Theta + 9\eta$  for all  $\ell = 1, \dots, k$ , so that the data  $\vec{\mathbf{m}}_{\mathbf{v}}^{k+1}$ ,  $\rho$  and  $\boldsymbol{\beta}_0$  satisfy all the assumptions of Proposition 3.3. Now  $\mathbf{u} \in D_j \subset \mathcal{V}_{\mathbf{s}_{k+1}}^{10\eta}$  lies in the range of  $\Phi_{\rho, \vec{\mathbf{m}}_{\mathbf{v}}^{k+1}, \boldsymbol{\beta}_0}$ ; we set  $\mathbf{v}' = \Phi_{\rho, \vec{\mathbf{m}}_{\mathbf{v}}^{k+1}, \boldsymbol{\beta}_0}^{-1}(\mathbf{u}) \in S_1^{d-1}$ . Then by “(b) $\Rightarrow$ (a)” in Proposition 3.3, the ray  $\mathbf{q}_{\rho, \boldsymbol{\beta}_0}(\mathbf{v}') + \mathbb{R}_{>0}\mathbf{v}'$  hits  $\mathbf{m}_1 + \mathcal{B}_\rho^d$ ; after scattering in this ball we get a ray which hits  $\mathbf{m}_2 + \mathcal{B}_\rho^d$ , and so on for  $\mathbf{m}_3 + \mathcal{B}_\rho^d, \dots, \mathbf{m}_{k+1} + \mathcal{B}_\rho^d$ , and after the final scattering we get a ray with direction  $\mathbf{u}$ . (So far we make no claim for any  $j$  on whether or not the ray leaving  $\mathbf{m}_j + \mathcal{B}_\rho^d$  passes through any *other* ball in  $\mathcal{L} + \mathcal{B}_\rho^d$  before hitting  $\mathbf{m}_{j+1} + \mathcal{B}_\rho^d$ .)

Let  $\mathbf{v}'_\ell$  be the direction of the  $\ell$ th ray in the above sequence ( $\ell \in \{0, \dots, k+1\}$ , thus  $\mathbf{v}'_0 = \mathbf{v}'$ ,  $\mathbf{v}'_{k+1} = \mathbf{u}$ ), so that  $\mathbf{m}_k + \rho\boldsymbol{\beta}_{\mathbf{v}'_{k-1}}^+(\mathbf{v}'_k) + \mathbb{R}_{>0}\mathbf{v}'_k$  is the ray which leaves  $\mathbf{m}_k + \mathcal{B}_\rho^d$  and hits  $\mathbf{m}_{k+1} + \mathcal{B}_\rho^d$ . We have  $\varphi(\mathbf{s}_{k+1}, \mathbf{v}'_k) < \frac{1}{2}C\rho < \eta$  (cf. (4.11)) and thus

$$(4.14) \quad \varphi(\mathbf{s}_k, \mathbf{v}'_k) \geq \varphi(\mathbf{s}_k, \mathbf{s}_{k+1}) - \varphi(\mathbf{s}_{k+1}, \mathbf{v}'_k) > B_\Theta + 9\eta - \eta$$

so that  $\mathbf{v}'_k \in \mathcal{V}_{s_k}^\eta$ . Hence by “(a) $\Rightarrow$ (b)” in Proposition 3.3,  $\mathbf{v}'$  also belongs to the domain of  $\Phi_{\rho, \bar{\mathbf{m}}_v^k, \beta_0}$ , and  $\mathbf{v}'_k = \Phi_{\rho, \bar{\mathbf{m}}_v^k, \beta_0}(\mathbf{v}')$  and

$$(4.15) \quad \beta_{\mathbf{v}'_{k-1}}^+(\mathbf{v}'_k) = \bar{\beta}^+(\mathbf{v}'_k) = \bar{\beta}_{\rho, \bar{\mathbf{m}}_v^k, \beta_0}^+(\mathbf{v}'_k).$$

From  $\varphi(\mathbf{s}_{k+1}, \mathbf{v}_k) < \frac{1}{2}C\rho$  and  $\varphi(\mathbf{s}_{k+1}, \mathbf{v}'_k) < \frac{1}{2}C\rho$  we get  $\varphi(\mathbf{v}_k, \mathbf{v}'_k) < C\rho$ . Note that  $\mathbf{v}_k \in \cup_{i \in M_v^k} D_i$  since  $\mathbf{v} \in \mathcal{U}^{(k)}$ . Now if  $\mathbf{v}'_k$  lies outside  $\cup_{i \in M_v^k} D_i$  then there is some point on the geodesic between  $\mathbf{v}_k$  and  $\mathbf{v}'_k$  which lies in  $\partial(\cup_{i \in M_v^k} D_i)$ , and hence  $\mathbf{v}_k \in \partial_{C\rho}(\cup_{i \in M_v^k} D_i)$ , i.e. (ii) in Lemma 4.2 holds.

It remains to consider the case  $\mathbf{v}'_k \in \cup_{i \in M_v^k} D_i$ . As in the discussion preceding the lemma,  $\mathbf{v}'_k \in \cup_{i \in M_v^k} D_i$  and  $\mathbf{v}' = \Phi_{\rho, \bar{\mathbf{m}}_v^k, \beta_0}^{-1}(\mathbf{v}'_k)$  imply that  $\mathbf{v}' \in \mathcal{U}^{(k)}$  and that (4.10) holds with  $\mathbf{v}'_k$  in place of  $\mathbf{u}$ ; thus  $\mathbf{v}'_\ell = \mathbf{v}_\ell(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}', \mathbf{v}'; \rho))$  for all  $\ell = 1, \dots, k$ . We know from above that the ray  $\mathbf{m}_k + \rho\beta_{\mathbf{v}'_{k-1}}^+(\mathbf{v}'_k) + \mathbb{R}_{>0}\mathbf{v}'_k$  hits  $\mathbf{m}_{k+1} + \mathcal{B}_\rho^d$ ; let us choose  $\tau' > 0$  and  $\mathbf{w}'_{k+1} \in S_1^{d-1}$  so that the point of impact is

$$(4.16) \quad \mathbf{m}_k + \rho\beta_{\mathbf{v}'_{k-1}}^+(\mathbf{v}'_k) + \tau'\mathbf{v}'_k = \mathbf{m}_{k+1} + \rho\mathbf{w}'_{k+1}.$$

After scattering in the ball  $\mathbf{m}_{k+1} + \mathcal{B}_\rho^d$  we get a ray with direction  $\mathbf{u}$ . It follows that the ray

$$(4.17) \quad \mathbf{m}_k + \rho\beta_{\mathbf{v}'_{k-1}}^+(\mathbf{v}'_k) + \mathbb{R}_{>0}\mathbf{v}'_k$$

must intersect some *other* ball in  $\mathcal{L} + \mathcal{B}_\rho^d$  before it hits  $\mathbf{m}_{k+1} + \mathcal{B}_\rho^d$ , for otherwise we would conclude  $\mathbf{m}_{k+1}(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}', \mathbf{v}'; \rho)) = \mathbf{m}_{k+1}$  and  $\mathbf{v}_{k+1}(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}', \mathbf{v}'; \rho)) = \mathbf{u}$ , i.e. (4.12) would hold for our  $\mathbf{v}' \in \mathcal{U}^{(k)}$ , contrary to our assumption. Thus there is some point

$$(4.18) \quad \mathbf{m}' = \mathbf{m}_{k+1}(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}', \mathbf{v}'; \rho)) \in \mathcal{L}, \quad \mathbf{m}' \neq \mathbf{m}_{k+1},$$

and some  $t' \in (0, \tau')$  such that  $\mathbf{m}_k + \rho\beta_{\mathbf{v}'_{k-1}}^+(\mathbf{v}'_k) + t'\mathbf{v}'_k \in \mathbf{m}' + \mathcal{B}_\rho^d$ . Note that  $\|\mathbf{m}' - \mathbf{m}_{k+1}\| \geq 10/C > 10^3\rho$ , and hence  $t' < \tau' - 100\rho$ .

Using  $\varphi(\mathbf{v}_k, \mathbf{v}'_k) < C\rho < \eta$  and  $\mathbf{v}_k \in \mathcal{V}_{s_k}^{10\eta}$  we see that the  $(S_1^{d-1})$ -geodesic from  $\mathbf{v}_k$  to  $\mathbf{v}'_k$  is contained inside  $\mathcal{V}_{s_k}^\eta$ ; hence from  $\sup_{\mathbf{h} \in T^1(\mathcal{V}_{s_k}^\eta)} \|D_{\mathbf{h}}\bar{\beta}^+\| < C_\eta$  (see Proposition 3.3) we deduce

$$(4.19) \quad \|\bar{\beta}^+(\mathbf{v}_k) - \bar{\beta}^+(\mathbf{v}'_k)\| < C_\eta C\rho < \eta.$$

Similarly, writing  $\bar{\beta}_{k+1}^- := \bar{\beta}_{\rho, \bar{\mathbf{m}}_v^{k+1}, \beta_0}^-$  and using  $\mathbf{w}_{k+1} = \bar{\beta}_{k+1}^-(\mathbf{v}_{k+1})$ ,  $\mathbf{w}'_{k+1} = \bar{\beta}_{k+1}^-(\mathbf{u})$  and  $\mathbf{v}_{k+1}, \mathbf{u} \in D_j \subset \mathcal{V}_{s_{k+1}}^{10\eta}$  (thus  $\varphi(\mathbf{v}_{k+1}, \mathbf{u}) < \frac{\eta}{C_\eta}$ ), we obtain

$$(4.20) \quad \|\mathbf{w}_{k+1} - \mathbf{w}'_{k+1}\| < \eta.$$

Thus the line segment  $\mathbf{m}_k + \rho\bar{\beta}^+(\mathbf{v}'_k) + (0, \tau')\mathbf{v}'_k$  has both its endpoints at distance  $< \rho\eta$  from the corresponding endpoints of the line segment  $\mathbf{m}_k + \rho\bar{\beta}^+(\mathbf{v}_k) + (0, \tau_{k+1})\mathbf{v}_k$ . Hence  $|\tau' - \tau_{k+1}| < 2\rho\eta$ , and there exists some  $t > 0$  with  $|t - t'| < 2\rho\eta$  such that

$$(4.21) \quad \|(\mathbf{m}_k + \rho\bar{\beta}^+(\mathbf{v}_k) + t\mathbf{v}_k) - (\mathbf{m}_k + \rho\bar{\beta}^+(\mathbf{v}'_k) + t'\mathbf{v}'_k)\| < \rho\eta.$$

It follows from this that  $\mathbf{m}_k + \rho\bar{\beta}^+(\mathbf{v}_k) + t\mathbf{v}_k \in \mathbf{m}' + \mathcal{B}_{\tilde{\rho}}^d$ , with  $\tilde{\rho} = (1 + \eta)\rho$ , and therefore  $\tau_1'' := \tau_1(\mathbf{m}_k + \rho(\bar{\beta}^+(\mathbf{v}_k) + 3\mathbf{v}_k), \mathbf{v}_k, \tilde{\rho})$  satisfies  $\tau_1'' < t - 3\rho$ . Take  $\mathbf{m}'' \in \mathcal{L}$  and set  $\mathbf{w}_1'' = \mathbf{w}_1(\mathbf{m}_k + \rho(\bar{\beta}^+(\mathbf{v}_k) + 3\mathbf{v}_k), \mathbf{v}_k, \tilde{\rho}) \in S_1^{d-1}$ , so that

$$(4.22) \quad \mathbf{m}_k + \rho(\bar{\beta}^+(\mathbf{v}_k) + 3\mathbf{v}_k) + \tau_1''\mathbf{v}_k = \mathbf{m}'' + \tilde{\rho}\mathbf{w}_1''.$$

Note that if the ray  $\mathbf{m}_k + \rho \bar{\boldsymbol{\beta}}^+(\mathbf{v}_k) + \mathbb{R}_{>0} \mathbf{v}_k$  would also intersect the slightly smaller ball  $\mathbf{m}'' + \mathcal{B}_\rho^d$  then  $\tau_1(\mathbf{m}_k + \rho(\bar{\boldsymbol{\beta}}^+(\mathbf{v}_k) + 3\mathbf{v}_k), \mathbf{v}_k, \rho) < \tau_1'' + 3\rho$ , and we would get a contradiction:

$$(4.23) \quad \begin{aligned} \tau_{k+1} = \tau_1(\mathbf{m}_k + \rho(\bar{\boldsymbol{\beta}}^+(\mathbf{v}_k) + 3\mathbf{v}_k), \mathbf{v}_k, \rho) + 3\rho &< \tau_1'' + 6\rho < t + 3\rho < t' + 4\rho \\ &< \tau' - 96\rho < \tau_{k+1} - 95\rho. \end{aligned}$$

Thus the ray  $\mathbf{m}_k + \rho \bar{\boldsymbol{\beta}}^+(\mathbf{v}_k) + \mathbb{R}_{>0} \mathbf{v}_k$  intersects  $\mathbf{m}'' + \mathcal{B}_\rho^d$  but *not*  $\mathbf{m}'' + \mathcal{B}_\rho^d$ . This implies  $-\mathbf{w}_1'' K(\mathbf{v}_k) \in \mathfrak{U}_{2\sqrt{\eta}} \subset \mathfrak{U}_{\omega(\eta)}$ . Hence (iv) in Lemma 4.2 holds.  $\square$

We next give the analogue of Lemma 4.2 in the case  $k = 0$ . Let us define  $\tilde{\boldsymbol{\beta}}_0 \in C^1(\mathbb{S}_1^{d-1})$  by

$$(4.24) \quad \tilde{\boldsymbol{\beta}}_0(\mathbf{v}) = (1 + \eta)^{-1}(\boldsymbol{\beta}_0(\mathbf{v}) + (C + 3)\mathbf{v}).$$

**Lemma 4.3.** *Set  $\tilde{\rho} = (1 + \eta)\rho$ . For each  $\mathbf{v} \in \mathbb{S}_1^{d-1} \setminus \mathcal{U}^{(1)}$ , one of the following statements holds:*

- (i)  $\tau_1(\mathbf{v}) = \infty$ ,
- (ii)  $-\mathbf{w}_1(\mathbf{q} + \rho \boldsymbol{\beta}_0(\mathbf{v}), \mathbf{v}; \rho) K(\mathbf{v}) \in \mathfrak{U}_{\omega(\eta)}$ ,
- (iii)  $-\mathbf{w}_1(\mathbf{q} + \tilde{\rho} \tilde{\boldsymbol{\beta}}_0(\mathbf{v}), \mathbf{v}; \tilde{\rho}) K(\mathbf{v}) \in \mathfrak{U}_{\omega(\eta)}$ .

*Proof.* This is very similar to the proof of Lemma 4.2, but easier in certain respects.  $\square$

**4.3. Proof of Theorem 4.1.** Let  $n \geq 2$ ,  $\lambda_0, \boldsymbol{\beta}_0, f_0$  be given as in the statement of the theorem.

**4.3.1. Some reductions.** Set  $X = \mathbb{S}_1^{d-1} \times (\mathbb{R}_{>0} \times \mathbb{S}_1^{d-1})^n$  and let  $\nu$  be the Borel measure on  $X$  which appears in the right hand side of (4.4) in Theorem 4.1, i.e. set

$$(4.25) \quad \nu(M) = \int_M p_{\boldsymbol{\alpha}, \boldsymbol{\beta}_0}(\mathbf{v}_0, \xi_1, \mathbf{v}_1) \left( \prod_{j=1}^{n-1} p_{\mathbf{0}, \boldsymbol{\beta}_{\mathbf{v}_{j-1}}^+}(\mathbf{v}_j, \xi_{j+1}, \mathbf{v}_{j+1}) \right) d\lambda_0(\mathbf{v}_0) d\xi_1 \cdots d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_n)$$

for every Borel subset  $M \subset X$ . Note that repeated application of (4.3), and the analogous relation for  $p_{\boldsymbol{\alpha}, \boldsymbol{\beta}_0}(\mathbf{v}_0, \xi_1, \mathbf{v}_1)$ , yields  $\nu(X) = 1$ . Set

$$(4.26) \quad X_0 = \{(\mathbf{v}_0, \xi_1, \dots, \mathbf{v}_n) \in X : \varphi(\mathbf{v}_{j-1}, \mathbf{v}_j) > B_\Theta, \forall j \in \{1, 2, \dots, n\}\}.$$

This is an open subset of  $X$  with  $\nu(X \setminus X_0) = 0$ , i.e.  $\nu(X_0) = 1$ . Hence we may now assume, without loss of generality, that  $f_0$  has compact support contained inside  $X_0$ . The general case of Theorem 4.1 then follows by a standard approximation argument.

We define functions  $f_m : \mathbb{S}_1^{d-1} \times (\mathbb{R}_{>0} \times \mathbb{S}_1^{d-1})^{n-m} \rightarrow \mathbb{R}$  for  $m = 1, \dots, n-1$  by the following recursive formula:

$$(4.27) \quad \begin{aligned} f_m(\mathbf{v}_0, \xi_1, \mathbf{v}_1, \dots, \xi_{n-m}, \mathbf{v}_{n-m}) &:= \int_{\mathbb{S}_1^{d-1}} \int_{\mathbb{R}_{>0}} f_{m-1}(\mathbf{v}_0, \xi_1, \mathbf{v}_1, \dots, \mathbf{v}_{n-m}, \xi, \mathbf{v}) \\ &\quad \times p_{\mathbf{0}, \boldsymbol{\beta}_{\mathbf{v}_{n-m-1}}^+}(\mathbf{v}_{n-m}, \xi, \mathbf{v}) d\xi d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}). \end{aligned}$$

The point of this is that now the right hand side of (4.4) in Theorem 4.1 can be expressed as

$$(4.28) \quad \int_X f_0 d\nu = \int_{\mathbb{S}_1^{d-1}} \int_{\mathbb{R}_{>0}} \int_{\mathbb{S}_1^{d-1}} f_{n-1}(\mathbf{v}_0, \xi_1, \mathbf{v}_1) p_{\boldsymbol{\alpha}, \boldsymbol{\beta}_0}(\mathbf{v}_0, \xi_1, \mathbf{v}_1) d\lambda_0(\mathbf{v}_0) d\xi_1 d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_1).$$

Since  $f_0$  has compact support contained in  $X_0$  there exists a constant  $\delta > 0$  such that  $f_0(\mathbf{v}_0, \xi_1, \dots, \mathbf{v}_n) = 0$  unless  $\varphi(\mathbf{v}_{j-1}, \mathbf{v}_j) > B_\Theta + \delta$  for all  $j = 1, 2, \dots, n$ . Hence by induction each  $f_m$  has compact support, and we have  $f_m(\mathbf{v}_0, \xi_1, \dots, \mathbf{v}_{n-m}) = 0$  unless  $\varphi(\mathbf{v}_{j-1}, \mathbf{v}_j) > B_\Theta + \delta$  for all  $j = 1, 2, \dots, n-m$ . Using this fact together with Remark 2.1 (rewriting (4.27) via (4.1)) one shows that *each  $f_m$  is continuous*. Furthermore, by repeated use of (4.3) we see that  $\sup |f_m| \leq \sup |f_0|$  for all  $m = 0, 1, \dots, n-1$ .

Finally, by a standard approximation argument we may assume without loss of generality that  $\lambda_0$  has a continuous density, i.e. that

$$(4.29) \quad \lambda_0 = g \cdot \text{vol}_{\mathbb{S}_1^{d-1}} \quad \text{for some continuous function } g : \mathbb{S}_1^{d-1} \rightarrow \mathbb{R}_{\geq 0}.$$

Indeed, note that for  $\lambda_0 = g \cdot \text{vol}_{S_1^{d-1}}$ ,  $g \in L^1$ , the right hand side of (4.4) depends linearly on  $g$  and is bounded in absolute value by  $(\sup |f_0|) \cdot \|g\|_{L^1}$ ; the same is true for the left hand side of (4.4), for each fixed  $\rho$ . We can thus mimic the first paragraph in the proof of Theorem 1.2 in [19, Sec. 9.2].

4.3.2. *Choosing  $\eta, \rho_0$  and  $F_1, F_2, F_3$ .* Let  $\varepsilon > 0$  be given. On the next few pages we will describe how to choose auxiliary positive numbers  $\eta$  and  $\rho_0$ , as well as families  $F_1, F_2, F_3$  of measures and functions to use in applications of Theorem 2.3. The goal is to set up things so that we will be able to prove that the two sides in (4.4) in Theorem 4.1 differ by at most  $O(\varepsilon)$  for every  $\rho \in (0, \rho_0)$ ; see (4.40) below.

Recall the definition of  $\omega(\eta)$ , (4.7), and note that  $\lim_{\eta \rightarrow 0} \omega(\eta) = 0$ . As in the discussion leading to (2.44) and (2.45) we see that we may fix  $0 < \eta < \min(\frac{\pi - B_\Theta}{100}, \varepsilon)$  so small that  $2\omega(\eta) < \frac{\pi}{2}$  and so that the following two inequalities hold for all absolutely continuous probability measures  $\lambda$  on  $S_1^{d-1}$ , all continuous functions  $\beta : S_1^{d-1} \rightarrow \mathbb{R}^d$ , any  $\alpha' \in \{\mathbf{0}, \alpha\}$ , and any two measurable functions  $z_1, z_2 : \text{supp}(\lambda) \rightarrow \{0\} \times \overline{\mathcal{B}_1^{d-1}}$  satisfying  $\|z_1(\mathbf{v}) - z_2(\mathbf{v})\| \leq \eta$  for all  $\mathbf{v} \in \text{supp}(\lambda)$ :

$$(4.30) \quad \int_0^\infty \int_{(\mathbb{A}_{2\omega(\eta)})_\perp} \int_{S_1^{d-1}} \Phi_{\alpha'}(\xi, \mathbf{w}, (\beta(\mathbf{v})K(\mathbf{v}))_\perp) d\lambda(\mathbf{v}) d\mathbf{w} d\xi < \varepsilon;$$

$$(4.31) \quad \int_0^\infty \int_{\{0\} \times \mathcal{B}_1^{d-1}} \int_{\text{supp}(\lambda)} \left| \Phi_{\mathbf{0}}(\xi, \mathbf{w}, z_1(\mathbf{v})) - \Phi_{\mathbf{0}}(\xi, \mathbf{w}, z_2(\mathbf{v})) \right| d\lambda(\mathbf{v}) d\mathbf{w} d\xi < \varepsilon.$$

As in Section 4.2 we fix a partition  $S_1^{d-1} = \bigsqcup_{j=1}^N D_j$  of the sphere  $S_1^{d-1}$  into Borel subsets  $D_1, \dots, D_N$ , each of positive volume and boundary of measure zero, and with diameter  $< \eta/C_\eta$  with respect to the metric  $\varphi$ . Given  $\mathbf{s} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$  and any subset  $M \subset \{1, \dots, N\}$  with  $\lambda_{\mathbf{s}}(\cup_{j \in M} D_j) > 0$ , we let  $\lambda_{M, \mathbf{s}}$  be the restriction of the measure  $\lambda_{\mathbf{s}}$  (cf. Sec. 3) to  $\cup_{j \in M} D_j$ , rescaled to be a probability measure:

$$(4.32) \quad \lambda_{M, \mathbf{s}} = \lambda_{\mathbf{s}}(\cup_M D_j)^{-1} \cdot \lambda_{\mathbf{s}|(\cup_M D_j)}.$$

Let  $F_1$  be the following family of probability measures:

$$(4.33) \quad F_1 := \{\lambda_0\} \cup \{\lambda_{M, \mathbf{s}} : \emptyset \neq M \subset \{1, \dots, N\}, \mathbf{s} \in S_1^{d-1}, (\cup_M D_j) \subset \mathcal{V}_{\mathbf{s}}^{10\eta}\}.$$

Note that  $\lim_{h \rightarrow 0^+} \text{vol}_{S_1^{d-1}}(\partial_h D_j) = 0$  for each  $j$ , since  $D_j$  has boundary of measure zero. Also note that  $\mathcal{V}_{\mathbf{s}}^{10\eta}$  has compact closure in  $\mathcal{V}_{\mathbf{s}}$ , and recall (3.3). From these facts together with (4.29), it follows that  $F_1$  is equismooth.

Let us fix a continuous function  $H : S_1^{d-1} \times \mathbb{R}_{>0} \times S_1^{d-1} \rightarrow [0, 1]$  of the form  $H(\mathbf{v}_0, \xi, \mathbf{v}_1) = H_0(\varphi(\mathbf{v}_0, \mathbf{v}_1))$  such that  $H(\mathbf{v}_0, \xi, \mathbf{v}_1) = 1$  whenever  $\varphi(\mathbf{v}_0, \mathbf{v}_1) \leq B_\Theta$  or  $\varphi(\mathbf{v}_0, -\beta_{\mathbf{v}_0}^-(\mathbf{v}_1)) \geq \frac{\pi}{2} - \omega(\eta)$ , and  $H(\mathbf{v}_0, \xi, \mathbf{v}_1) = 0$  whenever  $\varphi(\mathbf{v}_0, -\beta_{\mathbf{v}_0}^-(\mathbf{v}_1)) \leq \frac{\pi}{2} - 2\omega(\eta)$ . This is possible since it follows from our assumptions on  $\Theta_1$  in Sec. 2.2 that  $\varphi(\mathbf{v}_0, -\beta_{\mathbf{v}_0}^-(\mathbf{v}_1))$  only depends on  $\varphi(\mathbf{v}_0, \mathbf{v}_1)$ , as a strictly decreasing  $C^1$  function. Now let  $F_2$  be the following family of functions on  $S_1^{d-1} \times \mathbb{R}_{>0} \times S_1^{d-1}$ ,

$$(4.34) \quad F_2 = \{H\} \cup \left\{ (\mathbf{v}, \xi, \mathbf{v}') \mapsto f_m(\mathbf{v}_0, \xi_1, \dots, \xi_{n-m-1}, \mathbf{v}, \xi, \mathbf{v}') : \right. \\ \left. 0 \leq m \leq n-1, (\mathbf{v}_0, \xi_1, \dots, \xi_{n-m-1}) \in (S_1^{d-1} \times \mathbb{R}_{>0})^{n-m-1} \right\}.$$

Using the fact that each  $f_m$  has compact support we see that  $F_2$  is equicontinuous and uniformly bounded.

Let  $C$  be fixed as in (4.5), and let, for any  $\mathbf{q}' \in \mathbb{R}^d$ :

$$(4.35) \quad F_3^{(\mathbf{q}')} = \left\{ \beta : S_1^{d-1} \rightarrow \mathbb{R}^d : \beta \text{ is } C^1, \sup \|\beta\| \leq 3C, \sup_{\mathbf{h} \in T^1(S_1^{d-1})} \|D_{\mathbf{h}}\beta\| \leq 2C + \frac{4}{\eta} + C_\eta, \right. \\ \left. \text{if } \mathbf{q}' \in \mathcal{L}: (\beta(\mathbf{v}) + \mathbb{R}_{>0}\mathbf{v}) \cap \mathcal{B}_1^d = \emptyset, \forall \mathbf{v} \in S_1^{d-1} \right\}.$$

Let  $\rho_0(n, \eta, C, \eta)$  be as in Proposition 3.3. Now fix  $0 < \rho_0 < \min\left(\frac{\eta}{C^3 C_\eta}, \rho_0(n, \eta, C, \eta)\right)$  so small that for  $\mathbf{q}' \in \{\mathbf{0}, \mathbf{q}\}$  and all  $\rho \in (0, 2\rho_0)$ ,  $\lambda \in F_1$ ,  $f \in F_2$  and  $\beta \in F_3^{(\mathbf{q}' )}$ , we have (here and below we write  $\alpha' = \mathbf{0}$  if  $\mathbf{q}' = \mathbf{0}$ ;  $\alpha' = \alpha$  if  $\mathbf{q}' = \mathbf{q}$ )

$$(4.36) \quad \left| \int_{S_1^{d-1}} f(\mathbf{v}, \rho^{d-1} \tau_1(\mathbf{q}' + \rho\beta(\mathbf{v}), \mathbf{v}; \rho), \mathbf{v}_1(\mathbf{q}' + \rho\beta(\mathbf{v}), \mathbf{v}; \rho)) d\lambda(\mathbf{v}) - \int_{S_1^{d-1}} \int_{\mathbb{R}_{>0}} \int_{S_1^{d-1}} f(\mathbf{v}, \xi, \mathbf{v}_1) p_{\alpha', \beta}(\mathbf{v}, \xi, \mathbf{v}_1) d\lambda(\mathbf{v}) d\xi d\text{vol}_{S_1^{d-1}}(\mathbf{v}_1) \right| < \varepsilon.$$

This is possible by Theorem 2.3. Let us shrink  $\rho_0$  further if necessary, so that (with  $K_\eta$  as in Lemma 3.1)

$$(4.37) \quad \text{vol}_{S_1^{d-1}}(\partial_{C\rho_0}(D_j)) < K_\eta^{-1} \varepsilon \text{vol}_{S_1^{d-1}}(D_j), \quad \forall j \in \{1, \dots, N\};$$

$$(4.38) \quad \forall \mathbf{v}, \mathbf{v}' \in S_1^{d-1} : \quad \varphi(\mathbf{v}, \mathbf{v}') \leq 2C^2 \rho_0 \implies |g(\mathbf{v}) - g(\mathbf{v}')| < \varepsilon;$$

and also, for all  $m \in \{0, \dots, n-1\}$  and all points  $(\mathbf{v}_0, \xi_1, \dots, \mathbf{v}_{n-m})$  and  $(\mathbf{v}'_0, \xi'_1, \dots, \mathbf{v}'_{n-m})$  in  $S_1^{d-1} \times (\mathbb{R}_{>0} \times S_1^{d-1})^{n-m}$ :

$$(4.39) \quad \begin{aligned} \varphi(\mathbf{v}'_j, \mathbf{v}_j) &\leq C^2 \rho_0 \quad (j = 0, \dots, n-m), & |\xi'_j - \xi_j| &\leq 2C\rho_0^d \quad (j = 1, \dots, n-m) \\ &\implies \left| f_m(\mathbf{v}'_0, \xi'_1, \dots, \mathbf{v}'_{n-m}) - f_m(\mathbf{v}_0, \xi_1, \dots, \mathbf{v}_{n-m}) \right| < \varepsilon. \end{aligned}$$

Here (4.39) can be achieved since each  $f_m$  is continuous with compact support.

Having thus fixed a choice of  $\rho_0$ , we now claim that for all  $\rho \in (0, \rho_0)$  we have

$$(4.40) \quad \left| \int_{S_1^{d-1}} f_0(\mathbf{v}_0, \rho^{d-1} \tau_1(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho), \dots, \mathbf{v}_n(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda_0(\mathbf{v}_0) - \int_X f_0 d\nu \right| \ll \varepsilon.$$

Here and in any  $\ll$  bound in the remainder of the proof, the implied constant depends only on  $f_0$ ,  $\lambda_0$ ,  $d$ ,  $n$ ,  $C$  and  $\Theta$  (the scattering map). Since  $\varepsilon > 0$  was arbitrary, the bound (4.40) will complete the proof of Theorem 4.1.

4.3.3. *Bounding  $\lambda_0(S_1^{d-1} \setminus \mathcal{U}^{(n)})$ .* Take an arbitrary  $\rho \in (0, \rho_0)$ , and keep  $\rho$  fixed for the rest of this proof. Note that (4.6) holds. We now define the subsets  $S_1^{d-1} = \mathcal{U}^{(0)} \supset \mathcal{U}^{(1)} \supset \mathcal{U}^{(2)} \supset \dots \supset \mathcal{U}^{(n)}$  as in Section 4.2. We will prove that  $\lambda_0(S_1^{d-1} \setminus \mathcal{U}^{(n)})$  is small.

Let us first make explicit the conclusion from (4.36) in the case  $f = H$ . In this case, by changing variables (via (2.25), (2.26)) in the triple integral in (4.36), and using the definition of  $H$ , we see that the triple integral is less than the expression in (4.30), and hence  $< \varepsilon$ . Note also that if  $\mathbf{v} \in S_1^{d-1}$  is such that  $\mathbf{w}_1 = \mathbf{w}_1(\mathbf{q}' + \rho\beta(\mathbf{v}), \mathbf{v}; \rho)$  satisfies  $-\mathbf{w}_1 K(\mathbf{v}) \in \mathfrak{U}_{\omega(\eta)}$ , then  $\mathbf{v}_1 = \mathbf{v}_1(\mathbf{q}' + \rho\beta(\mathbf{v}), \mathbf{v}; \rho) = \Theta_1(\mathbf{v}, \mathbf{w}_1)$  satisfies  $H(\mathbf{v}, \xi, \mathbf{v}_1) = 1$  for all  $\xi > 0$ . Hence (4.36) implies that for all  $\mathbf{q}' \in \{\mathbf{0}, \mathbf{q}\}$ ,  $\rho \in (0, 2\rho_0)$ ,  $\lambda \in F_1$  and  $\beta \in F_3^{(\mathbf{q}' )}$ , we have

$$(4.41) \quad \lambda(\{\mathbf{v} \in S_1^{d-1} : -\mathbf{w}_1(\mathbf{q}' + \rho\beta(\mathbf{v}), \mathbf{v}; \rho) K(\mathbf{v}) \in \mathfrak{U}_{\omega(\eta)}\}) < 2\varepsilon.$$

We now apply Lemma 4.3 to prove that  $\lambda_0(S_1^{d-1} \setminus \mathcal{U}^{(1)})$  is small. The set  $\{\mathbf{v} \in S_1^{d-1} : \tau_1 = \infty\}$  has measure zero with respect to  $\text{vol}_{S_1^{d-1}}$  (see Sec. 2.1), and hence also with respect to  $\lambda_0$ .

Note  $\tilde{\rho} < 2\rho_0$  and  $\beta_0, \tilde{\beta}_0 \in F_3^{(\mathbf{q})}$ , cf. (4.24) and (4.35); hence we may apply (4.41) with  $\mathbf{q}' = \mathbf{q}$ ,  $\lambda = \lambda_0$  and  $\beta_0, \rho$  resp.  $\tilde{\beta}_0, \tilde{\rho}$ . This implies that the set of  $\mathbf{v}$ 's which satisfy (ii) or (iii) in Lemma 4.3 has  $\lambda_0$ -measure  $< 4\varepsilon$ . Thus

$$(4.42) \quad \lambda_0(S_1^{d-1} \setminus \mathcal{U}^{(1)}) < 4\varepsilon.$$

Next take  $k \in \{1, \dots, n-1\}$ ; we will apply Lemma 4.2 to prove that  $\lambda_0(\mathcal{U}^{(k)} \setminus \mathcal{U}^{(k+1)})$  is small. We call any two vectors  $\mathbf{v}, \mathbf{v}' \in \mathcal{U}^{(k)}$  equivalent if and only if  $\tilde{\mathbf{m}}_{\mathbf{v}}^k = \tilde{\mathbf{m}}_{\mathbf{v}'}^k$ . Let  $\mathcal{U}' \subset \mathcal{U}^{(k)}$  be any fixed equivalence class for this relation. Then the subset  $\emptyset \neq M_{\mathbf{v}}^k \subset \{1, \dots, N\}$  and

the functions  $\Phi_{\rho, \tilde{\mathbf{m}}_{\mathbf{v}}, \beta_0}$ , and  $\bar{\beta}_{\rho, \tilde{\mathbf{m}}_{\mathbf{v}}, \beta_0}^+$  are independent of  $\mathbf{v} \in \mathcal{U}'$ , and we will write simply  $M$ ,  $\Phi$  and  $\bar{\beta}^+$  for these. Note in particular  $\mathcal{U}' = \Phi^{-1}(\cup_{i \in M} D_i)$ .

We need to modify  $\bar{\beta}^+$  to get a function in  $F_3^{(0)}$ . Let us fix  $C^1$  functions  $c_1, c_2 : \mathbb{R}_{>0} \rightarrow [0, 1]$  with the properties

(4.43)

$$\begin{aligned} c_1(\varphi) &= 0, \quad \forall \varphi \leq B_\Theta + 5\eta; & c_1(\varphi) &= 1, \quad \forall \varphi \geq B_\Theta + 6\eta; & |c_1'(\varphi)| &\leq 2\eta^{-1}, \quad \forall \varphi > 0; \\ c_2(\varphi) &= 1, \quad \forall \varphi \leq B_\Theta + 6\eta; & c_2(\varphi) &= 0, \quad \forall \varphi \geq B_\Theta + 7\eta; & |c_2'(\varphi)| &\leq 2\eta^{-1}, \quad \forall \varphi > 0. \end{aligned}$$

Now define  $\hat{\beta} = \hat{\beta}_{\rho, \tilde{\mathbf{m}}_{\mathbf{v}}, \beta_0} \in C^1(S_1^{d-1})$  by

$$(4.44) \quad \hat{\beta}(\mathbf{v}) = \hat{\beta}_{\rho, \tilde{\mathbf{m}}_{\mathbf{v}}, \beta_0}(\mathbf{v}) = \begin{cases} c_1(\varphi(\mathbf{v}, \mathbf{s}_k)) \cdot \bar{\beta}^+(\mathbf{v}) + c_2(\varphi(\mathbf{v}, \mathbf{s}_k)) \cdot \mathbf{v} & \text{if } \mathbf{v} \in \mathcal{V}_{\mathbf{s}_k}^\eta \\ \mathbf{v} & \text{otherwise.} \end{cases}$$

We then have

$$(4.45) \quad \hat{\beta}_{\rho, \tilde{\mathbf{m}}_{\mathbf{v}}, \beta_0}(\mathbf{v}) = \bar{\beta}_{\rho, \tilde{\mathbf{m}}_{\mathbf{v}}, \beta_0}^+(\mathbf{v}), \quad \forall \mathbf{v} \in \mathcal{V}_{\mathbf{s}_k}^{10\eta}.$$

Furthermore, for  $\mathbf{v} \in S_1^{d-1}$  with  $B_\Theta + 4\eta < \varphi(\mathbf{v}, \mathbf{s}_k) < B_\Theta + 8\eta$  we have, for any  $\mathbf{h} \in T_{\mathbf{v}}^1(S_1^{d-1})$ ,

$$(4.46) \quad \|D_{\mathbf{h}} \hat{\beta}(\mathbf{v})\| = \|(D_{\mathbf{h}} c_1(\varphi(\mathbf{v}, \mathbf{s}_k))) \cdot \bar{\beta}^+(\mathbf{v}) + c_1 \cdot D_{\mathbf{h}} \bar{\beta}^+(\mathbf{v}) + (D_{\mathbf{h}} c_2(\varphi(\mathbf{v}, \mathbf{s}_k))) \cdot \mathbf{v} + c_2 \cdot \mathbf{h}\| \leq \frac{4}{\eta} + C_\eta + 1$$

by Proposition 3.3 and (4.43). From this one easily deduces  $\hat{\beta}_{\rho, \tilde{\mathbf{m}}_{\mathbf{v}}, \beta_0} \in F_3^{(0)}$ .

Using Lemma 4.2,  $\cup_{i \in M} D_i \subset \mathcal{V}_{\mathbf{s}_k}^{10\eta}$  and (4.45) we have

$$(4.47) \quad \begin{aligned} \Phi(\mathcal{U}' \setminus \mathcal{U}^{(k+1)}) &\subset \partial_{C\rho}(\cup_{i \in M} D_i) \cup \{ \mathbf{v}' \in \cup_{i \in M} D_i : \tau_1(\rho \hat{\beta}(\mathbf{v}'), \mathbf{v}', \rho) = \infty \} \\ &\cup \{ \mathbf{v}' \in \cup_{i \in M} D_i : -\mathbf{w}_1(\rho \hat{\beta}(\mathbf{v}'), \mathbf{v}', \rho) K(\mathbf{v}') \in \mathfrak{U}_{\omega(\eta)} \} \\ &\cup \{ \mathbf{v}' \in \cup_{i \in M} D_i : -\mathbf{w}_1(\tilde{\rho} \tilde{\beta}(\mathbf{v}'), \mathbf{v}', \tilde{\rho}) K(\mathbf{v}') \in \mathfrak{U}_{\omega(\eta)} \}, \end{aligned}$$

where  $\tilde{\rho} = (1 + \eta)\rho$  and  $\tilde{\beta}(\mathbf{v}) = (1 + \eta)^{-1}(\hat{\beta}(\mathbf{v}) + 3\mathbf{v})$ . Here note that, by Lemma 3.1 and (4.37),

$$(4.48) \quad \begin{aligned} \lambda_{M, \mathbf{s}_k}(\partial_{C\rho}(\cup_{i \in M} D_i)) &\leq K_\eta \text{vol}_{S_1^{d-1}}(\cup_{i \in M} D_i)^{-1} \text{vol}_{S_1^{d-1}}(\partial_{C\rho}(\cup_{i \in M} D_i)) \\ &\leq K_\eta \text{vol}_{S_1^{d-1}}(\cup_{i \in M} D_i)^{-1} \sum_{i \in M} \text{vol}_{S_1^{d-1}}(\partial_{C\rho}(D_i)) \leq \varepsilon. \end{aligned}$$

The second set in the right hand side of (4.47) has  $\lambda_{M, \mathbf{s}_k}$ -measure zero. Next recall that (4.46) led to  $\hat{\beta} \in F_3^{(0)}$ ; by a similar argument we also verify  $\tilde{\beta} \in F_3^{(0)}$ . Hence from (4.41) applied with  $\mathbf{q}' = \mathbf{0}$ ,  $\lambda = \lambda_{M, \mathbf{s}_k}$  and  $\hat{\beta}$ ,  $\rho$ , resp.  $\tilde{\beta}$ ,  $\tilde{\rho}$ , we see that each of the last two sets in (4.47) have  $\lambda_{M, \mathbf{s}_k}$ -measure  $< 2\varepsilon$ . Hence

$$(4.49) \quad \lambda_{M, \mathbf{s}_k}(\Phi(\mathcal{U}' \setminus \mathcal{U}^{(k+1)})) < 5\varepsilon.$$

Set  $\mu = \text{vol}_{S_1^{d-1}}(\mathcal{U}')^{-1} \cdot (\text{vol}_{S_1^{d-1}})_{|\mathcal{U}'}$ . By Proposition 3.3, and our assumption  $\rho < \rho_0(n, \eta, C, \eta)$ ,  $\Phi|_{\mathcal{U}'}$  is a diffeomorphism from  $\mathcal{U}'$  onto  $\cup_M D_i$ , which transforms the measure  $\mu$  into  $c\tilde{g} \cdot \lambda_{M, \mathbf{s}_k}$  where  $c > 0$  is a constant and  $\tilde{g}$  is a continuous function from  $\cup_M D_i$  to  $[1 - \eta, 1 + \eta]$ . Using  $c \int_{\cup_M D_i} \tilde{g}(\mathbf{v}) d\lambda_{M, \mathbf{s}_k}(\mathbf{v}) = 1$  we find  $(1 + \eta)^{-1} \leq c \leq (1 - \eta)^{-1}$ , and thus  $c\tilde{g}(\mathbf{v}) \in [1 - 3\eta, 1 + 3\eta]$  for all  $\mathbf{v} \in \cup_M D_i$ . Hence

$$(4.50) \quad \mu(\mathcal{U}' \setminus \mathcal{U}^{(k+1)}) < (1 + 3\eta)5\varepsilon < 10\varepsilon.$$

In other words  $\text{vol}_{S_1^{d-1}}(\mathcal{U}' \setminus \mathcal{U}^{(k+1)}) < 10\varepsilon \text{vol}_{S_1^{d-1}}(\mathcal{U}')$ . Adding this over all equivalence classes  $\mathcal{U}' \subset \mathcal{U}^{(k)}$  we obtain  $\text{vol}_{S_1^{d-1}}(\mathcal{U}^{(k)} \setminus \mathcal{U}^{(k+1)}) < 10\varepsilon \text{vol}_{S_1^{d-1}}(\mathcal{U}^{(k)}) \ll \varepsilon$ . Hence since  $\lambda_0 = g \cdot \text{vol}_{S_1^{d-1}}$

with  $g$  bounded (cf. (4.29)), we have

$$(4.51) \quad \lambda_0(\mathcal{U}^{(k)} \setminus \mathcal{U}^{(k+1)}) \ll \varepsilon.$$

Adding this over  $k = 1, \dots, n-1$  and combining with (4.42) we finally obtain

$$(4.52) \quad \lambda_0(S_1^{d-1} \setminus \mathcal{U}^{(n)}) \ll \varepsilon.$$

4.3.4. *Conclusion: Proof of (4.40).* We intend to show that for each  $m \in \{0, \dots, n-2\}$  we have

$$(4.53) \quad \left| \int_{S_1^{d-1}} f_m(\mathbf{v}_0, \rho^{d-1} \tau_1(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho), \dots, \mathbf{v}_{n-m}(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda_0(\mathbf{v}_0) - \int_{S_1^{d-1}} f_{m+1}(\mathbf{v}_0, \rho^{d-1} \tau_1(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho), \dots, \mathbf{v}_{n-m-1}(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda_0(\mathbf{v}_0) \right| \ll \varepsilon.$$

This will imply (4.40) (and thus complete our proof of Theorem 4.1), for note that (4.36) applied with  $\mathbf{q}' = \mathbf{q}$ ,  $\lambda = \lambda_0$ ,  $f = f_{n-1}$  and  $\beta = \beta_0$  gives, in view of (4.28):

$$(4.54) \quad \left| \int_{S_1^{d-1}} f_{n-1}(\mathbf{v}_0, \rho^{d-1} \tau_1(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho), \mathbf{v}_1(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda_0(\mathbf{v}_0) - \int_X f_0 d\nu \right| < \varepsilon.$$

Combining (4.53) for  $m = 0, 1, \dots, n-2$  and (4.54) we indeed obtain (4.40), as desired.

Now to prove (4.53) we fix  $m \in \{0, \dots, n-2\}$  and set  $k = n-m-1 \in \{1, \dots, n-1\}$ . In the following we will use the shorthand notation  $\mathbf{v}_\ell = \mathbf{v}_\ell(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho)$ ,  $\tau_\ell = \tau_\ell(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho)$ ,  $\mathbf{m}_\ell = \mathbf{m}_\ell(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho)$ ,  $\mathbf{s}_\ell = \mathbf{m}_\ell - \mathbf{m}_{\ell-1}$ , and  $\widehat{\mathbf{s}}_\ell = \|\mathbf{s}_\ell\|^{-1} \mathbf{s}_\ell$ . As before we call two vectors  $\mathbf{v}, \mathbf{v}' \in \mathcal{U}^{(k)}$  equivalent if and only if  $\vec{\mathbf{m}}_{\mathbf{v}}^k = \vec{\mathbf{m}}_{\mathbf{v}'}^k$ . Let  $\mathcal{U}' \subset \mathcal{U}^{(k)}$  be any fixed equivalence class for this relation; thus by construction  $\mathbf{m}_0, \mathbf{m}_1, \dots, \mathbf{m}_k$  are *constant* as  $\mathbf{v}_0$  varies through  $\mathcal{U}'$ . We write  $M \subset \{1, \dots, N\}$  and  $\Phi, \overline{\beta}^+, \widehat{\beta}$  for the index set and functions corresponding to our fixed class  $\mathcal{U}'$ , as in the discussion just above (4.47). For each  $\mathbf{v}_0 \in \mathcal{U}'$  we have  $|\tau_\ell - \|\mathbf{s}_\ell\|| \leq 2\rho$  and  $\varphi(\mathbf{v}_{\ell-1}, \mathbf{s}_\ell) < \frac{1}{2}C\rho$  for all  $\ell = 2, \dots, k$  (cf. (4.11)), and similarly using  $\sup \|\beta_0\| \leq C$  we get  $|\tau_1 - \|\mathbf{s}_1\|| \leq (C+1)\rho$  and  $\varphi(\mathbf{v}_0, \mathbf{s}_1) < \frac{1}{2}C(C+1)\rho$ . Hence by (4.39) we have the following approximation result for the  $\mathcal{U}'$ -contribution to the left integral in (4.53):

$$(4.55) \quad \left| \int_{\mathcal{U}'} f_m(\mathbf{v}_0, \rho^{d-1} \tau_1, \mathbf{v}_1, \dots, \mathbf{v}_k, \rho^{d-1} \tau_{k+1}, \mathbf{v}_{k+1}) d\lambda_0(\mathbf{v}_0) - \int_{\mathcal{U}'} f_m(\widehat{\mathbf{s}}_1, \rho^{d-1} \|\mathbf{s}_1\|, \dots, \widehat{\mathbf{s}}_k, \rho^{d-1} \|\mathbf{s}_k\|, \mathbf{v}_k, \rho^{d-1} \tau_{k+1}, \mathbf{v}_{k+1}) d\lambda_0(\mathbf{v}_0) \right| < \varepsilon \lambda_0(\mathcal{U}').$$

Recall  $\lambda_0 = g \cdot \text{vol}_{S_1^{d-1}}$ , and set  $\mu = \text{vol}_{S_1^{d-1}}(\mathcal{U}')^{-1} \cdot (\text{vol}_{S_1^{d-1}})_{|\mathcal{U}'}$ . From above we have  $\varphi(\mathbf{v}_0, \mathbf{s}_1) < C^2\rho$  for all  $\mathbf{v}_0 \in \mathcal{U}'$ ; by our assumption (4.38) this implies

$$(4.56) \quad \left| g(\mathbf{v}_0) - \lambda_0(\mathcal{U}') \text{vol}_{S_1^{d-1}}(\mathcal{U}')^{-1} \right| = \left| g(\mathbf{v}_0) - \int_{\mathcal{U}'} g(\mathbf{v}) d\mu(\mathbf{v}) \right| < \varepsilon, \quad \forall \mathbf{v}_0 \in \mathcal{U}'.$$

Hence replacing “ $d\lambda_0(\mathbf{v}_0)$ ” with “ $\lambda_0(\mathcal{U}') d\mu(\mathbf{v}_0)$ ” in the last integral in (4.55) causes an error  $\leq \varepsilon(\sup |f_m|) \text{vol}_{S_1^{d-1}}(\mathcal{U}')$ . Furthermore, by Proposition 3.3,  $\mathbf{v}_0 \mapsto \mathbf{v}_k = \mathbf{v}_k(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho) = \Phi(\mathbf{v}_0)$  is a diffeomorphism from  $\mathcal{U}'$  onto  $\cup_M D_i$  which, as we saw in the discussion preceding (4.50), transforms the measure  $\mu$  into  $c\tilde{g} \cdot \lambda_{M, \mathbf{s}_k}$  where  $c\tilde{g}(\mathbf{v}) \in [1-3\eta, 1+3\eta]$  for all  $\mathbf{v} \in \cup_M D_i$ . We also have  $\tau_{k+1} = \tau_1(\mathbf{m}_k + \rho\widehat{\beta}(\mathbf{v}_k), \mathbf{v}_k; \rho) = \tau_1(\rho\widehat{\beta}(\mathbf{v}_k), \mathbf{v}_k; \rho)$  for all  $\mathbf{v}_0 \in \mathcal{U}'$ , and similarly

$\mathbf{v}_{k+1} = \mathbf{v}_1(\rho\widehat{\beta}(\mathbf{v}_k), \mathbf{v}_k; \rho)$ . Hence, using also  $\eta < \varepsilon$ ,

$$(4.57) \quad \left| \int_{\mathcal{U}'} f_m(\widehat{\mathbf{s}}_1, \rho^{d-1}\|\mathbf{s}_1\|, \dots, \widehat{\mathbf{s}}_k, \rho^{d-1}\|\mathbf{s}_k\|, \mathbf{v}_k, \rho^{d-1}\tau_{k+1}, \mathbf{v}_{k+1}) d\lambda_0(\mathbf{v}_0) \right. \\ \left. - \lambda_0(\mathcal{U}') \int_{\cup_M D_i} f_m(\widehat{\mathbf{s}}_1, \dots, \rho^{d-1}\|\mathbf{s}_k\|, \mathbf{v}, \rho^{d-1}\tau_1(\rho\widehat{\beta}(\mathbf{v}), \mathbf{v}; \rho), \mathbf{v}_1(\rho\widehat{\beta}(\mathbf{v}), \mathbf{v}; \rho)) d\lambda_{M, \mathbf{s}_k}(\mathbf{v}) \right| \\ \leq \varepsilon(\sup |f_m|) \text{vol}_{\mathbb{S}_1^{d-1}}(\mathcal{U}') + 3\eta(\sup |f_m|)\lambda_0(\mathcal{U}') \ll \varepsilon(\text{vol}_{\mathbb{S}_1^{d-1}}(\mathcal{U}') + \lambda_0(\mathcal{U}')).$$

Now  $\lambda_{M, \mathbf{s}_k} \in F_1$ ,  $f_m(\widehat{\mathbf{s}}_1, \dots, \rho^{d-1}\|\mathbf{s}_k\|, \cdot, \cdot, \cdot) \in F_2$  and  $\widehat{\beta} \in F_3^{(0)}$ , so that by (4.36),

$$(4.58) \quad \left| \lambda_0(\mathcal{U}') \int_{\cup_M D_i} f_m(\widehat{\mathbf{s}}_1, \dots, \rho^{d-1}\|\mathbf{s}_k\|, \mathbf{v}, \rho^{d-1}\tau_1(\rho\widehat{\beta}(\mathbf{v}), \mathbf{v}; \rho), \mathbf{v}_1(\rho\widehat{\beta}(\mathbf{v}), \mathbf{v}; \rho)) d\lambda_{M, \mathbf{s}_k}(\mathbf{v}) \right. \\ \left. - \lambda_0(\mathcal{U}') \int_{\cup_M D_i} \int_{\mathbb{R}_{>0}} \int_{\mathbb{S}_1^{d-1}} f_m(\widehat{\mathbf{s}}_1, \rho^{d-1}\|\mathbf{s}_1\|, \widehat{\mathbf{s}}_2, \dots, \widehat{\mathbf{s}}_k, \rho^{d-1}\|\mathbf{s}_k\|, \mathbf{v}, \xi, \mathbf{v}') \right. \\ \left. \times p_{\mathbf{0}, \widehat{\beta}}(\mathbf{v}, \xi, \mathbf{v}') d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}') d\xi d\lambda_{M, \mathbf{s}_k}(\mathbf{v}) \right| < \varepsilon\lambda_0(\mathcal{U}').$$

(We have  $\lambda_{M, \mathbf{s}_k}(A) = 0$  for any  $A \subset \mathbb{S}_1^{d-1}$  disjoint from  $\cup_M D_i$ , so that we can indeed take the domain of integration for  $\mathbf{v}$  in the triple integral to be  $\cup_M D_i$  instead of  $\mathbb{S}_1^{d-1}$ .) Next we wish to replace  $\widehat{\beta}$  with  $\beta_{\widehat{\mathbf{s}}_k}^+$  in the triple integral in (4.58), so that it can be rewritten in terms of  $f_{m+1}$  using (4.27). Using (2.25) and (4.1), we see that the error caused by this replacement is bounded above by  $\lambda_0(\mathcal{U}') \sup |f_m|$  times the integral in (4.31), with  $\lambda = \lambda_{M, \mathbf{s}_k}$ ,  $\mathbf{z}_1(\mathbf{v}) = (\widehat{\beta}(\mathbf{v})K(\mathbf{v}))_{\perp}$  and  $\mathbf{z}_2(\mathbf{v}) = (\beta_{\widehat{\mathbf{s}}_k}^+(\mathbf{v})K(\mathbf{v}))_{\perp}$ . Note that for all  $\mathbf{v} \in \cup_M D_i \subset \mathcal{V}_{\mathbf{s}_k}^{10\eta}$  we have, by (4.45) and Proposition 3.3 and our assumption  $\rho < \rho_0(n, \eta, C, \eta)$ :

$$(4.59) \quad \|\mathbf{z}_1(\mathbf{v}) - \mathbf{z}_2(\mathbf{v})\| \leq \|\widehat{\beta}(\mathbf{v}) - \beta_{\widehat{\mathbf{s}}_k}^+(\mathbf{v})\| = \|\overline{\beta}^+(\mathbf{v}) - \beta_{\widehat{\mathbf{s}}_k}^+(\mathbf{v})\| < \eta.$$

Hence the inequality in (4.31) is valid for our choices of  $\lambda, \mathbf{z}_1, \mathbf{z}_2$ , and we obtain:

$$(4.60) \quad \left| \lambda_0(\mathcal{U}') \int_{\cup_M D_i} \int_{\mathbb{R}_{>0}} \int_{\mathbb{S}_1^{d-1}} f_m(\widehat{\mathbf{s}}_1, \rho^{d-1}\|\mathbf{s}_1\|, \widehat{\mathbf{s}}_2, \dots, \widehat{\mathbf{s}}_k, \rho^{d-1}\|\mathbf{s}_k\|, \mathbf{v}, \xi, \mathbf{v}') \right. \\ \left. \times p_{\mathbf{0}, \widehat{\beta}}(\mathbf{v}, \xi, \mathbf{v}') d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}') d\xi d\lambda_{M, \mathbf{s}_k}(\mathbf{v}) \right. \\ \left. - \lambda_0(\mathcal{U}') \int_{\cup_M D_i} f_{m+1}(\widehat{\mathbf{s}}_1, \rho^{d-1}\|\mathbf{s}_1\|, \widehat{\mathbf{s}}_2, \dots, \widehat{\mathbf{s}}_k, \rho^{d-1}\|\mathbf{s}_k\|, \mathbf{v}) d\lambda_{M, \mathbf{s}_k}(\mathbf{v}) \right| \ll \varepsilon\lambda_0(\mathcal{U}').$$

Next, by imitating the argument which led to (4.57), and then the argument which led to (4.55), we obtain

$$(4.61) \quad \left| \lambda_0(\mathcal{U}') \int_{\cup_M D_i} f_{m+1}(\widehat{\mathbf{s}}_1, \rho^{d-1}\|\mathbf{s}_1\|, \widehat{\mathbf{s}}_2, \dots, \widehat{\mathbf{s}}_k, \rho^{d-1}\|\mathbf{s}_k\|, \mathbf{v}) d\lambda_{M, \mathbf{s}_k}(\mathbf{v}) \right. \\ \left. - \int_{\mathcal{U}'} f_{m+1}(\mathbf{v}_0, \rho^{d-1}\tau_1(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho), \dots, \mathbf{v}_k(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda_0(\mathbf{v}_0) \right| \\ \ll \varepsilon(\text{vol}_{\mathbb{S}_1^{d-1}}(\mathcal{U}') + \lambda_0(\mathcal{U}')).$$

Combining (4.55), (4.57), (4.58), (4.60) and (4.61), and adding over all the equivalence classes  $\mathcal{U}' \subset \mathcal{U}^{(k)}$ , using  $\text{vol}_{\mathbb{S}_1^{d-1}}(\mathcal{U}^{(k)}) \leq \text{vol}(\mathbb{S}_1^{d-1}) \ll 1$  and  $\lambda_0(\mathcal{U}^{(k)}) \leq 1$ , we get

$$(4.62) \quad \left| \int_{\mathcal{U}^{(k)}} f_m(\mathbf{v}_0, \rho^{d-1}\tau_1(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho), \dots, \mathbf{v}_{k+1}(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda_0(\mathbf{v}_0) \right. \\ \left. - \int_{\mathcal{U}^{(k)}} f_{m+1}(\mathbf{v}_0, \rho^{d-1}\tau_1(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho), \dots, \mathbf{v}_k(\mathbf{q}_{\rho, \beta_0}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda_0(\mathbf{v}_0) \right| \ll \varepsilon.$$

This implies (4.53), since  $\lambda_0(\mathbb{S}_1^{d-1} \setminus \mathcal{U}^{(k)}) \ll \varepsilon$  (from (4.52)).

Since (4.40) follows from (4.53) and (4.54), the proof of Theorem 4.1 is now complete.  $\square$



**4.4. Proof of Theorem 1.1.** The generalization of Theorem 1.1 to the case of a general scattering map and a more general initial condition  $(\mathbf{q}_0 + \rho\boldsymbol{\beta}(\mathbf{v}_0), \mathbf{v}_0)$  is as follows. We set

$$(4.63) \quad \mathcal{B}_n := \{(\mathbf{S}_1, \dots, \mathbf{S}_n) \in (\mathbb{R}^d \setminus \{\mathbf{0}\})^n : \varphi(\mathbf{S}_j, \mathbf{S}_{j+1}) > B_\Theta \ (j = 1, \dots, n-1)\}.$$

**Theorem 4.4.** *Fix a lattice  $\mathcal{L} = \mathbb{Z}^d M_0$  and a point  $\mathbf{q}_0 \in \mathbb{R}^d$ , set  $\boldsymbol{\alpha} = -\mathbf{q}_0 M_0^{-1}$ , and let  $\boldsymbol{\beta} : \mathbb{S}_1^{d-1} \rightarrow \mathbb{R}^d$  be a  $C^1$  function. If  $\mathbf{q}_0 \in \mathcal{L}$  we assume that  $(\boldsymbol{\beta}(\mathbf{v}) + \mathbb{R}_{>0}\mathbf{v}) \cap \mathcal{B}_1^d = \emptyset$  for all  $\mathbf{v} \in \mathbb{S}_1^{d-1}$ . Then for each  $n \in \mathbb{Z}_{>0}$  there exists a (Borel measurable) function  $P_{\boldsymbol{\alpha}, \boldsymbol{\beta}}^{(n)} : \mathcal{B}_n \rightarrow \mathbb{R}_{\geq 0}$  such that, for any Borel probability measure  $\lambda$  on  $\mathbb{S}_1^{d-1}$  which is absolutely continuous with respect to  $\text{vol}_{\mathbb{S}_1^{d-1}}$ , and for any set  $\mathcal{A} \subset \mathbb{R}^{nd}$  with boundary of Lebesgue measure zero,*

$$(4.64) \quad \begin{aligned} & \lim_{\rho \rightarrow 0} \lambda(\{\mathbf{v}_0 \in \mathbb{S}_1^{d-1} : (s_1(\mathbf{q}_0 + \rho\boldsymbol{\beta}(\mathbf{v}_0), \mathbf{v}_0; \rho), \dots, s_n(\mathbf{q}_0 + \rho\boldsymbol{\beta}(\mathbf{v}_0), \mathbf{v}_0; \rho)) \in \rho^{-(d-1)}\mathcal{A}\}) \\ &= \int_{\mathcal{B}_n \cap \mathcal{A}} P_{\boldsymbol{\alpha}, \boldsymbol{\beta}}^{(n)}(\mathbf{S}_1, \dots, \mathbf{S}_n) \lambda'(\widehat{\mathbf{S}}_1) d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_1) \cdots d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_n), \end{aligned}$$

where  $\lambda' \in L^1(\mathbb{S}_1^{d-1})$  is the Radon-Nikodym derivative of  $\lambda$  with respect to  $\text{vol}_{\mathbb{S}_1^{d-1}}$ . Furthermore, there is a function  $\Psi : \mathcal{B}_3 \rightarrow \mathbb{R}_{\geq 0}$  which only depends on the scattering map, such that

$$(4.65) \quad P_{\boldsymbol{\alpha}, \boldsymbol{\beta}}^{(n)}(\mathbf{S}_1, \dots, \mathbf{S}_n) = P_{\boldsymbol{\alpha}, \boldsymbol{\beta}}^{(2)}(\mathbf{S}_1, \mathbf{S}_2) \prod_{j=3}^n \Psi(\mathbf{S}_{j-2}, \mathbf{S}_{j-1}, \mathbf{S}_j)$$

for all  $n \geq 3$  and all  $(\mathbf{S}_1, \dots, \mathbf{S}_n) \in \mathcal{B}_n$ .

Explicit formulas for  $P_{\boldsymbol{\alpha}, \boldsymbol{\beta}}^{(2)}$  and  $\Psi$  are given in the proof below. As in (1.12) we define the probability measure corresponding to (4.64) by

$$(4.66) \quad \mu_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}^{(n)}(\mathcal{A}) := \int_{\mathcal{B}_n \cap \mathcal{A}} P_{\boldsymbol{\alpha}, \boldsymbol{\beta}}^{(n)}(\mathbf{S}_1, \dots, \mathbf{S}_n) \lambda'(\widehat{\mathbf{S}}_1) d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_1) \cdots d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_n).$$

*Proof of Theorem 4.4.* By a standard approximation argument, using the absolute continuity of the limit measure and the assumption that  $\mathcal{A}$  has boundary of Lebesgue measure zero, one finds that it suffices to prove the corresponding statement for bounded continuous functions, i.e. to prove that for each bounded continuous function  $g : \mathbb{R}^{nd} \rightarrow \mathbb{R}_{\geq 0}$  we have

$$(4.67) \quad \begin{aligned} & \lim_{\rho \rightarrow 0} \int_{\mathbb{S}_1^{d-1}} g(\rho^{d-1}s_1(\mathbf{q}_0 + \rho\boldsymbol{\beta}(\mathbf{v}_0), \mathbf{v}_0; \rho), \dots, \rho^{d-1}s_n(\mathbf{q}_0 + \rho\boldsymbol{\beta}(\mathbf{v}_0), \mathbf{v}_0; \rho)) d\lambda(\mathbf{v}_0) \\ &= \int_{\mathcal{B}_n} g(\mathbf{S}_1, \dots, \mathbf{S}_n) P_{\boldsymbol{\alpha}, \boldsymbol{\beta}}^{(n)}(\mathbf{S}_1, \dots, \mathbf{S}_n) \lambda'(\widehat{\mathbf{S}}_1) d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_1) \cdots d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_n). \end{aligned}$$

This, however, is a direct consequence of Theorem 4.1, applied with  $f_0(\mathbf{v}_0, \xi_1, \mathbf{v}_1, \dots, \xi_n, \mathbf{v}_n) := g(\xi_1\mathbf{v}_0, \xi_2\mathbf{v}_1, \dots, \xi_n\mathbf{v}_{n-1})$ . Indeed, substituting  $\xi_j\mathbf{v}_{j-1} = \mathbf{S}_j$  in the right hand side of (4.4) in Theorem 4.1 gives exactly the right hand side of (4.67), with

$$(4.68) \quad \begin{aligned} P_{\boldsymbol{\alpha}, \boldsymbol{\beta}}^{(n)}(\mathbf{S}_1, \dots, \mathbf{S}_n) &= \left( \prod_{j=1}^n \|\mathbf{S}_j\|^{1-d} \right) p_{\boldsymbol{\alpha}, \boldsymbol{\beta}}(\widehat{\mathbf{S}}_1, \|\mathbf{S}_1\|, \widehat{\mathbf{S}}_2) \prod_{j=1}^{n-2} p_{\mathbf{0}, \boldsymbol{\beta}_{\widehat{\mathbf{S}}_j^+}}(\widehat{\mathbf{S}}_{j+1}, \|\mathbf{S}_{j+1}\|, \widehat{\mathbf{S}}_{j+2}) \\ &\quad \times \int_{\mathbb{S}_1^{d-1}} p_{\mathbf{0}, \boldsymbol{\beta}_{\widehat{\mathbf{S}}_{n-1}^+}}(\widehat{\mathbf{S}}_n, \|\mathbf{S}_n\|, \widehat{\mathbf{S}}) d\text{vol}_{\mathbb{S}_1^{d-1}}(\widehat{\mathbf{S}}) \end{aligned}$$

if  $n \geq 2$ , and

$$(4.69) \quad P_{\boldsymbol{\alpha}, \boldsymbol{\beta}}^{(1)}(\mathbf{S}_1) = \|\mathbf{S}_1\|^{1-d} \int_{\mathbb{S}_1^{d-1}} p_{\boldsymbol{\alpha}, \boldsymbol{\beta}}(\widehat{\mathbf{S}}_1, \|\mathbf{S}_1\|, \widehat{\mathbf{S}}) d\text{vol}_{\mathbb{S}_1^{d-1}}(\widehat{\mathbf{S}})$$

if  $n = 1$ .

Next we claim that (4.65) holds (for  $d \geq 3$ ) if we set

$$(4.70) \quad I(\mathbf{S}_1, \mathbf{S}_2) = \int_{S_1^{d-1}} p_{\mathbf{0}, \beta_{\widehat{\mathbf{S}}_1}^+}(\widehat{\mathbf{S}}_2, \|\mathbf{S}_2\|, \widehat{\mathbf{S}}) d\text{vol}_{S_1^{d-1}}(\widehat{\mathbf{S}})$$

and then define

$$(4.71) \quad \Psi(\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3) := \|\mathbf{S}_3\|^{1-d} p_{\mathbf{0}, \beta_{\widehat{\mathbf{S}}_1}^+}(\widehat{\mathbf{S}}_2, \|\mathbf{S}_2\|, \widehat{\mathbf{S}}_3) \frac{I(\mathbf{S}_2, \mathbf{S}_3)}{I(\mathbf{S}_1, \mathbf{S}_2)}$$

if  $I(\mathbf{S}_1, \mathbf{S}_2) \neq 0$ , while setting  $\Psi(\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3) = 0$  whenever  $I(\mathbf{S}_1, \mathbf{S}_2) = 0$ . To prove this claim it suffices to check

$$(4.72) \quad P_{\alpha, \beta}^{(n)}(\mathbf{S}_1, \dots, \mathbf{S}_n) = P_{\alpha, \beta}^{(n-1)}(\mathbf{S}_1, \dots, \mathbf{S}_{n-1}) \Psi(\mathbf{S}_{n-2}, \mathbf{S}_{n-1}, \mathbf{S}_n)$$

for all  $n \geq 3$  and all  $(\mathbf{S}_1, \dots, \mathbf{S}_n) \in \mathcal{B}_n$ .

If  $I(\mathbf{S}_{n-2}, \mathbf{S}_{n-1}) \neq 0$  then (4.72) is clear from (4.68) and (4.71). Now assume  $I(\mathbf{S}_{n-2}, \mathbf{S}_{n-1}) = 0$ . The function  $f : \widehat{\mathbf{S}} \mapsto p_{\mathbf{0}, \beta_{\widehat{\mathbf{S}}_1}^+}(\widehat{\mathbf{S}}_{n-1}, \|\mathbf{S}_{n-1}\|, \widehat{\mathbf{S}})$  is non-negative on all  $S_1^{d-1}$  and vanishes outside the open subset  $\mathcal{V}_{\widehat{\mathbf{S}}_{n-1}}$ . If  $d \geq 3$  then  $f$  is continuous on  $\mathcal{V}_{\widehat{\mathbf{S}}_{n-1}}$  (see Remark 4.1), so that  $I(\mathbf{S}_{n-2}, \mathbf{S}_{n-1}) = 0$  implies that  $f(\widehat{\mathbf{S}}) = 0$  holds for all  $\widehat{\mathbf{S}} \in \mathcal{V}_{\widehat{\mathbf{S}}_{n-1}}$ ; in particular we have  $P_{\alpha, \beta}^{(n)}(\mathbf{S}_1, \dots, \mathbf{S}_n) = 0$  as well as  $P_{\alpha, \beta}^{(n-1)}(\mathbf{S}_1, \dots, \mathbf{S}_{n-1}) = 0$ , and hence (4.72) holds.

To treat the remaining case,  $d = 2$ , it is simplest to first *modify* the function  $P_{\alpha, \beta}^{(n)}$  given by (4.68) by setting it to be zero at any point  $(\mathbf{S}_1, \dots, \mathbf{S}_n) \in \mathcal{B}_n$  such that  $\widehat{\mathbf{S}}_{j+2} = \Theta_1(\widehat{\mathbf{S}}_{j+1}, -\beta_{\widehat{\mathbf{S}}_j}^+(\widehat{\mathbf{S}}_{j+1})R_{\mathbf{S}_{j+1}})$  for some  $1 \leq j \leq n-2$ , where  $R_{\mathbf{S}} \in \text{O}(2)$  denotes reflection in the line  $\mathbb{R}\mathbf{S}$ . (This alteration only concerns a subset of  $\mathcal{B}_n$  of Lebesgue measure zero, and hence it does not affect the validity of (4.64).) Now the proof can be completed as before, using the fact that when  $d = 2$ , the function  $f$  considered in the last paragraph is continuous on  $\mathcal{V}_{\widehat{\mathbf{S}}_{n-1}}$  except possibly at the single point  $\widehat{\mathbf{S}} = \Theta_1(\widehat{\mathbf{S}}_{n-1}, -\beta_{\widehat{\mathbf{S}}_{n-2}}^+(\widehat{\mathbf{S}}_{n-1})R_{\mathbf{S}_{n-1}})$ , again by Remark 4.1.  $\square$

*Remark 4.5.* Note that  $\mu_{\alpha, \beta, \lambda}^{(n+1)}(\mathcal{A} \times \mathbb{R}^d) = \mu_{\alpha, \beta, \lambda}^{(n)}(\mathcal{A})$  for every measurable subset  $\mathcal{A} \subset \mathbb{R}^{nd}$ . This follows from (4.66), (4.68) and (4.69), using (4.3).

*Remark 4.6.* Recall that for  $\alpha \in \mathbb{R}^d \setminus \mathbb{Q}^d$  the function  $p_{\alpha, \beta}(\mathbf{v}_0, \xi_1, \mathbf{v}_1)$  is independent of both  $\alpha$  and  $\beta$ , cf. Remark 2.4. Hence by (4.68), (4.69), the same is true for the function  $P_{\alpha, \beta}^{(n)} =: P^{(n)}$ .

*Remark 4.7.* If  $d \geq 3$  then  $P^{(n)}$  is continuous on all of  $\mathcal{B}_n$ , and if also  $\sup \|\beta\| \leq 1$  then  $P_{\alpha, \beta}^{(n)}$  is continuous on  $\mathcal{B}_n$  for every  $\alpha \in \mathbb{R}^d$ . Next suppose  $d = 2$ , and let  $\mathcal{B}'_n$  be the following dense open subset of  $\mathcal{B}_n$ :

$$(4.73) \quad \mathcal{B}'_n := \{(\mathbf{S}_1, \dots, \mathbf{S}_n) \in \mathcal{B}_n : \beta_{e_1}^-(\widehat{\mathbf{S}}_2 K(\widehat{\mathbf{S}}_1))_{\perp} \neq (\beta(\widehat{\mathbf{S}}_1) K(\widehat{\mathbf{S}}_1))_{\perp}, \\ \widehat{\mathbf{S}}_{j+2} \neq \Theta_1(\widehat{\mathbf{S}}_{j+1}, -\beta_{\widehat{\mathbf{S}}_j}^+(\widehat{\mathbf{S}}_{j+1})R_{\mathbf{S}_{j+1}}) \ (j = 1, \dots, n-2)\}.$$

Then  $P^{(n)}$  is continuous on all of  $\mathcal{B}'_n$ , and if  $\sup \|\beta\| \leq 1$  then  $P_{\alpha, \beta}^{(n)}$  is continuous on  $\mathcal{B}'_n$  for every  $\alpha \in \mathbb{R}^d$ . These statements follow from (4.68), (4.69), Remark 2.4 and Remark 4.1.

**4.5. Proof of Theorem 1.3.** We now prove the ‘‘macroscopic’’ version of Theorem 1.1, i.e. Theorem 1.3, in the case of a general scattering map. The precise statement is as follows. Let  $\mathcal{B}_n$  be as in (4.63).

**Theorem 4.5.** *Let  $\Lambda$  be a Borel probability measure on  $\text{T}^1(\mathbb{R}^d)$  which is absolutely continuous with respect to Lebesgue measure. Then, for each  $n \in \mathbb{Z}_{>0}$ , and for any set  $\mathcal{A} \subset \mathbb{R}^d \times \mathbb{R}^{nd}$  with*

boundary of Lebesgue measure zero,

$$(4.74) \quad \lim_{\rho \rightarrow 0} \Lambda(\{(\mathbf{Q}_0, \mathbf{V}_0) \in \mathbb{T}^1(\rho^{d-1}\mathcal{K}_\rho) : (\mathbf{Q}_0, \mathbf{S}_1(\mathbf{Q}_0, \mathbf{V}_0; \rho), \dots, \mathbf{S}_n(\mathbf{Q}_0, \mathbf{V}_0; \rho)) \in \mathcal{A}\}) \\ = \int_{\mathcal{A} \cap (\mathbb{R}^d \times \mathcal{B}_n)} P^{(n)}(\mathbf{S}_1, \dots, \mathbf{S}_n) \Lambda'(\mathbf{Q}_0, \widehat{\mathbf{S}}_1) d\text{vol}_{\mathbb{R}^d}(\mathbf{Q}_0) d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_1) \cdots d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_n),$$

with  $P^{(n)}$  as in Remark 4.6, and where  $\Lambda'$  is the Radon-Nikodym derivative of  $\Lambda$  with respect to  $\text{vol}_{\mathbb{R}^d} \times \text{vol}_{\mathbb{S}_1^{d-1}}$ .

The probability measure corresponding to the above limiting distribution is defined by

$$(4.75) \quad \mu_\Lambda^{(n)}(\mathcal{A}) := \int_{\mathcal{A} \cap (\mathbb{R}^d \times \mathcal{B}_n)} P^{(n)}(\mathbf{S}_1, \dots, \mathbf{S}_n) \Lambda'(\mathbf{Q}_0, \widehat{\mathbf{S}}_1) d\text{vol}_{\mathbb{R}^d}(\mathbf{Q}_0) d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_1) \cdots d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_n).$$

*Proof of Theorem 4.5.* We will use the same basic technique as in [19, Sec. 9.2]. Note that by a standard approximation argument, using the fact that  $\mu_\Lambda^{(n)}$  is absolutely continuous with respect to Lebesgue measure, it suffices to prove the corresponding statement for continuous functions of compact support, i.e. to prove that for any continuous function  $g : \mathbb{R}^d \times (\mathbb{R}^d)^n \rightarrow \mathbb{R}$  of compact support we have

$$(4.76) \quad \int_{(\mathbf{Q}_0, \mathbf{V}_0) \in \mathbb{T}^1(\rho^{d-1}\mathcal{K}_\rho)} g(\mathbf{Q}_0, \mathbf{S}_1(\mathbf{Q}_0, \mathbf{V}_0; \rho), \dots, \mathbf{S}_n(\mathbf{Q}_0, \mathbf{V}_0; \rho)) \\ \times \Lambda'(\mathbf{Q}_0, \mathbf{V}_0) d\text{vol}_{\mathbb{R}^d}(\mathbf{Q}_0) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{V}_0) \rightarrow \int_{\mathbb{R}^d \times \mathbb{R}^{nd}} g d\mu_\Lambda^{(n)}$$

as  $\rho \rightarrow 0$ . By a further approximation argument, we may also assume  $\Lambda'$  to be continuous and of compact support (keeping  $\Lambda' \geq 0$  and  $\int_{\mathbb{T}^1(\mathbb{R}^d)} \Lambda' d\text{vol}_{\mathbb{R}^d} d\text{vol}_{\mathbb{S}_1^{d-1}} = 1$ ).

Recalling  $\mathbf{S}_n(\mathbf{Q}_0, \mathbf{V}_0; \rho) = \rho^{d-1} \mathbf{s}_n(\rho^{-(d-1)}\mathbf{Q}_0, \mathbf{V}_0; \rho)$  where  $\mathbf{s}_n$  is  $\mathcal{L}$ -periodic in its first variable, the left hand side of (4.76) may be expressed as

$$(4.77) \quad \int_{F \cap \mathcal{K}_\rho} \int_{\mathbb{S}_1^{d-1}} \left\{ \rho^{d(d-1)} \sum_{\mathbf{q} \in \mathbf{q}_0 + \mathcal{L}} g(\rho^{d-1}\mathbf{q}, \rho^{d-1}\mathbf{s}_1(\mathbf{q}_0, \mathbf{v}_0; \rho), \dots, \rho^{d-1}\mathbf{s}_n(\mathbf{q}_0, \mathbf{v}_0; \rho)) \right. \\ \left. \times \Lambda'(\rho^{d-1}\mathbf{q}, \mathbf{v}_0) \right\} d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_0) d\text{vol}_{\mathbb{R}^d}(\mathbf{q}_0),$$

where  $F \subset \mathbb{R}^d$  is a fundamental parallelogram for  $\mathbb{R}^d/\mathcal{L}$ . Because of our assumptions on  $g$  and  $\Lambda'$  we have

$$\rho^{d(d-1)} \sum_{\mathbf{q} \in \mathbf{q}_0 + \mathcal{L}} g(\rho^{d-1}\mathbf{q}, \mathbf{a}_1, \dots, \mathbf{a}_n) \Lambda'(\rho^{d-1}\mathbf{q}, \mathbf{v}_0) \rightarrow \int_{\mathbb{R}^d} g(\mathbf{q}, \mathbf{a}_1, \dots, \mathbf{a}_n) \Lambda'(\mathbf{q}, \mathbf{v}_0) d\text{vol}_{\mathbb{R}^d}(\mathbf{q})$$

as  $\rho \rightarrow 0$ , uniformly over all  $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^d$ ,  $\mathbf{v}_0 \in \mathbb{S}_1^{d-1}$  and  $\mathbf{q}_0$  in compact subsets of  $\mathbb{R}^d$ . Hence, using also Fubini's theorem, (4.77) equals

$$(4.78) \quad \int_F \int_{\mathbb{R}^d} \int_{\mathbb{S}_1^{d-1}} g(\mathbf{q}, \rho^{d-1}\mathbf{s}_1(\mathbf{q}_0, \mathbf{v}_0; \rho), \dots, \rho^{d-1}\mathbf{s}_n(\mathbf{q}_0, \mathbf{v}_0; \rho)) \\ \times \Lambda'(\mathbf{q}, \mathbf{v}_0) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_0) d\text{vol}_{\mathbb{R}^d}(\mathbf{q}) d\text{vol}_{\mathbb{R}^d}(\mathbf{q}_0) + o(1),$$

where the innermost integral should be interpreted as zero whenever  $\mathbf{q}_0 \notin \mathcal{K}_\rho$ . Now by Theorem 4.4 reformulated in the context of continuous test functions, cf. (4.67), for almost all  $(\mathbf{q}_0, \mathbf{q}) \in F \times \mathbb{R}^d$  the innermost integral in (4.78) tends to

$$(4.79) \quad \int_{\mathcal{B}_n} g(\mathbf{q}, \mathbf{S}_1, \dots, \mathbf{S}_n) P^{(n)}(\mathbf{S}_1, \dots, \mathbf{S}_n) \Lambda'(\mathbf{q}, \widehat{\mathbf{S}}_1) d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_1) \cdots d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_n).$$

Hence (4.76) follows by applying Lebesgue's Bounded Convergence Theorem to (4.78) (with  $(\mathbf{q}_0, \mathbf{q}) \mapsto (\sup |g|) \int_{\mathbb{S}_1^{d-1}} \Lambda'(\mathbf{q}, \mathbf{v}_0) d\mathbf{v}_0$  as a majorant function), and using  $\text{vol}(F) = 1$ .  $\square$

5. CONVERGENCE TO THE STOCHASTIC PROCESS  $\Xi(t)$ 

In this section we will prove Theorem 1.2, generalized to the situation of a general scattering map, and with  $(\mathbf{q}(t), \mathbf{v}(t))$  being the orbit of the flow  $\varphi_t$  for initial data of the form  $(\mathbf{q}_0 + \rho\boldsymbol{\beta}(\mathbf{v}_0), \mathbf{v}_0)$ .

The probability  $\mathbb{P}_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}^{(n)}(\Xi(t_1) \in \mathcal{D}_1, \dots, \Xi(t_M) \in \mathcal{D}_M)$  is defined in a similar way as before: We again write  $T_0 := 0$ ,  $T_n := \sum_{j=1}^n \|\mathbf{S}_j\|$ , and given any  $\mathbf{n} = (n_1, \dots, n_M) \in \mathbb{Z}_{\geq 0}^M$  we set

$$(5.1) \quad \mathbb{P}_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}^{(n)}(\Xi(t_1) \in \mathcal{D}_1, \dots, \Xi(t_M) \in \mathcal{D}_M \text{ and } T_{n_1} \leq t_1 < T_{n_1+1}, \dots, T_{n_M} \leq t_M < T_{n_M+1}) \\ := \mu_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}^{(n+1)}(\{(\mathbf{S}_1, \dots, \mathbf{S}_{n+1}) : \Xi_{n_j}(t_j) \in \mathcal{D}_j, T_{n_j} \leq t_j < T_{n_j+1} \ (j = 1, \dots, M)\})$$

with  $n = \max(n_1, \dots, n_M)$  and  $\Xi_n(t) := (\sum_{j=1}^n \mathbf{S}_j + (t - T_n)\widehat{\mathbf{S}}_{n+1}, \widehat{\mathbf{S}}_{n+1})$ , and with the measure  $\mu_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}^{(n)}$  now being given by Theorem 4.4 and (4.66). (Thus  $\mu_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}^{(n)}$  depends on the given scattering map and the given function  $\boldsymbol{\beta}$ .) We then set

$$(5.2) \quad \mathbb{P}_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}(\Xi(t_1) \in \mathcal{D}_1, \dots, \Xi(t_M) \in \mathcal{D}_M) \\ := \sum_{\mathbf{n} \in \mathbb{Z}_{\geq 0}^M} \mathbb{P}_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}^{(n)}(\Xi(t_1) \in \mathcal{D}_1, \dots, \Xi(t_M) \in \mathcal{D}_M \text{ and } T_{n_1} \leq t_1 < T_{n_1+1}, \dots, T_{n_M} \leq t_M < T_{n_M+1}).$$

Recall that given any set  $\mathcal{D} \subset \mathbb{T}^1(\mathbb{R}^d)$  we say that  $t \geq 0$  is  $\mathcal{D}$ -admissible if  $(t\widehat{\mathbf{S}}_1, \widehat{\mathbf{S}}_1) \notin \partial\mathcal{D}$  holds for  $(\text{vol}_{\mathbb{S}_1^{d-1}})$ -almost all  $\widehat{\mathbf{S}}_1 \in \mathbb{S}_1^{d-1}$ , and we write  $\text{adm}(\mathcal{D})$  for the set of all  $\mathcal{D}$ -admissible numbers  $t \geq 0$ .

The generalization of Theorem 1.2 now reads as follows.

**Theorem 5.1.** *Fix a lattice  $\mathcal{L} = \mathbb{Z}^d M_0$  and a point  $\mathbf{q}_0 \in \mathbb{R}^d$ , let  $\lambda$  be a Borel probability measure on  $\mathbb{S}_1^{d-1}$  which is absolutely continuous with respect to Lebesgue measure, and let  $\boldsymbol{\beta} : \mathbb{S}_1^{d-1} \rightarrow \mathbb{R}^d$  be a  $C^1$  function. If  $\mathbf{q}_0 \in \mathcal{L}$  we assume that  $(\boldsymbol{\beta}(\mathbf{v}) + \mathbb{R}_{>0}\mathbf{v}) \cap \mathcal{B}_1^d = \emptyset$  for all  $\mathbf{v} \in \mathbb{S}_1^{d-1}$ . Then, for any subsets  $\mathcal{D}_1, \dots, \mathcal{D}_M \subset \mathbb{T}^1(\mathbb{R}^d)$  with boundary of Lebesgue measure zero, and any numbers  $t_j \in \text{adm}(\mathcal{D}_j)$  ( $j = 1, \dots, M$ ),*

$$(5.3) \quad \lim_{\rho \rightarrow 0} \lambda(\{\mathbf{v}_0 \in \mathbb{S}_1^{d-1} : (\rho^{d-1}\mathbf{q}(\rho^{-(d-1)}t_j), \mathbf{v}(\rho^{-(d-1)}t_j)) \in \mathcal{D}_j, \ j = 1, \dots, M\}) \\ = \mathbb{P}_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}(\Xi(t_1) \in \mathcal{D}_1, \dots, \Xi(t_M) \in \mathcal{D}_M).$$

The convergence is uniform for  $(t_1, \dots, t_M)$  in compact subsets of  $\text{adm}(\mathcal{D}_1) \times \dots \times \text{adm}(\mathcal{D}_M)$ .

**5.1. Four lemmas.** As a preparation for the proof of Theorem 5.1 we will require the following four lemmas.

**Lemma 5.2.** *Given any  $t > 0$  and  $\varepsilon > 0$  there exist  $\rho_0 > 0$  and  $N \in \mathbb{Z}_{>0}$  such that, for all  $\rho < \rho_0$ ,*

$$(5.4) \quad \lambda(\{\mathbf{v}_0 \in \mathbb{S}_1^{d-1} : \|\mathbf{s}_1\| + \|\mathbf{s}_2\| + \dots + \|\mathbf{s}_N\| \leq \rho^{-(d-1)}t\}) < \varepsilon,$$

where  $\mathbf{s}_k = \mathbf{s}_k(\mathbf{q}_0 + \rho\boldsymbol{\beta}(\mathbf{v}_0), \mathbf{v}_0; \rho)$ .

*Proof.* For  $t > 0$  and  $N \in \mathbb{Z}_{>0}$  we have by Theorem 4.1 (coupled with a simple approximation argument)

$$(5.5) \quad \lim_{\rho \rightarrow 0} \lambda(\{\mathbf{v}_0 \in \mathbb{S}_1^{d-1} : \|\mathbf{s}_1\| + \|\mathbf{s}_2\| + \dots + \|\mathbf{s}_N\| \leq \rho^{-(d-1)}t\}) \\ = \nu(\{(\mathbf{v}_0, \xi_1, \mathbf{v}_1, \dots, \xi_N, \mathbf{v}_N) \in \mathbb{S}_1^{d-1} \times (\mathbb{R}_{>0} \times \mathbb{S}_1^{d-1})^N : \sum_{j=1}^N \xi_j < t\}),$$

where  $\nu$  is the measure defined in (4.25). Transforming the integral in (4.25) via (4.1) and (2.25), and then using the fact that  $\Phi_{\boldsymbol{\alpha}}(\xi, \mathbf{w}, \mathbf{z})$  and  $\Phi_0(\xi, \mathbf{w}, \mathbf{z})$  are uniformly bounded by 1,

we see that (5.5) is bounded from above by

$$(5.6) \quad \text{vol}(\mathcal{B}_1^{d-1})^N \int_{\substack{\xi_1 + \dots + \xi_N < t \\ \xi_1, \dots, \xi_N > 0}} d\xi_1 \cdots d\xi_N = \text{vol}(\mathcal{B}_1^{d-1})^N \frac{t^N}{N!}.$$

The lemma follows from the fact that this expression tends to zero as  $N \rightarrow \infty$ , for any fixed  $t > 0$ .  $\square$

Given any  $\mathcal{D} \subset \mathbb{T}^1(\mathbb{R}^d) = \mathbb{R}^d \times \mathbb{S}_1^{d-1}$  and  $\delta > 0$ , we write  $\partial_\delta \mathcal{D}$  and  $\mathcal{N}_\delta \mathcal{D}$  for the  $\delta$ -neighborhoods of  $\partial \mathcal{D}$  and  $\mathcal{D}$ , respectively, viz.:

$$(5.7) \quad \partial_\delta \mathcal{D} = \{(\mathbf{p}, \mathbf{v}) \in \mathbb{R}^d \times \mathbb{S}_1^{d-1} : \exists (\mathbf{p}_1, \mathbf{v}_1) \in \partial \mathcal{D} : \|\mathbf{p}_1 - \mathbf{p}\| + \varphi(\mathbf{v}_1, \mathbf{v}) < \delta\}.$$

$$(5.8) \quad \mathcal{N}_\delta \mathcal{D} = \mathcal{D} \cup \partial_\delta \mathcal{D} = \{(\mathbf{p}, \mathbf{v}) \in \mathbb{R}^d \times \mathbb{S}_1^{d-1} : \exists (\mathbf{p}_1, \mathbf{v}_1) \in \mathcal{D} : \|\mathbf{p}_1 - \mathbf{p}\| + \varphi(\mathbf{v}_1, \mathbf{v}) < \delta\}.$$

**Lemma 5.3.** *Let  $\mathcal{D} \subset \mathbb{T}^1(\mathbb{R}^d)$  and  $t \in \text{adm}(\mathcal{D})$ . Then*

$$(5.9) \quad \lim_{\delta \rightarrow 0} \text{vol}_{\mathbb{S}_1^{d-1}}(\{\widehat{\mathbf{S}}_1 \in \mathbb{S}_1^{d-1} : (t\widehat{\mathbf{S}}_1, \widehat{\mathbf{S}}_1) \in \partial_\delta \mathcal{D}\}) = 0.$$

*Proof.* By assumption the set  $M = \{t\widehat{\mathbf{S}}_1 \in \mathbb{S}_1^{d-1} : (t\widehat{\mathbf{S}}_1, \widehat{\mathbf{S}}_1) \in \partial \mathcal{D}\}$  has volume zero; hence for any given  $\varepsilon > 0$  there is an open set  $U \subset \mathbb{S}_1^{d-1}$  with  $M \subset U$  and  $\text{vol}_{\mathbb{S}_1^{d-1}}(U) < \varepsilon$ . Then  $C = \{(t\widehat{\mathbf{S}}_1, \widehat{\mathbf{S}}_1) : \widehat{\mathbf{S}}_1 \in \mathbb{S}_1^{d-1} \setminus U\}$  is a compact subset of  $\mathbb{T}^1(\mathbb{R}^d)$ , disjoint from  $\partial \mathcal{D}$ . Since  $\partial \mathcal{D}$  is closed, there is some  $\delta_0 > 0$  such that  $\|\mathbf{p}_1 - \mathbf{p}_2\| + \varphi(\mathbf{v}_1, \mathbf{v}_2) > \delta_0$  for all  $(\mathbf{p}_1, \mathbf{v}_1) \in C$ ,  $(\mathbf{p}_2, \mathbf{v}_2) \in \partial \mathcal{D}$ . Hence for each  $\delta \leq \delta_0$  the set  $M_\delta = \{t\widehat{\mathbf{S}}_1 \in \mathbb{S}_1^{d-1} : (t\widehat{\mathbf{S}}_1, \widehat{\mathbf{S}}_1) \in \partial_\delta \mathcal{D}\}$  is contained in  $U$ , and  $\text{vol}_{\mathbb{S}_1^{d-1}}(M_\delta) \leq \text{vol}_{\mathbb{S}_1^{d-1}}(U) < \varepsilon$ .  $\square$

**Lemma 5.4.** *Given any  $n \geq 1$ ,  $t \geq 0$  and  $\varepsilon > 0$ , there exists some  $\delta > 0$  such that, for every measurable subset  $\mathcal{D} \subset \mathbb{R}^d \times \mathbb{S}_1^{d-1}$  with  $[\text{vol}_{\mathbb{R}^d} \times \text{vol}_{\mathbb{S}_1^{d-1}}](\mathcal{D}) \leq \delta$ ,*

$$(5.10) \quad [(\text{vol}_{\mathbb{R}^d})^n \times \text{vol}_{\mathbb{S}_1^{d-1}}] \left( \left\{ (\mathbf{S}_1, \dots, \mathbf{S}_n, \widehat{\mathbf{S}}_{n+1}) \in \mathbb{R}^{nd} \times \mathbb{S}_1^{d-1} : \sum_{k=1}^n \|\mathbf{S}_k\| < t, \right. \right. \\ \left. \left. \left( \sum_{k=1}^n \mathbf{S}_k + \left( t - \sum_{k=1}^n \|\mathbf{S}_k\| \right) \widehat{\mathbf{S}}_{n+1}, \widehat{\mathbf{S}}_{n+1} \right) \in \mathcal{D} \right\} \right) < \varepsilon.$$

*Proof.* We may assume  $t > 0$  since otherwise the left hand side of (5.10) is zero. Let us first treat the case  $n = 1$ . Set

$$(5.11) \quad U_t = \{(\mathbf{S}_1, \widehat{\mathbf{S}}_2) \in \mathcal{B}_t^d \times \mathbb{S}_1^{d-1} : \mathbf{S}_1 \notin \mathbb{R}_{\geq 0} \widehat{\mathbf{S}}_2\}$$

and let  $\Theta_t$  be the map

$$(5.12) \quad \Theta_t : U_t \rightarrow \mathbb{R}^d \times \mathbb{S}_1^{d-1}; \quad (\mathbf{S}_1, \widehat{\mathbf{S}}_2) \mapsto (\mathbf{S}_1 + (t - \|\mathbf{S}_1\|)\widehat{\mathbf{S}}_2, \widehat{\mathbf{S}}_2).$$

Then since  $\mathcal{B}_t^d \times \mathbb{S}_1^{d-1} \setminus U_t$  has measure zero, the left hand side of (5.10) equals  $\text{vol}(\Theta_t^{-1}(\mathcal{D}))$ , where we write  $\text{vol} := \text{vol}_{\mathbb{R}^d} \times \text{vol}_{\mathbb{S}_1^{d-1}}$ . But one verifies that  $\Theta_t$  is a diffeomorphism from  $U_t$  onto  $\mathcal{B}_t^d \times \mathbb{S}_1^{d-1}$ . Hence  $\text{vol} \circ \Theta_t^{-1}$  is a bounded measure on  $\mathbb{R}^d \times \mathbb{S}_1^{d-1}$  which is absolutely continuous with respect to  $\text{vol}$ . This implies the desired claim (cf. [25, Thm. 6.11]).

In the remaining case  $n \geq 2$ , the left hand side of (5.10) can be expressed as

$$(5.13) \quad \int_{\substack{(\mathbf{S}_1, \dots, \mathbf{S}_{n-1}) \in \mathbb{R}^{(n-1)d} \\ T_{n-1} < t}} \text{vol} \left( \Theta_{t-T_{n-1}}^{-1} \left( \mathcal{D} - \sum_{j=1}^{n-1} \mathbf{S}_j \right) \right) d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_1) \cdots d\text{vol}_{\mathbb{R}^d}(\mathbf{S}_{n-1}),$$

where  $\mathcal{D} - \mathbf{q} := \{(\mathbf{p} - \mathbf{q}, \mathbf{v}) : (\mathbf{p}, \mathbf{v}) \in \mathcal{D}\}$  (also recall  $T_{n-1} := \sum_{j=1}^{n-1} \|\mathbf{S}_j\|$ ). Hence it suffices to prove that for any given  $\varepsilon_1 > 0$  we can choose  $\delta > 0$  so small that for every measurable subset  $\mathcal{D} \subset \mathbb{R}^d \times \mathbb{S}_1^{d-1}$  with  $\text{vol}(\mathcal{D}) \leq \delta$  and for every  $t' \in (0, t]$ , we have  $\text{vol}(\Theta_{t'}^{-1}(\mathcal{D})) < \varepsilon_1$ . To prove this, for  $t'$  so small that  $\text{vol}(U_{t'}) < \varepsilon_1$  we use the a priori bound  $\text{vol}(\Theta_{t'}^{-1}(\mathcal{D})) \leq \text{vol}(U_{t'})$ ; the remaining  $t'$ -interval can then be treated using the above discussion of  $\text{vol} \circ \Theta_t^{-1}$ , together

with the relation  $\Theta_{t'}^{-1} = L_{t'/t} \circ \Theta_t^{-1} \circ L_{t/t'}$  where  $L_x : \mathbb{R}^d \times \mathbb{S}_1^{d-1} \rightarrow \mathbb{R}^d \times \mathbb{S}_1^{d-1}$  is the map  $(\mathbf{p}, \mathbf{v}) \mapsto (x\mathbf{p}, \mathbf{v})$ .  $\square$

Given any  $\mathbf{n} \in \mathbb{Z}_{\geq 0}^M$  and subsets  $\mathcal{D}'_1, \dots, \mathcal{D}'_M \subset \mathbb{T}^1(\mathbb{R}^d)$ , we set  $n = \max(n_1, \dots, n_M)$  and

$$(5.14) \quad \mathcal{A}_{\mathbf{n}}(\mathcal{D}'_1, \dots, \mathcal{D}'_M) := \{(\mathbf{S}_1, \dots, \mathbf{S}_{n+1}) : \Xi_{n_j}(t_j) \in \mathcal{D}'_j, T_{n_j} \leq t_j < T_{n_j+1} \ (j = 1, \dots, M)\},$$

a subset of  $\mathbb{R}^{(n+1)d}$ .

**Lemma 5.5.** *Let  $\mathcal{D}_1, \dots, \mathcal{D}_M$  and  $t_1, \dots, t_M$  be as in Theorem 5.1, and assume furthermore that each set  $\mathcal{D}_j$  is bounded. Then for each  $\mathbf{n} \in \mathbb{Z}_{\geq 0}^M$  and  $\varepsilon > 0$ , there exists a choice of subsets  $\mathcal{D}'_j, \mathcal{D}''_j \subset \mathbb{T}^1(\mathbb{R}^d)$  with  $\mathcal{D}'_j \subset \mathcal{D}_j \subset \mathcal{D}''_j$  ( $j = 1, \dots, M$ ), such that the following holds:*

- (i) *There is some  $\delta > 0$  such that  $\mathcal{N}_\delta \mathcal{D}'_j \subset \mathcal{D}_j$  and  $\mathcal{N}_\delta \mathcal{D}_j \subset \mathcal{D}''_j$ .*
- (ii)  *$\mu_{\alpha, \beta, \lambda}^{(n+1)}(\mathcal{A}_{\mathbf{n}}(\mathcal{D}''_1, \dots, \mathcal{D}''_M)) - \varepsilon < \mu_{\alpha, \beta, \lambda}^{(n+1)}(\mathcal{A}_{\mathbf{n}}(\mathcal{D}_1, \dots, \mathcal{D}_M)) < \mu_{\alpha, \beta, \lambda}^{(n+1)}(\mathcal{A}_{\mathbf{n}}(\mathcal{D}'_1, \dots, \mathcal{D}'_M)) + \varepsilon$ .*
- (iii) *Both  $\mathcal{A}_{\mathbf{n}}(\mathcal{D}'_1, \dots, \mathcal{D}'_M)$  and  $\mathcal{A}_{\mathbf{n}}(\mathcal{D}''_1, \dots, \mathcal{D}''_M)$  have boundaries of Lebesgue measure zero.*

*Proof.* Given any  $\delta > 0$  we let  $\mathcal{F}_1$  be the family of all cubes of the form  $C = \delta \mathbf{m} + [0, \delta]^d$  with  $\mathbf{m} \in \mathbb{Z}^d$ , and fix  $\mathcal{F}_2$  to be a finite family of closed subsets  $D \subset \mathbb{S}_1^{d-1}$  with diameter  $< \delta$  (with respect to the metric  $\varphi$ ) and boundary of measure zero, such that  $\cup_{D \in \mathcal{F}_2} D = \mathbb{S}_1^{d-1}$ . Let  $\mathcal{F}$  be the family of all sets  $B = C \times D \subset \mathbb{T}^1(\mathbb{R}^d)$  with  $C \in \mathcal{F}_1$  and  $D \in \mathcal{F}_2$ . Now for each  $j \in \{1, \dots, M\}$  we define  $\mathcal{D}'_j$  as the union of all  $B \in \mathcal{F}$  with  $\mathcal{N}_\delta B \subset \mathcal{D}_j$ ; and define  $\mathcal{D}''_j$  as the union of all  $B \in \mathcal{F}$  with  $B \cap \mathcal{N}_\delta \mathcal{D}_j \neq \emptyset$ . Note that both these unions are finite, since  $\mathcal{D}_j$  is bounded, and (i) holds by construction. Also note that  $\mathcal{D}_j \setminus \mathcal{D}'_j \subset \partial_{(d+2)\delta} \mathcal{D}_j$  and  $\mathcal{D}''_j \setminus \mathcal{D}_j \subset \partial_{(d+2)\delta} \mathcal{D}_j$ .

We now prove that (ii) holds provided that  $\delta$  is sufficiently small. Every  $(\mathbf{S}_1, \dots, \mathbf{S}_{n+1}) \in \mathcal{A}_{\mathbf{n}}(\mathcal{D}_1, \dots, \mathcal{D}_M) \setminus \mathcal{A}_{\mathbf{n}}(\mathcal{D}'_1, \dots, \mathcal{D}'_M)$  satisfies  $\Xi_{n_j}(t_j) \in \mathcal{D}_j \setminus \mathcal{D}'_j$  for some  $j$ , and hence, using also Remark 4.5 repeatedly,

$$(5.15) \quad \begin{aligned} & \mu_{\alpha, \beta, \lambda}^{(n+1)}\left(\mathcal{A}_{\mathbf{n}}(\mathcal{D}_1, \dots, \mathcal{D}_M) \setminus \mathcal{A}_{\mathbf{n}}(\mathcal{D}'_1, \dots, \mathcal{D}'_M)\right) \\ & \leq \sum_{j=1}^M \mu_{\alpha, \beta, \lambda}^{(n_j+1)}\left(\{(\mathbf{S}_1, \dots, \mathbf{S}_{n_j+1}) \in \mathbb{R}^{(n_j+1)d} : \Xi_{n_j}(t_j) \in \partial_{(d+2)\delta} \mathcal{D}_j, T_{n_j} \leq t_j\}\right). \end{aligned}$$

For any  $j \in \{1, \dots, M\}$ , if  $n_j = 0$  then (via (4.69)) the  $j$ th term in the above sum equals

$$(5.16) \quad \lambda(\{\widehat{\mathbf{S}}_1 \in \mathbb{S}_1^{d-1} : (t_j \widehat{\mathbf{S}}_1, \widehat{\mathbf{S}}_1) \in \partial_{(d+2)\delta} \mathcal{D}_j\}),$$

which tends to zero as  $\delta \rightarrow 0$  by Lemma 5.3. On the other hand if  $n_j \geq 1$ , then the  $j$ th term in (5.15) equals  $\mu_{\alpha, \beta, \lambda}^{(n_j+1)}(\varpi^{-1}(\mathcal{M}))$  where  $\varpi$  is the projection

$$(5.17) \quad \varpi : \mathbb{R}^{n_j d} \times (\mathbb{R}^d \setminus \{\mathbf{0}\}) \rightarrow \mathbb{R}^{n_j d} \times \mathbb{S}_1^{d-1}; \quad (\mathbf{S}_1, \dots, \mathbf{S}_{n+1}) \mapsto (\mathbf{S}_1, \dots, \mathbf{S}_n, \widehat{\mathbf{S}}_{n+1}),$$

and  $\mathcal{M}$  is the same set as in the left hand side of (5.10), with  $n_j, t_j$  and  $\partial_{(d+2)\delta} \mathcal{D}_j$  in place of  $n, t, \mathcal{D}$ . Now  $\mu_{\alpha, \beta, \lambda}^{(n_j+1)} \circ \varpi^{-1}$  is a Borel probability measure on  $\mathbb{R}^{n_j d} \times \mathbb{S}_1^{d-1}$  which is absolutely continuous with respect to  $(\text{vol}_{\mathbb{R}^d})^{n_j} \times \text{vol}_{\mathbb{S}_1^{d-1}}$ . Also

$$(5.18) \quad [\text{vol}_{\mathbb{R}^d} \times \text{vol}_{\mathbb{S}_1^{d-1}}](\partial_{(d+2)\delta} \mathcal{D}_j) \rightarrow 0 \quad \text{as } \delta \rightarrow 0,$$

since  $\mathcal{D}_j$  is bounded and has boundary of measure zero. Hence by Lemma 5.4,

$$(5.19) \quad \mu_{\alpha, \beta, \lambda}^{(n_j+1)}(\varpi^{-1}(\mathcal{M})) \rightarrow 0 \quad \text{as } \delta \rightarrow 0.$$

In conclusion we see that for  $\delta$  sufficiently small, the sum in (5.15) must be  $< \varepsilon$ , so that the second inequality in (ii) holds. Similarly, using  $\mathcal{D}''_j \setminus \mathcal{D}_j \subset \partial_{(d+2)\delta} \mathcal{D}_j$ , one proves that also the first inequality in (ii) must hold for  $\delta$  sufficiently small.

The property (iii) is proved in a similar way, making use the fact that each set  $\mathcal{D}'_j$  and  $\mathcal{D}''_j$  is a finite union of sets  $B \in \mathcal{F}$ , and that, by construction, each such set  $B$  satisfies  $\text{vol}_{S_1^{d-1}}(\{\widehat{\mathcal{S}}_1 \in S_1^{d-1} : (t\widehat{\mathcal{S}}_1, \widehat{\mathcal{S}}_1) \in \partial B\}) = 0$  and  $[\text{vol}_{\mathbb{R}^d} \times \text{vol}_{S_1^{d-1}}](\partial B) = 0$ .  $\square$

## 5.2. Proof of Theorem 5.1.

5.2.1. *Convergence.* Note that both sides of (5.3) remain unchanged if we replace  $\mathcal{D}_j$  with  $\{(\mathbf{q}, \mathbf{v}) \in \mathcal{D}_j : \|\mathbf{q}\| < t_j + 1\}$  for each  $j = 1, \dots, M$ . Hence from now on we may assume without loss of generality that *each set  $\mathcal{D}_j$  is bounded*. Writing  $\mathbf{v}_k = \mathbf{v}_k(\mathbf{q}_0 + \rho\boldsymbol{\beta}(\mathbf{v}_0), \mathbf{v}_0; \rho)$ ,  $\mathbf{s}_k = \mathbf{s}_k(\mathbf{q}_0 + \rho\boldsymbol{\beta}(\mathbf{v}_0), \mathbf{v}_0; \rho)$  and  $T_n = T_n(\mathbf{q}_0 + \rho\boldsymbol{\beta}(\mathbf{v}_0), \mathbf{v}_0; \rho) := \rho^{d-1} \sum_{k=1}^n \|\mathbf{s}_k\|$ , the left hand side of (5.3) may be expressed as  $\lim_{\rho \rightarrow 0} \sum_{\mathbf{n} \in \mathbb{Z}_{\geq 0}^M} \lambda(\mathcal{S}_{\mathbf{n}, \rho})$ , where

$$(5.20) \quad \mathcal{S}_{\mathbf{n}, \rho} = \left\{ \mathbf{v}_0 \in S_1^{d-1} : \left( \rho^{d-1} \mathbf{q}_0 + \rho^d \boldsymbol{\beta}(\mathbf{v}_0) + \sum_{k=1}^{n_j} \rho^{d-1} \mathbf{s}_k + (t_j - T_{n_j}) \mathbf{v}_{n_j}, \mathbf{v}_{n_j} \right) \in \mathcal{D}_j, \right. \\ \left. T_{n_j} \leq t_j < T_{n_j+1} \ (j = 1, \dots, M) \right\}.$$

The right hand side of (5.3) is, by definition,

$$(5.21) \quad \mathbb{P}_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}(\Xi(t_1) \in \mathcal{D}_1, \dots, \Xi(t_M) \in \mathcal{D}_M) = \sum_{\mathbf{n} \in \mathbb{Z}_{\geq 0}^M} \mu_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}^{(n+1)}(\mathcal{A}_{\mathbf{n}}(\mathcal{D}_1, \dots, \mathcal{D}_M)).$$

By Lemma 5.2, for any given  $\varepsilon > 0$  there is a finite subset  $S \subset \mathbb{Z}_{\geq 0}^M$  such that  $\sum_{\mathbf{n} \in \mathbb{Z}_{\geq 0}^M \setminus S} \lambda(\mathcal{S}_{\mathbf{n}, \rho}) < \varepsilon$  holds for all sufficiently small  $\rho > 0$ . Hence to prove (5.3), and also deduce that the sum in (5.21) is indeed convergent, it suffices to prove that for every *fixed*  $\mathbf{n} \in \mathbb{Z}_{\geq 0}^M$ , we have

$$(5.22) \quad \lim_{\rho \rightarrow 0} \lambda(\mathcal{S}_{\mathbf{n}, \rho}) = \mu_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}^{(n+1)}(\mathcal{A}_{\mathbf{n}}(\mathcal{D}_1, \dots, \mathcal{D}_M)).$$

To prove (5.22), let  $\varepsilon > 0$  be given. Take  $\mathcal{D}'_j, \mathcal{D}''_j \subset T^1(\mathbb{R}^d)$  and  $\delta > 0$  so that all the claims in Lemma 5.5 hold. By Lemma 5.5 (i), if  $\rho > 0$  is so small that  $\|\rho^{d-1} \mathbf{q}_0 + \rho^d \boldsymbol{\beta}(\mathbf{v}_0)\| < \delta$  for all  $\mathbf{v}_0 \in S_1^{d-1}$ , then

$$(5.23) \quad \left\{ \mathbf{v}_0 \in S_1^{d-1} : (\mathbf{s}_1, \dots, \mathbf{s}_{n+1}) \in \rho^{-(d-1)} \mathcal{A}_{\mathbf{n}}(\mathcal{D}'_1, \dots, \mathcal{D}'_M) \right\} \\ \subset \mathcal{S}_{\mathbf{n}, \rho} \subset \left\{ \mathbf{v}_0 \in S_1^{d-1} : (\mathbf{s}_1, \dots, \mathbf{s}_{n+1}) \in \rho^{-(d-1)} \mathcal{A}_{\mathbf{n}}(\mathcal{D}''_1, \dots, \mathcal{D}''_M) \right\}.$$

Hence by Theorem 4.4, using also Lemma 5.5 (ii) and (iii), we have

$$(5.24) \quad \limsup_{\rho \rightarrow 0} \lambda(\mathcal{S}_{\mathbf{n}, \rho}) \leq \mu_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}^{(n+1)}(\mathcal{A}_{\mathbf{n}}(\mathcal{D}''_1, \dots, \mathcal{D}''_M)) < \mu_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}^{(n+1)}(\mathcal{A}_{\mathbf{n}}(\mathcal{D}_1, \dots, \mathcal{D}_M)) + \varepsilon;$$

$$(5.25) \quad \liminf_{\rho \rightarrow 0} \lambda(\mathcal{S}_{\mathbf{n}, \rho}) \geq \mu_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}^{(n+1)}(\mathcal{A}_{\mathbf{n}}(\mathcal{D}'_1, \dots, \mathcal{D}'_M)) > \mu_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}^{(n+1)}(\mathcal{A}_{\mathbf{n}}(\mathcal{D}_1, \dots, \mathcal{D}_M)) - \varepsilon.$$

Since this is true for each  $\varepsilon > 0$  we have now proved (5.22).

5.2.2. *Uniformity.* We now turn to the statement about uniformity in (5.3). By similar arguments as above one proves that the right hand side of (5.3) is continuous as a function of  $(t_1, \dots, t_M)$  at each point of  $\text{adm}(\mathcal{D}_1) \times \dots \times \text{adm}(\mathcal{D}_M)$ . Hence the desired statement about uniform convergence for  $(t_1, \dots, t_M)$  in compact subsets of  $\text{adm}(\mathcal{D}_1) \times \dots \times \text{adm}(\mathcal{D}_M)$  will follow if we can prove that

$$(5.26) \quad \lim_{\rho \rightarrow 0} \lambda(\{ \mathbf{v}_0 \in S_1^{d-1} : (\rho^{d-1} \mathbf{q}(\rho^{-(d-1)} t_j(\rho)), \mathbf{v}(\rho^{-(d-1)} t_j(\rho))) \in \mathcal{D}_j, \ j = 1, \dots, M \}) \\ = \mathbb{P}_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}(\Xi(t_1) \in \mathcal{D}_1, \dots, \Xi(t_M) \in \mathcal{D}_M).$$

holds for any functions  $t_j(\rho)$  from  $\mathbb{R}_{>0}$  to  $\mathbb{R}_{\geq 0}$  satisfying  $t_j = \lim_{\rho \rightarrow 0} t_j(\rho) \in \text{adm}(\mathcal{D}_j)$ . By Lemma 5.2, we see that it suffices to prove that (5.22) holds for any fixed  $\mathbf{n} \in \mathbb{Z}_{\geq 0}^M$  and with  $\mathcal{S}_{\mathbf{n}, \rho}$  redefined using (5.20) with each “ $t_j$ ” replaced by “ $t_j(\rho)$ ”.

To this end, let us define  $\mathcal{E}_\rho$  to be the set of those  $\mathbf{v}_0 \in \mathbb{S}_1^{d-1}$  for which there is some  $j$  such that exactly one of the two numbers  $t_j(\rho)$ ,  $t_j$  lies in  $[T_{n_j}, T_{n_j+1})$ . The point of this is that if we keep  $\rho > 0$  so small that  $\|\rho^{d-1}\mathbf{q}_0 + \rho^d\boldsymbol{\beta}(\mathbf{v}_0)\| + |t_j(\rho) - t_j| < \delta$  for all  $j$  and  $\mathbf{v}_0$ , then for our redefined  $\mathcal{S}_{n,\rho}$  the inclusions in (5.23) *can only fail for vectors*  $\mathbf{v}_0 \in \mathcal{E}_\rho$ . Now  $\lim_{\rho \rightarrow 0} \lambda(\mathcal{E}_\rho) = 0$ , as we see by applying Theorem 4.4 to the sets

$$(5.27) \quad \mathcal{A}_\delta := \cup_{j=1}^M \cup_{k \in \{n_j, n_{j+1}\} \setminus \{0\}} \{(\mathbf{S}_1, \dots, \mathbf{S}_{n+1}) \in \mathbb{R}^{(n+1)d} : T_k \in (t_j - \delta, t_j + \delta)\}$$

for  $\delta = \frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \dots$  (Indeed, note that  $\lim_{\delta \rightarrow 0} \mu_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}^{(n+1)}(\mathcal{A}_\delta) = 0$ , since  $\mu_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}^{(n+1)}$  is bounded and absolutely continuous with respect to  $(\text{vol}_{\mathbb{R}^d})^{n+1}$ .) Using these observations, the proof of (5.22) carries over to the present situation. This completes the proof of Theorem 5.1.  $\square$

**5.3. A counterexample.** We now give an example to show that the condition  $t_j \in \text{adm}(\mathcal{D}_j)$  in the statement of Theorem 5.1 (or Theorem 1.2) cannot be disposed with. Suppose  $M = 1$ ,  $t_1 > 0$  and  $\mathcal{D}_1 = \mathcal{B}_{t_1}^d \times \mathbb{S}_1^{d-1}$ . Then as in the above proof we have

$$(5.28) \quad \sum_{n \geq 1} \lambda(\mathcal{S}_{(n), \rho}) \rightarrow \sum_{n \geq 1} \mu_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}^{(n+1)}(\mathcal{A}_{(n)}(\mathcal{D}_1)) \quad \text{as } \rho \rightarrow 0.$$

However, in general we have

$$(5.29) \quad \lambda(\mathcal{S}_{(0), \rho}) \not\rightarrow \mu_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \lambda}^{(1)}(\mathcal{A}_{(0)}(\mathcal{D}_1)) \quad \text{as } \rho \rightarrow 0.$$

Indeed, there are many choices of  $\mathbf{q}_0$ ,  $\boldsymbol{\beta}$  and  $\lambda$  such that  $\rho^{d-1}\mathbf{q}_0 + \rho^d\boldsymbol{\beta}(\mathbf{v}_0) + t_1\mathbf{v}_0 \in \mathcal{B}_{t_1}^d$  holds for all sufficiently small  $\rho > 0$  and all  $\mathbf{v}_0$  in the support of  $\lambda$ , and in this case we have

$$(5.30) \quad \begin{aligned} \lambda(\mathcal{S}_{(0), \rho}) &= \lambda(\{\mathbf{v}_0 \in \mathbb{S}_1^{d-1} : (\rho^{d-1}\mathbf{q}_0 + \rho^d\boldsymbol{\beta}(\mathbf{v}_0) + t_1\mathbf{v}_0, \mathbf{v}_0) \in \mathcal{D}_1, \|\mathbf{s}_1\| > \rho^{-(d+1)}t_1\}) \\ &= \lambda(\{\mathbf{v}_0 \in \mathbb{S}_1^{d-1} : \rho^{d-1}\tau_1(\mathbf{q}_0 + \rho\boldsymbol{\beta}(\mathbf{v}_0), \mathbf{v}_0; \rho) > t_1\}) \rightarrow \int_{t_1}^{\infty} \Phi_{\boldsymbol{\alpha}, \boldsymbol{\beta}}(\xi) d\xi, \end{aligned}$$

where  $\boldsymbol{\alpha} = -\mathbf{q}_0 M_0^{-1}$ , cf. [19, Cor. 4.2]. This limit is in general non-zero. On the other hand we have  $\mathcal{A}_{(0)}(\mathcal{D}_1) = \emptyset$  since  $(t_1\widehat{\mathbf{S}}_1, \widehat{\mathbf{S}}_1) \notin \mathcal{D}_1$  for all  $\widehat{\mathbf{S}}_1 \in \mathbb{S}_1^{d-1}$ , and this proves (5.29). Combining (5.28) and (5.29) we see that the limit relation (5.3) *fails* for our  $M = 1$ ,  $\mathcal{D}_1$ ,  $t_1$  and many choices of  $\mathbf{q}_0, \boldsymbol{\beta}, \lambda$ .

Note that we may take  $\boldsymbol{\beta} \equiv \mathbf{0}$  above, i.e. there are many choices of  $\mathbf{q}_0$  and  $\lambda$  such that  $\rho^{d-1}\mathbf{q}_0 + t_1\mathbf{v}_0 \in \mathcal{B}_{t_1}^d$  holds for all sufficiently small  $\rho > 0$  and all  $\mathbf{v}_0$  in the support of  $\lambda$ ; thus the above example applies in particular to the situation in Theorem 1.2.

**5.4. Macroscopic initial conditions.** Finally we discuss the proof of Theorem 1.4, again in the setting of a general scattering map. (The statement of the theorem remains the same. In the definition of the limiting stochastic process we use  $\mu_{\Lambda}^{(n)}$  from (4.75).)

The basic strategy of the proof is to mimic the proof of Theorem 5.1 given in sections 5.1–5.2, using Theorem 4.5 in place of Theorem 4.4. We will here only point out the main differences.

Both Lemma 5.3 and Lemma 5.4 can in the present case be replaced by the much simpler

**Lemma 5.6.** *Given any  $n \geq 0$ ,  $t \geq 0$  and  $\varepsilon > 0$ , there exists some  $\delta > 0$  such that, for every measurable subset  $\mathcal{D} \subset \mathbb{R}^d \times \mathbb{S}_1^{d-1}$  with  $[\text{vol}_{\mathbb{R}^d} \times \text{vol}_{\mathbb{S}_1^{d-1}}](\mathcal{D}) \leq \delta$ ,*

$$(5.31) \quad \begin{aligned} & [(\text{vol}_{\mathbb{R}^d})^n \times \text{vol}_{\mathbb{S}_1^{d-1}}] \left( \left\{ (\mathbf{Q}_0, \mathbf{S}_1, \dots, \mathbf{S}_n, \widehat{\mathbf{S}}_{n+1}) \in \mathbb{R}^d \times \mathbb{R}^{nd} \times \mathbb{S}_1^{d-1} : \sum_{k=1}^n \|\mathbf{S}_k\| \leq t, \right. \right. \\ & \left. \left. (\mathbf{Q}_0 + \sum_{k=1}^n \mathbf{S}_k + (t - \sum_{k=1}^n \|\mathbf{S}_k\|)\widehat{\mathbf{S}}_{n+1}, \widehat{\mathbf{S}}_{n+1}) \in \mathcal{D} \right\} \right) < \varepsilon. \end{aligned}$$

*Proof.* By Fubini's Theorem the left hand side is equal to  $[\text{vol}_{\mathbb{R}^d} \times \text{vol}_{\mathbb{S}_1^{d-1}}](\mathcal{D})$  times a finite constant which only depends on  $n$  and  $t$ .  $\square$



We replace the definition (5.14) by

(5.32)

$$\mathcal{A}_n(\mathcal{D}'_1, \dots, \mathcal{D}'_M) := \{(\mathbf{Q}_0, \mathbf{S}_1, \dots, \mathbf{S}_{n+1}) : \Xi_{n_j}(t_j) \in \mathcal{D}'_j, T_{n_j} \leq t_j < T_{n_{j+1}} (j = 1, \dots, M)\},$$

a subset of  $\mathbb{R}^d \times \mathbb{R}^{(n+1)d}$ . (Recall that  $\Xi_n(t)$  is now given by (1.28).)

Now Lemma 5.5 and the discussion in Section 5.2 carries over with only few and obvious changes. By a simple approximation argument we may assume from the start that  $\Lambda$  has compact support; using this we may then also assume without loss of generality that each set  $\mathcal{D}_j$  is bounded, as before. The proof of the convergence (Sec. 5.2.1) takes a somewhat simpler shape in the present case, since (5.20) is now replaced by the exact identity

$$(5.33) \quad \mathcal{S}_{n,\rho} = \left\{ (\mathbf{Q}_0, \mathbf{V}_0) \in \mathbb{T}^1(\rho^{d-1}\mathcal{K}_\rho) : (\mathbf{Q}_0, \mathbf{S}_1, \dots, \mathbf{S}_{n+1}) \in \mathcal{A}_n(\mathcal{D}_1, \dots, \mathcal{D}_M) \right\},$$

with  $\mathbf{S}_k = \mathbf{S}_k(\mathbf{Q}_0, \mathbf{V}_0; \rho)$ .

## 6. A CONTINUOUS-TIME MARKOV PROCESS

As mentioned in Section 1.3, the operator  $L_t$  describing the dynamics of a particle cloud in the Boltzmann-Grad limit does not form a semigroup, and thus the stochastic process  $\Xi(t)$  is not Markovian. To overcome this difficulty, we set

$$(6.1) \quad \mathcal{X} := \{(\mathbf{Q}, \mathbf{V}, \mathcal{T}, \mathbf{V}_+) \in \mathbb{T}^1(\mathbb{R}^d) \times \mathbb{R}_{\geq 0} \times \mathbb{S}_1^{d-1} : \varphi(\mathbf{V}, \mathbf{V}_+) > B_\Theta\}$$

and extend phase space by the map

$$(6.2) \quad \mathcal{R} : \mathbb{T}^1(\rho^{d-1}\mathcal{K}_\rho) \rightarrow \mathcal{X}, \quad (\mathbf{Q}, \mathbf{V}) \mapsto (\mathbf{Q}, \mathbf{V}, \mathcal{T}, \mathbf{V}_+),$$

where

$$(6.3) \quad \mathcal{T} = \mathcal{T}(\mathbf{Q}, \mathbf{V}) = \|\mathbf{S}_1(\mathbf{Q}, \mathbf{V}; \rho)\|$$

represents the free path length until the next scatterer, and  $\mathbf{V}_+ = \mathbf{V}_+(\mathbf{Q}, \mathbf{V})$  is the velocity after the collision.

For random initial data  $(\mathbf{Q}_0, \mathbf{V}_0)$ , distributed as before with respect to an absolutely continuous probability measure  $\Lambda$ , the dynamics in this extended phase space is again described, in the limit  $\rho \rightarrow 0$ , by a stochastic process  $\widehat{\Xi}$  (cf. Theorem 6.1 below) which is Markovian (Proposition 6.3).

### 6.1. Transcription of Theorem 1.4. Set

$$(6.4) \quad \mathcal{X}^{(n)} = \{(\mathbf{Q}_0, \mathbf{v}_0, \xi_1, \dots, \xi_n, \mathbf{v}_n) \in \mathbb{T}^1(\mathbb{R}^d) \times (\mathbb{R}_{\geq 0} \times \mathbb{S}_1^{d-1})^n : \varphi(\mathbf{v}_{j-1}, \mathbf{v}_j) > B_\Theta, j = 1, \dots, n\}$$

(thus  $\mathcal{X}^{(1)} = \mathcal{X}$ ) and define the volume measure  $\sigma_n$  on  $\mathcal{X}^{(n)}$  by

$$(6.5) \quad d\sigma_n(\mathbf{Q}_0, \mathbf{v}_0, \xi_1, \dots, \xi_n, \mathbf{v}_n) = d\text{vol}_{\mathbb{R}^d}(\mathbf{Q}_0) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_0) d\xi_1 \cdots d\xi_n d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_n).$$

Given a probability density  $f \in L^1(\mathcal{X}, \sigma_1)$  we define for every  $n \geq 1$  the probability measure  $\nu_f^{(n)}$  on  $\mathcal{X}^{(n)}$  by

$$(6.6) \quad \nu_f^{(n)}(\mathcal{A}) := \int_{\mathcal{A}} f(\mathbf{Q}_0, \mathbf{v}_0, \xi_1, \mathbf{v}_1) \prod_{j=1}^{n-1} p_{\mathbf{0}, \beta_{\mathbf{v}_{j-1}}^+}(\mathbf{v}_j, \xi_{j+1}, \mathbf{v}_{j+1}) d\sigma_n(\mathbf{Q}_0, \mathbf{v}_0, \xi_1, \dots, \xi_n, \mathbf{v}_n),$$

for any Borel subset  $\mathcal{A} \subset \mathcal{X}^{(n)}$ . We will use the shorthand notation  $T_n = \xi_1 + \dots + \xi_n$ . Let us define

$$(6.7) \quad \widehat{\Xi}_n(t) := \left( \mathbf{Q}_0 + \sum_{j=1}^n \xi_j \mathbf{v}_{j-1} + (t - T_n) \mathbf{v}_n, \mathbf{v}_n, T_{n+1} - t, \mathbf{v}_{n+1} \right).$$

The stochastic process  $\widehat{\Xi}(t)$  is now characterized via the probability

$$(6.8) \quad \mathbb{P}_f(\widehat{\Xi}(t_1) \in \mathcal{D}_1, \dots, \widehat{\Xi}(t_M) \in \mathcal{D}_M) \\ := \sum_{\mathbf{n} \in \mathbb{Z}_{\geq 0}^M} \mathbb{P}_f^{(\mathbf{n})}(\widehat{\Xi}(t_1) \in \mathcal{D}_1, \dots, \widehat{\Xi}(t_M) \in \mathcal{D}_M \text{ and } T_{n_1} \leq t_1 < T_{n_1+1}, \dots, T_{n_M} \leq t_M < T_{n_M+1}),$$

where  $\mathcal{D}_j \subset \mathcal{X}$  are Borel subsets, and

$$(6.9) \quad \mathbb{P}_f^{(\mathbf{n})}(\widehat{\Xi}(t_1) \in \mathcal{D}_1, \dots, \widehat{\Xi}(t_M) \in \mathcal{D}_M \text{ and } T_{n_1} \leq t_1 < T_{n_1+1}, \dots, T_{n_M} \leq t_M < T_{n_M+1}) \\ := \nu_f^{(\mathbf{n}+1)}(\{(\mathbf{Q}_0, \mathbf{v}_0, \xi_1, \mathbf{v}_1, \dots, \xi_{n+1}, \mathbf{v}_{n+1}) : \widehat{\Xi}_{n_j}(t_j) \in \mathcal{D}_j, T_{n_j} \leq t_j \ (j = 1, \dots, M)\}),$$

with  $n := \max(n_1, \dots, n_M)$ . Note that we have automatically  $t_j \leq T_{n_j+1}$  in the right hand side of (6.9), by the definition of  $\mathcal{X}$ , and the subset corresponding to  $t_j = T_{n_j+1}$  of course has measure zero with respect to  $\nu_f^{(\mathbf{n}+1)}$ .

The following theorem is an extension of Theorem 1.4. Set  $\widehat{F}_t := \mathcal{R} \circ F_t : \mathbb{T}^1(\rho^{d-1}\mathcal{K}_\rho) \rightarrow \mathcal{X}$ .

**Theorem 6.1.** *Fix a lattice  $\mathcal{L}$  and let  $\Lambda$  be a Borel probability measure on  $\mathbb{T}^1(\mathbb{R}^d)$  which is absolutely continuous with respect to Lebesgue measure. Then, for any  $t_1, \dots, t_M \in \mathbb{R}_{\geq 0}$ , and any subsets  $\mathcal{D}_1, \dots, \mathcal{D}_M \subset \mathcal{X}$  with  $\sigma_1(\partial\mathcal{D}_1) = \dots = \sigma_1(\partial\mathcal{D}_M) = 0$ ,*

$$(6.10) \quad \lim_{\rho \rightarrow 0} \Lambda(\{(\mathbf{Q}_0, \mathbf{V}_0) \in \mathbb{T}^1(\rho^{d-1}\mathcal{K}_\rho) : \widehat{F}_{t_1}(\mathbf{Q}_0, \mathbf{V}_0) \in \mathcal{D}_1, \dots, \widehat{F}_{t_M}(\mathbf{Q}_0, \mathbf{V}_0) \in \mathcal{D}_M\}) \\ = \mathbb{P}_f(\widehat{\Xi}(t_1) \in \mathcal{D}_1, \dots, \widehat{\Xi}(t_M) \in \mathcal{D}_M)$$

where

$$(6.11) \quad f(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) = \Lambda'(\mathbf{Q}, \mathbf{V})p(\mathbf{V}, \xi, \mathbf{V}_+),$$

for  $p(\mathbf{V}, \xi, \mathbf{V}_+)$  as in Remark 2.4. The convergence is uniform for  $t_1, \dots, t_M$  in compact subsets of  $\mathbb{R}_{\geq 0}$ .

*Proof.* This is analogous to the proof of Theorem 1.4, using in place of Theorem 4.5 the fact that, for any subset  $\mathcal{A} \subset \mathcal{X}^{(n)}$  with  $\sigma_n(\partial\mathcal{A}) = 0$ ,

$$(6.12) \quad \lim_{\rho \rightarrow 0} \Lambda(\{(\mathbf{Q}_0, \mathbf{V}_0) \in \mathbb{T}^1(\rho^{d-1}\mathcal{K}_\rho) : (\mathbf{Q}_0, \mathbf{V}_0, \rho^{d-1}\tau_1(\rho^{1-d}\mathbf{Q}_0, \mathbf{V}_0; \rho), \mathbf{v}_1(\rho^{1-d}\mathbf{Q}_0, \mathbf{V}_0; \rho), \dots, \\ \rho^{d-1}\tau_n(\rho^{1-d}\mathbf{Q}_0, \mathbf{V}_0; \rho), \mathbf{v}_n(\rho^{1-d}\mathbf{Q}_0, \mathbf{V}_0; \rho)) \in \mathcal{A}\}) \\ = \int_{\mathcal{A}} p(\mathbf{v}_0, \xi_1, \mathbf{v}_1) \prod_{j=1}^{n-1} p_{\mathbf{0}, \beta_{\mathbf{v}_{j-1}}^+}(\mathbf{v}_j, \xi_{j+1}, \mathbf{v}_{j+1}) \Lambda'(\mathbf{Q}_0, \mathbf{v}_0) d\sigma_n(\mathbf{Q}_0, \mathbf{v}_0, \xi_1, \dots, \xi_n, \mathbf{v}_n).$$

The proof of this statement is almost identical to that of Theorem 4.5, with Theorem 4.4 replaced by Theorem 4.1.

The second main ingredient of the proof of Theorem 6.1 is a substitute of Lemma 5.6, which we formulate as follows. Given any  $n \geq 0$ ,  $t \geq 0$ ,  $\varepsilon > 0$ , there exists some  $\delta > 0$  such that, for every measurable subset  $\mathcal{D} \subset \mathcal{X}$  with  $\sigma_1(\mathcal{D}) \leq \delta$ ,

$$(6.13) \quad \sigma_{n+1}\left(\left\{(\mathbf{Q}_0, \mathbf{v}_0, \xi_1, \dots, \xi_{n+1}, \mathbf{v}_{n+1}) \in \mathcal{X}^{(n+1)} : \widehat{\Xi}_n(t) \in \mathcal{D}, T_n \leq t\right\}\right) < \varepsilon.$$

To establish (6.13), note that the left hand side equals

$$(6.14) \quad \frac{t^n}{n!} \text{vol}_{\mathbb{S}_1^{d-1}}(\mathcal{V}_{\mathbf{e}_1})^n \sigma_1(\mathcal{D}).$$

□

**6.2. A semigroup of propagators.** We write  $L^1_{\text{loc}}(\mathcal{X}, \sigma_1)$  for the space of measurable functions  $f : \mathcal{X} \rightarrow \mathbb{R}$  satisfying  $\int_C |f| d\sigma_1 < \infty$  for any compact subset  $C \subset \bar{\mathcal{X}}$ . In order to show that  $\widehat{\Xi}(t)$  is a Markov process, for each  $t \geq 0$  we introduce the operator

$$(6.15) \quad K_t : L^1_{\text{loc}}(\mathcal{X}, \sigma_1) \rightarrow L^1_{\text{loc}}(\mathcal{X}, \sigma_1)$$

by the relation

$$(6.16) \quad \int_{\mathcal{D}} [K_t f](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) d\sigma_1(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) = \mathbb{P}_f(\widehat{\Xi}(t) \in \mathcal{D})$$

for every  $f \in L^1_{\text{loc}}(\mathcal{X}, \sigma_1)$  and every bounded Borel subset  $\mathcal{D} \subset \mathcal{X}$ . The right hand side of (6.16) is defined by extending the definition of  $\mathbb{P}_f$  in (6.6), (6.8), (6.9) for  $f \in L^1_{\text{loc}}(\mathcal{X}, \sigma_1)$ : Note that for a given  $f \in L^1_{\text{loc}}(\mathcal{X}, \sigma_1)$ , (6.6) defines  $\nu_f^{(n)}$  as a signed Borel measure on any bounded open subset of  $\mathcal{X}^{(n)}$ , and then (6.8), (6.9) define  $\mathcal{D} \mapsto \mathbb{P}_f(\widehat{\Xi}(t) \in \mathcal{D})$  as a signed Borel measure on any bounded open subset of  $\mathcal{X}$ . Since this signed measure is absolutely continuous with respect to  $\sigma_1$ , there exists a unique  $K_t f \in L^1_{\text{loc}}(\mathcal{X}, \sigma_1)$  such that (6.16) holds for each bounded Borel subset  $\mathcal{D} \subset \mathcal{X}$ .

If  $f \in L^1(\mathcal{X}, \sigma_1)$  then  $\mathcal{D} \mapsto \mathbb{P}_f(\widehat{\Xi}(t) \in \mathcal{D})$  is in fact a signed Borel measure on all of  $\mathcal{X}$ , of total variation  $\leq \|f\|_{L^1}$ . Hence the restriction of  $K_t$  to  $L^1(\mathcal{X}, \sigma_1)$  maps into  $L^1(\mathcal{X}, \sigma_1)$ , and gives a bounded linear operator on  $L^1(\mathcal{X}, \sigma_1)$  of norm  $\leq 1$ .

Note that  $K_t$  commutes with the translation operators  $\{T_{\mathbf{R}} : \mathbf{R} \in \mathbb{R}^d\}$ ,

$$(6.17) \quad [T_{\mathbf{R}} f](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) := f(\mathbf{Q} - \mathbf{R}, \mathbf{V}, \xi, \mathbf{V}_+),$$

and with the rotation operators  $\{R_K : K \in O(d)\}$ ,

$$(6.18) \quad [R_K f](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) := f(\mathbf{Q}K, \mathbf{V}K, \xi, \mathbf{V}_+K),$$

cf. Remark 4.2.

**Proposition 6.2.** *Suppose that the flow  $F_t$  preserves the Liouville measure  $\nu$  (cf. Remark 2.3). Then the function  $f(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) = p(\mathbf{V}, \xi, \mathbf{V}_+)$  satisfies*

$$(6.19) \quad K_t f = f$$

for all  $t \geq 0$ .

*Proof.* Eq. (6.10) implies that for any bounded set  $\mathcal{D}$  with boundary of Lebesgue measure zero,

$$(6.20) \quad \nu(\{(\mathbf{Q}_0, \mathbf{V}_0) \in T^1(\rho^{d-1}\mathcal{K}_\rho) : \widehat{F}_t(\mathbf{Q}_0, \mathbf{V}_0) \in \mathcal{D}\}) \rightarrow \mathbb{P}_f(\widehat{\Xi}(t) \in \mathcal{D})$$

as  $\rho \rightarrow 0$ , with  $f$  as assumed above. The  $F_t$ -invariance of  $\nu$  implies that the left hand side of (6.20) equals

$$(6.21) \quad \nu(\{(\mathbf{Q}_0, \mathbf{V}_0) \in T^1(\rho^{d-1}\mathcal{K}_\rho) : (\mathbf{Q}_0, \mathbf{V}_0, \mathcal{T}(\mathbf{Q}_0, \mathbf{V}_0), \mathbf{V}_+(\mathbf{Q}_0, \mathbf{V}_0)) \in \mathcal{D}\}),$$

for all  $t \geq 0$ , and hence

$$(6.22) \quad \mathbb{P}_f(\widehat{\Xi}(t) \in \mathcal{D}) = \mathbb{P}_f(\widehat{\Xi}(0) \in \mathcal{D}) = \nu_f^{(1)}(\mathcal{D}) = \int_{\mathcal{D}} f(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) d\sigma_1(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+).$$

Hence by (6.16),  $K_t f = f \in L^1_{\text{loc}}(\mathcal{X}, \sigma_1)$ .  $\square$

Using (6.16) we can write, more explicitly,

$$(6.23) \quad K_t = \sum_{n=0}^{\infty} K_t^{(n)},$$

where

$$(6.24) \quad \int_{\mathcal{D}} [K_t^{(n)} f](\mathbf{Q}, \mathbf{v}_n, \xi_{n+1}, \mathbf{v}_{n+1}) d\text{vol}_{\mathbb{R}^d}(\mathbf{Q}) d\text{vol}_{S_1^{d-1}}(\mathbf{v}_n) d\xi_{n+1} d\text{vol}_{S_1^{d-1}}(\mathbf{v}_{n+1}) \\ = \nu_f^{(n+1)}(\{(\mathbf{Q}_0, \mathbf{v}_0, \xi_1, \mathbf{v}_1, \dots, \xi_{n+1}, \mathbf{v}_{n+1}) : \widehat{\Xi}_n(t) \in \mathcal{D}, T_n \leq t\}).$$

So in the case  $n = 0$ ,

$$(6.25) \quad [K_t^{(0)} f](\mathbf{Q}, \mathbf{v}_0, \xi, \mathbf{v}_1) = f(\mathbf{Q} - t\mathbf{v}_0, \mathbf{v}_0, \xi + t, \mathbf{v}_1),$$

and for  $n \geq 1$ ,

$$(6.26) \quad [K_t^{(n)} f](\mathbf{Q}, \mathbf{v}_n, \xi, \mathbf{v}_{n+1}) = \int_{T_n \leq t} f\left(\mathbf{Q} - \left(\sum_{j=1}^n \xi_j \mathbf{v}_{j-1} + (t - T_n) \mathbf{v}_n\right), \mathbf{v}_0, \xi_1, \mathbf{v}_1\right) \\ \times \prod_{j=1}^n p_{\mathbf{0}, \beta_{\mathbf{v}_{j-1}}^+}(\mathbf{v}_j, \xi_{j+1}, \mathbf{v}_{j+1}) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_0) d\xi_1 \cdots d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_{n-1}) d\xi_n$$

with the shorthand  $\xi_{n+1} := \xi + t - T_n$ . We remark that when restricting  $K_t$  and  $K_t^{(n)}$  to  $L^1(\mathcal{X}, \sigma_1)$ , the right hand side of (6.26) can be estimated from above as in the proof of Lemma 5.2, yielding

$$(6.27) \quad \|K_t^{(n)}\|_{L^1} \leq \frac{t^{n-1} \text{vol}(\mathcal{B}_1^{d-1})^{n-1}}{(n-1)!} \quad \text{for } n \geq 1,$$

and hence the sum (6.23) is uniformly operator convergent on  $L^1(\mathcal{X}, \sigma_1)$ .

The following proposition implies  $\widehat{\Xi}(t)$  is Markovian.

**Proposition 6.3.** *The family  $\{K_t : t \geq 0\}$  forms a semigroup on  $L_{\text{loc}}^1(\mathcal{X}, \sigma_1)$ , and a contraction semigroup on  $L^1(\mathcal{X}, \sigma_1)$ .*

*Proof.* Note that, for  $f \in L_{\text{loc}}^1(\mathcal{X}, \sigma_1)$ ,  $0 \leq s \leq t$ ,  $0 \leq m < n$ ,

$$(6.28) \quad [K_{t-s}^{(n-m)} K_s^{(m)} f](\mathbf{Q}, \mathbf{v}_n, \xi, \mathbf{v}_{n+1}) = \int_{\substack{T_n \leq t \\ T_m \leq s < T_{m+1}}} f\left(\mathbf{Q} - \left(\sum_{j=1}^n \xi_j \mathbf{v}_{j-1} + (t - T_n) \mathbf{v}_n\right), \mathbf{v}_0, \xi_1, \mathbf{v}_1\right) \\ \times \prod_{j=1}^n p_{\mathbf{0}, \beta_{\mathbf{v}_{j-1}}^+}(\mathbf{v}_j, \xi_{j+1}, \mathbf{v}_{j+1}) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_0) d\xi_1 \cdots d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_{n-1}) d\xi_n,$$

and for  $m = n > 0$ ,

$$(6.29) \quad [K_{t-s}^{(0)} K_s^{(n)} f](\mathbf{Q}, \mathbf{v}_n, \xi, \mathbf{v}_{n+1}) = \int_{T_n \leq s} f\left(\mathbf{Q} - \left(\sum_{j=1}^n \xi_j \mathbf{v}_{j-1} + (t - T_n) \mathbf{v}_n\right), \mathbf{v}_0, \xi_1, \mathbf{v}_1\right) \\ \times \prod_{j=1}^n p_{\mathbf{0}, \beta_{\mathbf{v}_{j-1}}^+}(\mathbf{v}_j, \xi_{j+1}, \mathbf{v}_{j+1}) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_0) d\xi_1 \cdots d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_{n-1}) d\xi_n.$$

Therefore

$$(6.30) \quad \sum_{m=0}^n K_{t-s}^{(n-m)} K_s^{(m)} = K_t^{(n)}$$

and thus

$$(6.31) \quad K_{t-s} K_s = \sum_{m,n=0}^{\infty} K_{t-s}^{(m)} K_s^{(n)} = \sum_{n=0}^{\infty} \sum_{m=0}^n K_{t-s}^{(n-m)} K_s^{(m)} = K_t.$$

This proves the semigroup property.

We now consider the action restricted to  $L^1(\mathcal{X}, \sigma_1)$ . Since we have already noted that  $\|K_t\|_{L^1} \leq 1$  for all  $t$ , it only remains to prove that for any given  $f \in L^1(\mathcal{X}, \sigma_1)$  the map  $\mathbb{R}_{\geq 0} \ni t \mapsto K_t f \in L^1(\mathcal{X}, \sigma_1)$  is continuous. In view of (6.23) and (6.27) it suffices to prove that the map  $t \mapsto K_t^{(n)} f$  is continuous, for each  $n \geq 0$ . This is clear for  $n = 0$ , thus we now assume  $n \geq 1$ . Given any  $s \geq 0$  and  $h > 0$  we split the integral (6.26) for  $t = s + h$  as

$$(6.32) \quad K_{s+h}^{(n)} f = I_1 + I_2,$$

where  $I_1$  corresponds to  $\xi_1 < h$  and  $I_2$  corresponds to  $\xi_1 > h$ . Repeated use of Remark 4.4 gives

$$(6.33) \quad \|I_1\|_{L^1} \leq \int_{\mathcal{X} \cap \{\xi < h\}} |f(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)| d\sigma_1(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+).$$

Furthermore, (6.28) with  $m = 0$  gives  $I_2 = K_s^{(n)}[K_h^{(0)}f]$ , and hence

$$(6.34) \quad \|K_{s+h}^{(n)}f - K_s^{(n)}f\| \leq \|I_1\| + \|K_s^{(n)}([K_h^{(0)}f] - f)\| \leq \|I_1\| + \|[K_h^{(0)}f] - f\| \rightarrow 0$$

as  $h \rightarrow 0^+$ , uniformly with respect to  $s \geq 0$ . This proves the desired continuity.  $\square$

**6.3. The Fokker-Planck-Kolmogorov equation.** We will now derive the Fokker-Planck-Kolmogorov equation of the Markov process  $\widehat{\Xi}(t)$ .

We introduce convenient spaces of continuous functions on  $\mathcal{X}$  and  $\mathbb{R}_{\geq 0} \times \mathcal{X}$  as follows. Recall the definition of  $J(\mathbf{v}_1, \mathbf{v}_2)$  in Remark 2.5; we consider  $J$  as a function  $\mathcal{X} \rightarrow \mathbb{R}_{> 0}$  by letting  $J(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) := J(\mathbf{V}, \mathbf{V}_+)$ . We also write  $\mathbf{Q} = (Q_1, \dots, Q_d) \in \mathbb{R}^d$ . Now set

$$(6.35) \quad \begin{aligned} \|f\|_J &:= \sup_{\mathcal{X}} |f|/J && \text{(for } f : \mathcal{X} \rightarrow \mathbb{R}\text{);} \\ C_J(\mathcal{X}) &:= \{f \in C(\mathcal{X}) : \|f\|_J < \infty\}; \\ C_J^1(\mathcal{X}) &:= \{f \in C_J(\mathcal{X}) : \partial_{Q_1}f, \partial_{Q_2}f, \dots, \partial_{Q_d}f, \partial_\xi f \in C_J(\mathcal{X})\}; \\ C_J(\mathbb{R}_{\geq 0} \times \mathcal{X}) &:= \{f \in C(\mathbb{R}_{\geq 0} \times \mathcal{X}) : \sup_{t \in [0, T]} \|f(t, \cdot)\|_J < \infty, \forall T > 0\}; \\ C_J^1(\mathbb{R}_{\geq 0} \times \mathcal{X}) &:= \{f \in C_J(\mathbb{R}_{\geq 0} \times \mathcal{X}) : \partial_t f, \partial_{Q_1}f, \partial_{Q_2}f, \dots, \partial_{Q_d}f, \partial_\xi f \\ &\quad \in C_J(\mathbb{R}_{\geq 0} \times \mathcal{X})\}. \end{aligned}$$

If  $d = 2$  then we impose from now on the following extra assumption on the scattering map:

$$(6.36) \quad \vartheta'_1(\varphi) \neq \vartheta'_2(\varphi) \text{ for (Lebesgue-)almost every } \varphi \in (-\frac{\pi}{2}, \frac{\pi}{2}),$$

where  $\vartheta_1, \vartheta_2$  are defined via (2.21). This assumption is always fulfilled in the case when the flow  $F_t$  preserves the Liouville measure (cf. Remark 2.3).

**Theorem 6.4.** *For any  $f_0 \in C_J^1(\mathcal{X})$ , the function  $f(t, \mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) := K_t f_0(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)$  belongs to  $C_J^1(\mathbb{R}_{\geq 0} \times \mathcal{X})$  and is the unique solution in  $C_J^1(\mathbb{R}_{\geq 0} \times \mathcal{X})$  of the differential equation*

$$(6.37) \quad \begin{aligned} &\left[ \partial_t + \mathbf{V} \cdot \nabla_{\mathbf{Q}} - \partial_\xi \right] f(t, \mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) \\ &= \int_{S_1^{d-1}} f(t, \mathbf{Q}, \mathbf{v}_0, 0, \mathbf{V}) p_{\mathbf{0}, \beta_{\mathbf{v}_0}^+}(\mathbf{V}, \xi, \mathbf{V}_+) d\text{vol}_{S_1^{d-1}}(\mathbf{v}_0) \end{aligned}$$

with  $f(0, \mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) \equiv f_0(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)$ .

*Remark 6.1.* The equation (6.37) can also be expressed as

$$(6.38) \quad \partial_t f(t, \mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) = [Zf(t, \cdot)](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)$$

where the operator  $Z$  (acting on functions on  $\mathcal{X}$ ) is defined by

$$(6.39) \quad \begin{aligned} [Zg](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) &= [\partial_\xi - \mathbf{V} \cdot \nabla_{\mathbf{Q}}]g(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) \\ &\quad + \int_{S_1^{d-1}} g(\mathbf{Q}, \mathbf{v}_0, 0, \mathbf{V}) p_{\mathbf{0}, \beta_{\mathbf{v}_0}^+}(\mathbf{V}, \xi, \mathbf{V}_+) d\text{vol}_{S_1^{d-1}}(\mathbf{v}_0). \end{aligned}$$

Thus on a formal level we have  $K_t = e^{tZ}$ .

To prepare for the proof of Theorem 6.4 we first establish a series of lemmas. For any  $\mathbf{v} \in S_1^{d-1}$  and  $\delta > 0$  we write  $\mathcal{N}_\delta(\mathbf{v})$  for the closed  $\delta$ -neighborhood of  $\mathbf{v}$  in  $S_1^{d-1}$ , i.e.

$$(6.40) \quad \mathcal{N}_\delta(\mathbf{v}) = \{\mathbf{w} \in S_1^{d-1} : \varphi(\mathbf{w}, \mathbf{v}) \leq \delta\}.$$

Given  $n \geq 1$  and  $\mathbf{v}_n \in \mathbb{S}_1^{d-1}$  we write

$$(6.41) \quad \mathcal{V}_{\mathbf{v}_n}^{[n]} := \{(\mathbf{v}_0, \dots, \mathbf{v}_{n-1}) \in (\mathbb{S}_1^{d-1})^n : \varphi(\mathbf{v}_j, \mathbf{v}_{j+1}) > B_\Theta, j = 0, \dots, n-1\}.$$

The point of this is that the integrand in (6.26) vanishes for all  $(\mathbf{v}_0, \dots, \mathbf{v}_{n-1})$  outside  $\mathcal{V}_{\mathbf{v}_n}^{[n]}$ .

**Lemma 6.5.** *Given  $n \geq 1$ ,  $\varepsilon > 0$  and  $\mathbf{v}_n^0, \mathbf{v}_{n+1}^0 \in \mathbb{S}_1^{d-1}$  with  $\varphi(\mathbf{v}_n^0, \mathbf{v}_{n+1}^0) > B_\Theta$ , there exist some  $\delta > 0$  and a compact subset  $D \subset \cap_{\mathbf{v}_n \in \mathcal{N}_\delta(\mathbf{v}_n^0)} \mathcal{V}_{\mathbf{v}_n}^{[n]}$  such that for all  $\mathbf{v}_n \in \mathcal{N}_\delta(\mathbf{v}_n^0)$ ,  $\mathbf{v}_{n+1} \in \mathcal{N}_\delta(\mathbf{v}_{n+1}^0)$ , the following holds:*

- (i)  $\varphi(\mathbf{v}_n, \mathbf{v}_{n+1}) > B_\Theta$ ;
- (ii)  $\int_{\mathcal{V}_{\mathbf{v}_n}^{[n]} \setminus D} \prod_{j=0}^{n-1} J(\mathbf{v}_j, \mathbf{v}_{j+1}) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_0) \cdots d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_{n-1}) < \varepsilon$ ;
- (iii) if  $d = 2$  then  $\beta_{\mathbf{e}_1}^-(\mathbf{v}_{j+1}K(\mathbf{v}_j))_\perp \neq (\beta_{\mathbf{v}_{j-1}}^+(\mathbf{v}_j)K(\mathbf{v}_j))_\perp$  for all  $j \in \{1, \dots, n\}$  and all  $(\mathbf{v}_0, \dots, \mathbf{v}_{n-1}) \in D$ .

*Proof of Lemma 6.5 if  $d \geq 3$ .* Note  $\int_{\mathcal{V}_{\mathbf{v}}} J(\mathbf{v}', \mathbf{v}) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}') = \text{vol}(\mathcal{B}_1^{d-1})$  (cf. (2.25)–(2.28)); thus the integral in (ii) with  $D = \emptyset$  is absolutely convergent. Note also that if  $K \in \text{SO}(d)$  is any rotation with  $\mathbf{v}_n K = \mathbf{v}_n^0$ , the left hand side in (ii) can be rewritten as

$$(6.42) \quad J(\mathbf{v}_n, \mathbf{v}_{n+1}) \int_{\mathcal{V}_{\mathbf{v}_n^0}^{[n]} \setminus DK} \left( \prod_{j=0}^{n-2} J(\mathbf{v}_j, \mathbf{v}_{j+1}) \right) J(\mathbf{v}_{n-1}, \mathbf{v}_n^0) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_0) \cdots d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_{n-1}).$$

Now we may fix a compact subset  $D \subset \mathcal{V}_{\mathbf{v}_n^0}^{[n]}$  such that (ii) holds when  $\mathbf{v}_n = \mathbf{v}_n^0$  and  $\mathbf{v}_{n+1} = \mathbf{v}_{n+1}^0$ , and then by continuity we may choose  $\delta > 0$  so small that (i) and (ii) hold for all  $\mathbf{v}_n \in \mathcal{N}_\delta(\mathbf{v}_n^0)$ ,  $\mathbf{v}_{n+1} \in \mathcal{N}_\delta(\mathbf{v}_{n+1}^0)$ .  $\square$

*Proof of Lemma 6.5 if  $d = 2$ .* First take  $\delta$  and  $D$  as in the previous proof; it then suffices to show that we can make (iii) hold by removing an open subset of arbitrarily small  $(\text{vol}_{\mathbb{S}_1^2}^n -)$  volume from  $D$  and possibly shrinking  $\delta$ .

First consider  $j = n$  in (iii); to deal with this case it suffices to prove

$$(6.43) \quad \lim_{\delta \rightarrow 0} \text{vol}_{\mathbb{S}_1^1} \left( \bigcup_{(\mathbf{v}_n, \mathbf{v}_{n+1}) \in \mathcal{N}_\delta(\mathbf{v}_n^0) \times \mathcal{N}_\delta(\mathbf{v}_{n+1}^0)} S_{\mathbf{v}_n, \mathbf{v}_{n+1}} \right) = 0,$$

where

$$(6.44) \quad S_{\mathbf{v}_n, \mathbf{v}_{n+1}} := \{ \mathbf{v}_{n-1} \in \mathcal{V}_{\mathbf{v}_n} : \beta_{\mathbf{e}_1}^-(\mathbf{v}_{n+1}K(\mathbf{v}_n))_\perp = (\beta_{\mathbf{v}_{n-1}}^+(\mathbf{v}_n)K(\mathbf{v}_n))_\perp \}.$$

Set  $K_\theta := \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$ ; then (6.43) will follow if we can prove

$$(6.45) \quad \lim_{\delta \rightarrow 0} \text{vol}_{\mathbb{S}_1^1} \left( \bigcup_{|\theta| \leq \delta} \left( \bigcup_{\mathbf{v}_{n+1} \in \mathcal{N}_{2\delta}(\mathbf{v}_{n+1}^0)} S_{\mathbf{v}_n^0, \mathbf{v}_{n+1}} \right) K_\theta \right) = 0.$$

Here the inner union equals (cf. Remark 4.1)

$$(6.46) \quad \{ \mathbf{v}_{n-1} \in \mathcal{V}_{\mathbf{v}_n^0} : \Theta_1(\mathbf{e}_1, -\beta_{\mathbf{v}_{n-1}K(\mathbf{v}_n^0)}^+(\mathbf{e}_1)R_{\mathbf{e}_1}) \in \mathcal{N}_{2\delta}(\mathbf{v}_{n+1}^0K(\mathbf{v}_n^0)) \},$$

where  $R_{\mathbf{e}_1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$  is the reflection in the line  $\mathbb{R}\mathbf{e}_1$ . Since (6.46) is a closed subset of  $\mathcal{V}_{\mathbf{v}_n^0}$  for every  $\delta > 0$ , it suffices to prove that the volume of (6.46) tends to zero as  $\delta \rightarrow 0$ ; and since  $\Theta_1(\mathbf{e}_1, \cdot)$  is a  $C^1$  diffeomorphism, this will follow if we can show

$$(6.47) \quad \lim_{\delta \rightarrow 0} \text{vol}_{\mathbb{S}_1^1} (\{ \mathbf{V} \in \mathcal{V}_{\mathbf{e}_1} : \beta_{\mathbf{V}}^+(\mathbf{e}_1) \in \mathcal{N}_\delta(\mathbf{W}) \}) = 0, \quad \forall \mathbf{W} \in \mathbb{S}_1^{d-1}.$$

But for  $\mathbf{V} = -(\cos \varphi)\mathbf{e}_1 + (\sin \varphi)\mathbf{e}_2$  with  $|\varphi| < \pi - B_\Theta$  we have, by a computation,

$$(6.48) \quad \beta_{\mathbf{V}}^+(\mathbf{e}_1) = (\cos \nu(\varphi))\mathbf{e}_1 + (\sin \nu(\varphi))\mathbf{e}_2 \quad \text{with} \quad \nu(\varphi) = \vartheta_2(\vartheta_1^{-1}(\varphi)) - \varphi,$$

where  $\vartheta_1^{-1}$  and  $\vartheta_2$  are as in Remark 2.3 and Remark 2.5 (thus the function  $\nu(\varphi)$  is  $C^1$  for all  $|\varphi| < \pi - B_\Theta$ ). Now, by (6.36), the set  $\{\nu'(\varphi) = 0\}$  has Lebesgue measure zero. This implies (6.47), and hence also (6.43).

Next, to deal with the case  $j = n - 1$  (thus  $n \geq 2$ ) in (iii), it suffices to prove

$$(6.49) \quad \lim_{\delta \rightarrow 0} \text{vol}_{\mathbb{S}_1^2}^2 \left( \bigcup_{\mathbf{v}_n \in \mathcal{N}_\delta(\mathbf{v}_n^0)} S_{\mathbf{v}_n} \right) = 0,$$

where

$$(6.50) \quad S_{\mathbf{v}_n} := \{(\mathbf{v}_{n-2}, \mathbf{v}_{n-1}) \in \mathcal{V}_{\mathbf{v}_n}^{[2]} : \beta_{\mathbf{e}_1}^-(\mathbf{v}_n K(\mathbf{v}_{n-1}))_\perp = (\beta_{\mathbf{v}_{n-2}}^+(\mathbf{v}_{n-1}) K(\mathbf{v}_{n-1}))_\perp\}.$$

However, (6.49) follows from the fact that  $S_{\mathbf{v}_n^0}$  is a closed subset of  $\mathcal{V}_{\mathbf{v}_n^0}^{[2]}$  with  $\text{vol}_{\mathbb{S}_1^2}^2(\mathcal{V}_{\mathbf{v}_n^0}^{[2]}) = 0$ , using rotational invariance in a similar way as in (6.45).

Finally, the case  $j \leq n - 2$  is easy, since

$$(6.51) \quad \text{vol}_{\mathbb{S}_1^3}^3(\{(\mathbf{v}_{j-1}, \mathbf{v}_j, \mathbf{v}_{j+1}) \in (\mathbb{S}_1^1)^3 : (\mathbf{v}_{j-1}, \mathbf{v}_j) \in S_{\mathbf{v}_{j+1}}\}) = 0,$$

with  $S_{\mathbf{v}_{j+1}}$  defined in analogy with (6.50).  $\square$

**Lemma 6.6.** *If  $f_0 \in C_J(\mathcal{X})$  then the function  $f(t, \mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) := K_t f_0(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)$  belongs to  $C_J(\mathbb{R}_{\geq 0} \times \mathcal{X})$ .*

*Proof.* From (6.26) we obtain, using Remarks 2.5, 4.3 and the fact that  $0 \leq \Phi_0 \leq 1$ ,

$$(6.52) \quad \begin{aligned} |[K_t^{(n)} f_0](\mathbf{Q}, \mathbf{v}_n, \xi, \mathbf{v}_{n+1})| &\leq \|f_0\|_J \frac{t^n}{n!} \int_{\mathcal{V}_{\mathbf{v}_n}^{[n]}} \prod_{j=0}^n J(\mathbf{v}_j, \mathbf{v}_{j+1}) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_0) \cdots d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_{n-1}) \\ &= \|f_0\|_J \frac{t^n}{n!} \text{vol}(\mathcal{B}_1^{d-1})^n J(\mathbf{v}_n, \mathbf{v}_{n+1}), \quad \forall n \geq 1. \end{aligned}$$

From (6.25) we see that the same bound is also true for  $n = 0$ . It follows that the sum  $[K_t f_0](\mathbf{Q}, \mathbf{v}_n, \xi, \mathbf{v}_{n+1}) = \sum_{n=0}^\infty [K_t^{(n)} f_0](\mathbf{Q}, \mathbf{v}_n, \xi, \mathbf{v}_{n+1})$  is uniformly absolutely convergent for  $t$  and  $(\mathbf{v}_n, \mathbf{v}_{n+1})$  in compacta (and  $\mathbf{Q}, \xi$  unrestricted), and we have  $\|K_t f_0\|_J \leq e^{t \text{vol}(\mathcal{B}_1^{d-1})} \|f_0\|_J$  for each  $t \geq 0$ . It now only remains to prove that  $f$  is continuous on  $\mathbb{R}_{\geq 0} \times \mathcal{X}$ , and for this it suffices to prove that each function  $K_t^{(n)} f_0$  is continuous on  $\mathbb{R}_{\geq 0} \times \mathcal{X}$ . The case  $n = 0$  is trivial, thus from now on we fix some  $n \geq 1$ .

Let  $(t^0, \mathbf{Q}^0, \mathbf{v}_n^0, \xi^0, \mathbf{v}_{n+1}^0) \in \mathbb{R}_{\geq 0} \times \mathcal{X}$  and  $\varepsilon > 0$  be given. Let  $\delta > 0$  and  $D \subset \mathcal{V}_{\mathbf{v}_n^0}^{[n]}$  be as in Lemma 6.5, with  $\|f_0\|_J^{-1} \frac{n!}{(t^0+1)^n} \varepsilon$  in the place of  $\varepsilon$ . By (6.26) and the bounds leading to (6.52) we then get that, for all  $(t, \mathbf{Q}, \mathbf{v}_n, \xi, \mathbf{v}_{n+1}) \in \mathbb{R}_{\geq 0} \times \mathcal{X}$  with  $t \leq t^0 + 1$ ,  $\mathbf{v}_n \in \mathcal{N}_\delta(\mathbf{v}_n^0)$  and  $\mathbf{v}_{n+1} \in \mathcal{N}_\delta(\mathbf{v}_{n+1}^0)$ , the value of  $[K_t^{(n)} f_0](\mathbf{Q}, \mathbf{v}_n, \xi, \mathbf{v}_{n+1})$  differs by at most  $\varepsilon$  from

$$(6.53) \quad \begin{aligned} K_{D,t}^{(n)} f_0(\mathbf{Q}, \mathbf{v}_n, \xi, \mathbf{v}_{n+1}) &:= \int_{T_n \leq t} \int_D f_0 \left( \mathbf{Q} - \left( \sum_{j=1}^n \xi_j \mathbf{v}_{j-1} + (t - T_n) \mathbf{v}_n \right), \mathbf{v}_0, \xi_1, \mathbf{v}_1 \right) \\ &\quad \times \prod_{j=1}^n p_{0, \beta_{\mathbf{v}_{j-1}}^+}(\mathbf{v}_j, \xi_{j+1}, \mathbf{v}_{j+1}) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_0) \cdots d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_{n-1}) d\xi_1 \cdots d\xi_n. \end{aligned}$$

By Remark 4.1 (using Lemma 6.5 (iii) if  $d = 2$ ), the integrand in (6.53) depends jointly continuously on all the variables  $(\xi_1, \xi_2, \dots, \xi_n) \in (\mathbb{R}_{\geq 0})^n$ ,  $(\mathbf{v}_0, \dots, \mathbf{v}_{n-1}) \in D$  and  $(t, \mathbf{Q}, \mathbf{v}_n, \xi, \mathbf{v}_{n+1}) \in \mathbb{R}_{\geq 0} \times \mathcal{X}$ , so long as  $T_n \leq t$ ,  $\mathbf{v}_n \in \mathcal{N}_\delta(\mathbf{v}_n^0)$  and  $\mathbf{v}_{n+1} \in \mathcal{N}_\delta(\mathbf{v}_{n+1}^0)$ . Hence since the domain of integration in (6.53) is compact, we have for all  $(t, \mathbf{Q}, \mathbf{v}_n, \xi, \mathbf{v}_{n+1}) \in \mathbb{R}_{\geq 0} \times \mathcal{X}$  sufficiently near  $(t^0, \mathbf{Q}^0, \mathbf{v}_n^0, \xi^0, \mathbf{v}_{n+1}^0)$ :

$$(6.54) \quad |K_{D,t}^{(n)} f_0(\mathbf{Q}, \mathbf{v}_n, \xi, \mathbf{v}_{n+1}) - K_{D,t^0}^{(n)} f_0(\mathbf{Q}^0, \mathbf{v}_n^0, \xi^0, \mathbf{v}_{n+1}^0)| < \varepsilon,$$

and thus

$$(6.55) \quad \left| [K_t^{(n)} f_0](\mathbf{Q}, \mathbf{v}_n, \xi, \mathbf{v}_{n+1}) - [K_{t^0}^{(n)} f_0](\mathbf{Q}^0, \mathbf{v}_n^0, \xi^0, \mathbf{v}_{n+1}^0) \right| < 3\varepsilon.$$

Hence  $[K_t^{(n)} f_0](\mathbf{Q}, \mathbf{v}_n, \xi, \mathbf{v}_{n+1})$  is indeed continuous, since  $\varepsilon > 0$  was arbitrary.  $\square$

**Lemma 6.7.** *If  $f_0 \in C_J^1(\mathcal{X})$  and  $f(t, \mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) := K_t f_0(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)$ , then the derivatives  $\partial_{Q_j} f$  ( $j = 1, \dots, d$ ) and  $\partial_\xi f$  all exist and belong to  $C_J(\mathbb{R}_{\geq 0} \times \mathcal{X})$ .*

*Proof.* We start by considering  $\partial_\xi K_t^{(n)} f_0$ . First assume  $n \geq 2$ . To differentiate (6.26) with respect to  $\xi$  we first move the integral over  $\xi_1$  to the innermost position; this integral will then appear as  $\int_0^{t-T_{2,n}} \dots d\xi_1$  where  $T_{2,n} := \sum_{j=2}^n \xi_j$ , and it may be differentiated with respect to  $\xi$  by using the differentiability assumption on  $f_0$  and the fact that  $p_{\mathbf{0}, \beta_{\mathbf{v}_{n-1}}^+}(\mathbf{v}_n, \cdot, \mathbf{v}_{n+1})$  depends continuously on its second argument (except when  $d = 2$  and  $\beta_{\mathbf{e}_1}^-(\mathbf{v}_{n+1} K(\mathbf{v}_n))_\perp = (\beta_{\mathbf{v}_{n-1}}^+(\mathbf{v}_n) K(\mathbf{v}_n))_\perp$ ). Hence we obtain, at least formally:

$$\begin{aligned}
& [\partial_\xi K_t^{(n)} f_0](\mathbf{Q}, \mathbf{v}_n, \xi, \mathbf{v}_{n+1}) \\
&= \int_{\mathcal{V}_{\mathbf{v}_n}^{[n]}} \int_{\substack{\xi_2, \dots, \xi_n \geq 0 \\ T_{2,n} \leq t}} f_0 \left( \mathbf{Q} - \sum_{j=2}^n \xi_j \mathbf{v}_{j-1} - (t - T_{2,n}) \mathbf{v}_n, \mathbf{v}_0, 0, \mathbf{v}_1 \right) \prod_{j=1}^{n-1} p_{\mathbf{0}, \beta_{\mathbf{v}_{j-1}}^+}(\mathbf{v}_j, \xi_{j+1}, \mathbf{v}_{j+1}) \\
&\quad \times p_{\mathbf{0}, \beta_{\mathbf{v}_{n-1}}^+}(\mathbf{v}_n, \xi + t - T_{2,n}, \mathbf{v}_{n+1}) d\xi_2 \dots d\xi_n d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_0) \dots d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_{n-1}) \\
&- \int_{\mathcal{V}_{\mathbf{v}_n}^{[n]}} \int_{\substack{\xi_2, \dots, \xi_n \geq 0 \\ T_{2,n} \leq t}} f_0 \left( \mathbf{Q} - (t - T_{2,n}) \mathbf{v}_0 - \sum_{j=2}^n \xi_j \mathbf{v}_{j-1}, \mathbf{v}_0, t - T_{2,n}, \mathbf{v}_1 \right) \prod_{j=1}^{n-1} p_{\mathbf{0}, \beta_{\mathbf{v}_{j-1}}^+}(\mathbf{v}_j, \xi_{j+1}, \mathbf{v}_{j+1}) \\
(6.56) \quad &\quad \times p_{\mathbf{0}, \beta_{\mathbf{v}_{n-1}}^+}(\mathbf{v}_n, \xi, \mathbf{v}_{n+1}) d\xi_2 \dots d\xi_n d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_0) \dots d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_{n-1}) \\
&+ \int_{\mathcal{V}_{\mathbf{v}_n}^{[n]}} \int_{\substack{\xi_1, \dots, \xi_n \geq 0 \\ T_n \leq t}} \left[ (\mathbf{v}_n - \mathbf{v}_0) \cdot \nabla_{\mathbf{Q}} f_0 + \partial_\xi f_0 \right] \left( \mathbf{Q} - \left( \sum_{j=1}^n \xi_j \mathbf{v}_{j-1} + (t - T_n) \mathbf{v}_n \right), \mathbf{v}_0, \xi_1, \mathbf{v}_1 \right) \\
&\quad \times \prod_{j=1}^n p_{\mathbf{0}, \beta_{\mathbf{v}_{j-1}}^+}(\mathbf{v}_j, \xi_{j+1}, \mathbf{v}_{j+1}) d\xi_1 \dots d\xi_n d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_0) \dots d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_{n-1}).
\end{aligned}$$

Here  $\partial_\xi f_0$  and all components of  $\nabla_{\mathbf{Q}} f_0$  lie in  $C_J(\mathcal{X})$ , since  $f_0 \in C_J^1(\mathcal{X})$ ; hence as in Lemma 6.6 one shows that the right hand side of (6.56) is a continuous function of  $(t, \mathbf{Q}, \mathbf{v}_n, \xi, \mathbf{v}_{n+1}) \in \mathbb{R}_{\geq 0} \times \mathcal{X}$ . Now to validate the formula (6.56) it suffices to prove that (6.56) holds true after integration with respect to  $\xi$  over an arbitrary finite interval  $[a, b] \subset \mathbb{R}_{\geq 0}$ , and this is easily verified using Fubini's theorem. The case  $n \geq 1$  is very similar, and the case  $n = 0$  is trivial.

From (6.56) and its analogues for  $n = 1, 0$ , one obtains a pointwise bound on  $\partial_\xi K_t^{(n)} f$  similar to (6.52), involving  $\|f\|_J$ ,  $\|\partial_\xi f\|_J$  and  $\|\partial_{Q_j} f\|_J$  for  $j = 1, \dots, d$ . This bound immediately implies that  $\partial_\xi f = \sum_{n=0}^\infty \partial_\xi K_t^{(n)} f_0$ , with uniform absolute convergence on compact subsets of  $\mathcal{X}$ , and  $\partial_\xi f \in C_J(\mathbb{R}_{\geq 0} \times \mathcal{X})$ , as desired.

The proof of  $\partial_{Q_j} f \in C_J(\mathbb{R}_{\geq 0} \times \mathcal{X})$  follows the same steps but with simpler formulas; in fact  $\partial_{Q_j} K_t^{(n)} f_0 = K_t^{(n)} \partial_{Q_j} f_0$  holds for each  $n \geq 0$ .  $\square$

**Lemma 6.8.** *If  $f_0 \in C_J^1(\mathcal{X})$  and  $f(t, \mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) := K_t f_0(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)$ , then also  $\partial_t f$  exists and belongs to  $C_J(\mathbb{R}_{\geq 0} \times \mathcal{X})$ , and the equation (6.37) holds throughout  $\mathbb{R}_{\geq 0} \times \mathcal{X}$ .*

*Proof.* Fix any point  $(t, \mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) \in \mathbb{R}_{\geq 0} \times \mathcal{X}$ . Using  $K_{t+h} = K_h K_t$  and the bound (6.52) (with  $h$  in place of  $t$ ) added over all  $n \geq 2$ , we get as  $h \rightarrow 0^+$ :

$$\begin{aligned}
(6.57) \quad f(t+h, \mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) &= \sum_{n=0}^\infty [K_h^{(n)} K_t f_0](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) \\
&= f(t, \mathbf{Q} - h\mathbf{V}, \mathbf{V}, \xi + h, \mathbf{V}_+) + \int_0^h \int_{\mathcal{V}_{\mathbf{v}_0}} f(t, \mathbf{Q} - (\xi_1 \mathbf{v}_0 + (h - \xi_1) \mathbf{V}), \mathbf{v}_0, \xi_1, \mathbf{V}) \\
&\quad \times p_{\mathbf{0}, \beta_{\mathbf{v}_0}^+}(\mathbf{V}, \xi + h - \xi_1, \mathbf{V}_+) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_0) d\xi_1 + O(h^2).
\end{aligned}$$



Using  $f(t, \cdot) \in C_J^1(\mathcal{X})$  (cf. Lemma 6.7), and treating the integral term with a continuity argument as in the proof of Lemma 6.6, we get (cf. (6.39))

$$(6.58) \quad = f(t, \mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) + h[Zf(t, \cdot)](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) + o(h).$$

A similar argument shows that the function  $[Zf(t, \cdot)](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)$  lies in  $C_J(\mathbb{R}_{\geq 0} \times \mathcal{X})$ . Letting  $h \rightarrow 0^+$  in (6.58) we conclude that the right derivative  $\partial_t^+ f(t, \mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)$  exists and equals  $[Zf(t, \cdot)](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)$ . Since the latter function is continuous, the relation now follows for the two-sided derivative, i.e. (6.38) ( $\Leftrightarrow$  (6.37)) holds, and  $\partial_t f \in C_J(\mathbb{R}_{\geq 0} \times \mathcal{X})$ .  $\square$

**Lemma 6.9.** *Let  $\{f_h\}_{h>0}$  be a family of functions in  $C_J(\mathcal{X})$  satisfying  $\limsup_{h \rightarrow 0} \|f_h\|_J < \infty$  and  $f_h(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) \rightarrow 0$  as  $h \rightarrow 0$ , uniformly over  $(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)$  in compact subsets of  $\mathcal{X}$ . Then for each  $(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) \in \mathcal{X}$  we have  $[K_t f_h](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) \rightarrow 0$  as  $h \rightarrow 0$ , uniformly over  $t$  in compact subsets of  $\mathbb{R}_{\geq 0}$ .*

*Proof.* Take  $C, h_0 > 0$  so that  $\|f_h\|_J \leq C$  for all  $h \in (0, h_0]$ . Given some  $(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) \in \mathcal{X}$  and  $T > 0, \eta > 0, \varepsilon > 0$ , after possibly shrinking  $h_0$  we may assume that  $|f_h(\mathbf{Q}', \mathbf{V}', \xi', \mathbf{V}_+)| < \varepsilon J(\mathbf{V}', \mathbf{V}_+)$  for all  $h \in (0, h_0]$  and all  $(\mathbf{Q}', \mathbf{V}', \xi', \mathbf{V}_+) \in \mathcal{X}$  with  $\|\mathbf{Q}' - \mathbf{Q}\| \leq T, \xi' \leq \xi + T$  and  $\mathbf{V}' \in \overline{\mathcal{V}_{\mathbf{V}_+}^\eta}$  (cf. (3.1)). Then for each  $n \geq 1, t \in [0, T]$  and  $h \in (0, h_0]$ , by mimicking (6.52) but splitting the integral over  $\mathcal{V}_{\mathbf{V}}^{[n]}$  according to the two cases  $\mathbf{v}_0 \in \mathcal{V}_{\mathbf{v}_1}^\eta$  and  $\mathbf{v}_0 \notin \mathcal{V}_{\mathbf{v}_1}^\eta$ , we obtain (with  $C_1 := \text{vol}(\mathcal{B}_1^{d-1})$ )

$$(6.59) \quad |[K_t^{(n)} f_h](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)| \leq \frac{C_1^n t^n}{n!} \left\{ \varepsilon + \frac{C}{C_1} \int_{\mathcal{V}_{\mathbf{e}_1} \setminus \mathcal{V}_{\mathbf{e}_1}^\eta} J(\mathbf{v}_0, \mathbf{e}_1) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_0) \right\} J(\mathbf{V}, \mathbf{V}_+).$$

This bound also holds for  $n = 0$  (without the  $\int_{\mathcal{V}_{\mathbf{e}_1} \setminus \mathcal{V}_{\mathbf{e}_1}^\eta}$ -term), so long as  $\eta < \varphi(\mathbf{V}, \mathbf{V}_+) - B_\Theta$ . The expression within the brackets in (6.59) can be made arbitrarily small by taking  $\varepsilon$  and  $\eta$  to be sufficiently small. Now the desired conclusion follows by adding (6.59) over  $n \geq 0$ .  $\square$

**Lemma 6.10.** *We have  $K_t Z f_0 = Z K_t f_0$ , for all  $f_0 \in C_J^1(\mathcal{X})$  and  $t \geq 0$ .*

*Proof.* Let  $f_0 \in C_J^1(\mathcal{X})$  be given. Then by Lemma 6.8 (cf. (6.38)) we have  $[\partial_t K_t f_0](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) \equiv [Z K_t f_0](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)$ , and this function belongs to  $C_J(\mathbb{R}_{\geq 0} \times \mathcal{X})$ . Hence

$$(6.60) \quad h^{-1}[K_h f_0 - f_0](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) = h^{-1} \int_0^h [Z K_t f_0](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) dt$$

for all  $h > 0$  and all  $(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) \in \mathcal{X}$ , and this expression tends to  $[Z f_0](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)$  as  $h \rightarrow 0^+$ , uniformly over  $(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)$  in compact subsets of  $\mathcal{X}$ . Using (6.60) we also get  $\sup_{0 < h \leq 1} \|h^{-1}[K_h f_0 - f_0]\|_J \leq \sup_{0 < t \leq 1} \|Z K_t f_0\|_J < \infty$ . Hence by Lemma 6.9 we have

$$(6.61) \quad \lim_{h \rightarrow 0^+} h^{-1}[K_t[K_h f_0 - f_0]](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) = [K_t Z f_0](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+),$$

for every  $(t, \mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) \in \mathbb{R}_{\geq 0} \times \mathcal{X}$ . Using  $K_t K_h f_0 = K_{t+h} f_0$  and (6.38) this implies

$$(6.62) \quad [K_t Z f_0](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) = [\partial_t^+ K_t f_0](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) = [Z K_t f_0](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+).$$

$\square$

*Proof of Theorem 6.4.* By Lemma 6.7 and Lemma 6.8, it only remains to prove the uniqueness. Thus assume that  $f \in C_J^1(\mathbb{R}_{\geq 0} \times \mathcal{X})$  satisfies (6.37); we then need to prove that  $f(t, \mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) = [K_t f(0, \cdot)](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)$  for all  $(t, \mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) \in \mathbb{R}_{\geq 0} \times \mathcal{X}$ .

Fix an arbitrary point  $(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) \in \mathcal{X}$  and some  $a > 0$ , and set

$$F(s) = [K_{a-s} f(s, \cdot)](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) \quad \text{for } s \in [0, a].$$

Then  $F(s)$  is continuous, by Lemma 6.9. We will prove that the right derivative  $D^+ F(s)$  vanishes for each  $s \in (0, a)$ . This will conclude the proof, since it implies  $F(0) = F(a)$ , which is the desired relation at  $(t, \mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) = (a, \mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)$ .

Thus fix  $s \in (0, a)$ . By Lemma 6.8 applied with  $f_0 = f(s, \cdot)$  we have

$$(6.63) \quad \lim_{h \rightarrow 0^+} \frac{[K_{a-s-h} - K_{a-s}][f(s, \cdot)](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)}{h} = -[ZK_{a-s}f(s, \cdot)](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+).$$

Next, using the identity

$$(6.64) \quad \frac{f(s+h, \mathbf{Q}', \mathbf{V}', \xi', \mathbf{V}'_+) - f(s, \mathbf{Q}', \mathbf{V}', \xi', \mathbf{V}'_+)}{h} = \frac{1}{h} \int_s^{s+h} \partial_t f(t, \mathbf{Q}', \mathbf{V}', \xi', \mathbf{V}'_+) dt$$

and  $\partial_t f \in C_J(\mathbb{R}_{\geq 0} \times \mathcal{X})$  we see that  $h^{-1}[f(s+h, \cdot) - f(s, \cdot)]$  has uniformly bounded  $\|\cdot\|_J$ -norm for  $h \in (0, a-s]$ , and approaches  $\partial_s f(s, \cdot)$  uniformly on compact subsets of  $\mathcal{X}$  as  $h \rightarrow 0^+$ . Hence, using Lemma 6.9 and the fact that the function  $\mathbb{R}_{\geq 0} \ni b \mapsto [K_b \partial_s f(s, \cdot)](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)$  is continuous (cf. Lemma 6.6), we obtain, at our fixed point  $(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) \in \mathcal{X}$ :

$$(6.65) \quad \lim_{h \rightarrow 0^+} \frac{K_{a-s-h}[f(s+h, \cdot) - f(s, \cdot)](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+)}{h} = [K_{a-s} \partial_s f(s, \cdot)](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+).$$

But we are assuming that  $f$  satisfies (6.38); thus  $[\partial_s f](s, \cdot) \equiv Zf(s, \cdot)$ . Hence, when adding (6.63) and (6.65) and using Lemma 6.10, we obtain  $D^+F(s) = 0$ , as desired.  $\square$

*Remark 6.2.* Proposition 6.2 and Theorem 6.4 imply that if the flow  $F_t$  preserves the Liouville measure, then

$$(6.66) \quad -\partial_\xi p(\mathbf{V}, \xi, \mathbf{V}_+) = \int_{\mathbb{S}_1^{d-1}} J(\mathbf{v}_0, \mathbf{V}) p_{\mathbf{0}, \beta_{\mathbf{v}_0}^+}(\mathbf{V}, \xi, \mathbf{V}_+) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{v}_0),$$

since  $p(\mathbf{v}_0, 0, \mathbf{V}) = J(\mathbf{v}_0, \mathbf{V})$ . Using (2.23) and (2.27) this equation can be reformulated as

$$(6.67) \quad -\partial_\xi \Phi(\xi, \mathbf{w}) = \int_{\{0\} \times \mathcal{B}^{d-1}} \Phi_0(\xi, \mathbf{w}, \mathbf{z}') dz',$$

with  $\Phi(\xi, \mathbf{w}) := \Phi_\alpha(\xi, \mathbf{w}, \mathbf{z})$  for  $\alpha \notin \mathbb{Q}^d$  as defined in Remark 2.1. For an alternative proof of Eq. (6.66) working directly from the definition of  $\Phi_\alpha$ , see [19, (8.32), (8.37)] and note  $\Phi_0(\xi, \mathbf{w}, \mathbf{z}') = \Phi_0(\xi, \mathbf{z}', \mathbf{w})$ .

*Remark 6.3.* Given any  $\bar{f}_0 \in L^1(\mathbb{T}^1(\mathbb{R}^d))$  we define  $f_0 \in L^1(\mathcal{X}, \sigma_1)$  by  $f_0(\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) = \bar{f}_0(\mathbf{Q}, \mathbf{V})p(\mathbf{V}, \xi, \mathbf{V}_+)$ . We then have the relation

$$(6.68) \quad [L_t \bar{f}_0](\mathbf{Q}, \mathbf{V}) = \int_0^\infty \int_{\mathbb{S}_1^{d-1}} [K_t f_0](\mathbf{Q}, \mathbf{V}, \xi, \mathbf{V}_+) d\text{vol}_{\mathbb{S}_1^{d-1}}(\mathbf{V}_+) d\xi,$$

where  $L_t$  is the propagator of the original stochastic process  $\Xi(t)$ , cf. Section 1.3. If we furthermore impose that  $\bar{f}_0$  is bounded continuous, with bounded continuous derivatives  $\partial_{Q_j} \bar{f}_0$  for  $j = 1, \dots, d$ , then  $f_0$  satisfies the assumption of Theorem 6.4, so that  $K_t f_0$  satisfies the Fokker-Planck-Kolmogorov equation, (6.37).

## REFERENCES

- [1] P. Balint and I.P. Toth, Exponential decay of correlations in multi-dimensional dispersing billiards, *Ann. Henri Poincaré* **9** (2008) 1309–1369.
- [2] P.M. Bleher, Statistical properties of two-dimensional periodic Lorentz gas with infinite horizon. *J. Statist. Phys.* **66** (1992), 315–373.
- [3] F.P. Boca and A. Zaharescu, The distribution of the free path lengths in the periodic two-dimensional Lorentz gas in the small-scatterer limit, *Commun. Math. Phys.* **269** (2007), 425–471.
- [4] C. Boldrighini, L.A. Bunimovich and Y.G. Sinai, On the Boltzmann equation for the Lorentz gas. *J. Statist. Phys.* **32** (1983), 477–501.
- [5] J. Bourgain, F. Golse and B. Wennberg, On the distribution of free path lengths for the periodic Lorentz gas. *Comm. Math. Phys.* **190** (1998), 491–508.
- [6] L.A. Bunimovich and Y.G. Sinai, Statistical properties of Lorentz gas with periodic configuration of scatterers. *Comm. Math. Phys.* **78** (1980/81), 479–497.
- [7] E. Caglioti and F. Golse, On the distribution of free path lengths for the periodic Lorentz gas. III. *Comm. Math. Phys.* **236** (2003), 199–221.

- [8] E. Caglioti and F. Golse, The Boltzmann-Grad limit of the periodic Lorentz gas in two space dimensions, *C. R. Math. Acad. Sci. Paris* **346** (2008) 477–482.
- [9] N.I. Chernov, Statistical properties of the periodic Lorentz gas. Multidimensional case. *J. Statist. Phys.* **74** (1994), 11–53.
- [10] P. Dahlqvist, The Lyapunov exponent in the Sinai billiard in the small scatterer limit. *Nonlinearity* **10** (1997), 159–173.
- [11] E.B. Davies, *One-parameter semigroups*, Academic Press Inc., 1980.
- [12] D. Dolgopyat, D. Szasz and T. Varju, Recurrence properties of the planar Lorentz process, *Duke Math. J.* **142** (2008) 241–281.
- [13] G. Gallavotti, Divergences and approach to equilibrium in the Lorentz and the Wind-tree-models, *Physical Review* **185** (1969), 308–322.
- [14] F. Golse, The periodic Lorentz gas in the Boltzmann-Grad limit. *Proc. ICM 2006, Madrid, Spain*, 183–201.
- [15] F. Golse, On the periodic Lorentz gas and the Lorentz kinetic equation, *Ann. Fac. Sci. Toulouse Math.* (6) **17** (2008) 735–749.
- [16] F. Golse and B. Wennberg, On the distribution of free path lengths for the periodic Lorentz gas. II. *M2AN Math. Model. Numer. Anal.* **34** (2000), 1151–1163.
- [17] M. W. Hirsch, *Differential Topology*, Springer-Verlag, New York, 1994.
- [18] H. Lorentz, Le mouvement des électrons dans les métaux, *Arch. Néerl.* **10** (1905), 336–371.
- [19] J. Marklof and A. Strömbergsson, The distribution of free path lengths in the periodic Lorentz gas and related lattice point problems, arXiv:0706.4395, to appear in *Annals of Math.*
- [20] J. Marklof and A. Strömbergsson, Kinetic transport in the two-dimensional periodic Lorentz gas. *Nonlinearity* **21** (2008) 1413–1422.
- [21] J. Marklof and A. Strömbergsson, The periodic Lorentz gas in the Boltzmann-Grad limit: Asymptotic estimates. In preparation.
- [22] I. Melbourne and M. Nicol, Almost sure invariance principle for nonuniformly hyperbolic systems. *Comm. Math. Phys.* **260** (2005), 131–146.
- [23] I. Melbourne and M. Nicol, A vector-valued almost sure invariance principle for hyperbolic dynamical systems, *Ann. Probab.* **37** (2009), 478–505.
- [24] R.G. Newton, *Scattering theory of waves and particles*. Second edition. *Texts and Monographs in Physics*. Springer-Verlag, New York-Berlin, 1982.
- [25] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, 1987.
- [26] D. Szasz and T. Varju, Limit laws and recurrence for the planar Lorentz process with infinite horizon, *J. Stat. Phys.* **129** (2007) 59–80.
- [27] H. Spohn, The Lorentz process converges to a random flight process, *Comm. Math. Phys.* **60** (1978), 277–290.

SCHOOL OF MATHEMATICS, UNIVERSITY OF BRISTOL, BRISTOL BS8 1TW, U.K.  
j.marklof@bristol.ac.uk

DEPARTMENT OF MATHEMATICS, BOX 480, UPPSALA UNIVERSITY, SE-75106 UPPSALA, SWEDEN  
astrombe@math.uu.se