

# On the free path length distribution for linear motion in an $n$ -dimensional box

Samuel Holmin

Pär Kurlberg

Daniel Månsso

February 22, 2017

We consider the distribution of free path lengths, or the distance between consecutive bounces of random particles, in an  $n$ -dimensional rectangular box. If each particle travels a distance  $R$ , then, as  $R \rightarrow \infty$  the free path lengths coincides with the distribution of the length of the intersection of a random line with the box (for a natural ensemble of random lines) and we give an explicit formula (piecewise real analytic) for the probability density function in dimension two and three.

In dimension two we also consider a closely related model where each particle is allowed to bounce  $N$  times, as  $N \rightarrow \infty$ , and give an explicit (again piecewise real analytic) formula for its probability density function.

Further, in both models we can recover the side lengths of the box from the location of the discontinuities of the probability density functions.

## 1. Introduction

We consider billiard dynamics on a rectangular domain, i.e., point shaped “balls” moving with linear motion with specular reflections at the boundary, and similarly for rectangular box shaped domains in three dimensions. We wish to determine the distribution of free path lengths of ensembles of trajectories defined by selecting a starting point and direction at random.

The question seems quite natural and interesting on its own, but we mention that it originated from the study of electromagnetic fields in “reverberation chambers” under the assumption of highly directional antennas [9]. Briefly, the connection is as follows (we refer to the forthcoming paper [5] for more details): given an ideal highly directional antenna and a highly transient signal, then the wave pulse dynamics is essentially the same as a point shaped billiard ball traveling inside a chamber, with specular reflection at the boundary. Signal loss is dominated by (linear) “spreading” of the electromagnetic field and by absorption occurring at each interaction (“bounce”) with the walls. The first simple model we use in this paper neglects absorption effects, and models signal loss from spreading by simply terminating the motion of the ball after it has travelled a certain

large distance. The second model only takes into account signal loss from absorption, and completely neglects spreading; here the motion is terminated after the ball has bounced a certain number of times.

We remark that the distribution of free path lengths is very well studied in the context of the Lorentz gas — here a point particle interacts with hard spherical obstacles, either placed randomly, or regularly on Euclidean lattices; recently quasicrystal configurations have also been studied (cf. [2–4, 7, 10, 11, 13, 15, 16].)

Let  $R > 0$  be large and let a rectangular  $n$ -dimensional box  $K \subseteq \mathbb{R}^n$  be given, where  $n \geq 2$ . We send off a large number  $M > 0$  of particles, each with a random initial position  $p^{(i)} \in K$  chosen with respect to a given probability measure  $\mu$  on  $K$ , and each with a uniformly random initial direction  $v^{(i)} \in \mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$ ,  $i = 1, \dots, M$ , for a total distance  $R$  each. Each particle travels along straight lines, changing direction precisely when it hits the boundary of the box, where it reflects specularly. We record the distance travelled between each pair of consecutive bounces for each particle. (Note in particular that we obtain more bounce lengths from some particles than from others.) Let  $X_{M,R}$  be the uniformly distributed random variable on this finite set of bounce lengths of all the particles. More precisely, a random sample of  $X_{M,R}$  is obtained as follows: first take a random i.i.d. sample of points (with respect to the measure  $\mu$ )  $p^{(1)}, \dots, p^{(M)} \in K$ , and a random sample of directions  $v^{(1)}, \dots, v^{(M)} \in \mathbb{S}^{n-1}$  (with respect to the uniform measure). Each pair  $(p^{(i)}, v^{(i)})$  then defines a trajectory  $T^i$  of length  $R$ , and each such trajectory gives rise to a finite multiset  $B^i$  of lengths between consecutive bounces. Finally, with  $B = \bigcup_{i=1}^M B^i$  denoting the (multiset) union of bounce length multisets  $B^1, \dots, B^M$ , we select an element of  $B$  with the uniform distribution. (That is, with  $1_B$  denoting the integer valued set indicator function for  $B$ , and  $B' = \{x : 1_B(x) \geq 1\}$  we select the element  $b \in B'$  with probability  $1_B(b)/\sum_{x \in B'} 1_B(x)$ .)

We are interested in the distribution of  $X_{M,R}$  for large  $M$  and  $R$ , and this turns out to be closely related to a model arising from integral geometry. Namely, let  $d\ell$  denote the unique (up to a constant) translation- and rotation-invariant measure on the set of directed lines  $\ell$  in  $\mathbb{R}^n$ , and consider the restriction of this measure to the set of directed lines  $\ell$  intersecting  $K$ , normalized such that it becomes a probability measure. Denote by  $X$  the random variable  $X := \text{length}(\ell \cap K)$  where  $\ell$  is chosen at random using this measure.

**Theorem 1.** *For any dimension  $n \geq 2$ , and for any distribution  $\mu$  on the starting points, the random variable  $X_{M,R}$  converges in distribution to the random variable  $X$ , as we take  $R \rightarrow \infty$  followed by taking  $M \rightarrow \infty$ , or vice versa.*

The mean free path length has a quite simple geometric interpretation. We have

$$\mathbb{E}[X] = 2\pi \frac{|\mathbb{S}^{n-1}|}{|\mathbb{S}^n|} \frac{\text{Vol}(K)}{\text{Area}(K)} = 2\sqrt{\pi} \cdot \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \frac{\text{Vol}(K)}{\text{Area}(K)} \quad (2)$$

where  $\text{Area}(K)$  is the  $(n-1)$ -dimensional surface area of the box  $K$ ,  $\text{Vol}(K)$  is the volume of the box  $K$ ,  $\Gamma$  is the gamma function, and where  $|\mathbb{S}^{n-1}| = 2\pi^{n/2}/\Gamma(n/2)$  is the  $(n-1)$ -dimensional surface area of the sphere  $\mathbb{S}^{n-1} \subseteq \mathbb{R}^n$ . The formula in (2) has been

proven in a more general setting earlier (see e.g. formula (2.4) in [6]); for further details, see Section 1.1. For the convenience of the reader we give a short proof of formula (2) in our setting in Section 2.2.

Throughout the paper, we will write  $\text{pdf}_Z$  and  $\text{cdf}_Z$  for the probability density function and the cumulative distribution function of  $Z$ , respectively, for random variables  $Z$ . We next give explicit formulas for the probability density function of  $X$  in dimensions two and three.

**Theorem 3.** *For a box of dimension  $n = 2$  with side-lengths  $a \leq b$ , the probability density function of  $X$  is given by*

$$\text{pdf}_X(t) = \frac{1}{a+b} \cdot \begin{cases} 1, & \text{if } t < a, b \\ \frac{a^2 b}{t^2 \sqrt{t^2 - a^2}}, & \text{if } a < t < b \\ -1 + \frac{1}{t^2} \left( \frac{a^2 b}{\sqrt{t^2 - a^2}} + \frac{ab^2}{\sqrt{t^2 - b^2}} \right), & \text{if } a, b < t. \end{cases}$$

for  $0 < t < \sqrt{a^2 + b^2}$ .

**Remark 4.** We note that the probability density function in Theorem 3 is analytic on all open subintervals of  $(0, \sqrt{a^2 + b^2})$  not containing  $a$  or  $b$ . Moreover, it is constant on the interval  $(0, \min(a, b))$  and has singularities of type  $(t - a)^{-1/2}$  and  $(t - b)^{-1/2}$  just to the right of  $a$  and  $b$ , respectively. See Figure 1 for more details. For an explanation of these singularities, see Remark 26.

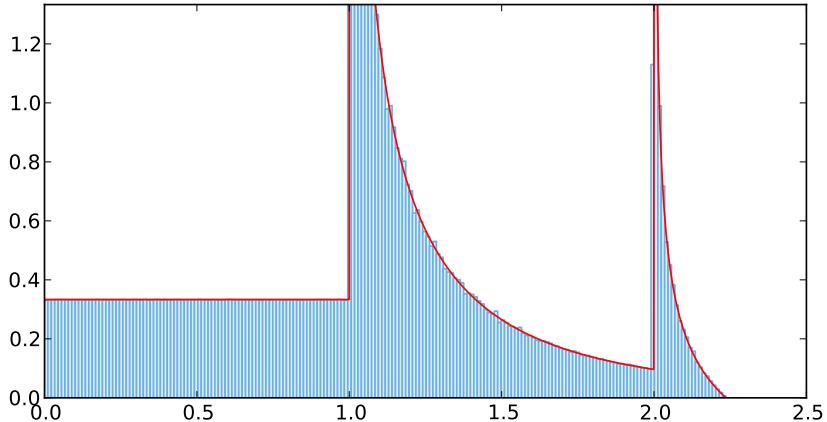


Figure 1: Simulation (blue histogram) vs explicit probability density function (red line) given by Theorem 3 for  $(a, b) = (1, 2)$ . (Simulation used  $10^5$  particles, each starting at the origin with a uniformly random direction, going for a total distance 1000 each.) The plot is cutoff at  $y = 1.3$  since  $\text{pdf}_X(t)$  tends to infinity as  $t \rightarrow 1^+$  and  $t \rightarrow 2^+$ .

**Theorem 5.** For a box of dimension  $n = 3$  with side-lengths  $a, b, c$ , the probability density function of  $X$  is given by

$$\text{pdf}_X(t) = \frac{F(a, b, c, t) + F(b, c, a, t) + F(c, a, b, t)}{3\pi t^3(ab + ac + bc)}$$

where  $F$  is the piecewise-defined function given by

$$F(a, b, c, t) = t^3(8a - 3t)$$

for  $0 < t < a$ , and by

$$F(a, b, c, t) = (6t^4 - a^4 + 6\pi a^2 bc) - 4(b+c)\sqrt{|t^2 - a^2|}(a^2 + 2t^2)$$

for  $a < t < \sqrt{a^2 + b^2}$ , and by

$$\begin{aligned} F(a, b, c, t) = & 6\pi a^2 bc + b^4 - 3t^4 - 6a^2 b^2 + \\ & \sqrt{|t^2 - a^2 - b^2|} 4c(a^2 + b^2 + 2t^2) + \\ & + 4a\sqrt{|t^2 - b^2|}(b^2 + 2t^2) - 12a^2 bc \cdot \arctan\left(\frac{\sqrt{|t^2 - a^2 - b^2|}}{b}\right) + \\ & - 4c\sqrt{|t^2 - a^2|}(a^2 + 2t^2) - 12ab^2 c \cdot \arctan\left(\frac{\sqrt{|t^2 - a^2 - b^2|}}{a}\right) \end{aligned}$$

for  $\sqrt{a^2 + b^2} < t < \sqrt{a^2 + b^2 + c^2}$ .

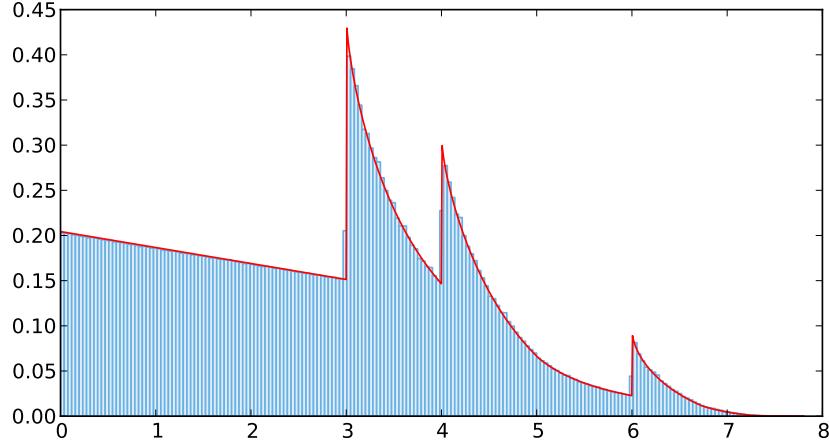


Figure 2: Simulation (blue histogram) vs explicit probability density function (red line) given by Theorem 5 for  $(a, b, c) = (3, 4, 6)$ . (Simulation used  $10^5$  particles, each starting at the origin with a uniformly random direction, going for a total distance 1000 each.) The fact that  $\text{pdf}_X(t)$  is not smooth at  $t = 5$  is barely noticeable.

**Remark 6.** We note that the probability density function in Theorem 5 is analytic on all open subintervals of  $(0, \sqrt{a^2 + b^2 + c^2})$  not containing any of the points

$$a, b, c, \sqrt{a^2 + b^2}, \sqrt{a^2 + c^2}, \sqrt{b^2 + c^2}.$$

Moreover, it is linear on the interval  $(0, \min(a, b, c))$  and has positive jump discontinuities at the points  $a, b, c$ . At the points  $\{\sqrt{a^2 + b^2}, \sqrt{a^2 + c^2}, \sqrt{b^2 + c^2}\} \setminus \{a, b, c\}$ , it is continuous and differentiable.

Note that the probability distribution  $X_{M,R}$  gives a larger “weight” to some particles than others, since some particles get more bounces than others for the same distance  $R$ . One could also consider a similar problem where we send off each particle for a certain number  $N > 0$  of bounces, and then consider the limit as  $M \rightarrow \infty$  followed by taking the limit  $N \rightarrow \infty$ , where  $M$  is the number of particles. This would give each particle the same “weight”. Denote the finite version of this distribution by  $Y_{M,N}$  and its limit distribution as  $M \rightarrow \infty$  and then  $N \rightarrow \infty$  by  $Y$ . With regard to the previous discussion about signal loss, we call the limit distribution  $X$  of  $X_{M,R}$  the **spreading model** and we call the limit distribution of  $Y_{M,N}$  the **absorption model**. Determining the probability density function of the absorption model appears to be the more difficult problem, and we give a formula only in dimension two:

**Theorem 7.** For a box of dimension  $n = 2$  with side-lengths  $a \leq b$ , the random variable  $Y_{M,N}$  converges in distribution to the random variable  $Y$ , as we take  $M \rightarrow \infty$  followed by taking  $N \rightarrow \infty$ , where the probability density function  $\text{pdf}_Y(t)$  is given by

$$\frac{2}{\pi} \left( \frac{2(a+b)}{(a^2+b^2)} - \frac{2ab}{(a^2+b^2)^{3/2}} \left( \tanh^{-1} \left( \frac{a}{\sqrt{a^2+b^2}} \right) + \tanh^{-1} \left( \frac{b}{\sqrt{a^2+b^2}} \right) \right) \right)$$

for  $0 < t < a, b$ , and by

$$\frac{2}{\pi} \left( \frac{a(b-\sqrt{t^2-a^2})}{t(b+\sqrt{t^2-a^2})\sqrt{t^2-a^2}} + \frac{2ab+2at-2a\sqrt{t^2-a^2}}{t(a^2+b^2)} + \frac{2ab \left( -\tanh^{-1} \left( \frac{t}{\sqrt{a^2+b^2}} \right) + \tanh^{-1} \left( \frac{\sqrt{t^2-a^2}\sqrt{a^2+b^2}}{tb} \right) - \tanh^{-1} \left( \frac{b}{\sqrt{a^2+b^2}} \right) \right)}{(a^2+b^2)^{3/2}} \right)$$

for  $a < t < b$ , and by

$$\frac{2}{\pi} \left( \frac{a(b-\sqrt{t^2-a^2})}{t(b+\sqrt{t^2-a^2})\sqrt{t^2-a^2}} + \frac{b(a-\sqrt{t^2-b^2})}{t(a+\sqrt{t^2-b^2})\sqrt{t^2-b^2}} + 2 \frac{2ab-a\sqrt{t^2-a^2}-b\sqrt{t^2-b^2}}{t(a^2+b^2)} + \frac{2ab \left( -2\tanh^{-1} \left( \frac{t}{\sqrt{a^2+b^2}} \right) + \tanh^{-1} \left( \frac{\sqrt{t^2-a^2}\sqrt{a^2+b^2}}{tb} \right) + \tanh^{-1} \left( \frac{\sqrt{t^2-b^2}\sqrt{a^2+b^2}}{ta} \right) \right)}{(a^2+b^2)^{3/2}} \right)$$

for  $a, b < t < \sqrt{a^2+b^2}$ .

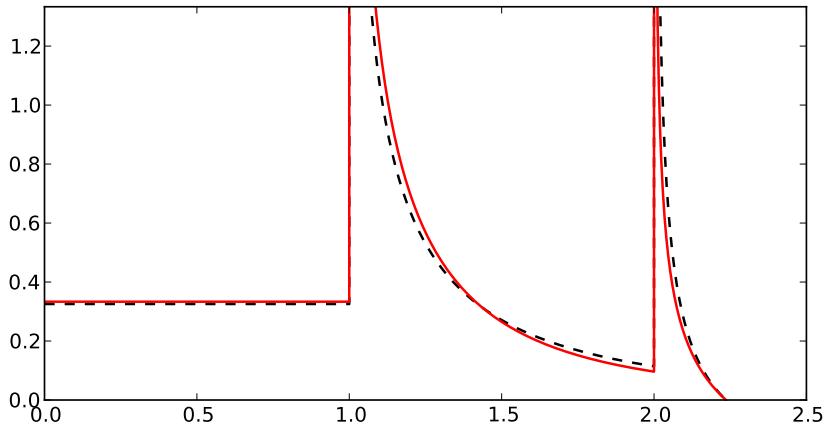


Figure 3: Probability density function for spreading model  $X$  (red line) from Theorem 3 vs absorption model (black dashed line) from Theorem 7, for  $(a, b) = (1, 2)$ .

See Figure 3 for a comparison between the probability density functions for the two different models in dimension 2.

**Remark 8.** It is not a priori obvious that the two limit distributions should differ, and it is natural to ask how much, if at all, they differ. We start by remarking that the expression for  $\text{pdf}_Y(t)$  does not simplify into the expression for  $\text{pdf}_X(t)$ ; indeed, for  $(a, b) = (1, 2)$  we have  $\text{pdf}_X(t) = 1/3$  but  $\text{pdf}_Y(t) \approx 0.32553$  on the interval  $(0, 1)$ . For very skew boxes, with  $a = 1$  and  $b \rightarrow \infty$ , it is straightforward to show that

$$\frac{\text{pdf}_Y(b/2)}{\text{pdf}_X(b/2)} \rightarrow \infty$$

as  $b \rightarrow \infty$ .

### 1.1. Discussion

Given a closed convex subset  $C \subset \mathbb{R}^n$  with nonempty interior it is possible to define a natural probability measure on the set of lines in  $\mathbb{R}^n$  that have nonempty intersection with  $C$ . The expected length of the intersection of a random line is then, up to a constant that only depends on  $n$ , given by  $\text{Vol}(C)/\text{Area}(C)$ ; this is known as Santalo's formula in the integral geometry and geometric probability literature (cf. [14, Ch. 3]).

A billiard flow on a manifold  $M$  with boundary  $\partial M$  gives rise to a billiard map (roughly speaking, the phase space  $\Omega$  is then the collection of inward facing unit vectors  $v$  at each point  $x \in \partial M$ ). Given  $(x, v) \in \Omega$  we define the associated free path as the distance the billiard particle, starting at  $x$  in the direction  $v$ , covers before colliding with  $\partial M$  again. As the billiard map carries a natural probability measure  $\nu$  we can view the free path as a random variable, and the mean free path is then just its expected value. Remarkably,

the mean free path (again up to a constant that only depends on the dimension) is then given by  $\text{Vol}(M)/\text{Vol}(\partial M)$  — even for non-convex billiards. This was deduced in the seventies at the Moscow seminar on dynamical systems directed by Sinai and Alekseev but was never published and hence rederived by a number of researchers. For further details and an interesting historical survey, see Chernov’s paper [6, Sec. 2].

In spirit our methods are closely related to the ones used by Barra-Gaspard [1] in their study of the level spacing distribution for quantum graphs, and this turns out to be given by the distribution of return times to a hypersurface of section of a linear flow on a torus. In particular, for graphs with a finite number of disconnected bonds of incommensurable lengths, the hypersurface of section is the “walls” of the torus, and the level spacings of the quantum graph is exactly the same same as the free path length distribution in our setting when all particles have the *same* starting velocity. (In particular, compare the numerator in (16) for  $v$  fixed with [1, Equation (49)].)

In [12], Marklof and Strömbergsson used the results by Barra-Gaspard to determine the gap distribution of the sequence of fractional parts of  $\{\log_b n\}_{n \in \mathbb{Z}^+}$ . The gap distribution depends on whether  $b$  is trancendental, rational or algebraic; quite remarkably the density function  $P(s)$  for these gaps share a number of qualitative features with the density function  $\text{pdf}_X(s)$  for free paths in our setting. Namely, the density functions both have compact support and are smooth apart from a finite number of jump discontinuities. Further, in some cases the density function is constant for  $s$  small; compare Figure 1 (here  $d = 2$ ) with [12, Figure 4] (here  $b = \sqrt{10}$ ). However, there are some important differences: for  $P(s)$ , left and right limits exist at the jump discontinuities, whereas for  $d = 2$ , the right limit of  $\text{pdf}_X(s)$  is  $+\infty$  at the jumps (cf. Figure 1.) Further, despite appearances,  $P(s)$  is not linear near  $s = 0$  (cf. [12, Figure 1] corresponding to  $b = e$ ) whereas for  $d = 3$ ,  $\text{pdf}_X(s)$  is indeed linear near  $s = 0$  (cf. Figure 2).

## 1.2. Acknowledgements

We would like to thank Z. Rudnick for some very helpful discussions, especially for suggesting the connection with integral geometry. We also thank J. Marklof for bringing references [1, 12] to our attention.

S.H. was partially supported by a grant from the Swedish Research Council (621-2011-5498). P.K. was partially supported by grants from the Göran Gustafsson Foundation for Research in Natural Sciences and Medicine, and the Swedish Research Council (621-2011-5498).

## 2. Proof of Theorem 1

In this section, we prove Theorem 1. For notational simplicity, we give the proof in dimension three; the general proof for  $n \geq 2$  dimensions is analogous.

Given a particle with initial position  $p$  and initial direction  $v$ , let  $N_{R,p,v}$  be the number of bounce lengths we get from that particle as it has travelled a total distance  $R > 0$ , and let  $N_{R,p,v}(t)$  be the number of such bounce lengths of length at most  $t \geq 0$ . The uniform

probability distribution on the set of bounce lengths of  $M$  particles with initial positions  $p^{(1)}, \dots, p^{(M)}$  and initial directions  $v^{(1)}, \dots, v^{(M)}$  has the cumulative distribution function

$$\text{cdf}_{X_{M,R}}(t) = \frac{\sum_{i=1}^M N_{R,p^{(i)},v^{(i)}}(t)}{\sum_{i=1}^M N_{R,p^{(i)},v^{(i)}}} = \frac{\frac{1}{M} \sum_{i=1}^M \frac{N_{R,p^{(i)},v^{(i)}}}{R} \frac{N_{R,p^{(i)},v^{(i)}}(t)}{N_{R,p^{(i)},v^{(i)}}}}{\frac{1}{M} \sum_{i=1}^M \frac{N_{R,p^{(i)},v^{(i)}}}{R}}. \quad (9)$$

(Note that the denominator is uniformly bounded from below, which follows from equation (11) below.) By the strong law of large numbers, the function (9) converges almost surely to

$$\frac{\int_K \int_{\mathbb{S}^2} \frac{N_{R,p,v}}{R} \frac{N_{R,p,v}(t)}{N_{R,p,v}} dS(v) d\mu(p)}{\int_K \int_{\mathbb{S}^2} \frac{N_{R,p,v}}{R} dS(v) d\mu(p)} \quad (10)$$

as  $M \rightarrow \infty$ , where  $d\mu$  is the probability measure with which we choose the starting points, and  $dS$  is the surface area measure on the sphere  $\mathbb{S}^2$ . By symmetry, we may restrict the inner integrals to  $\mathbb{S}_+^2 := \{(v_x, v_y, v_z) \in \mathbb{S}^2 : v_x, v_y, v_z > 0\}$ . We now look at the limit of (10) as  $R \rightarrow \infty$ , and we note that since the integrands are uniformly bounded, we may move the limit inside the integrals by the Lebesgue dominated convergence theorem. Fix one of the integrands, and denote it by  $f(R, p, v, t)$ . We will show that its limit  $g(p, v, t) := \lim_{R \rightarrow \infty} f(R, p, v, t)$  exists for all  $t$  and all directions  $v \in \mathbb{S}^2$ . Moreover, if  $p^{(i)}$  and  $v^{(i)}$  denote random variables corresponding to an initial position and an initial direction, respectively, as above, then

$$h(p^{(i)}, v^{(i)}, t) := \lim_{R \rightarrow \infty} \frac{N_{R,p^{(i)},v^{(i)}}}{R} \frac{N_{R,p^{(i)},v^{(i)}}(t)}{N_{R,p^{(i)},v^{(i)}}}$$

is a random variable with finite variance (and similarly for the terms in the denominator of (9); in particular recall it is uniformly bounded from below), and thus the strong law of large numbers gives that the limit of (9) as  $R \rightarrow \infty$ , and then  $M \rightarrow \infty$  almost surely equals (10). This shows that  $\lim_{M \rightarrow \infty} \lim_{R \rightarrow \infty} \text{cdf}_{X_{M,R}}(t)$  exists almost surely and is equal to  $\lim_{R \rightarrow \infty} \lim_{M \rightarrow \infty} \text{cdf}_{X_{M,R}}(t)$ .

Consider a particle with initial position  $p$  and initial direction  $v = (v_x, v_y, v_z) \in \mathbb{S}_+^2$ . By “unfolding” its motion with specular reflections on the walls of the box to the motion along a straight line in  $\mathbb{R}^n$  — see Figure 4 for a 2D illustration — we see that the particle’s set of bounce lengths is identical to the set of path lengths between consecutive intersections of the straight line segment  $\{p + tv : 0 \leq t \leq R\}$  with any of the planes  $x = na, y = nb, z = nc, n \in \mathbb{Z}$ . Thus we see that

$$N_{R,p,v} = R \frac{v_x}{a} + R \frac{v_y}{b} + R \frac{v_z}{c} + O(1) \quad (11)$$

for large  $R$ , and therefore

$$\frac{N_{R,p,v}}{R} \rightarrow \frac{v_x}{a} + \frac{v_y}{b} + \frac{v_z}{c} \quad (12)$$

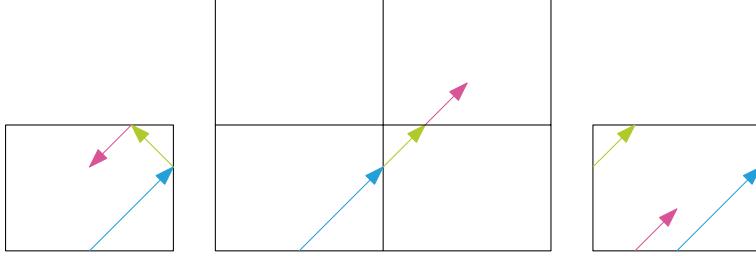


Figure 4: From left to right: Unfolding a motion with specular reflection in a 2D box to a motion the plane and then projecting back to the box.

as  $R \rightarrow \infty$ .

Now project the line  $\{p+tv : 0 \leq t \leq R\}$  to the torus  $\mathbb{R}^3/\Lambda$  where  $\Lambda = \{(n_1a, n_2b, n_3c) : n_1, n_2, n_3 \in \mathbb{Z}\}$  and let us identify the torus with the box  $K$ ; see Figure 4. Each bounce length corresponds to a line segment which starts in one of the three planes  $x = 0, y = 0$  or  $z = 0$  and runs in the direction  $v$  to one of the three planes  $x = a, y = b$  or  $z = c$ . There are  $R \frac{v_z}{c} + O(1)$  line segments which start from the plane  $z = 0$ , and thus the probability that a line segment starts from the plane  $z = 0$  is

$$\frac{\frac{v_z}{c}}{\frac{v_x}{a} + \frac{v_y}{b} + \frac{v_z}{c}}$$

as  $R \rightarrow \infty$ . By the ergodicity of the linear flow on tori (for almost all directions), the starting points of these line segments become uniformly distributed on the rectangle  $[0, a] \times [0, b] \times \{0\}$  for almost all  $v \in \mathbb{S}_+^2$  as  $R \rightarrow \infty$ ; from here we will assume that  $v$  is such a direction, and we will ignore the measure zero set of directions for which we do not have ergodicity. Consider one of these line segments and denote its length by  $T$  and its starting point by  $(x_0, y_0, 0)$ . For an arbitrary parameter  $t \geq 0$ , we have  $T \leq t$  if and only if  $tv_x \geq a - x_0$  or  $tv_y \geq b - y_0$  or  $tv_z \geq c$ ; the starting points  $(x_0, y_0) \in [0, a] \times [0, b]$  which satisfy this are precisely those outside the rectangle  $[0, a - tv_x] \times [0, b - tv_y]$  assuming that  $tv_z \leq c$  and otherwise it is the whole rectangle  $[0, a] \times [0, b]$ . The area of that region is

$$ab - (a - tv_x)(b - tv_y) \tag{13}$$

if  $a \geq tv_x, b \geq tv_y, c \geq tv_z$  and otherwise it is  $ab$ . Since the starting points  $(x_0, y_0)$  are uniformly distributed in the rectangle  $[0, a] \times [0, b]$  as  $R \rightarrow \infty$ , it follows that the probability that  $T \leq t$  is

$$1 - \frac{(a - tv_x)(b - tv_y)}{ab} \chi(a \geq tv_x, b \geq tv_y, c \geq tv_z),$$

where  $\chi(P)$  is the indicator function which is 1 whenever the condition  $P$  is true, and 0 otherwise. We get analogous expressions for the case when a line segment starts in the plane  $x = 0$  or  $y = 0$  instead. Thus the proportion of all line segments with length at

most  $t$  as  $R \rightarrow \infty$  is

$$\begin{aligned} \lim_{R \rightarrow \infty} \frac{N_{R,p,v}(t)}{N_{R,p,v}} &= \frac{\frac{v_x}{a}}{\frac{v_x}{a} + \frac{v_y}{b} + \frac{v_z}{c}} \left( 1 - \frac{(b - tv_y)(c - tv_z)}{bc} \chi(a \geq tv_x, b \geq tv_y, c \geq tv_z) \right) + \\ &\quad \frac{\frac{v_y}{b}}{\frac{v_x}{a} + \frac{v_y}{b} + \frac{v_z}{c}} \left( 1 - \frac{(a - tv_x)(c - tv_z)}{ac} \chi(a \geq tv_x, b \geq tv_y, c \geq tv_z) \right) + \\ &\quad \frac{\frac{v_z}{c}}{\frac{v_x}{a} + \frac{v_y}{b} + \frac{v_z}{c}} \left( 1 - \frac{(a - tv_x)(b - tv_y)}{ab} \chi(a \geq tv_x, b \geq tv_y, c \geq tv_z) \right) \end{aligned}$$

which can be written

$$1 - \frac{\chi(a \geq tv_x, b \geq tv_y, c \geq tv_z)}{abc(\frac{v_x}{a} + \frac{v_y}{b} + \frac{v_z}{c})} \times \left( v_x(b - tv_y)(c - tv_z) + v_y(a - tv_x)(c - tv_z) + v_z(a - tv_x)(b - tv_y) \right). \quad (14)$$

Recognizing that both integrands (12) and (14) are independent of the position  $p$ , we see that the limit of (10) as  $R \rightarrow \infty$  may be written as

$$\begin{aligned} \lim_{R \rightarrow \infty} \lim_{M \rightarrow \infty} \text{cdf}_{X_{M,R}}(t) &= 1 - \frac{1}{\int_{\mathbb{S}_+^2} (v_x bc + av_y c + abv_z) dS(v)} \times \\ &\quad \times \int_{\substack{v \in \mathbb{S}_+^2 \\ v_x \leq a/t \\ v_y \leq b/t \\ v_z \leq c/t}} ((abv_z + av_y c + v_x bc) - 2t(av_y v_z + v_x bv_z + v_x v_y c) + 3t^2 v_x v_y v_z) dS(v) \quad (15) \end{aligned}$$

for all  $t > 0$ . The corresponding formula in  $n$  dimensions is given by

$$\begin{aligned} \int_{\substack{v \in \mathbb{S}_+^{n-1} \\ v_i \leq a_i/t}} \left( \sum_{i=1}^n v_i \prod_{j \neq i} (a_j - tv_j) \right) dS(v) \\ \lim_{R \rightarrow \infty} \lim_{M \rightarrow \infty} \text{cdf}_{X_{M,R}}(t) = 1 - \frac{\text{for } i=1, \dots, n}{\left( \prod_{i=1}^n a_i \right) \int_{\mathbb{S}_+^{n-1}} \left( \sum_{i=1}^n \frac{v_i}{a_i} \right) dS(v)} \quad (16) \end{aligned}$$

for all  $t > 0$ , where the side-lengths of the box  $K$  are  $a_1, \dots, a_n$  and  $dS$  is the surface area measure on  $\mathbb{S}_+^{n-1} \cap [0, \infty)^n$ . (The denominator can be given explicitly by using Lemma 28 below.)

We have thus proved that the random variable  $X_{M,R}$  converges in distribution to a random variable with probability density function given by (16) as we take  $M \rightarrow \infty$  followed by taking  $R \rightarrow \infty$ , or alternatively, first taking  $R \rightarrow \infty$  followed by taking  $M \rightarrow \infty$ . It remains to prove that this distribution agrees with the distribution of the random variable  $X$  defined in the introduction.

## 2.1. Integral geometry

We start by recalling some standard facts from integral geometry (cf. [8, 14].) The set of directed straight lines  $\ell$  in  $\mathbb{R}^3$  can be parametrized by pairs  $(v, q)$  where  $v \in \mathbb{S}^2$  is a

unit vector pointing in the same direction as  $\ell$  and  $q \in v^\perp$  is the unique point in  $\ell$  which intersects the plane through the origin which is orthogonal to  $v$ . The unique translation-and rotation-invariant measure (up to a constant) on the set of directed straight lines in  $\mathbb{R}^3$  is  $d\ell := dA(q) dS(v)$  where  $dA$  is the surface measure on the plane through the origin orthogonal to  $v \in \mathbb{S}^2$ , and  $dS$  is the surface area measure on  $\mathbb{S}^2$ .

Consider the set  $L_{a,b,c}$  of directed straight lines in  $\mathbb{R}^3$  which intersect the box  $K$ . Now, since  $abv_z + av_y c + v_x bc$  is the area of the projection of the box  $K$  onto the plane  $v^\perp$  for  $v \in \mathbb{S}_+^2$ , it follows that the total measure of  $L_{a,b,c}$  with respect to  $d\ell$  is

$$C_{a,b,c} := 8 \int_{\mathbb{S}_+^2} (abv_z + av_y c + v_x bc) dS(v) = 2\pi(ab + ac + bc)$$

where we used symmetry, and the integral may be evaluated by switching to spherical coordinates. It follows that  $d\ell/C_{a,b,c}$  is a probability measure on the set of directed lines intersecting the box  $L_{a,b,c}$ . Let  $\ell$  be a random directed line with respect to this measure, and define the random variable  $X := \text{length}(\ell \cap K)$ , as in the introduction. Let us determine the probability that  $X \leq t$  for an arbitrary parameter  $t \geq 0$ . By symmetry it suffices to consider only directed lines with  $v \in \mathbb{S}_+^2$ . The set of all intersection points between the rectangle  $[0, a] \times [0, b] \times \{0\}$  and the lines  $\ell$  with  $X \leq t$  and direction  $v \in \mathbb{S}_+^2$  has area  $ab - (a - tv_x)(b - tv_y)\chi(a \geq tv_x, b \geq tv_y, c \geq tv_z)$ , as in (13), and its projection onto the plane  $v^\perp$  has area

$$v_z[ab - (a - tv_x)(b - tv_y)\chi(a \geq tv_x, b \geq tv_y, c \geq tv_z)].$$

By symmetry it follows that the area of the set of directed lines  $\ell \in L_{a,b,c}$  with  $X \leq t$  and direction  $v \in \mathbb{S}_+^2$  projected down to  $v^\perp$  is

$$\begin{aligned} U(v, t) := & v_x[bc - (b - tv_y)(c - tv_z)\chi(a \geq tv_x, b \geq tv_y, c \geq tv_z)] + \\ & v_y[ac - (a - tv_x)(c - tv_z)\chi(a \geq tv_x, b \geq tv_y, c \geq tv_z)] + \\ & v_z[ab - (a - tv_x)(b - tv_y)\chi(a \geq tv_x, b \geq tv_y, c \geq tv_z)], \end{aligned}$$

and it follows that

$$\text{Prob}[X \leq t] = \frac{1}{C_{a,b,c}} \int_{X \leq t} d\ell = \frac{8}{C_{a,b,c}} \int_{\mathbb{S}_+^2} U(v, t) dS(v),$$

which we see is identical to (15), and we have thus proved that  $X_{M,R}$  converges in distribution to  $X$  as we take  $M \rightarrow \infty$  and then  $R \rightarrow \infty$ . This concludes the proof of Theorem 1.

## 2.2. Computing the mean value

We will determine the mean value (2) of  $X$ ; to do this we exploit the integral geometry interpretation of the random variable  $X$ . By symmetry it suffices to restrict to directed lines  $\ell$  with  $v \in \mathbb{S}_+^2$ . For fixed  $v \in \mathbb{S}_+^2$ , denote by  $Q(v) = (K + \text{span}(v)) \cap v^\perp$  the set of  $q \in v^\perp$  such that the directed line  $\ell$  parametrized by  $(v, q)$  intersects  $K$ . We note that

$X dA(q)$  is a volume element of the box  $K$  for any fixed  $v \in \mathbb{S}_+^2$ , and thus integrating  $X dA(q)$  over all  $q$  yields the volume of the box. Hence the mean value is

$$\mathbb{E}[X] = \frac{8}{C_{a,b,c}} \int_{\mathbb{S}_+^2} \int_{Q(v)} X dA(q) dS(v) = \frac{8abc}{C_{a,b,c}} \int_{\mathbb{S}_+^2} dS(v) = \frac{2abc}{ab + ac + bc}.$$

In  $n$  dimensions we get a normalizing factor  $\frac{\text{Area}(K)}{2} \cdot 2^n \int_{\mathbb{S}_+^{n-1}} v_n dS(v)$ , so with the aid of the Lemma 28 in the Appendix, it follows that the mean value in  $n$  dimensions is

$$\mathbb{E}[X] = \frac{1}{2^n \frac{1}{\pi} \frac{|\mathbb{S}^n|}{2^n} \frac{\text{Area}(K)}{2}} 2^n \text{Vol}(K) \frac{|\mathbb{S}^{n-1}|}{2^n} = 2\pi \frac{|\mathbb{S}^{n-1}|}{|\mathbb{S}^n|} \frac{\text{Vol}(K)}{\text{Area}(K)}$$

where  $\text{Area}(K)$  is the  $(n-1)$ -dimensional surface area of the box  $K$ , and  $\text{Vol}(K)$  is the volume of the box  $K$ .

### 3. Proof of Theorem 3

Using formula (16) in dimension  $n = 2$ , we get

$$\text{cdf}_X(t) = 1 - \frac{\int_{v \in \mathbb{S}_+^1} \int_{\substack{v_x \leq a/t \\ v_y \leq b/t}} (v_x(b - tv_y) + v_y(a - tv_x)) dS(v)}{ab \int_{\mathbb{S}_+^1} \left( \frac{v_x}{a} + \frac{v_y}{b} \right) dS(v)}.$$

We use polar coordinates  $v_x = \cos \theta, v_y = \sin \theta$  so that  $dS(v) = d\theta$ . Then the above becomes

$$1 - \frac{\int_{\cos^{-1}(\min(a/t, 1))}^{\sin^{-1}(\min(b/t, 1))} (b \cos \theta + a \sin \theta - 2t \sin \theta \cos \theta) d\theta}{\int_0^{\pi/2} (b \cos \theta + a \sin \theta) d\theta} = 1 - \frac{1}{a+b} \left[ b \sin \theta - a \cos \theta + t \cos^2 \theta \right]_{\cos^{-1}(\min(a/t, 1))}^{\sin^{-1}(\min(b/t, 1))}. \quad (17)$$

The numerator of the second term may be written

$$\begin{aligned} & \chi(b < t) \left( b \cdot \frac{b}{t} - a \sqrt{1 - \frac{b^2}{t^2}} + t \left( 1 - \frac{b^2}{t^2} \right) \right) + \chi(b \geq t) (b - a \cdot 0 + t \cdot 0) + \\ & - \chi(a < t) \left( b \sqrt{1 - \frac{a^2}{t^2}} - a \cdot \frac{a}{t} + t \cdot \frac{a^2}{t^2} \right) - \chi(a \geq t) (b \cdot 0 - a + t) \end{aligned}$$

which can be simplified to

$$\chi(b < t) \left( t - b - a \sqrt{1 - \frac{b^2}{t^2}} \right) + \chi(a < t) \left( t - a - b \sqrt{1 - \frac{a^2}{t^2}} \right) + (a + b - t).$$

Inserting this into (17) and differentiating yields Theorem 3.

## 4. Proof of Theorem 5

We will evaluate the cumulative distribution function (15) and then differentiate. The denominator of the second term of (15) is

$$\int_{\mathbb{S}^2_+} (abv_z + av_y c + v_x bc) dS(v) = \frac{\pi}{4}(ab + ac + bc),$$

as may be evaluated by switching to spherical coordinates. Define

$$\begin{aligned} f(a, b, c) &:= bc \int_{\substack{v \in \mathbb{S}^2_+ \\ v_x \leq a/t \\ v_y \leq b/t \\ v_z \leq c/t}} v_x dS(v), \\ g(a, b, c) &:= -2tc \int_{\substack{v \in \mathbb{S}^2_+ \\ v_x \leq a/t \\ v_y \leq b/t \\ v_z \leq c/t}} v_x v_y dS(v), \\ h(a, b, c) &:= 3t^2 \int_{\substack{v \in \mathbb{S}^2_+ \\ v_x \leq a/t \\ v_y \leq b/t \\ v_z \leq c/t}} v_x v_y v_z dS(v). \end{aligned}$$

By symmetry, we have

$$\begin{aligned} f(c, a, b) &= ab \int_{\substack{v \in \mathbb{S}^2_+ \\ v_x \leq c/t \\ v_y \leq a/t \\ v_z \leq b/t}} v_x dS(v) = ab \int_{\substack{v \in \mathbb{S}^2_+ \\ v_x \leq a/t \\ v_y \leq b/t \\ v_z \leq c/t}} v_z dS(v), \\ f(b, c, a) &= ac \int_{\substack{v \in \mathbb{S}^2_+ \\ v_x \leq b/t \\ v_y \leq c/t \\ v_z \leq a/t}} v_x dS(v) = ac \int_{\substack{v \in \mathbb{S}^2_+ \\ v_x \leq a/t \\ v_y \leq b/t \\ v_z \leq c/t}} v_y dS(v), \\ g(c, a, b) &= -2tb \int_{\substack{v \in \mathbb{S}^2_+ \\ v_x \leq c/t \\ v_y \leq a/t \\ v_z \leq b/t}} v_x v_y dS(v) = -2tb \int_{\substack{v \in \mathbb{S}^2_+ \\ v_x \leq a/t \\ v_y \leq b/t \\ v_z \leq c/t}} v_x v_z dS(v), \\ g(b, c, a) &= -2ta \int_{\substack{v \in \mathbb{S}^2_+ \\ v_x \leq b/t \\ v_y \leq c/t \\ v_z \leq a/t}} v_x v_y dS(v) = -2ta \int_{\substack{v \in \mathbb{S}^2_+ \\ v_x \leq a/t \\ v_y \leq b/t \\ v_z \leq c/t}} v_y v_z dS(v), \end{aligned}$$

and thus we can write the numerator in the second term of (15) as

$$f(a, b, c) + f(c, a, b) + f(b, c, a) + g(a, b, c) + g(c, a, b) + g(b, c, a) + h(a, b, c).$$

Exploiting the symmetries, it suffices to evaluate  $h(a, b, c)$ ,  $g(a, b, c)$  and  $f(b, c, a)$  (note the order of the arguments to  $f$ ). We will evaluate these integrals by switching to spherical coordinates, but first we need to parametrize the part of the sphere inside the box  $0 \leq v_x \leq a/t, 0 \leq v_y \leq b/t, 0 \leq v_z \leq c/t$ .

**Lemma 18.** Fix  $t \in (0, \sqrt{a^2 + b^2 + c^2})$ . We have

$$\int_{\substack{v \in \mathbb{S}_+^2 \\ v_x \leq a/t \\ v_y \leq b/t \\ v_z \leq c/t}} F(v_x, v_y, v_z) dS(v) = \left( \int_{\theta_{min}}^{\theta_a} \int_0^{\pi/2} + \int_{\theta_a}^{\theta_{max}} \int_{\varphi_a}^{\pi/2} - \int_{\theta_b}^{\theta_{max}} \int_{\varphi_b}^{\pi/2} \right) \tilde{F}(\theta, \varphi) \sin \theta d\varphi d\theta$$

for any integrable function  $F : \mathbb{S}_+^2 \rightarrow \mathbb{R}$ , where  $\tilde{F}(\theta, \varphi) := F(\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$ , where

$$\begin{aligned} \theta_{min} &:= \cos^{-1} \left\{ \frac{c}{t} \right\}_1, \\ \theta_a &:= \max(\theta_{min}, \sin^{-1} \left\{ \frac{a}{t} \right\}_1), \\ \theta_b &:= \max(\theta_{min}, \sin^{-1} \left\{ \frac{b}{t} \right\}_1), \\ \theta_{max} &:= \sin^{-1} \left\{ \frac{\sqrt{a^2 + b^2}}{t} \right\}_1, \\ \varphi_a &:= \cos^{-1} \frac{a}{t \sin \theta} \quad (\text{whenever } a \leq t \sin \theta), \\ \varphi_b &:= \sin^{-1} \frac{b}{t \sin \theta} \quad (\text{whenever } b \leq t \sin \theta). \end{aligned}$$

and where we have used the shorthand  $\{u\}_1 := \min(u, 1)$ .

*Proof.* We will parametrize the set of points  $v = (v_x, v_y, v_z)$  on the sphere  $\mathbb{S}^2$  such that

$$\begin{aligned} 0 < v_x &\leq a/t, \\ 0 < v_y &\leq b/t, \\ 0 < v_z &\leq c/t. \end{aligned} \tag{19}$$

Switch to spherical coordinates  $v_x = \sin \theta \cos \varphi, v_y = \sin \theta \sin \varphi, v_z = \cos \theta$ . The non-negativity conditions of (19) are equivalent to the condition  $\theta, \varphi \in (0, \pi/2)$ . For such angles, the condition  $v_z \leq c/t$  is equivalent to

$$\cos^{-1} \left\{ \frac{c}{t} \right\}_1 \leq \theta,$$

and the conditions  $v_x \leq a/t, v_y \leq b/t$  are equivalent to

$$\cos^{-1} \left\{ \frac{a}{t \sin \theta} \right\}_1 \leq \varphi \leq \sin^{-1} \left\{ \frac{b}{t \sin \theta} \right\}_1. \tag{20}$$

The interval (20) is non-empty for precisely those  $\theta \in (0, \pi/2)$  such that  $\theta \leq \theta_{\max}$  since

$$1 \leq \left\{ \frac{a}{t \sin \theta} \right\}_1^2 + \left\{ \frac{b}{t \sin \theta} \right\}_1^2 \iff 1 \leq \left( \frac{a}{t \sin \theta} \right)^2 + \left( \frac{b}{t \sin \theta} \right)^2 \iff \\ \sin \theta \leq \frac{\sqrt{a^2 + b^2}}{t} \iff \theta \leq \sin^{-1} \left\{ \frac{\sqrt{a^2 + b^2}}{t} \right\}_1.$$

Thus we may restrict  $\theta$  to the interval given by the inequalities

$$\theta_{\min} \leq \theta \leq \theta_{\max}.$$

Note that we have  $\theta_{\min} \leq \theta_{\max}$  for all  $t \leq \sqrt{a^2 + b^2 + c^2}$  since

$$\theta_{\min} \leq \theta_{\max} \iff 1 \leq \left\{ \frac{c}{t} \right\}_1^2 + \left\{ \frac{\sqrt{a^2 + b^2}}{t} \right\}_1^2 \iff \\ 1 \leq \left( \frac{c}{t} \right)^2 + \left( \frac{\sqrt{a^2 + b^2}}{t} \right)^2 \iff t^2 \leq a^2 + b^2 + c^2.$$

We conclude that we can write

$$\int_{\substack{v \in \mathbb{S}_+^2 \\ v_x \leq a/t \\ v_y \leq b/t \\ v_z \leq c/t}} F(v_x, v_y, v_z) dS(v) = \int_{\theta_{\min}}^{\theta_{\max}} \int_{\cos^{-1} \left\{ \frac{a}{t \sin \theta} \right\}_1}^{\sin^{-1} \left\{ \frac{b}{t \sin \theta} \right\}_1} \tilde{F}(\theta, \varphi) \sin \theta d\varphi d\theta. \quad (21)$$

For  $\theta \in (0, \pi/2)$ , note that  $\cos^{-1} \frac{a}{t \sin \theta}$  is defined precisely when  $\sin^{-1} \left\{ \frac{a}{t} \right\}_1 \leq \theta$  and that  $\sin^{-1} \frac{b}{t \sin \theta}$  is defined precisely when  $\sin^{-1} \left\{ \frac{b}{t} \right\}_1 \leq \theta$ . We have  $\theta_{\min} < \theta_a$  if and only if  $t < \sqrt{a^2 + c^2}$ , and we have  $\theta_{\min} < \theta_b$  if and only if  $t < \sqrt{b^2 + c^2}$ . Moreover we note that we always have  $\theta_a, \theta_b \in [\theta_{\min}, \theta_{\max}]$ .

Let us rewrite the integration limits in the right-hand side of (21) in terms of  $\varphi_a$  and  $\varphi_b$ . A priori, we need to distinguish between the two cases  $\theta_a \leq \theta_b$  and  $\theta_b < \theta_a$ . If  $\theta_a \leq \theta_b$  then we get

$$\begin{aligned} \left( \int_{\theta_{\min}}^{\theta_{\max}} \int_{\cos^{-1} \left\{ \frac{x}{t \sin \theta} \right\}_1}^{\sin^{-1} \left\{ \frac{y}{t \sin \theta} \right\}_1} \right) &= \left( \int_{\theta_{\min}}^{\theta_a} \int_0^{\pi/2} + \int_{\theta_a}^{\theta_b} \int_{\varphi_a}^{\pi/2} + \int_{\theta_b}^{\theta_{\max}} \int_{\varphi_a}^{\varphi_b} \right) = \\ \left( \int_{\theta_{\min}}^{\theta_a} \int_0^{\pi/2} + \int_{\theta_a}^{\theta_{\max}} \int_{\varphi_a}^{\pi/2} - \int_{\theta_b}^{\theta_{\max}} \int_{\varphi_a}^{\pi/2} + \int_{\theta_b}^{\theta_{\max}} \int_{\varphi_a}^{\pi/2} - \int_{\theta_b}^{\theta_{\max}} \int_{\varphi_b}^{\pi/2} \right) &= \\ \left( \int_{\theta_{\min}}^{\theta_a} \int_0^{\pi/2} + \int_{\theta_a}^{\theta_{\max}} \int_{\varphi_a}^{\pi/2} - \int_{\theta_b}^{\theta_{\max}} \int_{\varphi_b}^{\pi/2} \right). \end{aligned} \quad (22)$$

If on the other hand  $\theta_b < \theta_a$  then

$$\begin{aligned} \left( \int_{\theta_{\min}}^{\theta_{\max}} \int_{\cos^{-1} \left\{ \frac{x}{t \sin \theta} \right\}_1}^{\sin^{-1} \left\{ \frac{y}{t \sin \theta} \right\}_1} \right) &= \left( \int_{\theta_{\min}}^{\theta_b} \int_0^{\pi/2} + \int_{\theta_b}^{\theta_a} \int_0^{\varphi_b} + \int_{\theta_a}^{\theta_{\max}} \int_{\varphi_b}^{\varphi_b} \right) = \\ \left( \int_{\theta_{\min}}^{\theta_b} \int_0^{\pi/2} + \int_{\theta_b}^{\theta_a} \int_0^{\pi/2} - \int_{\theta_b}^{\theta_a} \int_{\varphi_b}^{\pi/2} + \int_{\theta_a}^{\theta_{\max}} \int_{\varphi_a}^{\pi/2} - \int_{\theta_a}^{\theta_{\max}} \int_{\varphi_b}^{\pi/2} \right) & \end{aligned}$$

which we see is identical to (22). Combining (21) and (22) we get the conclusion of the lemma.  $\square$

Applying Lemma 18 we get

$$h(a, b, c) = 3t^2 \int_{\substack{v \in \mathbb{S}_+^2 \\ v_x \leq a/t \\ v_y \leq b/t \\ v_z \leq c/t}} v_x v_y v_z dS(v) =$$

$$3t^2 \left( \int_{\theta_{\min}}^{\theta_a} \int_0^{\pi/2} + \int_{\theta_a}^{\theta_{\max}} \int_{\varphi_a}^{\pi/2} - \int_{\theta_b}^{\theta_{\max}} \int_{\varphi_b}^{\pi/2} \right) (\sin^2 \theta \cos \theta \cos \varphi \sin \varphi) \sin \theta d\varphi d\theta.$$

An antiderivative of the integrand  $\cos \varphi \sin \varphi \cdot \sin^3 \theta \cos \theta$  with respect to  $\varphi$  is  $-\frac{1}{2} \cos^2 \varphi \sin^3 \theta \cos \theta$ , and thus the above is

$$3t^2 \left( \int_{\theta_{\min}}^{\theta_a} \cos^2 \varphi \Big|_{\varphi=0} + \int_{\theta_a}^{\theta_{\max}} \cos^2 \varphi \Big|_{\varphi=\varphi_a} - \int_{\theta_b}^{\theta_{\max}} \cos^2 \varphi \Big|_{\varphi=\varphi_b} \right) \frac{1}{2} \sin^3 \theta \cos \theta d\theta =$$

$$3t^2 \left( \int_{\theta_{\min}}^{\theta_a} 1 + \int_{\theta_a}^{\theta_{\max}} \frac{a^2}{t^2 \sin^2 \theta} + \int_{\theta_b}^{\theta_{\max}} \left( \frac{b^2}{t^2 \sin^2 \theta} - 1 \right) \right) \frac{1}{2} \sin^3 \theta \cos \theta d\theta =$$

$$\frac{3}{2} \left( \int_{\theta_{\min}}^{\theta_a} t^2 \sin^3 \theta \cos \theta d\theta + \int_{\theta_a}^{\theta_{\max}} a^2 \sin \theta \cos \theta d\theta + \int_{\theta_b}^{\theta_{\max}} (b^2 \sin \theta - t^2 \sin^3 \theta) \cos \theta d\theta \right) =$$

$$\frac{3}{2} \left( \left[ t^2 \frac{1}{4} \sin^4 \theta \right]_{\theta_{\min}}^{\theta_a} + \left[ a^2 \frac{1}{2} \sin^2 \theta \right]_{\theta_a}^{\theta_{\max}} + \left[ b^2 \frac{1}{2} \sin^2 \theta - t^2 \frac{1}{4} \sin^4 \theta \right]_{\theta_b}^{\theta_{\max}} \right). \quad (23)$$

Next consider

$$g(a, b, c) = -2tc \int_{\substack{v \in \mathbb{S}_+^2 \\ v_x \leq a/t \\ v_y \leq b/t \\ v_z \leq c/t}} v_x v_y dS(v) =$$

$$-2tc \left( \int_{\theta_{\min}}^{\theta_a} \int_0^{\pi/2} + \int_{\theta_a}^{\theta_{\max}} \int_{\varphi_a}^{\pi/2} - \int_{\theta_b}^{\theta_{\max}} \int_{\varphi_b}^{\pi/2} \right) (\sin^2 \theta \cos \varphi \sin \varphi) \sin \theta d\varphi d\theta.$$

An antiderivative of the integrand  $\cos \varphi \sin \varphi \cdot \sin^3 \theta$  with respect to  $\varphi$  is  $-\frac{1}{2} \cos^2 \varphi \sin^3 \theta$ ,

and thus the above is

$$\begin{aligned}
g(a, b, c) &= -2tc \int_{\substack{v \in \mathbb{S}_+^2 \\ v_x \leq a/t \\ v_y \leq b/t \\ v_z \leq c/t}} v_x v_y dS(v) = \\
&-tc \left( \int_{\theta_{\min}}^{\theta_a} \cos^2 \varphi \Big|_{\varphi=0} + \int_{\theta_a}^{\theta_{\max}} \cos^2 \varphi \Big|_{\varphi=\varphi_a} - \int_{\theta_b}^{\theta_{\max}} \cos^2 \varphi \Big|_{\varphi=\varphi_b} \right) \sin^3 \theta d\theta = \\
&-tc \left( \int_{\theta_{\min}}^{\theta_a} 1 + \int_{\theta_a}^{\theta_{\max}} \frac{a^2}{t^2 \sin^2 \theta} + \int_{\theta_b}^{\theta_{\max}} \left( \frac{b^2}{t^2 \sin^2 \theta} - 1 \right) \right) \sin^3 \theta d\theta = \\
&-tc \left( \int_{\theta_{\min}}^{\theta_a} \sin^3 \theta d\theta + \int_{\theta_a}^{\theta_{\max}} \frac{a^2 \sin \theta}{t^2} d\theta + \int_{\theta_b}^{\theta_{\max}} \left( \frac{b^2 \sin \theta}{t^2} - \sin^3 \theta \right) d\theta \right) = \\
&-tc \left( \left[ \frac{\cos^3 \theta}{3} - \cos \theta \right]_{\theta_{\min}}^{\theta_a} + \frac{a^2}{t^2} [-\cos \theta]_{\theta_a}^{\theta_{\max}} + \left[ -\frac{b^2 \cos \theta}{t^2} - \frac{\cos^3 \theta}{3} + \cos \theta \right]_{\theta_b}^{\theta_{\max}} \right). \quad (24)
\end{aligned}$$

We obtain  $g(b, c, a)$  and  $g(c, a, b)$  by switching the roles of  $a, b, c$  in (24). We remark that trying to obtain  $g(b, c, a)$  and  $g(c, a, b)$  directly, by integrating  $v_y v_z$  and  $v_x v_z$ , respectively, by first integrating with respect to  $\varphi$ , taking the limits  $\varphi \rightarrow \varphi_a$  and  $\varphi \rightarrow \varphi_b$ , and then finding an antiderivative with respect to  $\theta$ , seem to result in much more complicated expressions.

Finally consider

$$\begin{aligned}
f(b, c, a) &= ac \int_{\substack{v \in \mathbb{S}_+^2 \\ v_x \leq a/t \\ v_y \leq b/t \\ v_z \leq c/t}} v_y dS(v) = \\
&ac \left( \int_{\theta_{\min}}^{\theta_a} \int_0^{\pi/2} + \int_{\theta_a}^{\theta_{\max}} \int_{\varphi_a}^{\pi/2} - \int_{\theta_b}^{\theta_{\max}} \int_{\varphi_b}^{\pi/2} \right) (\sin \theta \sin \varphi) \sin \theta d\varphi d\theta.
\end{aligned}$$

An antiderivative of the integrand  $\sin \varphi \cdot \sin^2 \theta$  with respect to  $\varphi$  is  $-\cos \varphi \cdot \sin^2 \theta$ , and thus the above is

$$\begin{aligned}
&ac \left( \int_{\theta_{\min}}^{\theta_a} \cos \varphi \Big|_{\varphi=0} + \int_{\theta_a}^{\theta_{\max}} \cos \varphi \Big|_{\varphi=\varphi_a} - \int_{\theta_b}^{\theta_{\max}} \cos \varphi \Big|_{\varphi=\varphi_b} \right) \sin^2 \theta d\theta = \\
&ac \left( \int_{\theta_{\min}}^{\theta_a} 1 + \int_{\theta_a}^{\theta_{\max}} \frac{a}{t \sin \theta} - \int_{\theta_b}^{\theta_{\max}} \sqrt{1 - \frac{b^2}{t^2 \sin^2 \theta}} \right) \sin^2 \theta d\theta = \\
&ac \left( \int_{\theta_{\min}}^{\theta_a} \sin^2 \theta d\theta + \int_{\theta_a}^{\theta_{\max}} \frac{a \sin \theta}{t} d\theta - \int_{\theta_b}^{\theta_{\max}} \sqrt{\sin^2 \theta - \frac{b^2}{t^2}} \sin \theta d\theta \right) = \\
&ac \left( \frac{1}{2} [\theta - \sin \theta \cos \theta]_{\theta_{\min}}^{\theta_a} + \left[ \frac{-a \cos \theta}{t} \right]_{\theta_a}^{\theta_{\max}} - \int_{\theta_b}^{\theta_{\max}} \sqrt{1 - \frac{b^2}{t^2} - \cos^2 \theta} \sin \theta d\theta \right) \quad (25)
\end{aligned}$$

where the last integral inside the parentheses may be written as

$$\begin{aligned} & \left[ -\frac{1}{2} \left( \cos \theta \sqrt{1 - \frac{b^2}{t^2} - \cos^2 \theta} + \left(1 - \frac{b^2}{t^2}\right) \tan^{-1} \left( \frac{\cos \theta}{\sqrt{1 - \frac{b^2}{t^2} - \cos^2 \theta}} \right) \right) \right]_{\theta_b}^{\theta_{\max}} = \\ & \left[ -\frac{1}{2} \left( \cos \theta \sqrt{\sin^2 \theta - \frac{b^2}{t^2}} + \left(1 - \frac{b^2}{t^2}\right) \tan^{-1} \left( \frac{\cos \theta}{\sqrt{\sin^2 \theta - \frac{b^2}{t^2}}} \right) \right) \right]_{\theta_b}^{\theta_{\max}} \end{aligned}$$

whenever  $\theta_b < \pi/2$ , by using the fact that  $\frac{1}{2} \left( x \sqrt{c-x^2} + c \tan^{-1} \left( \frac{x}{\sqrt{c-x^2}} \right) \right)$  is an antiderivative of  $\sqrt{c-x^2}$  with respect to  $x$  when  $c$  is a constant. We obtain  $f(b, c, a)$  and  $f(c, a, b)$  by switching the roles of  $a, b, c$  in (25).

It remains to insert the limits  $\theta_{\min}, \theta_a, \theta_b, \theta_{\max}$  into the antiderivatives (23), (24) and (25) above. Noting that  $\theta_{\min}, \theta_a, \theta_b, \theta_{\max}$  are expressed in terms of piecewise-defined functions, the following manipulations will be useful. For any function  $\psi$ , we have

$$\begin{aligned} \psi(\theta_{\min}) &= \psi \left( \cos^{-1} \frac{c}{t} \right) \chi_c + \psi(\cos^{-1} 1)(1 - \chi_c) \\ &= \left( \psi \left( \cos^{-1} \frac{c}{t} \right) - \psi(0) \right) \chi_c + \psi(0) \end{aligned}$$

where  $\chi_c := \chi(t > c)$ . Similarly,

$$\psi(\theta_{\max}) = \left( \psi \left( \sin^{-1} \frac{\sqrt{a^2 + b^2}}{t} \right) - \psi(\pi/2) \right) \chi_{a,b} + \psi(\pi/2)$$

where  $\chi_{a,b} := \chi(\sqrt{a^2 + b^2} > t)$ , and

$$\begin{aligned} \psi(\theta_a) &= (1 - \chi_a) \psi(\pi/2) + (\chi_a - \chi_{a,c}) \psi \left( \sin^{-1} \frac{a}{t} \right) + \chi_{a,c} \psi \left( \cos^{-1} \frac{c}{t} \right) \\ &= \chi_{a,c} \cdot \left( \psi \left( \cos^{-1} \frac{c}{t} \right) - \psi \left( \sin^{-1} \frac{a}{t} \right) \right) + \chi_a \cdot \left( \psi \left( \sin^{-1} \frac{a}{t} \right) - \psi(\pi/2) \right) + \psi(\pi/2) \end{aligned}$$

and similarly,  $\psi(\theta_b)$  can be written as

$$\chi_{b,c} \cdot \left( \psi \left( \cos^{-1} \frac{c}{t} \right) - \psi \left( \sin^{-1} \frac{b}{t} \right) \right) + \chi_b \cdot \left( \psi \left( \sin^{-1} \frac{b}{t} \right) - \psi(\pi/2) \right) + \psi(\pi/2).$$

With this we can evaluate  $[\psi]_{\theta_{\min}}^{\theta_a}, [\psi]_{\theta_a}^{\theta_{\max}}, [\psi]_{\theta_b}^{\theta_{\max}}$ . But since we know that we will get a function symmetric with respect to the values  $a, b, c$ , it suffices to keep only those terms with  $\chi_a$  and  $\chi_{a,b}$ , say, and then the other terms may be evaluated by just switching the

order of  $a, b, c$ . Upon inserting the limits and differentiating, one obtains (after tedious calculations) that

$$\text{pdf}_X(t) = \frac{F(a, b, c, t) + F(b, c, a, t) + F(c, a, b, t)}{3\pi t^3(ab + ac + bc)}$$

where

$$\begin{aligned} F(a, b, c, t) &:= (8at^3 - 3t^4) + \\ \chi(t \geq a) &\left( (6t^4 - a^4 + 6\pi a^2 bc) - (8at^3 - 3t^4) - 4(b+c)\sqrt{|t^2 - a^2|}(a^2 + 2t^2) \right) + \\ \chi(t \geq \sqrt{a^2 + b^2}) &\left[ a^4 + b^4 - 9t^4 - 6a^2 b^2 + \sqrt{|t^2 - a^2 - b^2|} 4c(a^2 + b^2 + 2t^2) + \right. \\ &4a\sqrt{|t^2 - b^2|}(b^2 + 2t^2) - 12a^2 bc \cdot \arctan\left(\frac{\sqrt{|t^2 - a^2 - b^2|}}{b}\right) + \\ &\left. 4b\sqrt{|t^2 - a^2|}(a^2 + 2t^2) - 12ab^2 c \cdot \arctan\left(\frac{\sqrt{|t^2 - a^2 - b^2|}}{a}\right) \right]. \end{aligned}$$

Rewriting  $F$  as a piecewise function, we get Theorem (5).

## 5. Proof of Theorem 7

Consider the distribution of the random variable  $Y_{M,N}$ . Since we record the *same* number of bounces for each choice of angle  $\varphi$  we may replace the  $M$ -particle system with a one particle system  $Y_N$  as follows: randomly select, with uniform distribution, the angle  $\varphi$  and generate  $N$  bounce lengths and randomly select one of these bounce lengths (with uniform distribution); by the strong law of large numbers,  $Y_{M,N}$  converges in distribution to  $Y_N$  as  $M \rightarrow \infty$ .

We now determine the limit distribution of  $Y_N$ . As before, we first unfold the motion, and replace motion in a box with specular reflections on the walls with motion in  $\mathbb{R}^2$ ; see Figure 4. The path lengths between bounces is then the same as the lengths between the intersections with horizontal or vertical grid lines. To understand the spatial distribution, we project the dynamics to the torus  $\mathbb{R}^2/\Lambda$  where  $\Lambda$  is the lattice

$$\Lambda = \{(n_1 a, n_2 b) : n_1, n_2 \in \mathbb{Z}\},$$

and we may identify the torus with the rectangle  $[0, a] \times [0, b]$ .

Let us first consider the motion of a single particle with an arbitrary initial position, and direction of motion given by an angle  $\varphi$ . Taking symmetries into account, we may assume that  $\varphi \in [0, \pi/2]$ . (Note that  $\frac{d\varphi}{\pi/2}$  gives a probability measure on these angles.) If the particle travels a large distance  $R > 0$ , the number of intersections with horizontal, respectively vertical, grid lines is  $\frac{R \sin \varphi}{b} + O(1)$ , respectively  $\frac{R \cos \varphi}{a} + O(1)$ . Thus, in the limit  $R \rightarrow \infty$ , the probability of a line segment beginning at a horizontal (respectively

vertical) grid line is given by  $P_h$ , respectively  $P_v$  (here we suppress the dependence on  $\varphi$ ) where

$$P_h := \frac{\frac{\sin \varphi}{b}}{\frac{\sin \varphi}{b} + \frac{\cos \varphi}{a}}, \quad P_v := \frac{\frac{\cos \varphi}{a}}{\frac{\sin \varphi}{b} + \frac{\cos \varphi}{a}}.$$

The unfolded flow on the torus is ergodic for almost all  $\varphi$ , and thus the starting points of the line segments becomes uniformly distributed as  $R \rightarrow \infty$  for almost all  $\varphi$ .

Let

$$T = T(\varphi) := a / \cos \varphi.$$

Since  $\sin \varphi = \sqrt{T^2 - a^2} / T$ , we obtain that

$$P_h = \frac{\sqrt{T^2 - a^2}}{b + \sqrt{T^2 - a^2}}, \quad P_v = \frac{b}{b + \sqrt{T^2 - a^2}}.$$

Let  $\theta = \arctan b/a$  denote the angle of the diagonal in the box, and assume that  $0 \leq \varphi \leq \theta$ . We then observe the following regarding the line segment lengths.

First, if the segment begins at a horizontal line, it must end at a vertical line, and the possible lengths of these segment lie between 0 and  $T$ . We find that these lengths are uniformly distributed in  $[0, T]$  since the starting points of the segments are uniformly distributed.

On the other hand, if the line segment begins at a vertical line, it can either end at a vertical or horizontal line. Since the starting points are uniformly distributed, the former happens with probability

$$\frac{a \tan \varphi}{b} = \frac{a \frac{\sqrt{T^2 - a^2}}{a}}{b} = \frac{\sqrt{T^2 - a^2}}{b}$$

and the length of the segment is again uniformly distributed in  $[0, T]$ , whereas the latter happens with probability

$$\frac{b - a \tan \varphi}{b} = 1 - \frac{\sqrt{T^2 - a^2}}{b}$$

in which case the segment is always of length  $T$ .

Now,  $\varphi \in [0, \theta]$  implies that  $T \in [a, \sqrt{a^2 + b^2}]$ , and noting that

$$\frac{d\varphi}{dT} = \frac{a}{T \sqrt{T^2 - a^2}}$$

we find that the probability of observing a line segment of length  $t$  is the sum of a “singular part” (the segment begins and ends on vertical lines; note that all such segments have the *same* lengths) and a “smooth part” (the segment does not begin and end on vertical lines). Moreover, the smooth part contribution equals

$$\frac{1}{\pi/2} \int_{\max(a,t)}^{\sqrt{a^2 + b^2}} \frac{1}{T} \left( P_h + P_v \frac{a \tan \varphi}{b} \right) \frac{d\varphi}{dT} dT$$

which, on inserting (5), equals

$$\begin{aligned} \frac{1}{\pi/2} \int_{\max(a,t)}^{\sqrt{a^2+b^2}} \frac{1}{T} \cdot \left( \frac{\sqrt{T^2-a^2}}{b+\sqrt{T^2-a^2}} + \frac{b}{b+\sqrt{T^2-a^2}} \frac{a \tan \varphi}{b} \right) \cdot \frac{a}{T\sqrt{T^2-a^2}} dT = \\ \frac{1}{\pi/2} \int_{\max(a,t)}^{\sqrt{a^2+b^2}} \frac{1}{T} \cdot \left( \frac{\sqrt{T^2-a^2}}{b+\sqrt{T^2-a^2}} + \frac{b}{b+\sqrt{T^2-a^2}} \frac{\sqrt{T^2-a^2}}{b} \right) \cdot \frac{a}{T\sqrt{T^2-a^2}} dT = \\ \frac{1}{\pi/2} \int_{\max(a,t)}^{\sqrt{a^2+b^2}} \frac{2a}{b+\sqrt{T^2-a^2}} \cdot \frac{dT}{T^2}. \end{aligned}$$

On the other hand, the “singular part contribution”, provided  $t \geq a$ , to the probability of a segment having length  $t$  equals

$$\begin{aligned} \frac{P_v}{\pi/2} \cdot \frac{b-a \tan \varphi}{b} \cdot \frac{d\varphi}{dt} = \frac{1}{\pi/2} \cdot \frac{b}{b+\sqrt{t^2-a^2}} \cdot \left( 1 - \frac{\sqrt{t^2-a^2}}{b} \right) \cdot \frac{a}{t\sqrt{t^2-a^2}} = \\ \frac{1}{\pi/2} \cdot \frac{a}{t(b+\sqrt{t^2-a^2})\sqrt{t^2-a^2}} \cdot (b - \sqrt{t^2-a^2}). \end{aligned}$$

In case  $\theta \leq \varphi \leq \pi/2$ , a similar argument (we simple reverse the roles of  $a$  and  $b$ ) shows that the smooth contribution equals

$$\frac{1}{\pi/2} \int_{\max(b,t)}^{\sqrt{a^2+b^2}} \frac{2b}{a+\sqrt{T^2-b^2}} \cdot \frac{dT}{T^2}$$

and that the singular contribution (if  $t \geq b$ ) equals

$$\frac{1}{\pi/2} \cdot \frac{b}{t(a+\sqrt{t^2-b^2})\sqrt{t^2-b^2}} \cdot (a - \sqrt{t^2-b^2}).$$

Thus, if we let  $P_{\text{sing}}(t)$  denote the “singular contribution” to the probability density function we find the following: if  $t < a$ , then

$$P_{\text{sing}}(t) = 0$$

if  $t \in [a, b]$ , then

$$P_{\text{sing}}(t) = \frac{1}{\pi/2} \cdot \frac{a(b - \sqrt{t^2-a^2})}{t(b+\sqrt{t^2-a^2})\sqrt{t^2-a^2}}$$

and if  $t \in [b, \sqrt{a^2+b^2}]$ , then

$$P_{\text{sing}}(t) = \frac{1}{\pi/2} \cdot \left( \frac{a(b - \sqrt{t^2-a^2})}{t(b+\sqrt{t^2-a^2})\sqrt{t^2-a^2}} + \frac{b(a - \sqrt{t^2-b^2})}{t(a+\sqrt{t^2-b^2})\sqrt{t^2-b^2}} \right).$$

**Remark 26.** Note that  $P_{\text{sing}}$  has a singularity of type  $(t-a)^{-1/2}$  just to the right of  $t=a$  (and similarly just to the right of  $t=b$ ). In a sense this singularity arises from the singularity in the change of variables  $\varphi \mapsto T$  since  $\frac{d\varphi}{dT} = \frac{a}{T\sqrt{T^2-a^2}}$ . The reason for the singularities in the spreading model for  $n=2$  is similar, as the spreading model can be obtained from the absorption model by a smooth change of the angular measure.

Similarly, the ‘‘smooth part’’ of the contribution is (for  $t \in [0, \sqrt{a^2 + b^2}]$ ) given by

$$P_{\text{smooth}}(t) = \frac{1}{\pi/2} \left( \int_{\max(a,t)}^{\sqrt{a^2+b^2}} \frac{2a}{b + \sqrt{T^2 - a^2}} \cdot \frac{dT}{T^2} + \int_{\max(b,t)}^{\sqrt{a^2+b^2}} \frac{2b}{a + \sqrt{T^2 - b^2}} \cdot \frac{dT}{T^2} \right)$$

Hence the probability density function of the distribution of the segment length  $t$  is given by

$$\text{pdf}_Y(t) = P_{\text{sing}}(t) + P_{\text{smooth}}(t).$$

We will now evaluate  $P_{\text{smooth}}(t)$ . An antiderivative of  $\frac{2a}{b + \sqrt{T^2 - a^2}} \cdot \frac{1}{T^2}$  with respect to  $T$  for  $T \in (a, \sqrt{a^2 + b^2})$  is

$$\frac{2a(\sqrt{T^2 - a^2} - b)}{T(a^2 + b^2)} + \frac{2ab \left( \tanh^{-1} \left( \frac{T}{\sqrt{a^2 + b^2}} \right) - \tanh^{-1} \left( \frac{\sqrt{T^2 - a^2} \sqrt{a^2 + b^2}}{Tb} \right) \right)}{(a^2 + b^2)^{3/2}} \quad (27)$$

where  $\tanh^{-1}(z) = \frac{1}{2} \log \frac{1+z}{1-z}$  for  $|z| < 1$ . (A quick calculation shows that  $\frac{\sqrt{T^2 - a^2} \sqrt{a^2 + b^2}}{Tb} < 1$  whenever  $a < T < \sqrt{a^2 + b^2}$ .) We can rewrite (27) as

$$\frac{2a(\sqrt{T^2 - a^2} - b)}{T(a^2 + b^2)} + \frac{ab \log \left( \frac{(\sqrt{a^2 + b^2} + T)(Tb - \sqrt{T^2 - a^2} \sqrt{a^2 + b^2})}{(\sqrt{a^2 + b^2} - T)(Tb + \sqrt{T^2 - a^2} \sqrt{a^2 + b^2})} \right)}{(a^2 + b^2)^{3/2}}$$

By l’Hôpital’s rule we have

$$\lim_{T \rightarrow \sqrt{a^2 + b^2}^+} \frac{Tb - \sqrt{T^2 - a^2} \sqrt{a^2 + b^2}}{\sqrt{a^2 + b^2} - T} = \lim_{T \rightarrow \sqrt{a^2 + b^2}^+} \frac{b - \frac{T}{\sqrt{T^2 - a^2}} \sqrt{a^2 + b^2}}{-1} = \frac{a^2}{b}$$

so the limit of (27) as  $T \rightarrow \sqrt{a^2 + b^2}^+$  is

$$\frac{ab \log \left( \left( \frac{a^2}{b} \right) \cdot \frac{(\sqrt{a^2 + b^2} + \sqrt{a^2 + b^2})}{(b\sqrt{a^2 + b^2} + b\sqrt{a^2 + b^2})} \right)}{(a^2 + b^2)^{3/2}} = \frac{2ab \log \left( \frac{a}{b} \right)}{(a^2 + b^2)^{3/2}}.$$

The limit of (27) as  $T \rightarrow a^+$  is

$$\frac{-2b}{(a^2 + b^2)} + \frac{2ab \tanh^{-1} \left( \frac{a}{\sqrt{a^2 + b^2}} \right)}{(a^2 + b^2)^{3/2}}.$$

Thus, assuming  $a < b$ , we can write  $\frac{\pi}{2} P_{\text{smooth}}(t)$  as

$$\frac{2(a+b)}{(a^2 + b^2)} - \frac{2ab}{(a^2 + b^2)^{3/2}} \left( \tanh^{-1} \left( \frac{a}{\sqrt{a^2 + b^2}} \right) + \tanh^{-1} \left( \frac{b}{\sqrt{a^2 + b^2}} \right) \right)$$

if  $t < a, b$ , or as

$$\frac{2ab + 2at - 2a\sqrt{t^2 - a^2}}{t(a^2 + b^2)} + \frac{2ab\left(-\tanh^{-1}\left(\frac{t}{\sqrt{a^2+b^2}}\right) + \tanh^{-1}\left(\frac{\sqrt{t^2-a^2}\sqrt{a^2+b^2}}{tb}\right) - \tanh^{-1}\left(\frac{b}{\sqrt{a^2+b^2}}\right)\right)}{(a^2 + b^2)^{3/2}}$$

if  $a < t < b$  or as

$$2\frac{2ab - a\sqrt{t^2 - a^2} - b\sqrt{t^2 - b^2}}{t(a^2 + b^2)} + \frac{2ab\left(-2\tanh^{-1}\left(\frac{t}{\sqrt{a^2+b^2}}\right) + \tanh^{-1}\left(\frac{\sqrt{t^2-a^2}\sqrt{a^2+b^2}}{tb}\right) + \tanh^{-1}\left(\frac{\sqrt{t^2-b^2}\sqrt{a^2+b^2}}{ta}\right)\right)}{(a^2 + b^2)^{3/2}}$$

if  $a, b < t$ . Adding  $P_{\text{sing}}(t)$  to this, we get Theorem 7.

## A. Calculation of an integral

**Lemma 28.** Write  $|\mathbb{S}^{n-1}|$  for the  $(n-1)$ -dimensional surface area of the sphere  $\mathbb{S}^{n-1} \subseteq \mathbb{R}^n$ . Then we have

$$\int_{\mathbb{S}_+^{n-1}} v_n dS(v) = \frac{1}{\pi} \frac{|\mathbb{S}^n|}{2^n}.$$

where  $\mathbb{S}_+^{n-1} := \mathbb{S}^{n-1} \cap (0, \infty)^n$  is the part of the sphere  $\mathbb{S}^{n-1}$  with positive coordinates.

*Proof.* We may parametrize  $v = (v_1, \dots, v_n) \in \mathbb{S}_+^{n-1}$  with

$$\begin{aligned} v_1 &= \cos \theta_1 \\ v_2 &= \sin \theta_1 \cos \theta_2 \\ v_3 &= \sin \theta_1 \sin \theta_2 \cos \theta_3 \\ &\vdots \\ v_{n-1} &= \sin \theta_1 \cdots \sin \theta_{n-2} \cos \theta_{n-1} \\ v_n &= \sin \theta_1 \cdots \sin \theta_{n-2} \sin \theta_{n-1} \end{aligned}$$

for  $\theta_1, \dots, \theta_{n-1} \in (0, \pi/2)$ . We have the spherical area element

$$dS(v) = \sin^{n-2} \theta_1 \sin^{n-3} \theta_2 \cdots \sin \theta_{n-2} d\theta_1 \cdots d\theta_{n-1}.$$

Thus we get

$$\int_{\mathbb{S}_+^{n-1}} v_n dS(v) = \prod_{i=1}^{n-1} \int_0^{\pi/2} \sin^{n-1-i} \theta_i d\theta_i.$$

Introducing an additional integration variable  $\theta_n$ , we recognize the integrand as the spherical area element in  $n + 1$  dimensions, and thus the above is

$$\frac{1}{\int_0^{\pi/2} d\theta_n} \prod_{i=1}^n \int_0^{\pi/2} \sin^{n-1-i} \theta_i d\theta_i = \frac{1}{\pi/2} \frac{|\mathbb{S}^n|}{2^{n+1}}.$$

since  $\int_{\mathbb{S}_+^n} dS(v) = |\mathbb{S}^n|/2^{n+1}$ .  $\square$

## References

- [1] F. Barra and P. Gaspard. On the level spacing distribution in quantum graphs. *J. Statist. Phys.*, 101(1-2):283–319, 2000.
- [2] C. Boldrighini, L. A. Bunimovich, and Y. G. Sinai. On the Boltzmann equation for the Lorentz gas. *J. Statist. Phys.*, 32(3):477–501, 1983.
- [3] J. Bourgain, F. Golse, and B. Wennberg. On the distribution of free path lengths for the periodic Lorentz gas. *Comm. Math. Phys.*, 190(3):491–508, 1998.
- [4] L. A. Bunimovich and Y. G. Sinai. Statistical properties of Lorentz gas with periodic configuration of scatterers. *Comm. Math. Phys.*, 78(4):479–497, 1980/81.
- [5] M. Bäckström, S. Holmin, P. Kurlberg, D. Månsson. *Randomized Ray Tracing for Modeling UWB Transients in a Reverberation Chamber*. In preparation.
- [6] N. Chernov. Entropy, Lyapunov exponents, and mean free path for billiards. *J. Statist. Phys.*, 88(1-2):1–29, 1997.
- [7] F. Golse and B. Wennberg. On the distribution of free path lengths for the periodic Lorentz gas. II. *M2AN Math. Model. Numer. Anal.*, 34(6):1151–1163, 2000.
- [8] D. A. Klain and G.-C. Rota. *Introduction to geometric probability*. Lezioni Lincee. [Lincei Lectures]. Cambridge University Press, Cambridge, 1997.
- [9] D. Månsson, personal communication.
- [10] J. Marklof and A. Strömbergsson. The distribution of free path lengths in the periodic Lorentz gas and related lattice point problems. *Ann. of Math. (2)*, 172(3):1949–2033, 2010.
- [11] J. Marklof and A. Strömbergsson. The Boltzmann-Grad limit of the periodic Lorentz gas. *Ann. of Math. (2)*, 174(1):225–298, 2011.
- [12] J. Marklof and A. Strömbergsson. Gaps between logs. *Bull. Lond. Math. Soc.*, 45(6):1267–1280, 2013.
- [13] J. Marklof and A. Strömbergsson. Free path lengths in quasicrystals. *Comm. Math. Phys.*, 330(2):723–755, 2014.

- [14] L. A. Santaló. *Integral geometry and geometric probability*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, second edition, 2004. With a foreword by Mark Kac.
- [15] H. Spohn. The Lorentz process converges to a random flight process. *Comm. Math. Phys.*, 60(3):277–290, 1978.
- [16] B. Wennberg. Free path lengths in quasi crystals. *J. Stat. Phys.*, 147(5):981–990, 2012.