# Validity of Boltzmann Approximation for Kinetic Annihilation and Short Time Kinetics

Miloš Tašić

2020-2021

## Contents

**Abstract**

The aim of this project is to look a paper Karsten Matthies and Florian Theil, VALIDITY AND FAILURE OF THE BOLTZMANN APPROXIMATION OF KINETIC ANNIHILATION, with the purpose of giving an explanation that would be accessible to an undergraduate in mathematics and in doing so expanding on most proofs, so that they become more accessible. Additionally we will go into great detail so that I can use this paper as a reference later on. Furthermore we will talk about what the Boltzmann equation is trying to describe.

# 1 Introduction

We wish to describe the motion of a rarefied gas, consisting of a very large number of identical particles moving in $\mathbb{R}^3$ (or more generally $\mathbb{R}^3$). For $(t, u, v) \in [o, \infty) \times \mathbb{T}^3 \times \mathbb{R}^3$, we consider a function,

$$f(t, u, v),$$

describing the *density of particles* at time $t$, at a point $u$ with velocity $v$. In addition, we may also think of $f(t, u, v)$ as the probability of finding a particle with velocity $v$ at time $t$ at position $u$.

If no collisions occur, the velocity $v$ of each particle will remain constant in time. A particle with velocity $v$ and located at position $u$ at time $t = 0$ will move to $u + \tau v$ at a later time $\tau$. Therefore,

$$f(\tau, u, v) = f(0, u - \tau v, v).$$

In this case $f$ provides a solution to the linear transport equation,

$$\frac{\partial}{\partial t} f(t, u, v) + v \cdot \nabla_u f(t, u, v) = 0,$$

this is simply a result of the chain rule. The presence of collisions accounts for an additional quadratic term on the right hand side of the transport equation yielding the Boltzmann equation.

$$\frac{\partial}{\partial t} f(t, u, v) + v \cdot \nabla_u f(t, u, v) = Q[f, f] = Q_+[f, f] - Q_-[f, f].$$

where $Q$ is known as the *Collision Kernel* and $Q_+$ is the *Gain term* and $Q_-$ is the *Loss term*

## 1.1 Some simple examples and deducing Boltzmann equation

The Boltzmann equation describes the behaviour of a continuum in which there are infinitely many particles. Each particle under goes a Poisson number of collisions in time $t$. The Poisson parameter is given by $t\lambda(u, v)$ where $u \in \mathbb{T}^d$ and $v \in \mathbb{R}^d$. Hence the Poisson point process is an inhomogeneous one, depending on position and velocity. Finally we assume that for any two distinct particles the random variables that give the number if collisions in time $[0, T]$ are independent random variables. This is contrary to what one might expect, for say, a finite number of snooker balls on a snooker table, the number of collisions of one ball will be dependent on the number of collisions of another ball. However, in the limit as the number of snooker balls goes to infinity and their diameter goes to zero at some certain rate one expects to gain Independence. This is the so called *Stoßzahlansatz* or *Propagation of Chaos.* In fact this last assumption of the Stoßzahlansatz is exactly the thing that needs to be proved in the rigorous derivation of the Boltzmann equation and is the heart of our problem.

One can see that the Boltzmann equation just the time derivative of a Poisson distribution. We will loosely justify this for the case that $\lambda$ is independent of the $u$ variable, that is to say we are working the in spatially homogeneous setting. It is not too hard to believe that the *transport term* $v \cdot \nabla_u f$ is just an artifact of the chain rule once the spatial in-homogeneity is re-introduced.

First consider a Poisson point process which jumps between two states, that is to say the state space is $I = \{0, 1\}$. The holding times/ jump times are independent and identically distributed $\exp(\lambda)$ random variables. Let $X_t \in \{0, 1\}$ be the state of the point at time t. Recall the following basic theorem,

**Theorem 1.1.** The following are equivalent:

1. The process $(\tilde{X}_t)_{t\geq 0}$ taking values in $I = \mathbb{N}_0$ is a Poisson process with rate $\lambda$ if $S_1, S_2, \ldots$ are i.i.d. $\exp(\lambda)$ random variables and $Y_n = n$ where $S_n = J_n - J_{n-1}$ where $J_n$ is the n$^{th}$ jump time and $Y_n = \tilde{X}_{J_n}$

2. The process $(\tilde{X}_t)_{t\geq 0}$ has independent increments, $\tilde{X}_0 = 0$ and

$$\mathbb{P}(\tilde{X}_{t+h} - \tilde{X}_t = 0) = 1 - \lambda h + o(h) \text{ as } h \to 0$$
$$\mathbb{P}(\tilde{X}_{t+h} - \tilde{X}_t = 1) = \lambda h + o(h) \text{ as } h \to 0$$

   uniformly for $t \geq 0$

3. The process $(\tilde{X}_t)_{t\geq 0}$ has stationary and independent increments for all $t > 0$ and $\tilde{X}_t \sim \text{Poisson}(\lambda t)$

▲

What we have in our set is a process $X_t$ that is 1 if $\tilde{X}_t$ is an odd value and 0 if $\tilde{X}_t$ is an even value. The important part of the above theorem is part (2) and it still holds. Furthermore, we expect $X_t$ to have independent increments. From this we can conclude the following,

$$\mathbb{P}(X_{t+h} = 1) = \mathbb{P}(X_{t+h} = 1, \text{ no jumps in } [t, t+h]) + \mathbb{P}(X_{t+h} = 0, \text{ exactly one jump in } [t, t+h]) + o(h)$$

as $h \to 0$. By independent increments and part (2) of the previous theorem we know that,

$$\mathbb{P}(X_{t+h} = 1) = 1 - \lambda h \mathbb{P}(X_t = 1) + \lambda h \mathbb{P}(X_{t+h} = 0) + o(h)$$

and so from this it is easy to see that,

$$\frac{d}{dt} P_1(t) = \lambda P_0(t) - \lambda P_1(t)$$

where we define $P_1(t) = \mathbb{P}(X_t = 1)$ and $P_0(t) = \mathbb{P}(X_t = 0)$. Similarly one can show,

$$\frac{d}{dt} P_0(t) = \lambda P_1(t) - \lambda P_0(t)$$

Heuristically, the rate at which the probability of being in state 1 at time $t$ is equal to the chance of being in state 0 multiplied by the rate at which you leave state 0 (and hence go to state 1) minus the chance of being in state 1 multiplied by the rate if leaving state 1.

Instead of having $\{0, 1\}$ as being your state space let $\mathbb{N}_0$ be your state space and suppose one jumps from state $i$ to $j$ with rate $\lambda_{ij}$, i.e.

$$\mathbb{P}(X_t = j | X_0 = i) = 1 - e^{-\lambda_{ij} t}$$

that is to say the above law is distributed $exp(\lambda_{ij})$. As before define $P_i(t) = \mathbb{P}(X_t = i)$ then in a similar fashion as before one can show that,

$$\frac{d}{dt} P_i(t) = \sum_{j \in \mathbb{N}_0} \lambda_{ji} P_j(t) - \sum_{i \in \mathbb{N}_0} \lambda_{ij} P_i(t)$$

Considering simple functions and taking limits (i.e. using standard measure theory machinery) it is reasonable to expect if you have and uncountable state space and you jump from state $x$ to state $y$ with rate $K(x, y)$ that,

$$\frac{d}{dt} f(t, x) = \int K(y, x) f(t, y) \, \mathrm{d}y - \int K(x, y) f(x, t) \, \mathrm{d}y \tag{$\heartsuit$}$$

where $f(t,x) = \mathbb{P}(X_t = x)$ in a loose sense. Now consider $f(t,v)$ to be the density of particles at time $t$ with velocity $v$. We further assume this function is spatially homogeneous i.e. $f(t,v) = f(t,u,v)$ for all $t \geq 0$ and $v \in \mathbb{R}^d$. Next define,

$$v' = v + \omega \cdot (v_1 - v)\omega, \tag{$\star$}$$

$$v_1' = v_1 - \omega \cdot (v_1 - v)\omega, \tag{$\star\star$}$$

where $\omega \in \mathbb{S}^{d-1}$. When a particle with velocity $v$ and $v_1$ collide their new velocities are given by ($\star$) and ($\star\star$) (this will be explained in the next subsection). The converse also holds. The second term given in ($\heartsuit$) is the total rate at which the tagged particle changes velocity that is to say that,

$$\int K(x,y)\,\mathrm{d}y = \int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^d} ((v - v_1)\cdot\omega)_+ f(t,v_1)\,\mathrm{d}v_1\,\mathrm{d}\omega,$$

hence the name, *loss term*. Why this is the case we will explain later if time permits. The second term in ($\heartsuit$) can be thought of as the total rate of other particles colliding with our tagged particle and causing our tagged particle to change velocities and obtain a new velocity which is precisely $v$. But note unlike the loss term we cant have any collisions we must have specific collision velocities and angles in order to get our velocity $v$. That is why our colliding particles must have pre-collisionsal velocities $v'$ and $v_1'$ for any given $\omega$. Consequently we expect

$$\begin{aligned}\int K(y,x) f(t,y)\,\mathrm{d}y &= \int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^d} ((v' - v_1')\cdot\omega)_+ f(t,v_1') f(t,v')\,\mathrm{d}v_1\,\mathrm{d}\omega \\ &= \int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^d} ((v - v_1)\cdot\omega)_+ f(t,v_1') f(t,v')\,\mathrm{d}v_1\,\mathrm{d}\omega\end{aligned}$$

Thus we get the *homogeneous* Boltzmann equation

$$\frac{d}{dt} f(t,v) = \int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^d} ((v - v_1)\cdot\omega)_+ [f(t,v_1')f(t,v') - f(t,v)f(t,v_1)]\,\mathrm{d}v_1\,\mathrm{d}\omega$$

Now suppose that we loose the spatial homogeneity it is not too far fetched (due to the chain rule) to expect to see that,

$$\frac{\partial}{\partial t} f(t,u,v) + v\cdot\nabla_u f(t,u,v) = \int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^d} ((v - v_1)\cdot\omega)_+ [f(t,u,v_1')f(t,u,v') - f(t,u,v)f(t,u,v_1)]\,\mathrm{d}v_1\,\mathrm{d}\omega,$$

the *inhomogeneous* Boltzmann equation.

## 1.2 The Collision Kernel and Particle Ballistics

In this section we will explain where ($\star$) and ($\star\star$) come from at first and then try to build a heuristic explanation of where the collision a operator comes from. This section might be a bit annoying for those of you who like rigour as we won't be able to justify things rigorously but rather do an infinitesimal approach argument. The reason of course being because the rigorous explanation for the Boltzmann equation, for large times, is still an open question for most people.

### 1.2.1 Particle Ballistics

We model the behaviour of the particles on a microscopic level as *hard spheres* with fixed velocities (at least until they collide) and further impose that the collisions are *elastic*. By hard we me the spheres do not deform at the point of collision and always remain perfectly spherical no mater the collision and by elastic we mean no energy is lost, to say heat, upon collision.

If two hard spheres move in (for simplicity) in $\mathbb{R}^3$, by angle of collision we mean the unit vector $\omega$ parallel to the segment joining the two centres of two particles at the instant of collision. If two particles collide at an angle $\omega$ then the components of their velocities along $\omega$ will be exchanged, while the components perpendicular to $\omega$ will remain the same. Let $v$ and $v_1$ be the velocities before the collision occurs and $v'$ and $v_1'$ the respective velocities after the collision then we simply get as stated before,

$$v' = v + \omega \cdot (v_1 - v)\omega, \tag{$\star$}$$
$$v_1' = v_1 - \omega \cdot (v_1 - v)\omega, \tag{$\star\,\star$}$$

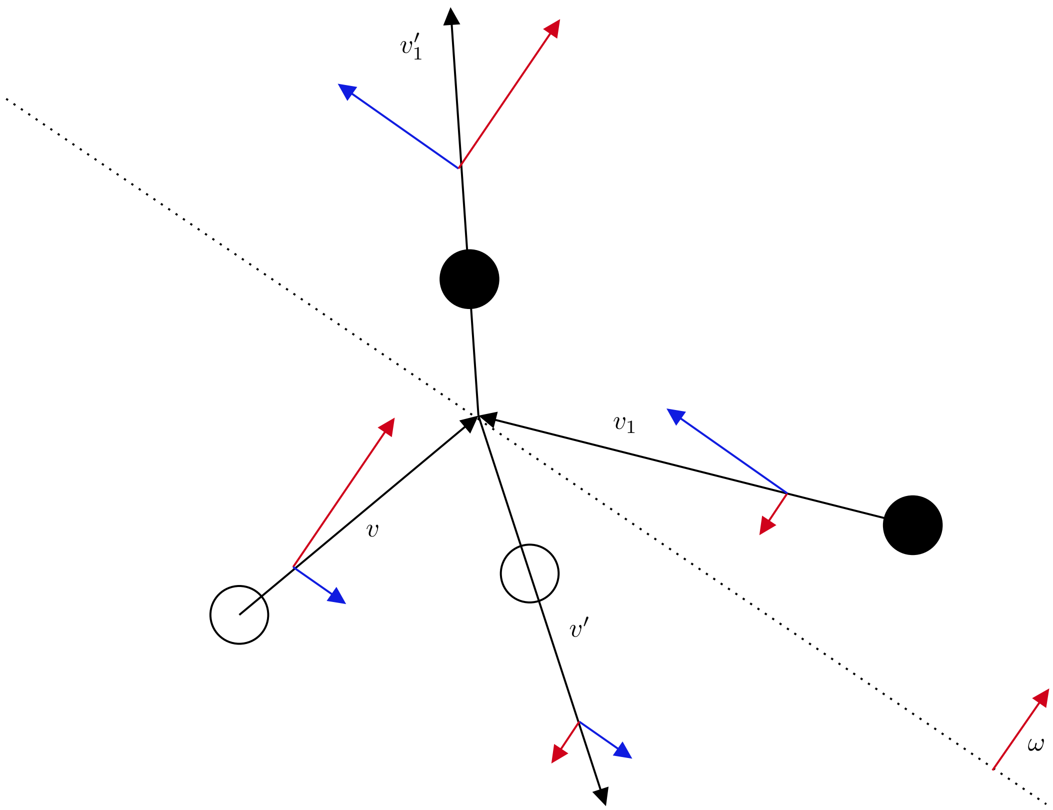



Fig: A figure depicting two particles colliding with velocities $v$ and $v'$. The velocities are broken down into their component relative to $\omega$. It is clear to see that after the collision only componets parallel to $\omega$ (i.e. the red arrows) are exchanged.

Notice that a change in the sign of $\omega$ does not affect the linear transformation given by ($\star$) and ($\star\star$). Next setting,

$$V = v - v_1, \quad V_0 = \frac{v + v_1}{2}, \quad \rho = \frac{|v - v_1|}{2},$$

an alternative construction of the incoming and outgoing velocities is illustrated in the next figure. Consider the sphere,

$$S_{vv_1} := \left\{ v* \in \mathbb{R}^3 \;\middle|\; |v * - V_0| = \rho \right\}$$

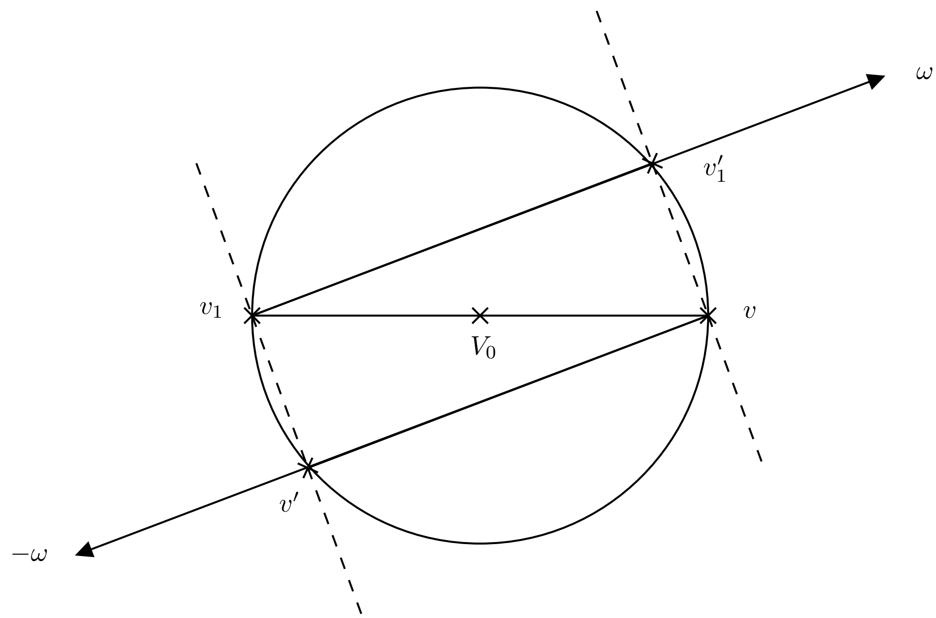



Fig: The surface of the ball $B(V_0, \rho)$ having segment $vv_1$ as a diameter.

If two particles have a collision at an angle $\omega$, the velocities $v'$ and $v_1'$ of the two outgoing particles are obtained as the intersection of the surface of the sphere $S_{vv_1}$ with the lines through $v$ and $v_1$ that are parallel to $\omega$.

#### 1.2.2 The Collision Kernel

So far we seem to have plucked the collision Kernel out of thin air. However, there are heuristic explanations of this Kernel and we endeavour now to give ourselves one.

Let $S^2 := \{n \in \mathbb{R}^3 \mid |n| = 1\}$ be the surface of the unit ball. To find the probability that the collision occurs at an angle $n$ it is convenient to take a different frame of reference. The reference we us is the one in which one particle is at rest, say the particle with velocity $v_1$, while the incoming particle has velocity $v_2 = v - v_1$. Furthermore, instead of considering two hard spheres of radius $r$, we equivalently think of a point hitting a hard sphere of radius $2r$.

# 2 Validity and Failure For Kinetic Annihilation for the Spatially Homogeneous Case

## 2.1 Introduction

In this Chapter we turn to the paper; "Validity and Failure of the Boltzmann approximation of Kinetic Annihilation" [1]. Our aim will be to show that hard ball dynamics undergoing kinetic annihilation, in the limit of the number of initial particles goes to infinity, can be described by the gainless Boltzmann equation. That is to say if $f_t(u, v)$ is viewed as the density of particles at position $u$ with velocity $v$ at time $t$ that undergo kinetic annihilation, then one expects that $f_t(u, v)$ satisfies the *gainless* Boltzmann equation,

$$\partial_t f + v \cdot \partial_u f = Q_-[f, f], \tag{1}$$

where, $Q_-[f, g] = -\kappa_d f(v) \int_{\mathbb{R}^d} |v - v'| \; \mathrm{d}g(v')$ and $\kappa_d$ is the volume of the $d - 1$ dimensional ball of radius 1.

Roughly speaking in both this chapter and the next, we will study some sort of empirical distribution that looks at the probabilistic behaviour of $N$ particles, with diameter $a$, living on the unit torus. These particles conform to the laws set out by Newton for non-accelerating particles and annihilate upon collision. Additionally we work with the supposition that $N$ is a random number and is given by the size of $\omega$, which is Poissonion with intensity $n$. Furthermore, the radius and intensity are linked via the equation,

$$na^{d-1} = 1, \tag{2}$$

which is know as the *Boltzmann-Grad Scaling limit*. We will then take $n$ to infinity and show this empirical distribution in some sense converges to some other idealised distribution, which is a distribution that is somehow induced by the gainless Boltzmann equation (1).

The content of this chapter will be to attack the discussed problem but from the point of view that the particles are homogeneous in the spatial dimension i.e. $f_t(u,v) = f_t(v)\ \forall u \in \mathbb{T}^d$. We'd like to show that for a tagged particle with position and velocity at time $t$ given by $(u(t), v(t))$ and $A \subseteq \mathbb{T}^d \times \mathbb{R}^d$, Borel, that,

$$\lim_{a\to 0} \text{Prob}((u(t), v(t)) \in A \text{ and the tagged particle is intact at time } t) = \frac{1}{f_0\left(\mathbb{R}^d\right)} \int_A \mathrm{d}u\ \mathrm{d}f_t(v). \tag{3}$$

It is beyond the intended scope of this work to discuss the following in detail, but if one has the following,

$$\lim_{a\to 0} \text{Prob}\left(\left|\frac{1}{n}\#\{i \mid (u_i(t), v_i(t)) \in A \text{ and particle } i \text{ is intact at time } t\} - \int_A \mathrm{d}u\ \mathrm{d}f_t(v)\} > \varepsilon\right|\right) = 0, \tag{4}$$

then by the strong law or large numbers (3) necessarily holds [4].

## 2.2 The Main Results

As is usually the case we need to introduce some definitions and unpack them, before we can proceed to state the main theorem of this chapter.

Suppose we have $N$ particles in our state space with initial values of the $i^{\text{th}}$ given by,

$$(u(i, t=0), v(i, t=0)) = (u_0(i), v_0(i)) \in \mathbb{T}^d \times \mathbb{R}^d.$$

As discussed these particles will follow Newton's laws and never accelerate and so $\forall i \in \{1, \dots N\}$ we have,

$$\frac{\partial}{\partial t} u(i,t) = v(i,t), \qquad \frac{\partial}{\partial t} v(i,t) = 0. \tag{5}$$

**Definition 2.1** (Particle Distance Metric)**.**
We define a metric that gives us the distance between any two particles at time $t$ but ignores initial intersections that have existed continuously up to time $t$.

$$d\left((u,v), (u', v'), t\right) = \begin{cases} 2a & \text{if } |u-u'|_{\mathbb{T}^d} < a \text{ and } |u - u' + s(v-v')|_{\mathbb{T}^d} < a\ \forall s \in [0,t) \\ |u-u'+t(v-v')|_{\mathbb{T}^d} & \text{else} \end{cases} \tag{6}$$

Where we define $|u|_{\mathbb{T}^d} := \inf_{k\in\mathbb{Z}^d} |u-k|$, which is the euclidean distance on the unit torus. ▲

**Remark 2.2.** The first condition fixes the distance to $2a$ for particles that are initially realised to be intersecting and remain intersecting up to time $t$. For example suppose two particles are intersecting at time zero and their velocities have a dot product sufficiently close to 1, then they will remain intersected for a very long time and hence have distance $2a$ at time $t$. ▲

**Definition 2.3** (Scattering State Indicator)**.**
We define a function which tells us whether a given particle $i$ has been annihilated by time $t$.

$$\beta_i^{(a)}(t) = \begin{cases} 1 & \text{if } d(z_i, z_j, s) > a\beta_j^{(a)}(s) \text{ and } |u_i^0 - u_j^0| > a \text{ for all } s \in [0,t), i \neq j \\ 0 & \text{otherwise} \end{cases}. \tag{7}$$

▲

Now we will define and comment on the tagged Poisson point process which is how we are going to define the initial data of our $N$ particles.

**Definition 2.4** (Tagged Poisson Point Process)**.**
Let $\Omega$ be a locally compact metric space. The tagged particle $z_1$ is an independent random variable with law $\mu/\mu(\Omega)$. The random variable $\tilde{z} \in \bigcup_{r=0}^{\infty} \Omega^r$ forms a Poisson point process with density $\mu \in M_+(\Omega)$ (non-negative Radon measures, i.e. positive elements of $\left(C_c^0(\Omega)\right)^*$) if

$$\operatorname{Prob}\left(\tilde{z} \in \Omega^r\right) = e^{-\mu(\Omega)} \frac{\mu(\Omega)^r}{r!}, \quad \operatorname{law}\left(\tilde{z}_i\right) = \mu/\mu(\Omega)$$

and $\tilde{z}_1, \ldots, \tilde{z}_r$ are independent. Now letting $N = r+1 \in \{1, 2, \ldots\}$, realizations of the tagged Poisson point process (tppp) are obtained by letting $z = (z_1, \ldots, z_N) = (z_1, \tilde{z})$, i.e. one obtains for symmetric $A \subseteq \bigcup_{N=1}^{\infty} \Omega^N$ that

$$\operatorname{Prob}_{\text{tppp}}\left((z_1, \tilde{z}) \in A\right) = \operatorname{Prob}_{\text{tppp}}\left((z_1, \ldots, z_N) \in A\right) = \frac{1}{\mu(\Omega) e^{\mu(\Omega)}} \sum_{N=1}^{\infty} \frac{1}{(N-1)!} \int_{A \cap \Omega^N} \mathrm{d}\mu\left(z_1\right) \ldots \mathrm{d}\mu\left(z_N\right)$$

▲

**Remark 2.5** (Comments on Definition 2.4)**.**

1. Note that $N$ is a random number. We have a tagged point $z_1$ and then according to the Poisson Point Process we generate $r = N-1$ extra particles.

2. The formula at the end of the definition is a way to get a distribution which has the tagged particle uniformly distributed w.r.t. the measure $\mu$ and the rest of the particles being obtained by a Poisson Point Process of intensity $\mu$.

3. If we take $A \subseteq \Omega^r$ for some fixed $r > 0$ then the infinite sum disappears and only the term $N = r+1$ survives.

4. Recall that the Poisson point process is *as random as it can be*, by this we mean that the distribution of a particle is uniform conditional on the event that, that the particle of interest exists.

5. Combining points 2. and 4. we can view the second formula in definition 2.3 as being the distribution for a Poisson point process conditional on the event that there is at least one particle.

6. The integral is only calculating the size of $A$ with respect to the measure $\mu$ in each "direction". For example let $B \subseteq \mathbb{T}^d \times \mathbb{R}^d$ and define $A = \bigcup_{r=0}^{\infty} B \times \Omega^{r-1}$ then

$$\begin{aligned}
\operatorname{Prob}_{\text{tppp}}\left((z_1, \tilde{z}) \in A\right) &= \frac{1}{\mu(\Omega) e^{\mu(\Omega)}} \sum_{N=1}^{\infty} \frac{1}{(N-1)!} \int_{A \cap \Omega^N} \mathrm{d}\mu\left(z_1\right) \ldots \mathrm{d}\mu\left(z_N\right) \\
&= \frac{1}{\mu(\Omega) e^{\mu(\Omega)}} \sum_{N=1}^{\infty} \frac{1}{(N-1)!} \int_{\Omega^{N-1}} \int_B \mathrm{d}\mu\left(z_1\right) \ldots \mathrm{d}\mu\left(z_N\right) \\
&= \frac{\mu(B)}{\mu(\Omega)} e^{-\mu(\Omega)} \sum_{N=1}^{\infty} \frac{1}{(N-1)!} \int_{\Omega^{N-1}} \mathrm{d}\mu\left(z_2\right) \ldots \mathrm{d}\mu\left(z_N\right) \\
&= \frac{\mu(B)}{\mu(\Omega)}
\end{aligned}$$

We have now given an example that illustrates that the tagged Poisson point process behaves like a Poisson point process conditional on there being at least one particle. ▲

**Theorem 2.6** (Validity of the Gainless Boltzmann Equation (Homogeneous Version))**.**
Suppose that $f_0 \in PM(\mathbb{R}^d)$, then further suppose that, $f_0$ has finite total mass and Kinetic energy

$$\int_{\mathbb{R}^d} (1+|v|)^2 \, \mathrm{d}f_0(v) = K_{\text{ini}} < \infty \tag{8}$$

and that $f_0$ does not concentrate any mass on lines,

$$\int_{\rho(v,\nu)} \mathrm{d}f_0\left(v'\right) = 0 \text{ for all } v \in \mathbb{R}^d, \nu \in S^{d-1}, \tag{9}$$

where $\rho(v,\nu) = v + \nu\mathbb{R}^d$. Let $\omega \subset \mathbb{T}^d \times \mathbb{R}^d$ be a realisation of the *tagged Poisson point process* with intensity $\mu = n(\mathbb{1}_{\mathbb{T}^d} \otimes f_0)$ where $\mathbb{1}_{\mathbb{T}^d}$ is the Lebesgue measure on the unit torus. Define $N := \#\omega$ and further suppose we have $N$ particles adhering to (5) with initial positions and velocities given by $\omega$. Then for all $t \in [0,\infty)$,

$$\lim_{a\to 0} \sup_{A \subset \mathbb{T}^d \times \mathbb{R}^d \text{ Borel}} \left| \text{Prob}_{\text{tppp}} \left( z(1,t) \in A \text{ and } \beta^{(a)}(1,t) = 1 \right) - \int_A \mathrm{d}u \; \mathrm{d}f_t(v) \right| = 0, \tag{10}$$

where $f : [0,\infty) \to PM\left(\mathbb{R}^d\right)$ is the unique solution of the homogeneous, gainless Boltzmann equation

$$\dot{f} = Q_-[f,f], \quad f_{t=0} = f_0 \tag{11}$$

with $Q_-[f,f](v) = -\int_{\mathbb{R}^d} \kappa_d \left|v - v'\right| f(v) \; \mathrm{d}f\left(v'\right)$, and $\kappa_d$ the volume of $d-1$ dimensional unit-ball. ▲

**Remark 2.7.**

1. In the next chapter we will consider $f_0 \in PM(\mathbb{T}^d \times \mathbb{R}^d)$ and hence drop the spatial homogeneity condition by taking a Semi-group approach.
2. The assumption given by (8) is pretty standard to use in the derivation of the Boltzmann Equation.
3. The assumption given by (9) seems to be new and is indeed necessary, we shall show this in the last subsection of this chapter. Note that this assumption does not exclude concentrations on lower dimensional subsets that aren't lines, e.g. spheres or planes.
4. The equation on line (11) seems to be missing a term in the gainless Boltzmann equation (cf. (1)). However, by making the assumption that we are working in the spatially homogeneous setting the solution the the Boltzmann equation is independent of the spatial variable and ergo, we drop the $\partial_u f$ term.
5. As we will see in a later lemma, by (11) we do indeed mean a strong solution.

▲

**Corollary 2.8.** The measures $\mathbf{1}_{\mathbb{T}^d} \otimes f_t$

$$\mathrm{d}\hat{f}_t^{(a)}(u,v) = \text{Prob}_{\text{tppp}} \left( z(1,t) \in [u, u+\mathrm{d}v) \times [v, v+\mathrm{d}v) \text{ and } \beta^{(a)}(1,t) = 1 \right)$$

are both absolutely continuous with respect to $\mathbf{1}_{\mathbb{T}^d} \otimes f_0$. Furthermore,

$$\lim_{a\to 0} \hat{f}_t^{(a)} = \mathbf{1}_{\mathbb{T}^d} \otimes f_t \tag{12}$$

in the $L^1\left(\mathbf{1}_{\mathbb{T}^d} \otimes f_0\right)$ norm. ▲

**Remark 2.9.**
This $[u, u+du) \times [v, v+dv)$ notation just means the probability the tagged particle has position and velocity in some subset of $\mathbb{T}^d \times \mathbb{R}^d$. ▲

## 2.3 Proof of the Main Results

For now we wish to look at the system of PDEs given by,

$$\frac{\partial f_k}{\partial t} = Q_-[f_{k-1}, f_k] = -\kappa_d f_{k-1}(v) \int_{\mathbb{R}^d} |v - v'| \; \mathrm{d}f_k(v'), \tag{13}$$

$$f_{t=0,k} = f_0 \tag{14}$$

In some sense you can think about $f_k$ only considering the behaviour of particles that have had up to $k$ collisions in its "history" and ignores all collisions after the $k^{\text{th}}$ "ancestor". Furthermore, we know exactly what $f_{t,k}$ is, it is precisely,

$$f_{t,k} = \exp\left( - \int_0^t L[f_{s,k-1}] \, \mathrm{d}s \right) f_0, \tag{15}$$

where,

$$L[f](v) = \kappa_d \int_{\mathbb{R}^d} |v - v'| \; \mathrm{d}f(v').$$

Indeed it is simple to check this. Firstly,

$$f_{0,k} = \exp\left( - \int_0^0 L[f_{s,k-1}] \, \mathrm{d}s \right) f_0 = f_0,$$

$$\frac{\partial f_k}{\partial t} = L[f_{s,k-1}] \exp\left( - \int_0^t L[f_{s,k-1}] \, \mathrm{d}s \right) f_0 = Q_-[f_{k-1}, f_k].$$

It is easy to see from the definition of $f_{t,k}$ that, $f_{t,k}$ is absolutely continuous with respect to $f_0$.

Now we introduce a Lemma, which along with another proposition, will let us piece everything together in the end. This Lemma will let us pass from the solution of the gainless Boltzmann to a "truncated" solution of the Boltzmann (i.e. $f_{t,k}$). Subsequently, a proposition, that we'll prove and state later, will allow us to pass from these "truncated solutions" to the so called *Idealised* distribution. This is important because a lot of the work we will concern ourselves with, is, to show the convergence of the distribution of the many body dynamics, called the *empirical* distribution, to the *idealised* distribution. Emerging from this convergence of distributions will come the proof of Theorem 2.5 (Validity of the Gainless Boltzmann Equation).

Before we sate the lemma we define the following spaces.

- Suppose $w \in C(\mathbb{R})$ and $w \geq 0$ be a weight. Then for a Radon measure $\mu \in M(\mathbb{R}^d)$ define,

$$\|\mu\|_w = \sup_{\substack{\phi \in BC^0(\mathbb{R}^d) \\ \|\phi\|_\infty \leq 1}} \int_{\mathbb{R}^d} |\phi(v) w(v)| \, \mathrm{d}\mu(v).$$

- Define $M_w = \{\mu \in M(\mathbb{R}^d) \mid \|\mu\|_w < \infty\}$.
- Define further $\|u\|_\rho = \sup_{t \in [0,\infty)} e^{-\rho t} \|u(t)\|_X$
- $\mathrm{C}^0_\rho([0,\infty), X) := \{u \in \mathrm{C}^0([0,\infty), X) \mid \|u\|_\rho = \sup_{t \in [0,\infty)} e^{-\rho t} \|u(t)\|_X < \infty\}$

One last thing before we state the Lemma is to recall a famous theorem that we will need n the proof.

**Theorem 2.10** (The Banach Fixed Point Theorem)**.**
Let $(X, d)$ be a *non-empty complete* metric space with a *contraction* mapping $T : X \to X$ (contraction simply meaning meaning Lipschitz with Lipschitz constant $L \in [0, 1)$). Then $T$ admits a *unique fixed point* $x^*$. Furthermore, let $x_0 \in X$ be arbitrary and define a sequence $(x_n)_{n \in \mathbb{N}}$ by having $T(x_n) = x_{n+1}$ for all $n \geq 0$. Then $x_n \to x^*$ in $(X, d)$. ▲

**Lemma 2.11** (Convergence of Truncated Solutions)**.**
Let $f_0 \in M_{(1+|v|)^2}$ then $f_k$ converges in $C^0_\rho\left([0,\infty), M_{1+|v|}\right)$ to $f$ for some $\rho > 0$ and $f \in C^1\left([0,\infty), M_{1+|v|}\right)$ is the unique solution of (11). Furthermore $f_t \in M_{(1+|v|)^2}$ for all $t \in [0,\infty)$. ▲

*Proof.* The vast majority of this proof is setting up a contraction and showing that it is indeed a contraction. The rest follows from the Banach Fixed Point Theorem (cf Theorem 2.10). Firstly, recall,

$$f_{t,k}(v) = e^{-\int_0^t L[f_{s,k-1}](v)\,\mathrm{d}s} f_0,$$

where,

$$L[f_{s,k-1}](v) = \kappa_d \int_{\mathbb{R}^d} |v - v'| \,\mathrm{d}f_{s,k-1}(v').$$

Since $f_0$ is a probability measure it is easy to see that $L[f_{s,k-1}]$ takes only non-negative values. Thus $f_{t,k}$ decreases in $t$ and similarly $\|f_{t,k}\|_{(1+|v|)^2}$ decreases as $t$ increases. Next let $\phi \in BC^0(\mathbb{R}^d)$ such that $\|\phi\| \leq 1$, then consider the map,

$$\begin{aligned} T : C^0_\rho\big([0,\infty), M_{1+|v|}\big) &\to C^0_\rho\big([0,\infty), M_{1+|v|}\big), \\ g_t &\mapsto e^{-\int_0^t L[g_s]\,\mathrm{d}s} f_0. \end{aligned}$$

In particular $T(f_{t,k}) = f_{t,k+1}$. Let $t \in [0,\infty)$ be arbitrary. Then,

$$\begin{aligned} e^{-\rho t}\|T(f_{t,k}) - T(f_{t,k-1})\|_{1+|v|} &= e^{-\rho t}\|f_{t,k+1} - f_{t,k}\|_{1+|v|} \\ &= e^{-\rho t}\left| \int_{\mathbb{R}^d} \phi(v)(1+|v|)(\mathrm{d}f_{t,k+1} - \mathrm{d}f_{t,k})(v) \right| \\ &= \int_{\mathbb{R}^d} \phi(v)(1+|v|)e^{-\rho t}\left| e^{-\int_0^t L[f_{s,k}](v)\,\mathrm{d}s} - e^{-\int_0^t L[f_{s,k-1}](v)\,\mathrm{d}s} \right| \mathrm{d}f_0(v). \end{aligned}$$

Then since $e^x \geq 1 + x$ for all $x \in \mathbb{R}$ we can easily show $|e^{-x} - 1| \leq |x|$. Thus,

$$\begin{aligned} \left| e^{-\int_0^t L[f_{s,k}](v)\,\mathrm{d}s} - e^{-\int_0^t L[f_{s,k-1}](v)\,\mathrm{d}s} \right| &= e^{-\int_0^t L[f_{s,k-1}](v)\,\mathrm{d}s} \left| e^{-\int_0^t L[f_{s,k}](v) - L[f_{s,k-1}](v)\,\mathrm{d}s} - 1 \right| \\ &\leq \left| e^{-\int_0^t L[f_{s,k}](v) - L[f_{s,k-1}](v)\,\mathrm{d}s} - 1 \right| \end{aligned}$$

since $f_{s,k-1} \geq 0$ and L is positive,

$$\begin{aligned} &\leq \left| \int_0^t L[f_{s,k}](v) - L[f_{s,k-1}](v)\,\mathrm{d}s \right| \\ &\leq \int_0^t \big| L[f_{s,k}](v) - L[f_{s,k-1}](v) \big| \,\mathrm{d}s. \end{aligned}$$

Returning to our original line of thought we have,

$$\begin{aligned} &e^{-\rho t}\|f_{t,k+1} - f_{t,k}\|_{1+|v|} \\ &\quad \leq \int_{\mathbb{R}^d} \phi(v)(1+|v|)e^{-\rho t} \int_0^t \big| L[f_{s,k}](v) - L[f_{s,k-1}](v) \big| \,\mathrm{d}s\,\mathrm{d}f_0(v), \end{aligned}$$

which by definition of $L$ leads us to see,

$$\leq \int_{\mathbb{R}^d} \phi(v)(1+|v|)e^{-\rho t}\kappa_d \int_0^t \int_{\mathbb{R}^d} |v - v'| |\,\mathrm{d}f_{s,k} - \mathrm{d}f_{s,k-1}|(v')\,\mathrm{d}s\,\mathrm{d}f_0(v),$$

then since $|v-v'| \leq 1+|v|+|v'|$ and using $\psi \equiv 1 \in BC^0(\mathbb{R}^d)$ we have,

$$\leq \int_{\mathbb{R}^d} \phi(v)(1+|v|)\kappa_d \int_0^t e^{-\rho(t-s)} \left[e^{-\rho s}\|f_{s,k}-f_{s,k-1}\|_{1+|v|} + e^{-\rho s}|v|\,\|f_{s,k}-f_{s,k-1}\|_1\right] \mathrm{d}s\,\mathrm{d}f_0(v)$$

and since in general we have $\|g\|_1 \leq \|g\|_{1+|v|}$, we get that,

$$\begin{aligned}
&\leq \int_{\mathbb{R}^d} \phi(v)(1+|v|)^2\kappa_d \sup_{0\leq s<\infty} \left(\int_0^t e^{-\rho(t-\tau)}\,\mathrm{d}\tau\right) e^{-\rho s}\|f_{s,k}-f_{s,k-1}\|_{1+|v|}\,\mathrm{d}f_0(v) \\
&\leq 2\kappa_d\|f_0\|_{(1+|v|)^2}\frac{1}{\rho}\left(1-e^{-\rho t}\right)\|f_{\cdot,k}-f_{\cdot,k-1}\|_\rho.
\end{aligned}$$

Thus when one sets $\rho > 2\kappa_d\|f_0\|_{(1+|v|)^2}$ one gets that,

$$\|T(f_{\cdot,k})-T(f_{\cdot,k-1})\|_\rho \leq \left(1-e^{-\rho t}\right)\|f_{\cdot,k}-f_{\cdot,k-1}\|_\rho,$$

from which we glean that $T$ is indeed a contraction on the Banach space $C^0_\rho([0,\infty), M_{1+|v|})$. Thus the Banach Fixed Point Theorem gives us that there exists a unique $f_\cdot \in C^0_\rho([0,\infty), M_{1+|v|})$ such that for all $t \in [0,\infty)$,

$$f_t = e^{-\int_0^t L[f_s]\,\mathrm{d}s} f_0.$$

Consequently, $f_t$ is differentiable in the classical sense and solves,

$$\frac{\partial f_t}{\partial t} = Q_-[f_t, f_t],$$

for all $t \in [0,\infty)$. Furthermore, $f_t$ is unique and $f_{t,k}$ $tof_t$ as $k \to \infty$ in $C^0_\rho([0,\infty), M_{1+|v|})$. Additionally, if $h \in C^1([0,T], M_{1+|v|})$ solves (11) then it must also follow that $h(t) = e^{-\int_0^t L[h(s)]\,\mathrm{d}s} f_0$ and so $h(t) = f_t$ for all $t \in [0,T]$ and hence we also have uniqueness in $C^1([0,T], M_{1+|v|})$.

Finally recall $0 \leq f_{t,k}(v) \leq f_0(v)$, as established at the beginning. Then $f_{t,k}(A) \to f_t(A)$ as $k \to \infty$ for all $A \in \mathbb{B}^d$. Suppose for contradiction there exists $A \in \mathbb{B}^d$ of positive measure such that $f_{t,k}(A) \not\to f_t(A)$ for some fixed $t$. Let $\varepsilon > 0$ such that,

$$e^{-\rho t}|f_{t,k}(A) - f_t(A)| \geq \varepsilon,$$

for infinitely many $k$. By noting that $1 \leq 1+|v|$ and considering $\phi \equiv 1$ we then see that,

$$e^{-\rho t}\|f_{t,k}-f_t\|_{1+|v|} \geq e^{-\rho t}|f_{t,k}(A)-f_t(A)| \geq \varepsilon,$$

for infinitely many $k$. In particular $\|f_{\cdot,k}-f_\cdot\|_\rho \geq \varepsilon$ for infinitely many $k$. This contradicts our apriori convergence. Thus indeed $f_{t,k}(A) \to f_t(A)$ as $k \to \infty$ for all $A \in \mathbb{B}^d$ and for all $t \in [0,\infty)$. So now we see that $0 \leq f_t \leq f_0$ for all $t \in [0,\infty)$. It is then nugacious to conclude that,

$$f_t \in M_{(1+|v|)^2}.$$

■

As seems to be the pattern so far the subscript $k$ seems to denote some sort of truncation. We introduce a function that is akin to the scattering state indicator function (cf. Definition 2.3). This truncated scattering state indicator function will only give us the scattering data of a chosen particle determined from going $k$ back generations back in the collision history of a the said particle.

**Definition 2.12** (Truncated Scattering State Indicator)**.**

$$\beta_k(i,t) = \begin{cases} 1 & \text{if } \operatorname{dist}(z_i, z_{i'}, s) \geq a\beta_{k-1}(i', s) \text{ for all } s \in [0,t), i' \neq i, \\ 0 & \text{else}, \end{cases} \tag{16}$$

$$\beta_1(i) \equiv 1 \tag{17}$$

▲

What will make more sense at the end of the next sub-subsection is that, $\beta_k(i,t)$ gives the scattering state of particle $i$ if the *collision tree* at time $t$ is allowed to have at most height $k$ and any collisions that would have distance greater than $k$ from the root particle $i$ would be ignored.

**Lemma 2.13.**
For all realisations of the many body collision processes, where the particles have initial conditions $(u_0, v_0) \in \bigcup_{N=0}^{\infty} \left(\mathbb{T}^d \times \ \mathbb{R}^d\right)^N$ both $\beta_k(i,t)$ and $\beta(i,t)$, are well defined and

$$\lim_{k\to\infty} \beta_k(i,t) = \beta(i,t)$$

pointwise in $i$ and uniformly in $t$. ▲

*Proof.* For the time being this is mainly a analytical question and not something to be too concerned with. If time permits I shall revisit this proof and add it in. ■

### 2.3.1 The Concept of Marked Trees

In this section we will introduce the notion of *Marked Trees*. These trees are objects that will allow us to trace the history of a given particle from some fixed time $t$. By fixing a time and a particle we generate a tree, where each neighbour of any given node will denote the particles that have collided with the given node. Additionally, each node will be marked (i.e. assigned data) that will give us information about the collision that caused the particle to occur in the tree.

**Definition 2.14** (Marked Trees)**.**
Let $\mathbb{N} = \{1, 2, \ldots\}$. The height of a node (or multi-index) $l \in \mathbb{N}^i$ is defined by $|l| := i$, the parent node of $l \in \mathbb{N}^i$ is $\bar{l} = (l_1, \ldots, l_{i-1})$. Let $\mathcal{F} = \cup_{i=1}^{\infty} \mathbb{N}^i$ be the set of multi-indices. We say that $m \subset \mathcal{F}$ is a *tree skeleton with root* $(m \in \mathcal{T})$, if

- $\#m < \infty$
- $m \cap \mathbb{N} = \{1\}$,
- $\bar{l} \in m$ for all $l \in m \backslash \{1\}$,
- $l - 1 \in m$ for all $l \in m$ such that $l \neq (*, \ldots, *, 1)$, where $l - 1 = l - (0, \ldots, 0, 1)$. We say that a tree $m$ has at most height $k$ $(m \in \mathcal{T}_k)$ if $m \cap \mathbb{N}^{k+1} = \emptyset$.

Let $Y = \left\{(u, v, s, \nu) \in \mathbb{T}^d \times \mathbb{R}^d \times [0, \infty) \times \mathbb{S}^{d-1}\right\}$ be the space of initial values and collision parameters. The set of *marked trees* is given by

$$\begin{aligned}
\mathcal{MT} = \Big\{(m, \phi) \; |m \in \mathcal{T}, \phi : m \to Y & \text{ with the property } s_l \in [s_{l-1}, s_{\bar{l}}] \\
& \text{and } \nu_l = \frac{1}{a}(u_{\bar{l}} - u_l + s_l(v_{\bar{l}} - v_l)) \text{ for all } l \in m \backslash \{1\} \\
& d\left((u_l, v_l), (u_{\bar{l}}, v_{\bar{l}}), s_l\right) = a \Big\}
\end{aligned}$$

where $s_{(*,\ldots,*,0)} = 0$. $\mathcal{MT}_k$ is obtained if $\mathcal{T}$ is replaced with $\mathcal{T}_k$. For each skeleton $m \in \mathcal{T}$ we define the set of marked trees with skeleton $m$

$$\mathcal{E}(m) = \{(\tilde{m}, \phi) \in \mathcal{MT} \mid \tilde{m} = m\}.$$

▲

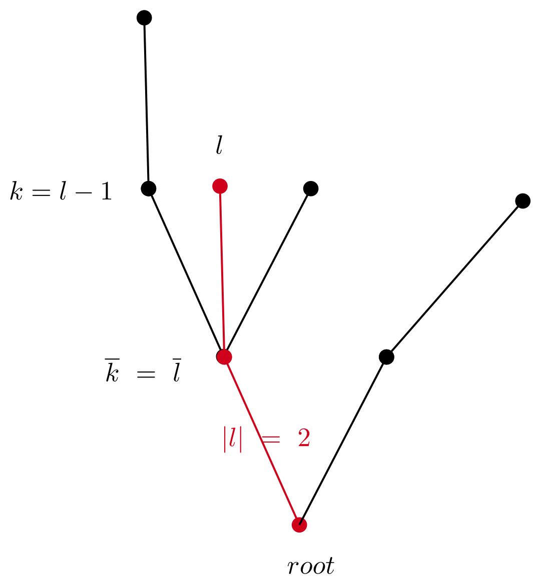



Fig 1: Illustration of $|l|$ , $l-1$ and $\bar{l}$ given $l$.

To unwind this definition we chose a particle and call it the root and fix a time $t$. After this we mark the root particle with its initial position and velocity. Then we look at all the particles that have collided with the root particle up to time $t$ and add them to the tree, these are all the neighbours of the root particle. We then order these neighbours of the root from left to right, where the left most particle is the one that first collides with the root particle and the right most is the last particle that collides with the root particle (up to time $t$). We then mark all the neighbours of the root with the following data, the initial position, velocity, collision time with the root and the "angle" of collision (i.e. $\nu$). Then we pretend the the neighbours of the root are the root and repeat the above process but replace the time $t$ by the particles respective collision time. This will generate an new tree of height one and we attach it to the already existing tree which contains our actual root. Note that we don't care if the particle has been scattered or not, whence even if the collision does not technically occur due to annihilation occurring in one of the two involved particles history, we still record the collision. A perhaps simpler way of looking at the particle dynamics when generating the marked tree is to think no annihilation occurs but rather particles pass through each other with no interaction and we record the instances when two particle touch. Finally note that we should really give the root particle an impact parameter $\nu$. However, physically this doesn't make much sense so instead we just give it a dummy variable $\nu^*$.

**Remark 2.15** (Simplifying the Markings)**.** Note it is fine for us to drop some of the data given for each node. Provided we have the initial position of the root node, we can work out the initial position of all particles in the tree by only knowing the particles velocity, impact parameter and collision time. Hence we will drop the initial position of the non-root particles and not lose any information. ▲

**Example 2.16** (Four interacting particles)**.**
In the below figure we have four particles labeled A, B, C and D. All the particles start at time $t = 0$ at the arrow tail and are, at time $t = 1$, at the arrow head. The black circles on the intersection of trajectories, represent the event that the two particles in question have touched. Each particle will generate its own collision tree which we have illustrated below.

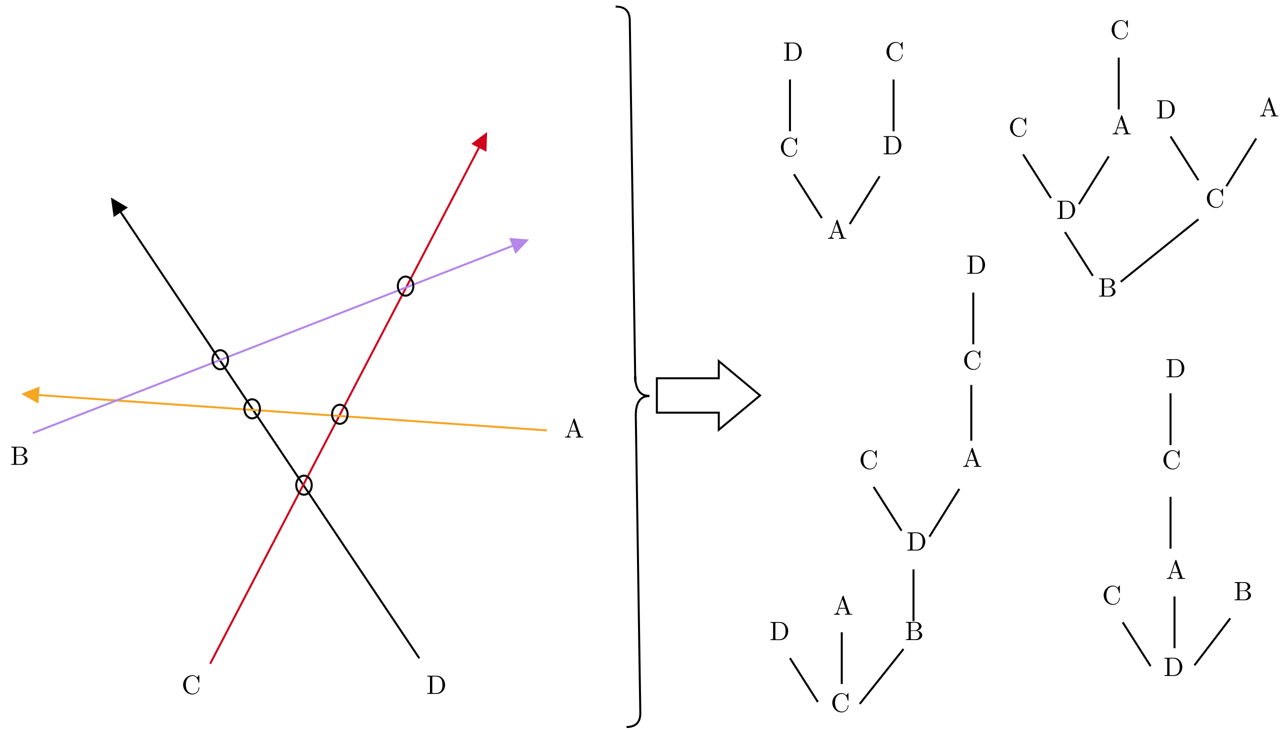



Fig 2: Illustration of four interacting particles and the trees that they generate.

Now Consider the top-left tree (i.e. the one with root A. The we will illustrate what the tree's markings would be. Suppose each particle has initial conditions $(u_A, v_A), \dots (u_D, v_D)$ then,

$$\begin{aligned} m =& \{1, (1,1), (1,2), (1,1,1), (1,2,1)\} \\ \phi =& \{(1, (u_A, v_A, t, \nu^*)), ((1,1), (u_C, v_C, s_{11}, \nu_{11})), ((1,2), (u_D, v_D, s_{12}, \nu_{12})) \\ & ((1,1,1), (u_D, v_D, s_{111}, \nu_{12})), ((1,2,1), (u_C, v_C, s_{121}, \nu_{121}))\} \end{aligned}$$

$$\begin{aligned} s_{11} &= \min\{s \in [0,t] \mid \text{dist}((u_A, v_A), (u_C, v_C), s) = a\} \\ s_{12} &= \min\{s \in [0,t] \mid \text{dist}((u_A, v_A), (u_D, v_D), s) = a\} \\ s_{111} = s_{121} &= \min\{s \in [0,t] \mid \text{dist}((u_C, v_C), (u_D, v_D), s) = a\} \end{aligned}$$

$$\begin{aligned} \nu_{11} &= \frac{1}{a}(u_A - u_C + s_{11}(v_A - v_C)) \\ \nu_{12} &= \frac{1}{a}(u_A - u_D + s_{12}(v_A - v_D)) \\ \nu_{111} = -\nu_{121} &= \frac{1}{a}(u_D - u_C + s_{111}(v_D - v_C)) \end{aligned}$$

▲

**Remark 2.17** (Tree Structure)**.**
It will later become important to note that the tree structure is important. Due to siblings in the tree having a precise order depending on their collision times, permuting any two particles in the tree without destroying the *graph theoretical* tree structure will result in the destruction of the *marked tree* structure. ▲

**Definition 2.18** (Mean Free Path)**.**
In the kinetic theory of gases, the *mean free path* of a particle is the average distance the particle travels between collisions with other moving particles. ▲

**Remark 2.19** (Mean Free Path and Boltzmann-Grad Scaling)**.**
From a heuristic point of view the Boltzmann-grad scaling limit ensures that the mean free path remains constant. To picture this as we increase the number of particles we increase the chance of particle collision occurring however by decreasing the diameter of the particles we decrease the chance of particle collisions occurring. Consequently, the Boltzmann-grad scaling strikes a balance between these two outcomes in the sense that the rate of decrease of radius and the rate of increase in the number in particles is sufficiently fast such that we don't get the mean free path either collapsing to zero or exploding to infinity. ▲

Before concluding this subsection we wish to return to the scattering state. Notice the leafs of a marked tree must have a scattering state value of 1. This is due to the fact that there have been no possible prior collisions for the leafs of the tree up until the assigned collision time. Furthermore, the scattering state of any node $l$, at time $t$, is 0 if any of its children have a scattering state of 1. Thus we deduce the simple formula,

$$\beta_l = \prod_{\substack{l' \in m \\ \bar{l'} = l}} (1 - \beta_{l'}) \tag{18}$$

### 2.3.2 Idealised Distribution

We touched upon earlier about some distributions that we shall show convergence for. In this subsection we shall introduce and analyse idealised distribution and their behaviour corresponds to the gainless Boltzmann equation. The main purpose of this subsection is to build up to a lemma which allows us to bridge the gap between truncated solutions and the idealised distributions. It is worth noting that for the time being we are taking a step away from the many body dynamics (which we may also call the *Microscopic dynamics*) and only focus on the "theoretical" side of the Boltzmann equation.

As discussed before (cf. 2.15) $\mathcal{E}(m) \subseteq \left(\mathbb{T}^d \times \mathbb{R}^d \times \mathbb{R}^+\right) \times \left(\mathbb{R}^d \times S^{d-1} \times \mathbb{R}^+\right)^{(\#m)-1}$. We can view $\mathcal{E}(m)$ as a *manifold* and hence induce a subspace topology with its own Borel subsets. Then let $\Omega_m \subseteq \mathcal{E}(m)$ be a Borel subset induced by the subspace topology of the manifold $\mathcal{E}(m)$, the we say $\Omega \subseteq \mathcal{MT}$ is a Borel subset if

$$\Omega = \bigcup_{m \in \mathcal{T}} \Omega_m. \tag{19}$$

**Definition 2.20** (Idealised Distribution)**.**
Let $\Omega \subseteq \mathcal{MT}$ be Borel and let $t \in [0, \infty)$ and $k \in \mathbb{N}_0$, then we define the idealised distribution to be given by,

$$P_{t,k}(\Omega) = \sum_{m \in \mathcal{T}_k} \int_{\Omega \cap \mathcal{E}(m)} e^{-\sum_{j<k} \Gamma_j(\Phi)} \mathrm{d}\lambda^m(\phi), \tag{20}$$

where,

$$\Gamma_j(\Phi) = \sum_{l \in m, |l| = j} \gamma_l(\Phi) \tag{21}$$

$$\gamma_l(\Phi) = \int_0^{s_l} L[f_0](v_l) \mathrm{d}s' = s_l L[f_0](v_l) \geq 0 \tag{22}$$

$$\lambda^m(\phi) = \mathbf{1}_{\mathbb{T}^d}(u_1) \otimes f_0(v_1) \otimes \delta(s_1 - t) \otimes \prod_{l \in m \setminus \{1\}} [((v_l - v_{\bar{l}}) \cdot \nu_l)_+ \chi_{[s_{l-1}, s_l]}(s_l) \mathrm{d}f_0(v_l) \mathrm{d}\nu_l \; \mathrm{d}s_l] \tag{23}$$

▲

**Remark 2.21** (Independence and Absolute Continuity)**.**

1. Let $\Omega \subseteq \mathcal{MT}_l$ be a Borel set and suppose $l \leq k$. Then

$$P_{t,k}(\Omega) = P_{t,l}(\Omega).$$

2. Notice that $\lambda^1 = \mathbb{1}_{\mathbb{T}^d} \otimes f_0 \otimes \delta(s_1 - t)$. It follows then that,

$$\begin{aligned} P_{t,1}(\Omega) &= \int_{\Omega \cap \mathcal{E}((1))} e^{-\sum_{j<1} \Gamma_j(\Phi)} \, \mathrm{d}(\mathbb{1}_{\mathbb{T}^d} \otimes f_0 \otimes \delta(s_1 - t)) \\ &= \int_{\Omega \cap \mathcal{E}((1))} \mathrm{d}(\mathbb{1}_{\mathbb{T}^d} \otimes f_0 \otimes \delta(s_1 - t)). \end{aligned}$$

   But then since $\mathcal{MT} \cap \mathcal{E}((1)) = \mathbb{T}^d \times \mathbb{R}^d \times \mathbb{R}^+$ it is clear to see that $P_{t,1} \in \mathrm{PM}(\mathcal{MT})$

▲

Before we continue we take a bit of a hiatus from trying to prove the main result. The reason being that to the untrained eye (i.e. me) this idealised distribution will later seem to be some sort of hocus pucos magic which some how makes everything work. However, it turns out the idealised distribution is the only thing it can be. We will now delve into the reasoning behind the idealised distribution.

The equations (20), (21), (22) and (23) give us the idealised distribution up to height $k$. What it is is really a Poisson probability in disguise. In particular an *inhomogeneous* Poisson probability. We will contary to mathematical tradition work partially towards the idealised distribution and partially backwards and meet somewhere in middle. We will discuss the tree $m = \{(1), (1,1), (1,2)\}$ and the idealised distribution $P_{t,2}$ as this is fairly simple and once this is all spelled out it is quite easily to see how everything generalises to arbitrary trees and idealised distributions.

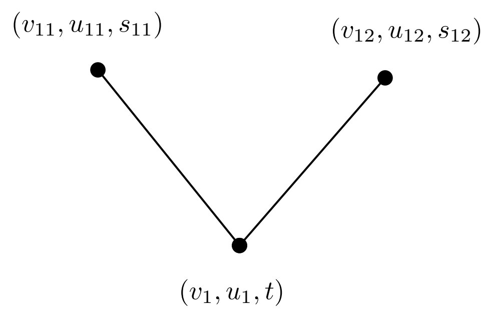



Fig: A illustration of the tree $m$

On the other hand, hidden inside the Boltzmann equation is a inhomogeneous Poisson distribution, one might say that the Boltzmann equation is simply the equation satisfied by differentiating the law with respect to time. That is to say a particle travailing for time $t$ and with velocity $v_1$ will have a Poisson$(t\lambda(v_1))$ number of collisions, where we define,

$$\lambda(v_1) = \kappa_d \int_{\mathbb{R}^d} |v_1 - v'| \, \mathrm{d}f_0(v')$$

This is all built into the Boltzmann equation. The following figure gives us two representations of the tree $m$ which will be useful the keep in mind when trying to understand the idealised distribution.

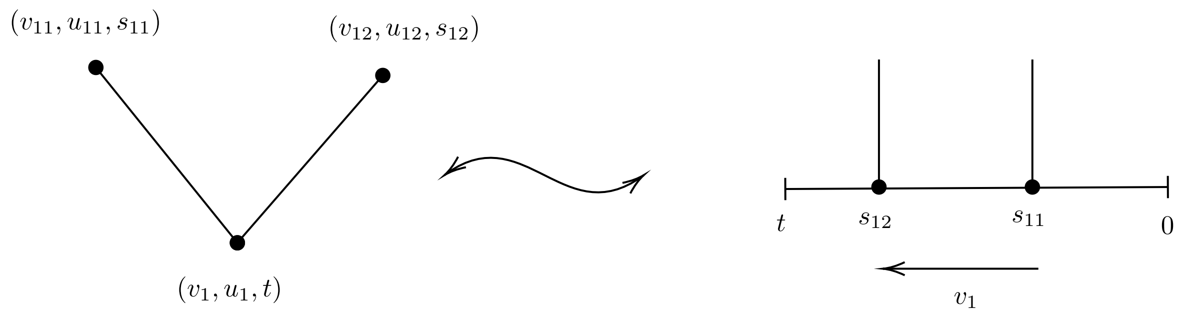



Fig: A illustration of the tree $m$ in two different ways

Now look at $\mathcal{E}(m) \subseteq \mathcal{MT}_2 \subseteq \mathcal{MT}$ and plug it into $P_{t,2}$ and so,

$$P_{t,2}(\mathcal{E}(m)) = \int_{\mathcal{E}(m)} e^{-\Gamma_1(\Phi)} \,\mathrm{d}\lambda^m(\Phi) \tag{$\diamond$}$$

Recalling that $\Gamma_1(\Phi) = t\kappa_d \int_{\mathbb{R}^d} |v_1 - v'| \,\mathrm{d}f_0(v')$ and integrating out $\nu_1$ and $\nu_2$ we see that $(\diamond)$ can be expressed as,

$$P_{t,2}(\mathcal{E}(m)) = \int_{\mathcal{E}(m)} e^{t\kappa_d \int_{\mathbb{R}^d} |v_1 - v'| \,\mathrm{d}f_0(v')} \kappa_d |v_1 - v_{11}| \mathbb{1}_{[0,t]}(s_{11}) \kappa_d |v_1 - v_{12}| \mathbb{1}_{[s_{11},t]}(s_{12}) \\ \mathrm{d}s_{12} \,\mathrm{d}s_{11} \,\mathrm{d}f_0(v_{12}) \,\mathrm{d}f_0(v_{11}) \,\mathrm{d}f_0(v_1) \,\mathrm{d}\mathcal{L}^d(u_1)$$

expanding the integral we get,

$$P_{t,2}(\mathcal{E}(m)) = \int_{\mathbb{T}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_0^t \int_0^t e^{t\kappa_d \int_{\mathbb{R}^d} |v_1 - v'| \,\mathrm{d}f_0(v')} \kappa_d |v_1 - v_{11}| \mathbb{1}_{[0,t]}(s_{11}) \kappa_d |v_1 - v_{12}| \mathbb{1}_{[s_{11},t]}(s_{12}) \\ \mathrm{d}s_{12} \,\mathrm{d}s_{11} \,\mathrm{d}f_0(v_{12}) \,\mathrm{d}f_0(v_{11}) \,\mathrm{d}f_0(v_1) \,\mathrm{d}\mathcal{L}^d(u_1)$$

Then by integrating out the variable $u_1$ and applying Fubini's theorem we can re-arrange the integrals to get,

$$P_{t,2}(\mathcal{E}(m)) = \int_{\mathbb{R}^d} \Bigg[ \int_0^t \int_0^t \left( \kappa_d \int_{\mathbb{R}^d} |v_1 - v'| \,\mathrm{d}f_0(v') \right)^2 e^{t\kappa_d \int_{\mathbb{R}^d} |v_1 - v'| \,\mathrm{d}f_0(v')} \\ \mathbb{1}_{[0,t]}(s_{11}) |v_1 - v_{12}| \mathbb{1}_{[s_{11},t]}(s_{12}) \,\mathrm{d}s_{12} \,\mathrm{d}s_{11} \Bigg] \mathrm{d}f_0(v_1) \tag{$\bullet$}$$

The item inside the square brackets is a Poisson probability when the velocity of the root particle is some fixed $v_1$.

Now we work from the Boltzmann equation and try to arrive a an identical equation. Suppose a particle is following the Boltzmann equation that has exactly two collisions at time $t_1, t_2 \in [0,t]$ and the particle has velocity $v_1$. Set $x_1 := t_1 \wedge t_2$ and $x_2 := t_1 \vee t_2$.

Thus consider,

$$\mathbb{P}\big(\text{particle has exactly 2 collisions, } x_1 \in [a_1, b_1], x_2 \in [a_2, b_2] \big| \text{ given root has velocity } v_1\big),$$

since Poisson processes have exponentially distributed increments

$$= \int_{a_1}^{b_1} \mathbb{P}\big(\text{particle has exactly 2 collisions }, x_2 \in [a_2, b_2] \big| v_1, x_1\big) \lambda e^{-\lambda x_1} \,\mathrm{d}x_1$$

since $x_2$ is defined to be after $x_1$ by definition we have that,

$$
\begin{aligned}
&= \int_{a_1}^{b_1} \int_{a_2 \vee x_1}^{b_2} \mathbb{P}\big(\text{particle has exactly 2 collisions } \big| v_1, x_1, x_2\big) \lambda e^{-\lambda x_1} \lambda e^{-\lambda(x_2 - x_1)} \, \mathrm{d}x_2 \, \mathrm{d}x_1 \\
&= \int_{a_1}^{b_1} \int_{a_2 \vee x_1}^{b_2} \mathbb{P}\big(\text{no collisions in } [x_2, t] \big| v_1, x_1, x_2\big) \lambda^2 e^{-\lambda x_2} \, \mathrm{d}x_2 \, \mathrm{d}x_1 \\
&= \int_{a_1}^{b_1} \int_{a_2 \vee x_1}^{b_2} \mathbb{P}\big(\text{no collisions in } [x_2, t]\big) \lambda^2 e^{-\lambda x_2} \, \mathrm{d}x_2 \, \mathrm{d}x_1 \\
&= \int_{a_1}^{b_1} \int_{a_2 \vee x_1}^{b_2} \lambda^2 e^{-\lambda t} \, \mathrm{d}x_2 \, \mathrm{d}x_1 \\
&= \int_0^t \int_0^t \mathbb{1}_{[a_1, b_1]}(x_1) \mathbb{1}_{[a_2 \vee x_1, b_2]}(x_2) \lambda^2 e^{-\lambda t} \, \mathrm{d}x_2 \, \mathrm{d}x_1
\end{aligned}
$$

Where we recall that $\lambda(v_1) = \kappa_d \int_{\mathbb{R}^d} |v_1 - v'| \, \mathrm{d}f_0(v')$ and so we have that,

$$
\begin{aligned}
&\mathbb{P}\big(\text{particle has exactly 2 collisions, } x_1 \in [a_1, b_1], x_2 \in [a_2, b_2]\big) \\
&\quad = \int_{\mathbb{R}^d} \int_0^t \int_0^t \mathbb{1}_{[a_1, b_1]}(x_1) \mathbb{1}_{[a_2 \vee x_1, b_2]}(x_2) \lambda(v_1)^2 e^{-\lambda(v_1) t} \, \mathrm{d}x_2 \, \mathrm{d}x_1 \, \mathrm{d}f_0(v_1).
\end{aligned} \tag{$\blacklozenge$}
$$

Since $x_1$ and $x_2$ play the same roles as $s_{11}$ and $s_{22}$ respectively along with the definition of $\lambda(v_1)$ we see that ($\blacklozenge$) agrees with ($\bullet$). Hence we can say that the idealised distribution applied to $\mathcal{E}(m)$ gives us the probability that a particle adhering to Boltzmann will have exactly 2 collisions. Similarly it is easy to see that for any tree with skeleton $\tilde{m} = \{(1), (1,1), (1,2), (1,3)\}$ the exponent of 2 will be be replaced by a 3 in $\bullet$ and $\blacklozenge$ and there will be an extra time integral and hence ($\bullet$) and ($\blacklozenge$) will still agree. It follows similarly that for any tree of height 2 or less both ($\bullet$) and ($\blacklozenge$) will agree. So far we have only spoken about idealised distributions with $k = 2$ this means we ignore, in the first example, the probabilities that the leafs of $m$ will have no collisions we will need $P_{t,3}$ to see these probabilities occur. So consider,

$$
\begin{aligned}
P_{t,3}(\mathcal{E}(m)) &= \int_{\mathcal{E}(m)} e^{-\Gamma_1(\Phi)} e^{-\Gamma_2(\Phi)} \, \mathrm{d}\lambda^m(\Phi) \\
&= \int_{\mathcal{E}(m)} \underbrace{e^{s_{11}\kappa_d \int_{\mathbb{R}^d} |v_{11} - v'| \, \mathrm{d}f_0(v')}}_{(i)} \underbrace{e^{s_{12}\kappa_d \int_{\mathbb{R}^d} |v_{12} - v'| \, \mathrm{d}f_0(v')}}_{(ii)} \underbrace{e^{-\Gamma_1(\Phi)} \, \mathrm{d}\lambda^m(\Phi)}_{(iii)}.
\end{aligned}
$$

$(i)$ is essentially the probability a Poisson($s_{11}\lambda(v_{11})$) random variable equals zero. That is to say the node $(1,1)$ has no collisions up to time $s_{11}$.

$(ii)$ is essentially the probability a Poisson($s_{12}\lambda(v_{12})$) random variable equals zero. That is to say the node $(1,2)$ has no collisions up to time $s_{12}$.

$(iii)$ is essentially $P_{t,2}(\mathcal{E}(m))$ which deals with the rest of the probabilities involved in seeing the set $\mathcal{E}(m)$ occur.

Now given any $m \in \mathcal{T}$ it is not hard to see what $P_{t,k}(\mathcal{E}(m))$ represents by iterating the above logic. Hopefully, now the idealised distribution is no longer some hocus pocus mumbo jumbo magic.

**Definition 2.22** (Recursive Random Variable)**.**
A random variable $x : \mathcal{T} \to \mathbb{R}$ is said to be recursive if there exists a family of functions, called *recurrence functions*, $h_b : \mathbb{R}^b \to \mathbb{R}, b \in \mathbb{N}$, which are invariant under permutations of the $b$ components in $\mathbb{R}^b$, such that for all $m \in \mathcal{T}$ with $b = r_1$, where $r_l = \#\left\{l' \in m \mid \overline{l'} = l\right\}$, the equation

$$
x(m) = h_{r_1}\left(x\left(m_1\right), \ldots, x\left(m_{r_1}\right)\right)
$$

holds, where

$$m_j = \left\{\left(1, l_3, \ldots, l_{|l|}\right) \mid l \in m \text{ such that } l_2 = j\right\} \in \mathcal{T}$$

is the $j$ -th sub-tree of m. ▲

This is to say a random variable is recursive if it can be expressed in as a function of itself but applied to the *children sub-trees*. We introduce these random variables as we can compute their expectation and because we want to know the expectation of the number of nodes in a tree.

**Example 2.23** (Two Recursive Random Variables).

$$\text{(number of nodes)} \qquad x^{\#}(m) = \#(m) = 1 + \sum_{j=1}^{r_1} x^{\#}(m_j), \tag{24}$$

$$\text{(scattering state of root)} \qquad x^{\beta}(m) = \beta_1(m) = \prod_{j=1}^{r_1}(1 - x^{\beta}(m_j)). \tag{25}$$

▲

In order to make life a lot simpler we will introduce a new measure $\bar{P}_{t,k}$ which is just $P_{t,k}$ but integrated over the impact parameter $\nu$. If we replace $Y$ by $\hat{Y} := \mathbb{R}^d \times [0, \infty)$ in the definition of marked tree yielding a new set of marked trees denoted $\widehat{\mathcal{MT}}$. A calculation yields that

$$\int_{\mathbb{S}^{d-1}} ((v - v') \cdot \nu)_+ \, \mathrm{d}\nu = \kappa_d |v - v'|.$$

Then for $\widehat{\Omega} \subseteq \widehat{\mathcal{MT}}$, henceforth we define

$$\bar{P}_{t,k}(\widehat{\Omega}) = \sum_{m \in \mathcal{T}_k} \int_{\widehat{\Omega} \cap \mathcal{E}(m)} e^{-\sum_{j<k} \Gamma_j(\Phi)} \mathrm{d}f_0\left(v_1\right) \otimes \delta\left(s_1 - t\right) \otimes \prod_{l \in m \setminus \{1\}} \mathrm{d}\bar{\lambda}_l(\phi), \tag{26}$$

where,

$$\bar{\lambda}_l(\phi) = f_0(v_1) \otimes \delta(s_1 - t) \otimes \prod_{l \in m \setminus \{1\}} [|(v_l - v_{\bar{l}})| \chi_{[s_{l-1}, s_l]}(s_l) \mathrm{d}f_0(v_l) \, \mathrm{d}s_l] \tag{27}$$

**Lemma 2.24** (Recursive Expectation of Recursive Random Variables).
Let $x$ be a recursive random variable with recurrence functions $h_b$. Then it follows that,

$$\begin{aligned} \int_{\mathcal{MT}} x(m) \, \mathrm{d}\bar{P}_{t,k}(\Phi) &= \int_{\mathbb{R}^d} e^{-\Gamma_1} \sum_{r=0}^{\infty} \int_0^t \int \kappa_d |v - v_1| \int_{s_1}^t \int \kappa_d |v - v_2| \cdots \int_{s_r}^t \int \kappa_d |v - v_r| \\ &\quad h_r(x(m_1) \ldots x(m_r)) \Big( \prod_{i=0}^{r-1} \mathrm{d}\bar{P}_{s_{r-i}, k-1}(\Phi_{r-i}) \, \mathrm{d}s_{r-i} \Big) \, \mathrm{d}f_0(v) \\ &= \int_{\mathbb{R}^d} e^{-\Gamma_1} \sum_{r=0}^{\infty} \frac{1}{r!} \int_0^t \int \kappa_d |v - v_1| \int_0^t \int \kappa_d |v - v_2| \cdots \int_0^t \int \kappa_d |v - v_r| \\ &\quad h_r(x(m_1) \ldots x(m_r)) \Big( \prod_{i=0}^{r-1} \mathrm{d}\bar{P}_{s_{r-i}, k-1}(\Phi_{r-i}) \, \mathrm{d}s_{r-i} \Big) \, \mathrm{d}f_0(v). \end{aligned} \tag{28}$$

Where $v$ is the velocity of the root particle and $v_j$ is the velocity of the root particle of the sub-tree $\Phi_r = (m_r, \phi)$ and $\int_{\mathbb{R}^d} \kappa_d \, |v - v'| \, t \, \mathrm{d}f_0\,(v')$. Finally when $r = 0$ we assign 1 to the value of the summation. ▲

**Remark 2.25.** A point of confusion for me was why do we have the d$s$ in (28)? Shouldn't they be absorbed into $\bar{P}_{t,k-1}$? Well, no. The reason for this is simply because the idea behind the proof is to look at the measure $\bar{\lambda}_{li}$ for $i \in \{1, \ldots r_i\}$. From here we will break down $\bar{\lambda}_{li}$ into its component measures (c.f. (27)) extract a d$s$ and hence convert $\bar{P}_{t,k}$ to $\bar{P}_{t,k-1}$. ▲

*Proof.* This is mainly a measure theoretic question and not something to be too concerned with. If time permits I shall revisit this proof and add it in. ■

Lemma 2.24 allows us to express the expectation of a recursive random variable in terms of the components that make up the random variable itself. But what is more important to note is the sub-trees introduced in the lemma have a *reduced* height and also we are able to reduce the the truncation height in our measure, that is to say we are able to avoid using $P_{t,k}$ and instead use $P_{t,k-1}$. This fact implies that for a recursive random variable $x$ its expectation can be expressed as an integral of some function with measure $P_{t,1}$. Equipped with last fact and realising that $x(m) = 1$ is a recursive random variable, we invoke that last equality of (28) to get that,

$$\begin{aligned}\int_{\mathcal{MT}} \mathrm{d}\bar{P}_{t,2}(\Phi) &= \int_{\mathbb{R}^d} e^{-\Gamma_1} \sum_{r=0}^{\infty} \frac{1}{r!} \int_0^t \int \kappa_d |v - v_1| \int_0^t \int \kappa_d |v - v_2| \cdots \int_0^t \int \kappa_d |v - v_r| \\ &\quad \left( \prod_{i=0}^{r-1} \mathrm{d}\bar{P}_{s_{r-i},1}(\Phi_{r-i}) \,\mathrm{d}s_{r-i} \right) \mathrm{d}f_0(v) \\ &= \int_{\mathbb{R}^d} e^{-\Gamma_1} \sum_{r=0}^{\infty} \frac{1}{r!} \left( \int_{\mathbb{R}^d} \kappa_d |v - v'| \,\mathrm{d}f_0(v') \right)^r \mathrm{d}f_0(v) \\ &= \int_{\mathbb{R}^d} e^{-\Gamma_1} \sum_{r=0}^{\infty} \frac{(\Gamma_1)^r}{r!} \,\mathrm{d}f_0(v) = 1.\end{aligned}$$

Then proceeding with induction in $k$ and using the above method for in our inductive step we deduce that $P_{t,k}$ is a probability measure for all $k \in \mathbb{N}$ and for all $t \in [0, \infty)$. Before we continue we will remind ourselves of *Grömwalls inequality.*

**Lemma 2.26** (Grömwalls Inequality for Continuous Functions)**.**
Let $\alpha, \beta, u : [0, t] \to \mathbb{R}$ and further suppose that $\beta$ and $u$ are continuous and $\alpha$ is bounded and non-decreasing, then if

$$u(x) \leq \alpha(x) + \int_0^x \beta(s) u(s) \,\mathrm{d}s.$$

Then,

$$u(x) \leq \alpha(x) e^{\int_0^x \beta(s) \,\mathrm{d}s}.$$

▲

Later on at the end of this chapter when we show convergence of measures it will be very useful to bound the expected number of nodes from above. This is the content of the next lemma and is obtainable due to 2.24.

**Lemma 2.27** (An Upper Bound on the Expected Number of Nodes)**.**
For a tree $m \in \mathcal{T}$ the number of non-root nodes is given by $X(m) = (\#m) - 1$. The expected value of $X$ with respect to the measure $P_{t,k}$ satisfies the estimate $\mathbb{E}(X) \leq K_{\text{ini}} \exp(\kappa_d K_{\text{ini}}\, t)$ with $K_{\text{ini}} = \int_{\mathbb{R}^d} \mathrm{d}f_0(v)(1 + |v|)^2$ as in (8). Furthermore, notice it is independent of $k$. ▲

*Proof.* First of all we define $F_{t,k} := \mathbb{E}(X | v_1 = v, m \in \mathcal{T}_k)$, i.e. the conditional expectation that if we know that the root has velocity $v$ and the tree generated by the root particle, up to time $t$, has height at most $k$ (so it belongs to $\mathcal{T}_k$). Then by definition of conditional expectation (w.r.t. $P_{t,k}$) it follows that,

$$\mathbb{E}(X) = \int X \, \mathrm{d}P_{t,k} = \int F_{t,k}(v) \, \mathrm{d}P_{t,k} = \int F_{t,k}(v) \, \mathrm{d}f_0(v).$$

Where for the last equality we observe that $F_{t,k}$ only depends on $v$. Then it is fairly simple to see that,

$$\mathbb{E}(X) \leq \sup_{k\in\mathbb{N}} \int F_{t,k}(v) \, \mathrm{d}f_0(v).$$

The next step is to notice that $X$ is a *recursive* random variable and then use Lemma 2.24. To see this note that, $X(m) = r + \sum_{i=1}^{r} X(m_i)$, where $r$ is the degree of the root particle and $m_i$ is the sub-tree whose root is a child of the root of $m$. Furthermore, denote the velocity of the root particle of $m_i$ by $v_i$. Now we are ready to invoke Lemma 2.24, doing this allows us to see what $\mathbb{E}(X)$ "looks like". Consequently, we can extract a $v$ dependent random variable whose integral with respect to $f_0(v)$ over $\mathbb{R}^d$ deliverers us $\mathbb{E}(X)$, in fact this random variables integral over any $A \in \mathcal{B}^d$ equals $\mathbb{E}(X\mathbb{1}_A)$. By definition of conditional expectation this random variable must indeed be $F_{t,k}(v)$. With this Lemma 2.24 nicely hands us the following equality,

$$F_{t,k}(v) = e^{-\Gamma_1} \sum_{r=0}^{\infty} \frac{1}{r!} \int_0^t \int \kappa_d |v - v_1| \cdots \int_0^t \int \kappa_d |v - v_r| \left( r + \sum_{i=1}^{r} X(m_i) \right) \mathrm{d}\bar{P}_{s_r,k-1} \, \mathrm{d}s_r \ldots \mathrm{d}\bar{P}_{s_1,k-1} \, \mathrm{d}s_1 \, \mathrm{d}f_0(v). \tag{29}$$

Where as before, $\Gamma_1(v) = L[f](v)t = \kappa_d t \int_{\mathbb{R}^d} |v - v'| \, \mathrm{d}f(v')$. Register that $\mathrm{d}\bar{P}_{s_i,k-1} \, \mathrm{d}s_i$ only measures the random variable $X(m_i)$ in a non-constant way and so one can see that,

$$\begin{aligned}
&\int_0^t \int \kappa_d |v - v_1| \cdots \int_0^t \int \kappa_d |v - v_r| X(m_i) \, \mathrm{d}\bar{P}_{s_r,k-1} \, \mathrm{d}s_r \ldots \mathrm{d}\bar{P}_{s_1,k-1} \, \mathrm{d}s_1 \, \mathrm{d}f_0(v) \\
&= \int_0^t \int \kappa_d |v - v_i| X(m_i) \, \mathrm{d}\bar{P}_{s_i,k-1} \, \mathrm{d}s_i \prod_{\substack{j=1 \\ j\neq i}}^{r} \int_0^t \int \kappa_d |v - v_j| \, \mathrm{d}\bar{P}_{s_j,k-1} \, \mathrm{d}s_j.
\end{aligned}$$

Analysing the inner integral in the product and using the same trick as at the beginning of the proof one notices that,

$$\int \kappa_d |v - v_j| \, \mathrm{d}\bar{P}_{s_j,k-1} = \int \kappa_d |v - v_j| \, \mathrm{d}f_0(v_j) = \frac{1}{t}\Gamma_1.$$

Moving to the inner integral of the first term and noticing that $\kappa_d |v - v_i| X(m_i)$ is still a recursive random variable and, again, by the definition of conditional expectation,

$$\int \kappa_d |v - v_i| X(m_i) \, \mathrm{d}\bar{P}_{s_i,k-1} = \int \kappa_d |v - v_i| F_{t,k}(v_i) \, \mathrm{d}f_0(v_i).$$

Hence after integrating over $t$,

$$\int_0^t \int \kappa_d |v - v_i| X(m_i) \, \mathrm{d}\bar{P}_{s_i,k-1} \, \mathrm{d}s_i \prod_{\substack{j=1 \\ j\neq i}}^{r} \int_0^t \int \kappa_d |v - v_j| \, \mathrm{d}\bar{P}_{s_j,k-1} \, \mathrm{d}s_j = \Gamma_1^{r-1} \int_0^t \int \kappa_d |v - v_i| F_{t,k}(v_i) \, \mathrm{d}f_0(v_i) \, \mathrm{d}s_i.$$

Now invoking (28) and inserting the above equality for each $r$, one sees that,

$$\begin{aligned}
F_{t,k}(v) &= e^{-\Gamma_1} \sum_{r=1}^{\infty} \left( r \frac{\Gamma_1^r}{r!} + \frac{\Gamma_1^{r-1}}{r!} \sum_{i=1}^{r} \int_0^t \int \kappa_d |v - v_i| F_{t,k}(v_i) \, \mathrm{d}f_0(v_i) \, \mathrm{d}s_i \right) \\
&= \Gamma_1 + \int_0^t \int_{\mathbb{R}^d} \kappa_d |v - v'| F_{t,k}(v') \, \mathrm{d}f_0(v') \, \mathrm{d}s.
\end{aligned}$$

Where the last equality comes from the definition of $e$ and the fact that the double integral in the sum is the same value for each $i \in \{1, \ldots r\}$.

We now introduce two new objects, a norm and an integral operator and our aim will be to set things up with these new objects in order to be able to apply *Grömwall's Inequality* and hence conclude the proof. Define:

- The norm $\|F\|_1 := \sup_{v\in\mathbb{R}^d} \dfrac{F(v)}{1+|v|}$,
- The integral operator $A_{f_0}F(v) := \kappa_d \int_{\mathbb{R}^d} |v-v'|F(v')\,\mathrm{d}f_0(v')$.

By our previous calculations it is easy to see that,

$$F_{t,k} = \Gamma_1 + \int_0^t A_{f_0}F_{s,k-1}\,\mathrm{d}s.$$

Then note that in general,

$$\begin{aligned}
\|A_{f_0}F\| &= \sup_{v\in\mathbb{R}^d} \frac{\kappa_d}{1+|v|}\int_{\mathbb{R}^d} |v-v'|F(v')\,\mathrm{d}f_0(v'),\\
&\le \sup_{v\in\mathbb{R}^d} \frac{\kappa_d\|F\|_1}{1+|v|}\int_{\mathbb{R}^d} |v-v'|(1+|v'|)\,\mathrm{d}f_0(v'),\\
&\le \kappa_d K_{\mathrm{ini}}\|F\|_1.
\end{aligned}$$

Where the first inequality comes from, $\dfrac{|v-v'|}{1+|v|} \le \dfrac{|v'|}{1+|v|} + \dfrac{|v|}{1+|v|} \le (1+|v'|)$. Also notice that,

$$\|\Gamma_1\|_1 = t\sup_v \frac{\kappa_d\int_{\mathbb{R}^d}\ \mathrm{d}f_0\,(v')\,|v-v'|}{1+|v|} \le t\kappa_d\int_{\mathbb{R}^d}\ \mathrm{d}f_0\,(v')\,(1+|v'|) \le t\kappa_d K_{\mathrm{ini}}.$$

Furthermore $F_{t,k}(v)$ is monotonically in $k$, as $P_{t,k}$ assigns the probability of trees of height greater than $k+1$ to trees of height $k$, reducing the number of expected nodes. Hence,

$$\int_0^t A_{f_0}F_{s,k-1}\,\mathrm{d}s \le \int_0^t A_{f_0}F_{s,k}\,\mathrm{d}s,$$

and so,

$$\|F_{t,k}\|_1 \le \kappa_d K_{\mathrm{ini}}\left(t + \int_0^t\ \mathrm{d}s\,\|F_{s,k}\|_1\right).$$

This puts us in a position to apply Grömwall's inequality and hence yielding,

$$\|F_{t,k}\|_1 \le \kappa_d K_{\mathrm{ini}} t e^{\kappa_d K_{\mathrm{init}}},$$

which due to $t \ge 0$ and $\kappa_d, K_{\mathrm{ini}} > 0$ implies that,

$$\|F_{t,k}\|_1 \le e^{\kappa_d K_{\mathrm{init}}}e^{\kappa_d K_{\mathrm{init}}} = e^{2\kappa_d K_{\mathrm{init}}}.$$

Finally, we invoke the definition of conditional expectation one last time to see that,

$$\mathbb{E}(X) \le \sup_{k\in\mathbb{N}}\int_{\mathbb{R}^d} F_{t,k}(v)\,\mathrm{d}f_0(v) \le \sup_{k\in\mathbb{N}}\|F_{t,k}\|_1\int_{\mathbb{R}^d}(1+|v|)\,\mathrm{d}f_0(v) \le K_{\mathrm{ini}}e^{2\kappa_d K_{\mathrm{init}}},$$

which is independent of $k$ and completes our proof. ■

We are now equipped with the right tools to complete our intermediate steps. As we eluded to before the strong law of large numbers bridges the gap between our main Theorem and the justification of the gainless-Boltzmann equation. Lemma 2.11 (convergence of truncated solutions) connects solutions to to the Boltzmann equation to the truncated solutions. Then the next Proposition will be the last intermediate step between our main theorem and the distributions that we will study. In essence it will say that the *truncated solutions* are in fact the *root marginals of the idealised distributions.*

**Proposition 2.28** (Root Marginal and Truncated Solution Identification).
Let $\sigma \in \{0,1\}, \Omega \subset \mathbb{R}^d$ Borel, $t \in [0,\infty)$ and $k \in \mathbb{N} \cup \{0\}$. Then the equation

$$P_{t,k+1}(v_1 \in \Omega \text{ and } \beta_1 = \sigma_1) = \int_\Omega [(1-\sigma_1)(\mathrm{d}f_0(v) - \mathrm{d}f_{t,k}(v)) + \sigma_1 \, \mathrm{d}f_{t,k}(v)] \tag{30}$$

holds, where $f_{t,k}$ is the solution of system (15). ▲

*Proof.* We shall proceed with induction on $k$. One point to raise is that the scattering sate (i.e. the value of $\sigma$) is only dependent on the tree structure, this is simply because in any tree the leafs must be unscattered as they have not undergone any collisions by definition. In turn it must be that the parents of any leaves must have been scattered and so on, this is the content of (18). The base case ($k = 0$) is easy but for the sake of completion we shall write it.

$$P_{t,1}(v_1 \in \Omega,\ \beta_1 = \sigma_1) = \begin{cases} P_{t,1}(v_1 \in \Omega) & \text{if } \sigma_1 = 1 \\ 0 & \text{if } \sigma_1 = 0 \end{cases}.$$

since trees of only one point can't possible scatter and hence must always be unscattered. Thus the second case is impossible resulting in,

$$P_{t,1}(v_1 \in \Omega,\ \beta_1 = \sigma_1) = P_{t,1}(v_1 \in \Omega) = \int_\Omega \mathrm{d}f_0(v_1) = \int_\Omega \mathrm{d}f_{t,0}(v_1).$$

As before we shall deal with the modified measure $\bar{P}_{t,k}$ as we can simply integrate the parameter $\nu$ and hence ignore it (so we are using (26) and (27)). Define the set of trees with scattering state $\sigma$ to be.

$$\mathscr{A}(\sigma) = \left\{ (m, \sigma') \,\middle|\, m \in \mathcal{T}_2, \sigma' : m \to \{0,1\} \text{ such that } \prod_{l' \in m \cap \mathbb{N}^2} (1 - \sigma'_l) = \sigma \right\}.$$

Now at first glance this seems absurd as all trees of height two must lie in $\mathscr{A}(0)$, since all leaves have scattering state 1. But this is not we are looking at. There is a subtle distinction; we are looking at trees of height at most two whose scattering state doesn't have to be one but instead have the freedom be either 0 or 1. The reason for this is that we want to look at arbitrary trees of height $k$ and then look at the sub-tree whose root is the child of the root of the original tree. This is how we plan to use induction. However, we can avoid this be actually looking at a tree of height $k$ and the after some gardening we reduce it to height two but we make sure to leave the scattering states unchanged. Thus looking at all trees that have height at most $k$ and then trimming them two height 2 will give us all trees of height at most 2 with all the possible different scattering states. To make life easier notice,

$$\begin{aligned} \mathcal{A}(0) &= \left\{ (m, \sigma') \mid m \in \mathcal{T}_2, \sigma' : m \to \{0,1\} \text{ such that } \sigma'_l = 1 \text{ for some } l \in m \cap \mathbb{N}^2 \right\}, \\ \mathcal{A}(1) &= \left\{ (m, \sigma') \mid m \in \mathcal{T}_2, \sigma' : m \to \{0,1\} \text{ such that } \sigma'_l = 0 \text{ for all } l \in m \cap \mathbb{N}^2 \right\} \end{aligned}$$

Our inductive hypothesis bestows upon us the following equality,

$$\int_0^t \int_\Omega \kappa_d |v - v_i| \, \mathrm{d}\bar{P}_{s_i,k} \, \mathrm{d}s = (1 - \sigma'_i) \int_0^t \int_\Omega \kappa_d |v - v_i| (\mathrm{d}f_0(v') - \mathrm{d}f_{t,k-1}(v')) \, \mathrm{d}s + \sigma'_i \int_0^t \int_\Omega \kappa_d |v - v_i| \, \mathrm{d}f_{t,k-1}(v') \, \mathrm{d}s.$$

With the above fact and invoking Lemma 2.24 where we set $x(m) \equiv 1$ we see that,

$$\begin{aligned} &P_{t,k+1}(v_1 \in \Omega,\ \beta_1 = \sigma) \\ &= \int_\Omega e^{-\Gamma_1} \sum_{r=0}^\infty \frac{1}{r!} \left( \prod_{i=1}^r \int_0^t \int \kappa_d |v_1 - v_i| \, \mathrm{d}\bar{P}_{s_i,k} \, \mathrm{d}s_i \right) \mathrm{d}f_0(v), \end{aligned}$$

we purposefully avoid the set of integration for the inner most integral, but it is a Borel subset of $\mathcal{MT}$ such that it contains precisely the trees of the form $\beta_1 = \sigma$,

$$= \sum_{(m,\sigma')\in\mathcal{A}(\sigma)} \int_{v_1\in\Omega} \Bigg( \frac{e^{-\Gamma_1}}{r_1(m)!} \prod_{l'\in m\cap\mathbb{N}^2} \bigg[ (1-\sigma'_{l'}) \int_0^t \int_{v'\in\mathbb{R}^d} \kappa_d |v_1 - v'| (\mathrm{d}f_0(v') - \mathrm{d}f_{s,k-1}(v'))\,\mathrm{d}s$$
$$+ \sigma'_{l'} \int_0^t \int_{v'\in\mathbb{R}^d} \kappa_d |v_1 - v'|\ \mathrm{d}f_{s,k-1}(v')\ \mathrm{d}s \bigg] \Bigg) \mathrm{d}f_0(v_1).$$

In particular,

$$= \int_{v_1\in\Omega} \sum_{(m,\sigma')\in\mathcal{A}(\sigma)} I(m,\sigma',v_1)\,\mathrm{d}f_0(v_1),$$

where,

$$\begin{aligned}
\mathrm{I}(m,\sigma',v_1) &= \frac{e^{-\Gamma_1}}{r_1(m)!} \prod_{l'\in m\cap\mathbb{N}^2} \bigg[ (1-\sigma'_{l'}) \int_0^t \int_{v'\in\mathbb{R}^d} \kappa_d |v_1 - v'| (\mathrm{d}f_0(v') - \mathrm{d}f_{s,k-1}(v'))\,\mathrm{d}s \\
&\quad + \sigma'_{l'} \int_0^t \int_{v'\in\mathbb{R}^d} \kappa_d |v_1 - v'|\ \mathrm{d}f_{s,k-1}(v')\ \mathrm{d}s \bigg] \\
&= \frac{e^{-\Gamma_1}}{r_1(m)!} \prod_{l'\in m\cap\mathbb{N}^2} \bigg[ (1-\sigma'_{l'}) \Big( \Gamma_1 - \int_0^t L[f_{s,k-1}](v_1)\,\mathrm{d}s \Big) + \sigma'_{l'} \int_0^t L[f_{s,k-1}](v_1)\,\mathrm{d}s \bigg],
\end{aligned}$$

since,

$$L[f_{s,k-1}](v) = \int_{v'\in\mathbb{R}^d} \kappa_d |v - v'|\,\mathrm{d}f_{s,k-1}(v').$$

Next we define,

$$J_k(\sigma, v_1) = \sum_{(m,\sigma')\in\mathcal{A}(\sigma)} I(m,\sigma',v_1),$$

consequently,

$$P_{t,k+1}(v_1\in\Omega \text{ and } \beta_1 = \sigma) = \int_{v_1\in\Omega} [(1-\sigma) J_k(0,v_1) + \sigma J_k(1,v_1)] \mathrm{d}f_0(v_1). \tag{31}$$

This last step is a bit subtle but if in $J_k(0,v_1)$ then all the terms in $\mathrm{I}(m,\sigma',v_1)$ that are proceeded by $(1-\sigma'_l)$ must disappear since at least one of the terms must be zero. Similarly in the case of $J_k(1,v_1)$ then $(1-\sigma'_l)$ must be zero for all $l$. Notice $\mathscr{A}(1)$ contains all the trees of height 2 such that every leaf has 0 as its scattering data or the singleton node, in which case the node must have scattering state 1. Hence it follows that, for any tree of height two there is only one copy of it in $\mathscr{A}(1)$ i.e. the copy which we assign every leaf the value 0. So,

$$\begin{aligned}
J_k(1,v_1) &= \sum_{(m,\sigma')\in\mathcal{A}(\sigma)} \frac{e^{-\Gamma_1}}{r_1(m)!} \prod_{l'\in m\cap\mathbb{N}^2} \Big( \Gamma_1 - \int_0^t L[f_{s,k-1}](v_1)\,\mathrm{d}s \Big) \\
&= \sum_{j=0}^{\infty} \frac{e^{-\Gamma_1}}{r_1(m)!} \Big( \Gamma_1 - \int_0^t L[f_{s,k-1}](v_1)\,\mathrm{d}s \Big)^{r_1(m)}, \\
&= \exp\Big( -\int_0^t L[f_{s,k-1}](v_1)\,\mathrm{d}s \Big).
\end{aligned} \tag{32}$$

Now it isn't too hard to see that,

$$\int_\Omega \mathrm{d}f_0(v) = P_{t,k}(v\in\Omega,\beta_1\in\{0,1\}) = P_{t,k}(v\in\Omega,\beta_1=0)+P_{t,k}(v\in\Omega,\beta_1=1)$$
$$= \int_\Omega J_k(0,v)\,\mathrm{d}f_0(v) + \int_\Omega J_k(1,v)\,\mathrm{d}f_0(v). \tag{33}$$

Where the last equality comes directly from (31). Now combining (32) and (33) and the fact that we are dealing with probability measures one sees that,

$$\int_\Omega J_k(0,v)\,\mathrm{d}f_0(v) = \int_\Omega 1-\exp\Big(-\int_0^t L[f_{s,k-1}](v_1)\,\mathrm{d}s\Big)\,\mathrm{d}f_0(v),$$

which taken into consideration along with (31) and (32) one concludes that,

$$P_{t,k+1}(v\in\Omega,\ \beta_1=\sigma)$$
$$= \int_\Omega \Big[(1-\sigma)\Big(1-\exp\Big(-\int_0^t L[f_{s,k-1}](v)\,\mathrm{d}s\Big)\Big)+\sigma\exp\Big(-\int_0^t L[f_{s,k-1}](v)\,\mathrm{d}s\Big)\Big]\,\mathrm{d}f_0(v),$$

recalling (15) gives us the result,

$$= \int_\Omega (1-\sigma)(\mathrm{d}f_0(v)-\mathrm{d}f_{t,k}(v))+\sigma\,\mathrm{d}f_{t,k}(v).$$

Thus we are done by induction. ■

**Remark 2.29.** We have now completed the link from solutions of the the Boltzmann equation to the idealised distributions. To lay it out nicely one last time,

1. We wish to show, in the Boltzmann grad scaling limit, that the distribution of a single hard spheres behaving according to newtons laws leads to the solution of the Boltzmann equation, this is the content of (3)
2. By the strong law of large numbers it is sufficient to show that the proportion of particles in a set (this is in essence the empirical distribution) converges to the solution of the Boltzmann equation,in probability. This is the content of (3).
3. By Lemma 2.11 it is enough to look at the *truncated* solutions instead of the full solution of the Boltzmann equation.
4. By Proposition 2.28 we many instead look the idealised distributions instead of the truncated solutions.
5. The content of the next few subsections is to, firstly, set up the empirical distribution in a useful way, secondly show that the empirical distribution converges to the idealised distribution and the by applications of the above points conclude the main theorem.

▲

### 2.3.3 The Empirical Distribution $\widehat{P}_{t,k}$

The content of this subsection will be to introduce the empirical distribution. We do this by returning our thoughts to the many body problem. Recall that the particles at time $t=0$ are scattered randomly in our domain. We don't know how many particles there will be. We know that there is at least one particle, the so called *tagged particle*, this particle has uniform probability distribution with respect to the underlying measure $\mu = n(\mathbb{1}_{\mathbb{T}^d}\otimes f_0)$. The rest of the particles are generated according to a Poisson point process with density $\mu = n(\mathbb{1}_{\mathbb{T}^d}\otimes f_0)$. It is worth noting that this $n$ in the definition of $\mu$ is precisely the $n$ in the Boltzmann-grad scaling. Every realisation, $\omega$, of the processes just described we will have $N = \#\omega$ particles. The processes described above is what we call the *Tagged Poisson Point Process.*

The empirical distribution $\widehat{P}_{t,k} \in PM(\mathcal{MT})$ is a probability measure that calculates the probabilities that a certain tree will appear given the root is the tagged particle.

**Lemma 2.30.**
Let $\Phi = (m, \phi) \in \mathcal{MT}_k$ and $i^*$ is the index of the root particle. Then $\beta^{(a)}(i^*, t) = \beta_k(i^*, s_1)$.

*Proof.* This proof is done by induction over $k$. However, for the time being it is sufficient to realise that that the scattering state of a tree of height $k$ is given by the *truncated* scattering state provided that the truncation is does not change the said tree in question in any way. ■

Before we jump into the formal definition of the empirical distribution, I'd like to jog our memories of the fact that, $\Phi(t, k) = (m(t, k), \phi) \in \mathcal{MT}_k$ where $m(t, k) \in \mathcal{T}_k$ is the tree skeleton and $\phi : m(t, k) \to Y$ specifies the values at each node. It is much easier to define and understand $\widehat{P}_{t,k}$ as an algorithm, so we shall do that.

**Definition 2.31** (Empirical distribution $\widehat{P}_{t,k}$)**.**
The empirical distribution $\widehat{P}_{t,k}$ tells us the probability of the the following occurring for some chosen set.

1. Let $\Omega \in \mathcal{MT}$ be some Borel set.
2. Generate a random number of particles whose quantity and initial data depends on the tagged Poisson point process.
3. Let the generated particles go there deterministic journey and the freeze time at $t$.
4. Then look at the tree whose root is the tagged particle and if its height is above $k$, trim the tree by removing all nodes whose height is above the value of $k$.
5. Then finally look at whether the tree belongs to $\Omega$.

In short (and abusing notation) the empirical distribution is given by

$$\widehat{P}_{t,k}(\Omega) := \mathrm{Prob}_{\mathrm{tppp}}((m(t, k), \phi) \in \Omega), \tag{34}$$

where $\Omega \in \mathcal{MT}$ and is a Borel set. ▲

**Remark 2.32.** Note for any given realisation $\omega$ of the tagged Poisson point process the skeleton tree of any particle $m$ will increase monotonically in size as either $t$ or $k$ increase. This is simply because increasing $t$ allows for more collisions to occur as there is more time to do so and increasing $k$ increase the height and hence allows for larger trees. Finally note that $\phi_l$ is independent of both $k$ and $t$. From this one can see that for any fixed time $t$ the probability given by $\widehat{P}_{t,k}$ restricted to height at most $j$ is given by $\widehat{P}_{t,j}$. More simply put,

$$\widehat{P}_{t,k}\left(\left(m(t,k) \cap \left(\bigcup_{i=1}^{j} \mathbb{N}^i\right), (\phi_l)_{|l|\leq j}\right) \in \Omega\right) = \widehat{P}_{t,j}\left(\left(m(t,j), (\phi_l)_{|l|\leq j}\right) \in \Omega\right), \tag{35}$$

for all $\Omega \subseteq \mathcal{MT}_j$ and for all $j \leq k$. Note that $\Omega \subseteq \mathcal{MT}_j$ is not strictly needed but any part of $\Omega$ lying outside of $\mathcal{MT}_j$ will have measure zero. ▲

We take a quick detour to state the *disintegration theorem* which is the cornerstone for the theory of marginal distributions for general measures.

**Definition 2.33** (Radon Space)**.**
A *Radon Space* is a topological space such that every Borel Probability measure is inner regular.

Note that every separable metric space on which every probability measure is a Radon measure is a Radon space. But also recall that a Polish space is a separable completely metrisable topological space. Furthermore, we work with $\mathcal{MT}$ and $\mathcal{MT}_k$ all of which under the natural subspace topology of $\mathbb{R}^b$ (for sufficient $b$), must in turn be Polish since $\mathbb{R}^b$ is. Then as it is a known fact that probability measures on the Borel $\sigma$-algebra of any polish space are Radon measures we can rest easily tonight knowing that $\mathcal{MT}$ and $\mathcal{MT}_k$ are Radon spaces.

**Theorem 2.34** (Disintegration Theorem)**.**
Suppose the following hold,

1. $X$ and $Y$ are Radon spaces,
2. $\mu \in PM(Y)$,
3. $\pi : Y \to X$ is a Borel measurable function.

Then define $\nu \in PM(X)$ by $\nu := \mu \circ \pi^{-1} = \pi_* \mu$. Then for $\nu$ almost every $x \in X$ there exists a unique $\mu_x \in PM(Y)$ such that,

1. $x \mapsto \mu_x(B)$ is Borel measurable for all $B \in \mathcal{B}(Y)$,
2. $\mu_x$ lives on the fibre $\pi^{-1}(x)$ for $\nu$ almost every $x \in X$. i.e. $\mu_x(Y \backslash \pi^{-1}(x)) = 0$ and $\mu_x(E) = \mu_x(E \cap \pi^{-1}(x))$.
3. For all Borel measurable functions $f : Y \mapsto [0, \infty]$,
$$\int_Y f(y)\,\mathrm{d}\mu(y) = \int_X \int_{\pi_{-1}(x)} f(y)\,\mathrm{d}\mu_x(y)\,\mathrm{d}\nu(x), \tag{36}$$
in particular, for all $E \in \mathcal{B}(Y)$
$$\mu(E) = \int_X \mu_x(E)\,\mathrm{d}\nu(x). \tag{37}$$

▲

Let $(m', \phi') \in \mathcal{MT}_{k-1}$ and let $\widehat{P}_{t,k}(\,\cdot\,|(m', \phi')) \in PM(\mathcal{MT}_k)$ be the conditional probability of $\widehat{P}_{t,k}$ in the sense that,

$$\widehat{P}_{t,k}(\Omega|(m', \phi')) := \widehat{P}_{t,k}((m(k), \phi) \in \Omega \mid m \cap \mathbb{N}^j = m' \cap \mathbb{N}^j\ \forall j \in \{1, \dots k-1\}, \phi_l = \phi'_l\ \forall l \in m \text{ s.t. } |l| < k), \tag{38}$$

which for notational convince we shall refer to as $\widehat{P}_{t,k}(\,\cdot\,|\Phi')$. Letting $Y = \mathcal{MT}_k$ and $X = \mathcal{MT}_k$ and $\pi : X \to Y$ such that $\pi : \Phi \mapsto \Phi'$ where $\Phi'$ is $\Phi$ with all the nodes of height $k$ removed. Then the disintegration theorem yields that when we set $\mu = \widehat{P}_{t,k}$ and $\nu = \widehat{P}_{t,k} \circ \pi^{-1}$ that there exists a unique probability measure on $Y$, $\mu_{\Phi'} = \widehat{P}_{t,k}(\,\cdot\,|\Phi')$. But also by (38) we have,

$$\widehat{P}_{t,k}(\Omega) = \int_{\mathcal{MT}_{k-1}} \widehat{P}_{t,k}(\Omega\,|\Phi')\,\mathrm{d}\nu. \tag{39}$$

However, $\nu = \widehat{P}_{t,k} \circ \pi^{-1}$ is actually $\widehat{P}_{t,k-1}$ in disguise! Why? Well let $\Omega_n = \Omega \cap \mathcal{MT}_n$, then for $\Omega \in \mathcal{B}(\mathcal{MT})$ then,

$$\begin{aligned}\widehat{P}_{t,k}(\pi^{-1}(\Phi') \in \Omega) &\overset{(i)}{=} \widehat{P}_{t,k}(\Phi' \in \Omega_{k-1}) \\ &\overset{(ii)}{=} \widehat{P}_{t,k-1}(\Phi' \in \Omega_{k-1}).\end{aligned}$$

Equality $(i)$ is due to the fact that if $\Phi' \Omega_{k-1}$ then $\pi^{-1}(\Phi')$ is the set of all possible trees of height $k$ such that the first $k-1$ levels are given by $\Phi'$. This on the set of trees whose first $k-1$ levels are $\Phi'$ we are looking at the probability of all of them. Equality $(ii)$ follows directly from (35). Combined with (39) gives the following formula,

$$\widehat{P}_{t,k}(\Omega) = \int_{\mathcal{MT}_{k-1}} \widehat{P}_{t,k}(\Omega \,|\Phi')\,\mathrm{d}\widehat{P}_{t,k-1}(\Phi').$$

Applying (36) $k-1$ times and setting $f$ to be $\widehat{P}_{t,j}(\Phi' \mid \cdot))$ for sufficient $j$ yields the following useful formula,

$$\widehat{P}_{t,k}(\Omega) = \int_{\mathcal{MT}_1}\int_{\mathcal{MT}_2}\cdots\int_{\mathcal{MT}_{k-1}} \widehat{P}_{t,k}(\Omega \,|\Phi')\,\mathrm{d}\widehat{P}t,k-1(\Phi_{k-1} \mid \Phi_{k-2})\ldots\mathrm{d}\widehat{P}t,2(\Phi_2 \mid \Phi_1)\,\mathrm{d}\widehat{P}_{t,1}(\Phi_1), \tag{40}$$

where,

$$\widehat{P}_{t,1} = (\mathbb{1}_{\mathbb{T}^d} \otimes f_0)(z_1) \in PM(\mathbb{T}^d \times \mathbb{R}^d). \tag{41}$$

**Remark 2.35.** Define,

$$\hat{f}^{(a)}_{t,k}(u,v) := \widehat{P}_{t,k}\big(z(1,t) \in [u,u+\mathrm{d}u) \times [v,v+\mathrm{d}v) \text{ and } \beta^{(a)}(1,t)=1\big),$$

then is absolutely continuous with respect to $\mathbb{1}_{\mathbb{T}^d} \otimes f_0$. This follows directly from (40) when we realise that $\Phi_1 = z(1,t)$ and recall (41). To be explicit, it is simple to see that

$$\hat{f}^{(a)}_{t,k}(\Omega) \leq \widehat{P}_{t,k}(\Omega) \overset{(40),(41)}{\leq} \mathbb{1}_{\mathbb{T}^d} \otimes f_0(\Omega).$$

▲

#### 2.3.4 The Notion of Good Trees

In this subsection we shall talk about in the form $z=(u,v) \in \mathbb{T}^d \times \mathbb{R}^d$ rather than $(v,\nu,t) \in \mathbb{R}^d \times \mathbb{S}^{d-1} \times \mathbb{R}_{\geq 0}$. It is safe to do this for any marked tree as given the data in either form we can obtain the other form of data. i.e. marked trees with either type of data are equivalent.

We will later like to manipulate (40) in order to derive a alternative formula for the empirical distribution in such a way that the empirical distribution, on paper, looks almost identical to the idealised distribution. However this manipulation is only able to work on the set of *good trees*. These good trees are crucial as they allow us to show convergence of the two distributions due to the similarity of the distributions on the set of good trees. The two reasons that the empirical distribution might not look like the idealised distribution and hence fail to converge are:

1. One particle might appear twice at different points in the collision tree. i.e. the map $z : m \to \mathbb{T}^d \times \mathbb{R}^d$ might not be injective. These are the infamous *recollisions*.

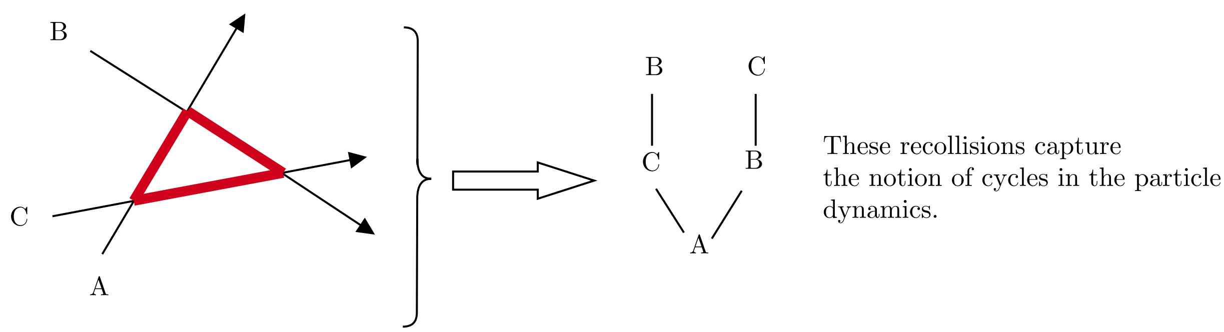



Fig 3: Illustration of how recollisions cause a loss of injectivity.

2. The other problem comes from particles whose past and future locations might intersect. This is due to the periodic boundary conditions. We call this effect *resonance*. It is worth noting that one can only avoid resonance for finite time, when the diameter of the particle is non-zero. This is due to the fact that toral translations are; periodic if the translation is parameterised by a rational vector and dense if the vector is irrational. Later on we will see that the assumption of non-resonance is essential and hence we can conclude that we have no hope of using this approach for infinite time but rather only for any finite time.

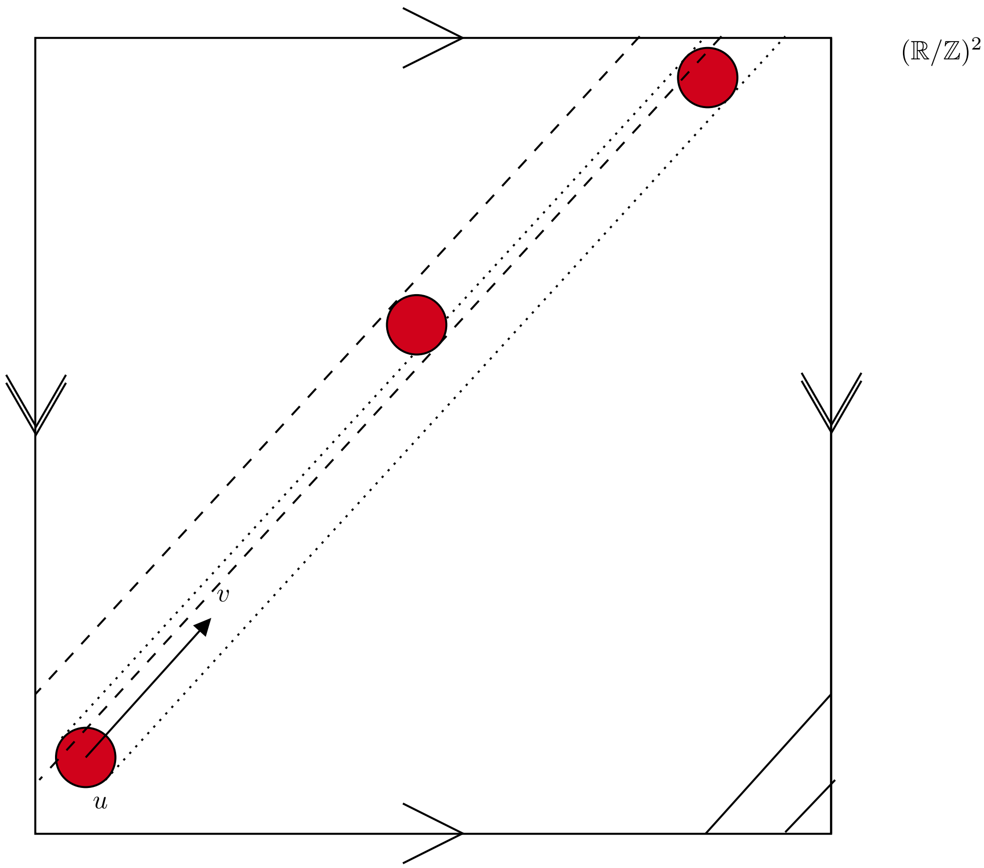



Fig 4: Particle with initial data $(u, v)$ will resonate given enough time.

This is trying to illustrate that the red particle with initial data $(u, v)$ will at first have a trajectory given by the dotted tube and then by the solid tube and finally by the dashed tube. We can see the dashed tube, which is the far future of the red particle and the dotted tube, which is the near future, intersect. Hence this red particle particle is resonating for the time frame given by the illustration.

Now that we have an idea of what kind of particles we want to avoid we introduce a more strict partial order on $\mathcal{T}$ in order to define rigorously the notion of good trees.

Let $l, l' \in m$ then we write $l < l'$ if any of the following hold;

1. $|l| < |l'|$,
2. $|l| = |l'|$ and $\bar{l} < \bar{l'}$,
3. $|\bar{l}| = |\bar{l'}|$ and $l_l < l'_l$.

This means that we can now give a partial order to the nodes of $m$ that are on the same height. If one draws a tree according to the nodes label's and does so in a sensible way this partial order essentially says that the nodes at the far left of any level of the tree are the "smallest" and the ones on the far right are the "largest".

**Definition 2.36** (Good Trees)**.**
Define the set of resonating velocities to be,

$$R(t,a) = \left\{ v \in \mathbb{R}^d \mid \min \left\{ |sv - \xi| \mid s \in [0,t], \xi \in \mathbb{Z}^d \backslash \{0\} \right\} \leq a \right\}. \tag{42}$$

Define for each $l \in m$ the set of colliding initial data to be,

$$C_l = \left\{ z' \in \mathbb{T}^d \times \mathbb{R}^d \middle| \min_{s' \in [0,s_l]} \mathrm{d}\left(z_l, z', s'\right) \leq a \right\}, \tag{43}$$

where d is as in the definition of the particle distance metric, definition 2.1.

Then for each $a_0 > 0$ the set of good trees $\mathcal{G}(a_0) \subseteq \mathcal{MT}$ consists of those trees $(m, \phi) \in \mathcal{MT}$ with the property that for all $0 < a \leq a_0$ and all $l \in m$ we have that,

$$v_l - v_{\bar{l}} \in \mathbb{R}^d \setminus R(t,a) \qquad \text{(all parent-child-pairs are non-resonant)}, \tag{44}$$

$$z_l \notin \bigcup_{\substack{l' < l \\ l' \neq \bar{l}}} C_{l'} \qquad \text{(no particle appears twice in the tree)}. \tag{45}$$

▲

**Remark 2.37** (Being a Resonant Velocity)**.**
Note $v \in R(t,a)$ if $v$ is close $v^*$ such that the components of $v^*$ are rationally dependent (i.e. dependent over $\mathbb{Z}$) and for some $\eta \in [-t,t] \setminus \{0\}$, $\eta v^* \in \mathbb{Z}^d$. Further note all periodic flows on $\mathbb{T}^d$ are the ones given by a velocity that is rationally dependent, so if $v^*$ is close to $v$ then the cylinder of $v$ almost superimposes the cylinder of $v^*$ and so if $V^*$ yields a flow whose period $\eta \in [0,t]$ then $v$ being sufficiently close to $v^*$ will cause $v$'s cylinder to overlap at time $\eta$ with itself. Finally note the distance quantified by the distance "close" is directly related to $a$ the diameter of the particles. ▲

It is worth noting that $C_l$ can be thought as an infinite union, over $\mathbb{R}^d$, of cylinder. That is to say that $C_l \cap (\mathbb{T}^d \times \{v'\})$ is a cylinder for each $v' \in \mathbb{R}^d$.

$$C_l \;\times\; \left((\mathbb{R}/\mathbb{Z})^d \;\times\; \{v'\}\right)$$

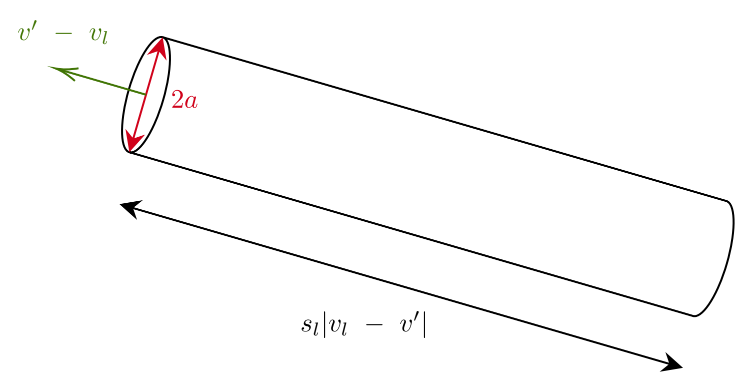



Fig 5: Illustration of a cylinder

**Remark 2.38** (A Useful Calculation)**.**
Later on it'll be useful to have the d-dimensional volume of cylinder. Thus we will do that here. Using the following figure,

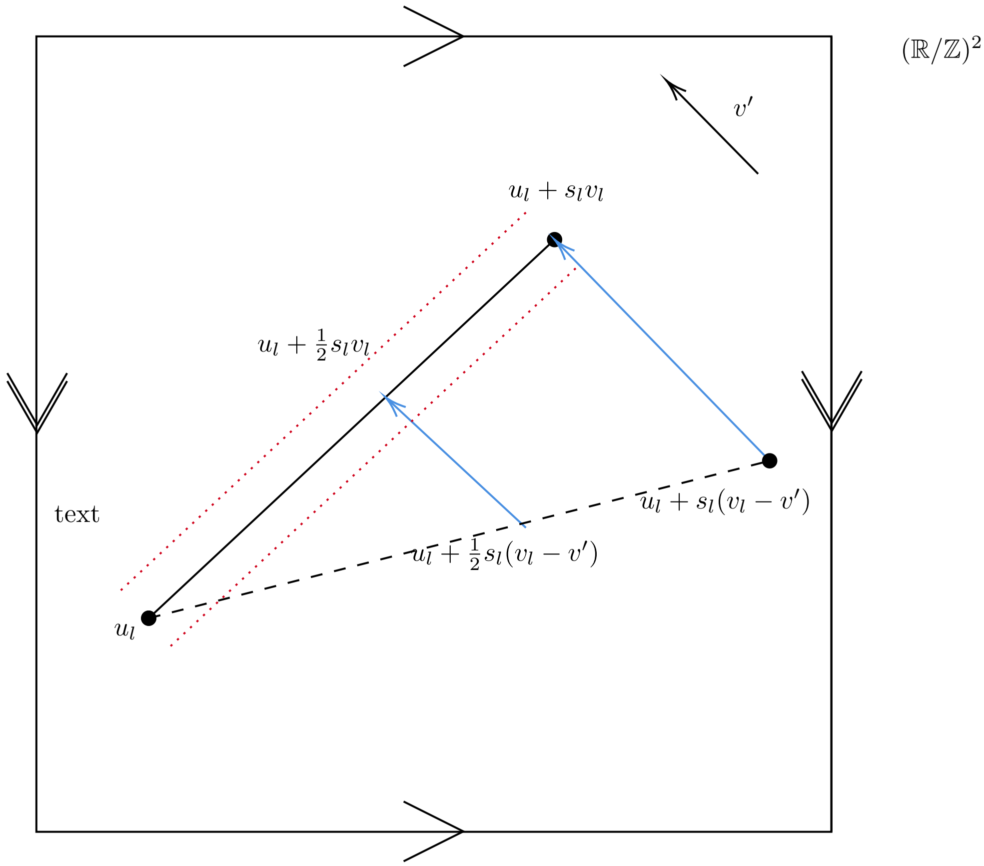



Fig 6: Illustration of how the volume of the cylinder $C_l \cap \left(\mathbb{T}^d \times \{v'\}\right)$ is found

The black line is the trajectory of the centre point of particle $l$ up to time $s_l$. The dashed line is the set of all points that with velocity $v'$ will collide with the particle centre within time $s_l$. Now noticing that a particle has diameter $a$ one realises that the dashed line can be stretched to a cylinder of radius $a$. Noting that a d-dimensional cylinder has volume given by the volume of a (d-1)-dimensional ball times the length, it is easy to see that,

$$\begin{aligned}\mathcal{H}^d(C_l \cap (\mathbb{T}^d \times \{v'\})) &= Vol(B_{d-1}(a))(|s_l(v_l - v')|),\\ &= a^{d-1}\kappa_d|v_l - v'|s_l,\\ &= \frac{\kappa_d}{n}|v_l - v'|s_l,\end{aligned} \tag{46}$$

where $B_{d-1}(a)$ is a (d-1)-dimensional ball of radius $a$ and the third equality is due to Boltzmann-Grad scaling. One final thing to note is that for (46) to hold we require $v' \in \mathbb{R}^d \setminus (v_l + R(t,a))$. This is done to ensure that we don't get overlaps akin to the one illustrated in figure 4, these overlaps would mess with our calculations and hence (46) would be too large. Be that as it may it is still useful to cognize that for $v' \in v_l + R(t,a)$,

$$\mathcal{H}^d(C_l \cap (\mathbb{T}^d \times \{v'\})) \leq \frac{\kappa_d}{n}|v_l - v'|s_l. \tag{47}$$

▲

#### 2.3.5 The Useful Formula for the Empirical Distribution

It is the goal of this subsection to find a better formula for both $\widehat{P}_{t,k}$ and $\widehat{P}_{t,k}(\,\cdot\,|\Phi_{k-1})$. In particular we'd like the formula of $\widehat{P}_{t,k}(\,\cdot\,|\Phi_{k-1})$ which in turn will give us a formula for $\widehat{P}_{t,k}$ which is very similar in structure to that of $P_{t,k}$. Handily this is possible to do on the set of good trees and even more conveniently we'll show the set of good trees is sufficiently big (i.e. is a set of full measure as in the scaling limit).

Before we begin on our escapade in search of a better formula for $\widehat{P}_{t,k}(\,\cdot\,|\Phi_{k-1})$ we should hark back to the days of Poisson point processes on $\mathbb{R}^d$. A Poisson point process in $\mathbb{R}^d$ has that if $B \in \mathcal{B}^d$ is finite in size and $\nu$ a random collection of points adhering to a Poisson point process of intensity $\lambda$ in $\mathbb{R}^d$ then,

$$|\nu \cap B| \sim Po(\lambda Vol(B)).$$

Furthermore, if $B_1, B_2 \in \mathcal{B}^d$ are disjoint then random variables $|\nu \cap B_1|$ and $|\nu \cap B_2|$ are independent and hence trivially,

$$\mathbb{P}(|\nu \cap B_1| = k || \nu \cap B_2| = 0) = \mathbb{P}(|\nu \cap B_1| = k) = \frac{(\lambda Vol(B_1))^k}{k!} e^{-\lambda Vol(B_1)}.$$

Next we will go through a simple exercise to remind ourselves of both $\text{Prob}_{\text{ppp}}$ and to remind ourselves of some basic probability.

**Example 2.39.** Let $\omega \in \mathbb{T}^d \times \mathbb{R}^d$ be a realisation of a Poisson point process with intensity $\mu \in M_+(\mathbb{T}^d \times \mathbb{R}^d)$. Then,

$$\begin{aligned}
\text{Prob}_{\text{ppp}}(|\omega| = r) &= \text{Prob}_{\text{ppp}}(\omega \in (\mathbb{T}^d \times \mathbb{R}^d)^r) \\
&= e^{-\mu(\mathbb{T}^d \times \mathbb{R}^d)} \frac{\mu^{\otimes r}((\mathbb{T}^d \times \mathbb{R}^d)^r)}{r!} \\
&= e^{-\mu(\mathbb{T}^d \times \mathbb{R}^d)} \frac{\mu((\mathbb{T}^d \times \mathbb{R}^d))^r}{r!}.
\end{aligned}$$

Then letting $A^{(r)} \subseteq (\mathbb{T}^d \times \mathbb{R}^d)^r$ and $\omega = (z_1, \ldots, z_N)$ is a vector of random length and each $z_i$ has law $\frac{\mu}{\mu(\mathbb{T}^d \times \mathbb{R}^d)}$ and are i.i.d. then,

$$\begin{aligned}
\text{Prob}_{\text{ppp}}(\omega \in A^{(r)}) &= \text{Prob}_{\text{ppp}}(\omega \in A^{(r)} || \omega| = r) \text{Prob}_{\text{ppp}}(|\omega| = r) \\
&= \frac{\mu(A^{(r)})^r}{\mu(\mathbb{T}^d \times \mathbb{R}^d)^r} e^{-\mu(\mathbb{T}^d \times \mathbb{R}^d)} \frac{\mu(\mathbb{T}^d \times \mathbb{R}^d)^r}{r!} \\
&= \frac{e^{-\mu(\mathbb{T}^d \times \mathbb{R}^d)}}{r!} \int_{A^{(r)} \cap (\mathbb{T}^d \times \mathbb{R}^d)^r} \mathrm{d}\mu(z_1) \ldots \mathrm{d}\mu(z_r).
\end{aligned}$$

Now if we are interested in more general sets, say, $A \in \bigcup_{N=0}^{\infty} (\mathbb{T}^d \times \mathbb{R}^d)^N$, then we can write $A$ in the following form $A = \bigsqcup_{r=0}^{\infty} A^{(r)}$ for suitably chosen $A^{(r)} \subseteq (\mathbb{T}^d \times \mathbb{R}^d)^r$. It is then a mere nugacity to see that,

$$\text{Prob}_{\text{ppp}}(\omega \in A) = e^{-\mu(\mathbb{T}^d \times \mathbb{R}^d)} \sum_{N=0}^{\infty} \frac{1}{N!} \int_{A \cap (\mathbb{T}^d \times \mathbb{R}^d)^N} \mathrm{d}\mu(z_1) \ldots \mathrm{d}\mu(z_r). \tag{48}$$

We now state the lemma that will be the driving force in our attempt to introduce a new formula for $\widehat{P}_{t,k}(\,\cdot\,|\Phi_{k-1})$.

**Lemma 2.40.** Let the random set $\omega \subset \mathbb{T}^d \times \mathbb{R}^d$ be distributed according to a Poisson point process with density $\mu$. Furthermore let, $\overline{\mathcal{C}}, \mathcal{C} \subset \mathbb{T}^d \times \mathbb{R}^d$ and $A \subset \cup_{r=0}^{\infty} (\mathcal{C} \backslash \overline{\mathcal{C}})^r$ be symmetric. Then we obtain the following formula for the conditional probability of the event $A$ :

$$\text{Prob}_{\text{ppp}}(\omega \cap \mathcal{C} \in A \mid \omega \cap \overline{\mathcal{C}} = \emptyset) = e^{(-\mu(\mathcal{C} \backslash \overline{\mathcal{C}}))} \sum_{r=0}^{\infty} \frac{1}{r!} \int_{A \cap \mathcal{C}^r} \mathrm{d}\mu^r(z), \tag{49}$$

where, $\mu^r = \mu^{\otimes r}$. ▲

*Proof.* The proof is omitted. All the technology needed for this proof can be found in the ramble at the beginning of this subsection. ■

The idea for the rest of this subsection is to apply Lemma 2.40 to $\text{Prob}_{\text{ppp}}$ and the notice that $\text{Prob}_{\text{tppp}}$ is a scalar multiple of $\text{Prob}_{\text{ppp}}$. It is key to note that the scalar multiple is dependent on $n$ (as in the $n$ in the Boltzmann Grad scaling) and that the constant tends to 1 at an exponential rate as $n$ tends to infinity.

If $\Omega \subseteq \mathcal{MT}$ then clearly $\Omega = \bigcup_{m\in\mathcal{T}}(\mathcal{E}(m)\cap\Omega)$. For simplicity we assume for the rest of this subsection that $\Omega \subseteq \mathcal{E}(m)$ for some $m \in \mathcal{T}$.

Recall the order defined on the nodes of $l \in m$. This allows us to write each $(z_l)_{l\in m} \in \Omega \subseteq \mathcal{E}(m)$ in the following form,

$$(z_l)_{l\in m} = (z_{l_1}, \dots, z_{l_{\#m}})$$

such that $l_1 < l_2 < \dots < l_{\#m}$. We denote the set of all such points,

$$A(\Omega) = \{(z_l)_{l\in m} | (m,\phi) \in \Omega\} \subseteq (\mathbb{T}^d \times \mathbb{R}^d)^{\#m}.$$

Next for a given tree $(z_l)_{l\in m}$ we define the collection of points in the tree that have height equal to $k$ as,

$$Z_k = (z_l)_{|l|=k} \in (\mathbb{T}^d \times \mathbb{R}^d)^{\#(m\cap\mathbb{N}^k)}.$$

Subsequently, for a given tree $\Phi \in \mathcal{MT}_{k-1}$ we define the set of trees that are in $\Omega$ but whose first $k-1$ levels are precisely $\Phi$ as,

$$A_k(\Omega,\Phi) = \left\{Z_k \in (\mathbb{T}^d \times \mathbb{R}^d)^{\#(m\cap\mathbb{N}^k)} \mid (Z_k,\Phi) \in \Omega\right\}.$$

Before we can apply Lemma 2.40 we must specify $\mathcal{C}$ and $\overline{\mathcal{C}}$. Fix $a_0 > 0$ and let $\Phi \in \mathcal{MT} \cap \mathcal{G}(a_0)$. Then define the set containing all the initial datum that would correspond with particles colliding with particles of height $k$ of $\Phi$. That is to say the set containing all possible particles that could be on level $k+1$ of $\Phi$.

$$\mathcal{C}_k = \bigcup_{l\in m\cap\mathbb{N}^k} C_l(\phi) \subseteq \mathbb{T}^d \times \mathbb{R}^d. \tag{50}$$

Now we define the set of all particles that could have collided with any of the nodes of $\Phi$ that have height strictly less than $k$.

$$\bar{\mathcal{C}_k} = \bigcup_{|l|<k} C_l(\phi) \subseteq \mathbb{T}^d \times \mathbb{R}^d. \tag{51}$$

Thus $\omega \cap \bar{\mathcal{C}_k} = \{z_l | |l| \leq k\}$. This at first glance seems wrong, one would expect the inequality to be strict. But this is not the case the particle of height $k$ will lie in the sets $C_l(\phi)$ where $|l| = k-1$ as the particles of height $k$ can collide with the particles of height $k-1$ by definition and so the assertion is indeed correct. Next due to the tree $\Phi$ being good it means that we can't have any particle appear twice in the tree and hence $\omega \cap \mathcal{C}_k(\Phi) \cap \bar{\mathcal{C}_k}(\Phi) = \emptyset$. Thus for all $\Omega \in \mathcal{MT} \cap \mathcal{G}(a_0)$ and for all $\Phi \in \mathcal{MT}_k \cap \mathcal{G}(a_0)$ one has that,

$$\widehat{P}_{t,k+1}(\Omega|\Phi) = \text{Prob}_{\text{tppp}}(\mathcal{C}_k(\Phi) \cap \omega \in sym(A_{k+1}(\Omega,\Phi)) \mid \omega \cap \mathcal{C}_k(\Phi) \cap \bar{\mathcal{C}_k}(\Phi) = \emptyset). \tag{52}$$

Where the set sym(A) is the symmetrisation of the set $A$ that is to say $(z_1, \dots, z_N) \in sym(A)$ if and only if for all $\pi \in S_N$ $(z_{\pi(1)}, \dots, z_{\pi(N)}) \in sym(A)$

**Remark 2.41.**
Prior to continuing it'd be nice to unpack (52). The left hand side of (52) tells us what is the probability of the tree of the tagged particle ending up in the good set $\Omega$, given the first $k$ levels of the tree are precisely the good tree $\Phi$. But this is exactly the same as saying what is the probability of the event $\mathcal{C}_k(\Phi) \cap \omega \in sym(A_{k+1}(\Omega,\Phi))$ occurring. Which in turn is equivalent to the right hand side of (52) since due to $\Phi$ being a good tree $\omega \cap \mathcal{C}_k(\Phi) \cap \bar{\mathcal{C}_k}(\Phi) = \emptyset$ is almost surely true and hence we can condition it out. ▲

If one stares at the definitions above for some finite time one will perceive that

$$A_{k+1}(\Omega,\Phi) \cap \left[\bigtimes_{N=1}^{\#(m\cap\mathbb{N}^{k+1})} \overline{\mathcal{C}_k}(\Phi)\right] = \emptyset.$$

This is true because both $\Phi$ and $\Omega$ are good. Notice that when $Z_{k+1} \in A_{k+1}(\Omega,\Phi) \subseteq (\mathbb{T}^d \times \mathbb{R}^d)^{\#(m\cap\mathbb{N}^{k+1})}$ then $(Z_{k+1},\Phi) \in \Omega$ and hence must be disjoint from all the potential data of the tree $\Phi$ resulting in $(Z_{k+1})_i \notin \overline{\mathcal{C}_k}(\Phi)$ for each $i$. Additionally, recall that $\Omega \subseteq \mathcal{E}(m)$ (also for ease let $r = \#(m \cap \mathbb{N}^{k+1})$). Then it follows, from the previous claim, that $A_{k+1}(\Omega,\Phi) \subseteq (\mathcal{C}_k(\Phi) \setminus \overline{\mathcal{C}_k}(\Phi))^r$. Thus by Lemma 2.40 we have that,

$$\mathrm{Prob}_{\mathrm{ppp}}(\mathcal{C}_k(\Phi) \cap \omega \in sym(A_{k+1}(\Omega,\Phi))|\omega \cap \mathcal{C}_k(\Phi) \cap \overline{\mathcal{C}_k}(\Phi) = \emptyset) = e^{-\widehat{\Gamma}_k(\Phi)} \frac{1}{r!} \int_{sym(A_{k+1}(\Omega,\Phi))} \mathrm{d}\mu^r(Z_{k+1}), \quad (53)$$

where,

$$\widehat{\Gamma}_k(\Phi) = \mu(\widehat{\mathcal{C}_k}(\Phi)), \quad (54)$$

and in turn $\widehat{\mathcal{C}_k}(\Phi) = \mathcal{C}_k(\Phi) \setminus \overline{\mathcal{C}_k}(\Phi)$. I now have to admit that it is not entirely clear to me how in (53) we can exchange the measure $\mathrm{Prob}_{\mathrm{ppp}}$ with the measure $\mathrm{Prob}_{\mathrm{tppp}}$. The difference between $\mathrm{Prob}_{\mathrm{ppp}}$ and $\mathrm{Prob}_{\mathrm{tppp}}$ is that the latter can be thought as the former measure conditional on there being at least one particle. The probability of such an event occurring is of the form $1 - e^{-C(\Omega)n}$ for some sufficient positive constant $C(Omega)$ which depends on $\Omega$. Thus the left hand side of (53) should in the scaling limit converge very quickly to the measure of the same event with the measure now being $\mathrm{Prob}_{\mathrm{tppp}}$. Thus in the scaling limit we have that,

$$\mathrm{Prob}_{\mathrm{tppp}}(\mathcal{C}_k(\Phi) \cap \omega \in sym(A_{k+1}(\Omega,\Phi))|\omega \cap \mathcal{C}_k(\Phi) \cap \overline{\mathcal{C}_k}(\Phi) = \emptyset) = e^{-\widehat{\Gamma}_k(\Phi)} \frac{1}{r!} \int_{sym(A_{k+1}(\Omega,\Phi))} \mathrm{d}\mu^r(Z_{k+1}). \quad (55)$$

Furthermore, for some fixed $n$ this equality is a very good approximation. Thus by (52) and (55) we have that,

$$\widehat{P}_{t,k+1}(\Omega|\Phi) = e^{-\widehat{\Gamma}_k(\Phi)} \frac{1}{r!} \int_{sym(A_{k+1}(\Omega,\Phi))} \mathrm{d}\mu^r(Z_{k+1}).$$

Noticing that permutations of the labels on $l \in m$ destroy the tree, we are no longer interested in the all $r$ nodes on level $k+1$ being in any position but only in some fixed position dictated by the tree structure. Thus we can drop both the symmetrisation and the $\frac{1}{r!}$ to get,

$$\widehat{P}_{t,k+1}(\Omega|\Phi) = e^{-\widehat{\Gamma}_k(\Phi)} \int_{A_{k+1}(\Omega,\Phi)} \mathrm{d}\mu^r(Z_{k+1}). \quad (56)$$

This equation (56) is precisely the kind of formula that we want. We now find a more useful expression for $\widehat{P}_{t,k}$ and then conclude this chapter with some comparisons between $ed$ and $id$.

**Lemma 2.42.** Let $\Omega \subset \mathcal{G}(a) \cap \mathcal{T}_k$ be a Borel set, then

$$\widehat{P}_{t,k}(\Omega) = \sum_{m\in\mathcal{T}_k} \int_{A(\Omega)} e^{-\sum_{j<k} \hat{\Gamma}_j(\Phi(z))} \mathrm{d}\mu^{\#m}(z). \quad (57)$$

▲

*Proof.* The proof is omitted but it is just induction over $k$ and plugging equation (56) into (40). ■

We are very close to getting $\widehat{P}_{t,k}$ to look like $P_{t,k}$, all that is left to do is apply a change in variables. We wish to change variables of the form $z = (u,v) \in \mathbb{T}^d \times \mathbb{R}^d$ to variables of the form $(v,\nu,t) \in \mathbb{R}^d \times \mathbb{S}^{d-1} \times \mathbb{R}_{\geq 0}$. Thus we are interested in the map,

$$(z_l)_{l\in m} \mapsto (u_1,v_1) \times (s_l,\nu_l,v_l)_{l\in m\setminus\{1\}}.$$

You will have to take my word for it that this map has Jacobian given by,

$$\det D_\Phi z(\Phi) = \prod_{l\in m\setminus\{1\}} \left(a^{d-1} \left(\nu_l \cdot (v_l - v_{\bar{l}})\right)_+\right).$$

Recall that $\nu = n(\mathbb{1}_{\mathbb{T}^d} \otimes f_0$. Then for each $m \in \mathcal{T}$ we have that,

$$\begin{aligned}
&\int_{A(\Omega)} e^{-\sum_{j<k} \widehat{\Gamma}_j(\Phi(z))} \mathrm{d}\mu^{\# m}(z) \\
&= \int_{\Omega} e^{-\sum_{j<k} \widehat{\Gamma}_j(\Phi)} \prod_{l \in m \setminus \{1\}} (na^{d-1}[(v_l - v_l) \cdot \nu_l]_+ \ \mathrm{d}f_0(v_l) \mathrm{d}\nu_l \ \mathrm{d}s_l \chi_{[0,s_l]}(s_l)) \mathrm{d}P_1(z_1) \\
&\stackrel{(2)}{=} \int_{\Omega} e^{-\sum_{j<k} \widehat{\Gamma}_j(\Phi)} \prod_{l \in m \setminus \{1\}} ([(v_l - v_{\bar{l}}) \cdot \nu_l)]_+ \mathrm{d}f_0(v_l) \mathrm{d}\nu_l \ \mathrm{d}s_l \chi_{[0,s_l]}(s_l)) \mathrm{d}P_1(z_1) \\
&= \int_{\Omega} e^{-\sum_{j<k} \widehat{\Gamma}_j(\Phi)} \mathrm{d}\lambda^m
\end{aligned}$$

and we are done! We have a formula for $\widehat{P}_{t,k}$ that looks almost identical to that of $P_{t,k}$, which is given by (23). Thus for any $\Omega \in \mathcal{G}(a_0)$ we have that,

$$\hat{P}_k(\Omega) = \sum_{m \in \mathcal{T}_k} \int_{\Omega \cap \mathcal{E}(m)} e^{-\sum_{j<k} \hat{\Gamma}_j(\Phi)} \mathrm{d}\lambda^m(\phi). \tag{58}$$

We now introduce the error function that gives us the error between $P_{t,k}$ and $\widehat{P}_{t,k}$. Define $e_{t,k}$, the *error function*, to be the function such that,

$$P_{t,k}(\Omega) = \widehat{P}_{t,k}(\Omega) - e_{t,k}(\Omega), \tag{59}$$

holds for all $\Omega \in \mathcal{G}(a_0)$. Thus the error function has the form,

$$e_{t,k}(\Omega) = \sum_{m \in \mathcal{T}_k} \int_{\Omega \cap \mathcal{E}(m)} \left( e^{-\sum_{j<k} \widehat{\Gamma_j}(\Phi)} - e^{-\sum_{j<k} \Gamma_j(\Phi)} \right) \mathrm{d}\lambda^m(\phi). \tag{60}$$

Finally, we talk about the sign of $e_{t,k}$.

**Lemma 2.43** (Sign of The Error)**.**
The error function $e_{t,k}(\Omega)$ is a non-negative measure. ▲

*Proof.* Let $\Omega \subseteq \mathcal{G}(a_0)$ and recollect that

$$\Gamma_k(\Phi) = \sum_{\substack{l \in m \\ |l| = k}} \gamma_l(\Phi), \qquad \text{where,} \qquad \gamma_l(\Phi) = s_l \kappa_d \int_{\mathbb{R}^d} |v_l - v'| \, \mathrm{d}f_0(v').$$

Our first step will be to show that $\mu(c_l) \leq \gamma_l(\Phi)$. Recall our useful calculation (46) and (47) found in Remark 2.38. With this now in our brains along with Fubini one has the following,

$$\begin{aligned}
\mu(C_l) &= n \int_{\mathbb{R}^d} \int_{C_l \cap (\mathbb{T}^d \times \{v'\})} \mathrm{d}\mathcal{H}^d \, \mathrm{d}f_0(v') \\
&= n \int_{\mathbb{R}^d \setminus (v_l + R(t,a))} \int_{C_l \cap (\mathbb{T}^d \times \{v'\})} \mathrm{d}\mathcal{H}^d \, \mathrm{d}f_0(v') + n \int_{v_l + R(t,a)} \int_{C_l \cap (\mathbb{T}^d \times \{v'\})} \mathrm{d}\mathcal{H}^d \, \mathrm{d}f_0(v'),
\end{aligned}$$

by (46)

$$= n \int_{\mathbb{R}^d \setminus (v_l + R(t,a))} \kappa_d s_l |v_l - v'| \, \mathrm{d}f_0(v') + n \int_{v_l + R(t,a)} \int_{C_l \cap (\mathbb{T}^d \times \{v'\})} \mathrm{d}\mathcal{H}^d \, \mathrm{d}f_0(v'),$$

and by (47),

$$
\begin{aligned}
&\leq n \int_{\mathbb{R}^d \setminus (v_l + R(t,a))} \kappa_d s_l |v_l - v'| \, \mathrm{d}f_0(v') + n \int_{v_l + R(t,a)} \kappa_d s_l |v_l - v'| \, \mathrm{d}f_0(v'), \\
&= \gamma_l(\Phi).
\end{aligned}
$$

Following which we have the following inequality,

$$
\Gamma_k(\Phi) = \sum_{\substack{l \in m \\ |l| = k}} \gamma_l(\Phi) \geq \sum_{\substack{l \in m \\ |l| = k}} \mu(C_l) \geq \mu(\mathcal{C}_k(\Phi)).
$$

Thus we are done when we combine (60) with the following observation,

$$
\Gamma_k(\Phi) \geq \mu(\mathcal{C}_k(\Phi)) \geq \mu(\widehat{\mathcal{C}_k}(\Phi) = \widehat{\Gamma}_k(\Phi)
$$

■

Formula (60) is the key formula that will allow us to quantify the difference between both the idealised and empirical distributions which will in turn allow us to show the convergence in the main theorem of this chapter.

#### 2.3.6 Total Variation of $P_{t,k} - \widehat{P}_{t,k}$

The title of this subsection is essentially the content of this chapter.

**Lemma 2.44.**
Suppose that $f_0 \in PM(\mathbb{R}^d)$, then further suppose that, $f_0$ has finite total mass and Kinetic energy

$$
\int_{\mathbb{R}^d} (1 + |v|)^2 \, \mathrm{d}f_0(v) = K_{\text{ini}} < \infty \tag{61}
$$

and that $f_0$ does not concentrate any mass on lines,

$$
\int_{\rho(v,\nu)} \mathrm{d}f_0\,(v') = 0 \text{ for all } v \in \mathbb{R}^d, \nu \in S^{d-1}, \tag{62}
$$

where $\rho(v,\nu) = v + \nu\mathbb{R}^d$. These are the assumptions in the main theorem, Theorem 2.6. It then follows that,

$$
\lim_{a \to 0} \int_{R(t,a)} (1 + |v|) \, \mathrm{d}f_0(v) \tag{63}
$$

▲

*Proof.* This is a fun proof. For each $\xi \in \mathbb{R}^d \setminus \{0\}$ and for each $a > 0$ we define the infinite cone around $\xi$ with aperture $a$ to be,

$$
M(\xi, a) = \left\{ v \in \mathbb{R}^d \mid (v \cdot \xi)^2 \geq (|\xi|^2 - a^2)|v|^2 \right\}.
$$

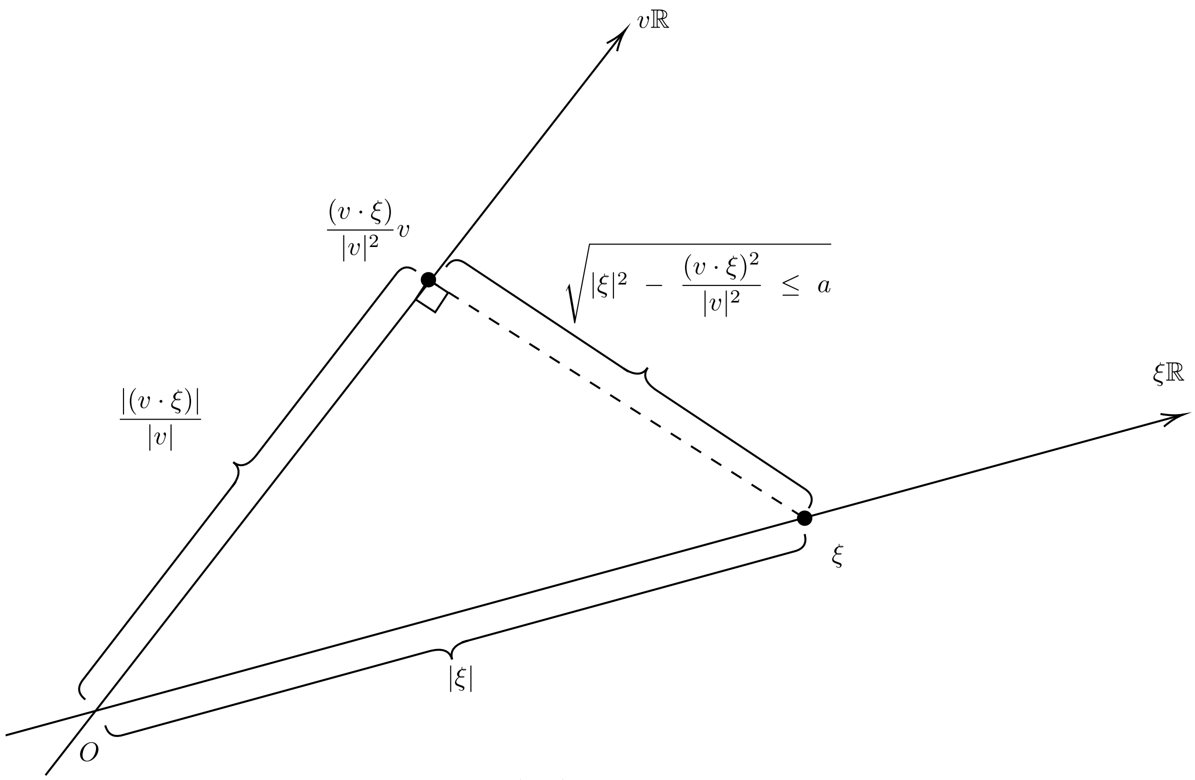



Fig: Illustration of the set $M(\xi, a)$ and showing it is indeed a cone.

We can see easily that the condition on the dotted line is equivalent to that imposed in the definition of $M(\xi, a)$. Next it is worth noting that, $M(\xi, a) \subseteq M(\frac{1}{|\xi|}\xi, a)$ whenever $|\xi| \geq 1$. To see this note if $v \in M(\xi, a)$ then,

$$(v \cdot \xi)^2 \geq (|\xi|^2 - a^2)|v|^2,$$

which is equivalent to

$$\frac{1}{|\xi|^2}(v \cdot \xi)^2 \geq \left(1 - \frac{a^2}{|\xi|^2}\right)|v|^2$$

and since $|\xi| \geq 1$ we are done since,

$$(1 - \frac{a^2}{|\xi|^2})|v|^2 \geq (1 - a^2)|v|^2.$$

In particular, this inclusion holds for all $\xi \in \mathbb{Z}^d \setminus \{0\}$. Next define,

$$c(a) := \left\{ \int_{M(\xi,a)} \mathrm{d}f_0(v) \,\middle|\, \xi \in \mathbb{Z}^d \setminus \{0\} \right\}.$$

We claim that $\limsup_{a\to 0} c(a) = 0$ and we support this claim by contradiction. Suppose for the sake of absurdity that $\limsup_{a\to 0} c(a) > 0$, then there exists a sequence $\overline{\xi}_i \in \mathbb{Z}^d \setminus \{0\}$ and $a_i > 0$ such that $a_i \to 0$ and

$$\lim_{i\to\infty} \int_{M\left(\overline{\xi}_i, a_i\right)} \mathrm{d}f_0(v) > 0$$

Then let $\xi_i$ be the normalisation of $\overline{\xi}_i$ and let $\xi_j$ be a converging subsequence of $\xi_i$ with limit $\xi$. Then since $M(\overline{\xi}_i, a_i) \subseteq M(\xi_i, a_i)$ and since $f_0$ is a probability measure we conclude that,

$$\lim_{i\to\infty} \int_{M(\xi_i,a_i)} \mathrm{d}f_0(v) > 0.$$

Furthermore, by definition we have,

$$\lim_{j\to\infty}\int_{M(\xi_j,a_j)} \mathrm{d}f_0(v) > 0.$$

In order to obtain our desired contradiction we claim that,

$$M(\xi_j,a_j) \subseteq M\left(\xi, a_j + |\xi-\xi_j| + 2\sqrt{|\xi-\xi_j|}\right). \tag{†}$$

This claim is very easy to see to be true when one draws a picture.

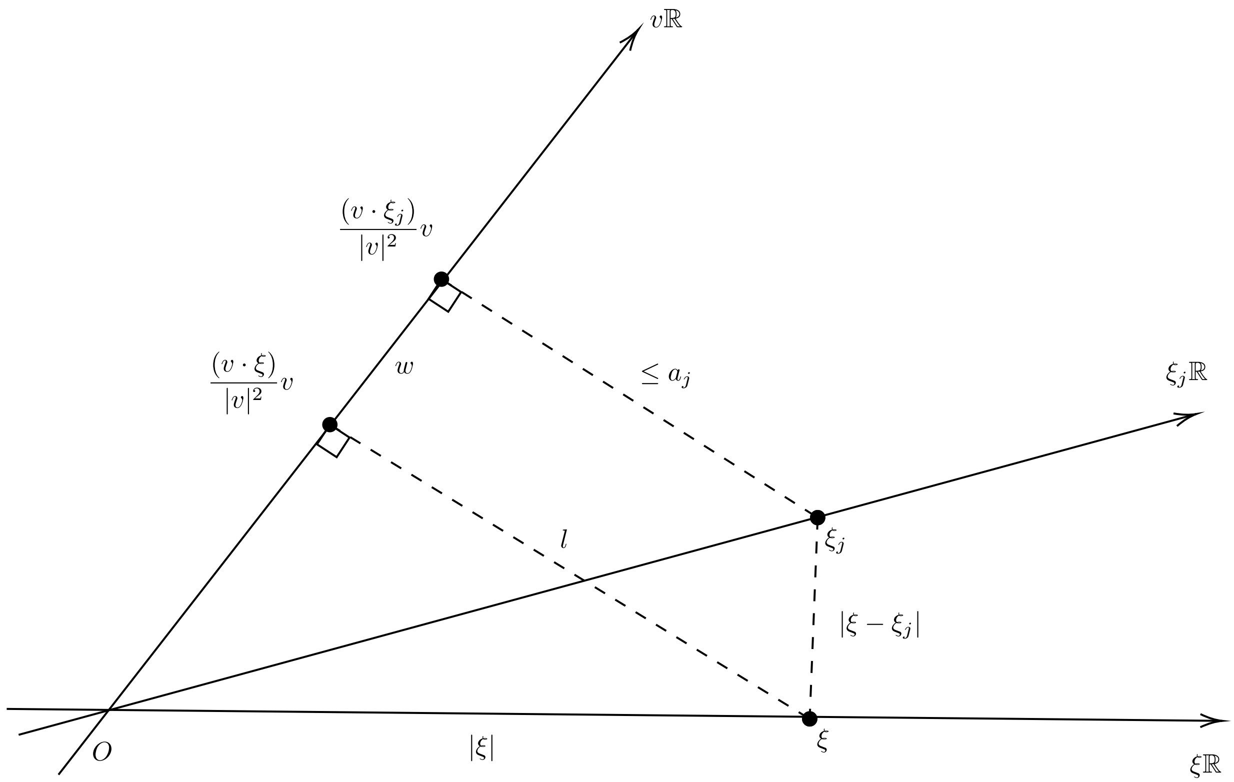



Fig: By rotating the centre of a cone slightly and increasing the aperture slightly we can absorb the original cone. (Note in this figure $l$ and $w$ are lengths and not vectors)

It is easy to see that $l \leq w + a_j + |\xi-\xi_j|$. However,

$$w = \left|\frac{(\xi-\xi_j)\cdot v}{|v|^2}v\right| \leq |\xi-\xi_j|,$$

but since $\xi,\xi_j\in\mathbb{S}^{d-1}$ we have $|\xi-\xi_j|\in[0,2]$. Then when we see that $x\leq 2\sqrt{x}$ for all $x\in[0,2]$ we realise that indeed (†) holds since we have just shown that $w\leq 2\sqrt{|\xi-\xi_j}$. Whence, we can conclude,

$$\lim_{j\to\infty}\int_{M(\xi_j,a_j+|\xi-\xi_j|+2(|\xi-\xi_j|)^{\frac{1}{2}})} \mathrm{d}f_0(v) \geq \lim_{j\to\infty}\int_{M(\xi_j,a_j)} \mathrm{d}f_0(v) > 0,$$

which in turn implies,

$$\limsup_{a\to 0}\int_{M(\xi,a)} \mathrm{d}f_0(v) > 0.$$

and thus there exists $a_j \to 0$ as $j \to 0$ such that,

$$\lim_{j\to\infty} \int_{M(\xi,a_j)} \mathrm{d}f_0(v) > 0$$

and by the dominated convergence theorem,

$$\lim_{j\to\infty} \int_{M(\xi,a_j)} \mathrm{d}f_0(v) = \lim_{j\to\infty} \int_{\xi\mathbb{R}} \mathrm{d}f_0(v).$$

This now is absurd! To see that it is absurd all we have to do is look at the assumption labeled (63). Hence we have shown $c(a) = o(a)$ as $a \to 0$ (i.e. $\lim_{a\to 0} c(a) = 0$). Next let

$$v \in R(t,a) = \{v \in \mathbb{R}^d \mid \min\{|sv - \xi| \mid s \in [0,t], \xi \in \mathbb{Z}^d \setminus \{0\}\} \le a\}$$

and further suppose that $|v| \le \overline{V}$. Thus there exists a $\xi in \mathbb{Z}^d \setminus \{0\}$ and $s \in [0,t]$ such that $|sv - \xi| \le a$ and since $|v| \le \overline{V}$ and $s \in [0,t]$ it is simple to see that $|\xi| \le a + t\overline{V}$. The net question is how many possible such $\xi$'s exist? Well since we have a bound on their speed we know that all such $\xi$ lie in the d-dimensional cube of side length $2a + 2t\overline{V}$ and so there are at most $(2a + 2t\overline{V})^d$ such $\xi$'s. Finally, by looking at the below diagram,

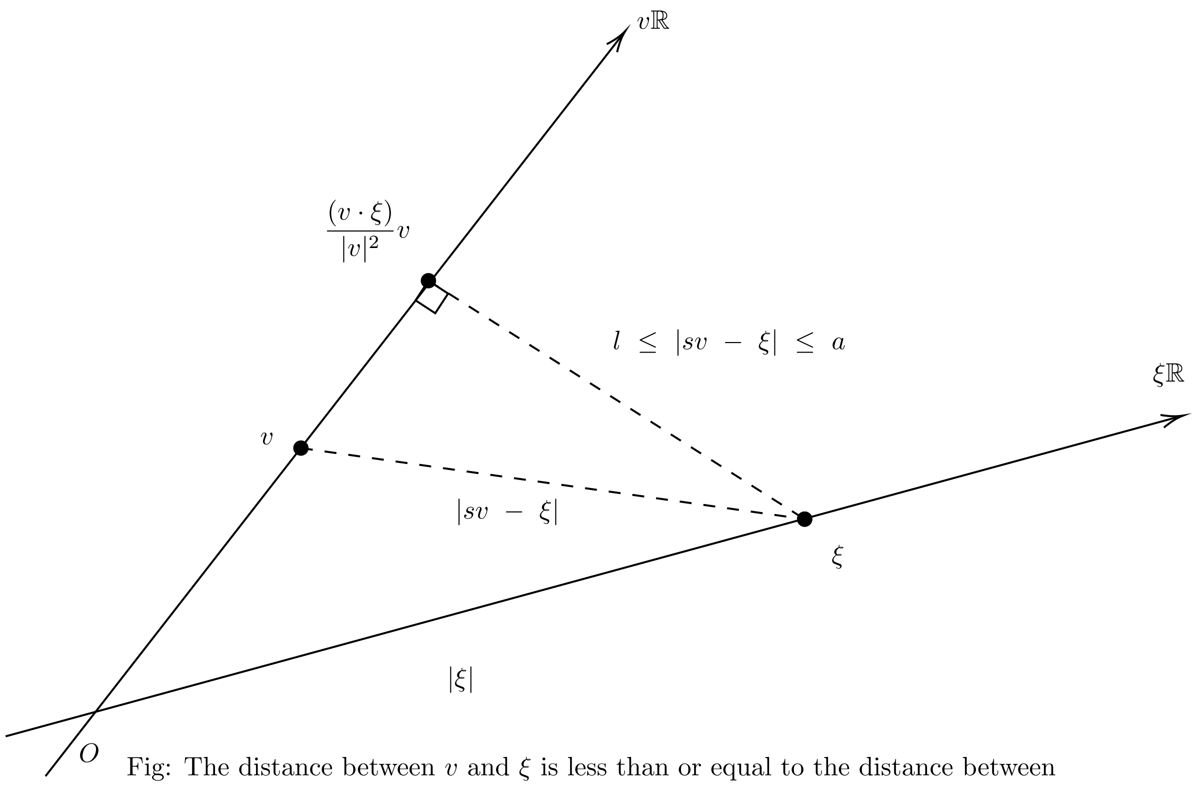



Fig: The distance between $v$ and $\xi$ is less than or equal to the distance between $\xi$ and the projection of $\xi$ onto $v\mathbb{R}$.

it is easy to see that $v \in M(\xi, a)$. Thus to glue everything together, each $v \in R(t,a)$ such that $v \le \overline{V}$ has at least one $\xi(v) \in \mathbb{Z}^d \setminus \{0\}$ such that $|\xi(v)| \le \overline{V}t + a$ and $v \in M(\xi(v), a)$ and there are at most $(2a + 2t\overline{V})^d$ such

$\xi(v)$'s. Thus,

$$
\begin{aligned}
\int_{R(t,a)} 1+|v'|\,\mathrm{d}f_0(v') &\le \int_{R(t,a)\cap\{|v'|\le \overline{V}\}} 1+|v'|\,\mathrm{d}f_0(v') + \int_{\{|v'|>\overline{V}\}} 1+|v'|\,\mathrm{d}f_0(v') \\
&\le (1+\overline{V})(2t\overline{V}+2a)^d c(a) + \int_{\{|v'|>\overline{V}\}} 1+|v'|\,\mathrm{d}f_0(v') \\
&\le (1+\overline{V})(2t\overline{V}+2a)^d c(a) + \frac{1}{\overline{V}}\int_{\{|v'|>\overline{V}\}} |v'|+|v'|^2\,\mathrm{d}f_0(v') \\
&\le (1+\overline{V})(2t\overline{V}+2a)^d c(a) + \frac{K_{\mathrm{ini}}}{\overline{V}}.
\end{aligned}
$$

Thus, choosing $\overline{V}$ sufficiently large first makes the second term sufficiently small then choose $a$ sufficiently small to make the first term sufficiently small since $c(a)\to 0$ as $a\to 0$. Hence,

$$
\lim_{a\to 0}\int_{R(t,a)} 1+|v'|\,\mathrm{d}f_0(v') = 0
$$

■

In a crucial theorem we will need to control the velocities and angles of particles on trees, this is purely technical. Thus we introduce a modified subset of good trees that satisfies all the properties that we need.

**Definition 2.45** (Well Behaved Good Trees)**.**
Let $\varepsilon(a)$ and $V(a)$ be monotone positive functions of $a$ such that $\lim_{a\to 0}\varepsilon(a)=0$ and $\lim_{a\to 0}V(a)=+\infty$. We define the set

$$
\begin{aligned}
\widehat{\mathcal{G}}(a_0) = \bigcap_{0<a<a_0}\Bigg\{(m,\phi)\in\mathcal{G}(a)\,\Bigg|\, &\min_{l\in m}|v_l - v_{\bar{l}}| \ge \varepsilon(a) \text{ and } |v|\le V(a) \\
&\text{and } \min_{l\in m}\min_{l'<l,\,l'\neq\bar{l}}\left(1-\left|\frac{v_l-v_{\bar{l}}}{|v_l-v_{\bar{l}}|}\cdot\frac{v_{l'}-v_{\bar{l}}}{|v_{l'}-v_{\bar{l}}|}\right|\right)\ge\varepsilon(a)\Bigg\}.
\end{aligned}
\tag{64}
$$

▲

**Remark 2.46** (Remarks of Definition 2.45)**.**
$\widehat{\mathcal{G}}(a_0)$ is the collection of trees such that:

- The trees are good for all $a\in(0,a_0)$,
- The velocity difference between any node and its parent node is sufficiently large.
- Particles are never too fast.
- The angle given by the angle induced by the vectors $v_l - v_{\bar{l}}$ and $v_{l'} - v_{\bar{l}}$ is sufficiently large.

▲

It turns out we don't need to consider all good trees, it is actually sufficient to consider the subset of all well behaved good trees (cf. Definition 2.45). This is shown in the following Lemma.

**Lemma 2.47.** For any $V(\cdot)$ and $\varepsilon(\cdot)$ that are monotone positive functions such that $\lim_{a_0\to 0}V(a_0)=+\infty$ and $\lim_{a_0\to 0}\varepsilon(a_0)=0$. Then we have,

$$
\lim_{a_0\to 0}\sup_{k\in\mathbb{N}} P_{t,k}\big(\widehat{\mathcal{G}}(a_0)\big) = 1 \tag{65}
$$

▲

*Proof.* Since $\varepsilon$ and $V$ are monotone functions one can, from their asymptotic behaviour, deduce that, $\varepsilon(\beta) \leq \varepsilon(\alpha)$, $V(\beta) \leq V(\alpha)$ and $\mathcal{G}(\alpha) \subseteq \mathcal{G}(\beta)$ for $\alpha \geq \beta$. Consequently $\widehat{\mathcal{G}}(\alpha) \subseteq \widehat{\mathcal{G}}(\beta)$ and so $\widehat{\mathcal{G}}(a_0)$ increases monotonically as $a_0 \to 0$. Then from upwards continuity of measures we get that,

$$\lim_{a_0 \to 0} P_{t,k}\big(\widehat{\mathcal{G}}(a_0)\big) = P_{t,k}\bigg(\bigcup_{a_0>0} \widehat{\mathcal{G}}(a_0)\bigg).$$

It thus follows that, for $(m, \phi) \in \bigcup_{a_0>0} \widehat{\mathcal{G}}(a_0) \subseteq \bigcup_{a_0>0} \mathcal{G}(a_0)$, then $z_l$ must adhere to, both,

$$v_l - v_{\bar{l}} \notin R(t,0), \tag{$\mathscr{A}$}$$

$$z_l \notin \bigcup_{\substack{l'<l \\ l' \neq \bar{l}}} C_{l'}(a_0 = 0), \tag{$\mathscr{B}$}$$

for all $l \in m$. We will show that both sets $\mathscr{A}$ and $\mathscr{B}$ have measure zero with respect to the measure $P_{t,k}$. Now it is not too hard to see that $R(t,0) = \{q \in \mathbb{Q}^d \mid |qt| \geq 1\}$ which is a countable union of half-lines. Now reminisce on the good old days of when you read the definition of $\lambda(\phi)$ (cf (20)) and realise the component of the measure $\lambda$ that measures the roots velocity is only $f_0$. However, the component of $\lambda$ that measures the velocity of any other particle in a tree looks like $((v_l - v_{\bar{l}}) \cdot \nu_l)_+ \, \mathrm{d}f_0(v_l)$. Thus when we consider the root particle we apply assumption (9) to a countable union of lines to conclude that each line has measure zero. When considering a particle away from the root we simply notice that $((v_l - v_{\bar{l}}) \cdot \nu_l)_+ \leq (1 + |v_l|) + (1 + |v_{\bar{l}}|)$ and apply Lemma 2.44. This establishes that the set in $\mathscr{A}$ is a null set.

Next we turn to show that the event $\mathscr{B}$ is almost surely true. This is given to us if we can show for any $l \in m$ we have that $P_{t,k}(C_l(a_0 = 0)) = 0$. It is now that we make an admission in that we have been ignoring something that was quite frankly a waste of space up until now. If one recalls in the definition of marked trees Definition 2.14 when we assign data to the node (i.e. mark it) we actually still give the node its initial data even though it can be obtained from the rest of the data. This has the the consequences that in the idealised distribution we should have a component measure that measures the location of the initial position of each particle in the spacial domain, which is $\mathbb{T}^d$. Thus in the product term of (23) we should have in each term the following measure $\mathbb{1}_{\mathbb{T}^d}(u_l)$. However this term does not effect anything and always integrates out to one hence it was nothing but a waste of space including it in (23), here however it'll be useful to show that $C_l(a_0)$ is a null set. For the sake of ease remember that,

$$C_l(a_0 = 0) = \left\{ z' \in \mathbb{T}^d \times \mathbb{R}^d \,\middle|\, \min_{s' \in [0, s_l]} d(z_l, z', s') = 0 \right\},$$

and after introducing the above mentioned "correction",

$$\lambda^m(\phi) = \mathbf{1}_{\mathbb{T}^d}(u_1) \otimes f_0(v_1) \otimes \delta(s_1 - t) \otimes \prod_{l \in m \setminus \{1\}} [((v_l - v_{\bar{l}}) \cdot \nu_l)_+ \chi_{[s_{l-1}, s_l]}(s_l) \mathrm{d}f_0(v_l) \mathrm{d}\nu_l \, \mathrm{d}s_l \, \mathrm{d}\mathbb{1}_{\mathbb{T}^d}(u_l)].$$

It then follows that,

$$P_{t,k}(C_l(a_0 = 0)) = \int_{C_l(a_0=0)} e^{-\sum_{j<l} \Gamma_j} \, \mathrm{d}\lambda^m,$$

since the set we are integrating over only depends on $z_l$ and $v_{\bar{l}}$ were for the purposes of the integral $v_{\bar{l}}$ is treated as a constant it is thus sufficient to consider;

$$\begin{aligned} \int_{C_l(a_0=0)} e^{-\sum_{j<l} \Gamma_j} |v_l - v_{\bar{l}}| \, \mathrm{d}\mathbb{1}_{\mathbb{T}^d}(u_l) \, \mathrm{d}f_0(v_l) &= \int_{\mathbb{R}^d} \int_{C_l(a_0=0) \cap (\mathbb{T}^d \times \{v_l\})} e^{-\sum_{j<l} \Gamma_j} |v_l - v_{\bar{l}}| \, \mathrm{d}\mathbb{1}_{\mathbb{T}^d}(u_l) \, \mathrm{d}f_0(v_l) \qquad (\dagger) \\ &\leq \int_{\mathbb{R}^d} \int_{C_l(a_0=0) \cap (\mathbb{T}^d \times \{v_l\})} |v_l - v_{\bar{l}}| \, \mathrm{d}\mathbb{1}_{\mathbb{T}^d}(u_l) \, \mathrm{d}f_0(v_l) \\ &\leq \int_{\mathbb{R}^d} \int_{C_l(a_0=0) \cap (\mathbb{T}^d \times \{v_l\})} |v_l| + |v_{\bar{l}}| \, \mathrm{d}\mathbb{1}_{\mathbb{T}^d}(u_l) \, \mathrm{d}f_0(v_l). \end{aligned}$$

Now the key step to notice is that $C_l(a_0 = 0) \cap (\mathbb{T}^d \times \{v_l\})$ is a finite d-1 dimensional plane (i.e. a d-dimensional cylinder of radius 0) and is in fact compact. Thus the first integral is a constant, namely, $|v_l| + |v_{\bar{l}}|$ times the d-dimensional Hausdorff measure of a d-1 dimensional surface which is quite nicely zero. Thus we have that the left hand side of (†) is indeed zero and so we have that ($\mathscr{B}$) is almost surely true. From this it follows that given any skeleton tree $m \in \mathcal{T}$ the probability that we violate either ($\mathscr{A}$) or ($\mathscr{B}$) is zero and so,

$$\widehat{P}_{t,k}\left(\mathcal{E}(m) \cap \bigcup_{a_0>0} \widehat{\mathcal{G}}(a_0)\right) = \widehat{P}_{t,k}(\mathcal{E}(m)).$$

Hence,

$$\lim_{a_0 \to 0} \widehat{P}_{t,k}(\widehat{\mathcal{G}}(a_0)) = \sum_{m \in \mathcal{T}} \widehat{P}_{t,k}\left(\mathcal{E}(m) \cap \bigcup_{a_0>0} \widehat{\mathcal{G}}(a_0)\right) = \sum_{m \in \mathcal{T}} \widehat{P}_{t,k}(\mathcal{E}(m)) = 1.$$

Now adding the supremum over $k$ is a simple exercise. ■

Now that we have shown that the subset of trees we want to consider is sufficiently large. Following this happy turn of events we want to show that the empirical distribution is convergent to the idealised distribution on this subset of trees. More over this convergence is "uniform" (i.e. in the total variation norm).

**Lemma 2.48** (Convergence in Total Variation)**.**
For any $V(\cdot)$ and $\varepsilon(\cdot)$ that are monotone positive functions such that $\lim_{a_0 \to 0} V(a_0) = +\infty$ and $\lim_{a_0 \to 0} \varepsilon(a_0) = 0$. It then follows that,

$$\lim_{a_0 \to 0} \lim_{a \to 0} \sup_{\substack{\Omega \in \mathcal{MT}, \text{Borel}, \\ k \geq 1}} \left| \widehat{P}_{t,k}(\Omega \cap \widehat{\mathcal{G}}(a_0)) - P_{t,k}(\Omega \cap \widehat{\mathcal{G}}(a_0)) \right| = 0 \tag{66}$$

▲

**Remark 2.49.** When I first read this I thought that $a_0$ was to do with particle diameter and the set of marked well behaved marked trees and hence was subsequently confused about the existence of the limit indexed by $a$. Foolishly I was wrong! In fact what the originals authors mean is that we first send the particle size to zero which is indexed by $a$ and then we send the set of well behaved marked trees to the set of all marked trees (well almost surely), this is indexed by $a_0$ ▲

*Proof.* Fix $a_0$ and let $\Omega \subseteq \widehat{\mathcal{G}}(a_0)$. From Lemma 2.27 (an upper bound on the expected number of nodes) we have that for the random variable defined by $X(m) = \#m - 1$ that,

$$\mathbb{E}(X + 1) \leq K_{\text{ini}} e^{\kappa_d K_{\text{ini}} t} + 1. \tag{67}$$

Since $X(m) \geq 0$ we can invoke Markov's inequality to get,

$$\sum_{\#m-1>r} P_{t,k}(\mathcal{E}(m)) = P_{t,k}(X(m) > r) < \frac{\mathbb{E}(X)}{r} \leq K_{\text{ini}} e^{\kappa_d K_{\text{ini}} t}, \tag{68}$$

Furthermore, it simply follows from the definition of a probability measure that,

$$1 = \sum_{m \in \mathbb{T}} P_{t,k}(\mathcal{E}(m)) = \sum_{\#m-1 \leq r} P_{t,k}(\mathcal{E}(m)) + \sum_{\#m-1>r} P_{t,k}(\mathcal{E}(m)) \leq \sum_{\#m-1 \leq r} P_{t,k}(\mathcal{E}(m)) + \frac{K_{\text{ini}}}{r} e^{\kappa_d K_{\text{ini}} t}. \tag{69}$$

Thus for any given $\delta$ we can find a $r$ such that $r \geq \dfrac{K_{\text{ini}}}{\delta} e^{\kappa_d K_{\text{ini}} t} + 1$. Consequently, we get that,

$$\sum_{\substack{\#m \leq r \\ m \in \mathcal{T}}} P_{t,k}(\mathcal{E}(m)) \geq 1 - \delta. \tag{70}$$

Now define,

$$
\begin{aligned}
I_1 &:= \sup_{k\geq 1} \sum_{\substack{m\in\mathcal{T}\\ \#m\leq r}} \left|\widehat{P}_{t,k}(\Omega\cap\mathcal{E}(m)) - P_{t,k}(\Omega\cap\mathcal{E}(m))\right|,\\
I_2 &:= \sup_{k\geq 1} \sum_{\substack{m\in\mathcal{T}\\ \#m> r}} \widehat{P}_{t,k}(\widehat{\mathcal{G}}(a_0)\cap\mathcal{E}(m)),\\
I_3 &:= \sup_{k\geq 1} \sum_{\substack{m\in\mathcal{T}\\ \#m> r}} P_{t,k}(\widehat{\mathcal{G}}(a_0)\cap\mathcal{E}(m)).
\end{aligned}
$$

Thus it follows that for all $r>0$, that,

$$
\lim_{a\to 0}\sup_{k\geq 1}\left|\widehat{P}_{t,k}(\Omega) - P_{t,k}(\Omega)\right| \leq \lim_{a\to 0}(I_1+I_2+I_3)
$$

and by showing that, $\lim_{a\to 0} I_1 = 0$ and that $\limsup_{a\to 0}(I_2+I_3)\leq\delta$ we will be done.
Case $I_1$ :
Fixing $r>0$, then there exist a finite number of trees whose number of nodes is at most $r$. Thus it suffices to show that,

$$
\lim_{a\to 0}\sup_{k\geq 1}\left|\widehat{P}_{t,k}(\Omega\cap\mathcal{E}(m)) - P_{t,k}(\Omega\cap\mathcal{E}(m))\right|,
$$

for all $m\in\mathcal{T}$ such that $\#m\leq r$. Then from the definitions of $P_{t,k}$ and $\widehat{P}_{t,k}$ and equation (59) we have that,

$$
0\leq e_{t,k}(\Omega\cap\mathcal{E}(m)) = \int_{\Omega\cap\mathcal{E}(m)} e^{-\sum_{j<k}\widehat{\Gamma}_j(\Phi)} - e^{-\sum_{j<k}\Gamma_j(\Phi)}\,\mathrm{d}\lambda^m(\Phi).
$$

Then in the proof of Lemma 2.43 we have that $\Gamma_j(\Phi)\geq\widehat{\Gamma}_j(\Phi)$ and so by simply factoring out we get that,

$$
e_{t,k}(\Omega\cap\mathcal{E}(m)) \leq \int_{\widehat{\mathcal{G}}(a_0)\cap\mathcal{E}(m)} e^{-\sum_{j<k}\Gamma_j(\Phi)}\left[e^{\sum_{j<k}(\Gamma_j(\Phi)-\widehat{\Gamma}_j(\Phi))} - 1\right]\mathrm{d}\lambda^m(\Phi).
$$

Then we can quickly prove this case if we show that there exists a $K(a_0,a)>0$ such that $\lim_{a\to 0}K(a_0,a)=0$ and for all $\Phi\in\mathcal{E}(m)\cap\widehat{\mathcal{G}}(a_0)$ and for all $j\in\mathbb{N}$ we have that,

$$
0\leq\Gamma_j(\Phi)-\widehat{\Gamma}_j(\Phi)\leq K(a_0,a) \tag{71}
$$

holds. Assuming equation (71) and invoking the fact that $e^x-1\leq xe^x$ for all $x\in\mathbb{R}$ and further since, $\int_{\widehat{\mathcal{G}}(a_0)\cap\mathcal{E}(m)} e^{-\sum_{j<k}\Gamma_j(\Phi)}\,\mathrm{d}\lambda^m(\Phi)\leq 1$ and $\sum_{j\leq k}1\leq\#m\leq r$ we have that,

$$
\begin{aligned}
e^{\sum_{j<k}(\Gamma_j(\Phi)-\widehat{\Gamma}_j(\Phi))} - 1 &\leq \sum_{j<k}(\Gamma_j(\Phi)-\widehat{\Gamma}_j(\Phi))e^{\sum_{j<k}(\Gamma_j(\Phi)-\widehat{\Gamma}_j(\Phi))}\\
&\leq rK(a_0,a)e^{rK(a_0,a)}.
\end{aligned}
$$

This results in that,

$$
e(\Omega\cap\mathcal{E}(m))\leq rK(a_0,a)e^{rK(a_0,a)}\int_{\widehat{\mathcal{G}}(a_0)\cap\mathcal{E}(m)} e^{-\sum_{j<k}\Gamma_j(\Phi)}\,\mathrm{d}\lambda^m(\Phi)\leq rK(a_0,a)e^{rK(a_0,a)}.
$$

which leads us to deduce that,

$$
0\leq\left|\widehat{P}_{t,k}(\Omega\cap\mathcal{E}(m)) - P_{t,k}(\Omega\cap\mathcal{E}(m))\right|\leq rK(a_0,a)e^{rK(a_0,a)}, \tag{72}
$$

and since the right hand side is independent of $k$ we can introduce the supremum, over k, without any problems. Thus given the estimate (71) we have that $\lim_{a\to 0} I_1 = 0$. We now proceed to prove equation (71). As $\mu = n(\mathbb{1}_{\mathbb{T}^d} \otimes f_0)$ and that $\widehat{\Gamma}_j(\Phi) = \mu(\widehat{\mathcal{C}}_j(\Phi))$ where $\widehat{\mathcal{C}}_j(\Phi) = \mathcal{C}_j(\Phi)) \setminus \bar{\mathcal{C}}_j(\Phi))$, $\mathcal{C}_j(\Phi) = \bigcup_{l\in m\cap\mathbb{N}^k} C_l(\phi)$ and $\bar{\mathcal{C}}_j(\Phi) = \bigcup_{|l|<k} C_l(\phi)$ and finally defining the errors to be,

$$\begin{aligned}
e_1 &= n\int_{\mathbb{R}^d} \mathcal{H}^d((\mathcal{C}_j(\phi)\cap\bar{\mathcal{C}}_j(\phi))\cap(\mathbb{T}^d\times\{v'\})\,\mathrm{d}f_0(v'),\\
e_2 &= n\sum_{|l|=j}\int_{v_l+R(t,a)} \mathcal{H}^d(C_j(\phi)\cap(\mathbb{T}^d\times\{v'\})\,\mathrm{d}f_0(v')\\
e_3 &= n\sum_{|l|=j}\int_{v_l+R(t,a)} \kappa_d|v_l-v'|t\,\mathrm{d}f_0(v') = n\sum_{|l|=j}\int_{R(t,a)} \kappa_d|v'|t\,\mathrm{d}f_0(v')
\end{aligned}$$

we have that,

$$\begin{aligned}
\widehat{\Gamma}_j(\Phi) &= n\int_{\mathbb{R}^d}\int_{\widehat{\mathcal{C}}_j(\Phi)\cap(\mathbb{T}^d\times\{v'\})} \mathrm{d}\mathcal{H}^d\,\mathrm{d}f_0(v')\\
&= n\int_{\mathbb{R}^d} \mathcal{H}^d(\widehat{\mathcal{C}}_l(\Phi)\cap(\mathbb{T}^d\times\{v'\})\,\mathrm{d}f_0(v')\\
&\geq n\sum_{|l|=j}\int_{\mathbb{R}^d} \mathcal{H}^d(C_j(\phi)\cap(\mathbb{T}^d\times\{v'\})\,\mathrm{d}f_0(v') - re_1
\end{aligned}$$

(In the original paper the constant $r$ in front of $e_1$ does not appear. I have to admit that I have no idea why this last inequality is the way that it is int the original paper. A simple application of the definition of a measure applied to the Hausdorff measure suggests the inequality should be the other way round. How the authors get this inequality is completely beyond my grasp. However since we know that $j \leq \#m \leq r$ we can apply the inclusion exclusion principle and hence fix this problem as we'll see later that $K(a_0, a)$ depending on $r$ will not cause any problems.)

$$= n\sum_{|l|=j}\int_{\mathbb{R}^d\setminus(v_l+R(t,a))} \mathcal{H}^d(C_l(\phi)\cap(\mathbb{T}^d\times\{v'\})\,\mathrm{d}f_0(v') - re_1 + e_2,$$

then by (46),

$$\begin{aligned}
&= n\sum_{|l|=j}\int_{\mathbb{R}^d\setminus(v_l+R(t,a))} \kappa_d|v_l-v'|t\,\mathrm{d}f_0(v') - re_1 + e_2,\\
&= \Gamma_j(\Phi) - re_1 + e_2 + e_3.
\end{aligned}$$

Now set $K(a_0, a) = re_1 - e_2 - e_3$. We will now proceed to show that $\lim_{a\to 0} e_j = 0$ for $j = 1, 2, 3$. From (46) and (47) we have that in general, $n\mathcal{H}^d(C_l(\phi)\cap(\mathbb{T}^d\times\{v'\}) \leq \kappa_d|v_l - v'|t$ and since $\sum_{|l|=j} 1 \leq \#m \leq r$ we get that,

$$e_2 + e_3 \leq 2\kappa_d r t\int_{R(t,a)} |v|\,\mathrm{d}f_0(v)$$

and so by Lemma 2.44 we have that $\lim_{a\to 0} e_2+e_3 = 0$. The hard part of this proof is to show that $\lim_{a\to 0} e_1 = 0$, which we will do now. We want to bound $\mathcal{H}^d((C_{\bar{l}}\cap C_{l'})\cap(\mathbb{T}^d\cap\{v'\})$. Firstly, fix $\bar{V}$ and suppose $|v'| \leq \bar{V}$. Then define,

$$\begin{aligned}
c(a_0, a, v') = \sup\Bigg\{\zeta(u', u'', v, v', v'', a)\Bigg|\, & u', u'' \in \mathbb{T}^d, v, v'' \in \mathbb{R}^d, |v' - v''| \geq \varepsilon(a_0)\\
&\text{and } |v|, |v''| \leq V(a_0) \text{ and } \left|\frac{v - v'}{|v - v'|}\cdot\frac{v'' - v'}{|v'' - v'|}\right| \leq 1 - \varepsilon(a_0)\Bigg\},
\end{aligned}$$

where we define,

$$\zeta(u', u'', v, v', v'', a) := \mathcal{H}^d\left(\left\{u \in \mathbb{T}^d \,\middle|\, \inf_{s\in[0,t]} |u - u' + s(v - v')|_{\mathbb{T}^d} \leq a \text{ and } \inf_{s\in[0,t]} |u - u'' + s(v - v'')|_{\mathbb{T}^d} \leq a\right\}\right).$$

Thus $c(a_0, a, v')$ is the maximum volume contained in the intersection of of two cylinders of diameter $a$ and centre lines $v - v'$ and $v - v''$ with the geometric constraints stated above. As stated we would like to bound the size of $\mathcal{H}^d((C_{\bar{l}} \cap C_{l'}) \cap (\mathbb{T}^d \cap \{v'\})$ and we do this in the following steps. Firstly bound the number of intersections two cylinders can have on $\mathbb{T}^d$ and then bound the volume of the intersection of any two cylinders. To tackle the first step recall that we assumed that $|v'| \leq \overline{V}$ and that $|v|, |v''| \leq V(a_0)$. It then follows that the longest possible length of a cylinder is bounded above by $(V(a_0) + \overline{V})t$. This means that the maximum number intersections of 2 cylinders on $\mathbb{T}^d$ is bounded above by $((V(a_0) + \overline{V})t + 1)^2$. To get an idea of why this bound is true consider a tube of length $(V(a_0) + \overline{V})t$, as in the figure below tilt the tube as little as possible in order for the tube to never self intersect on the torus.

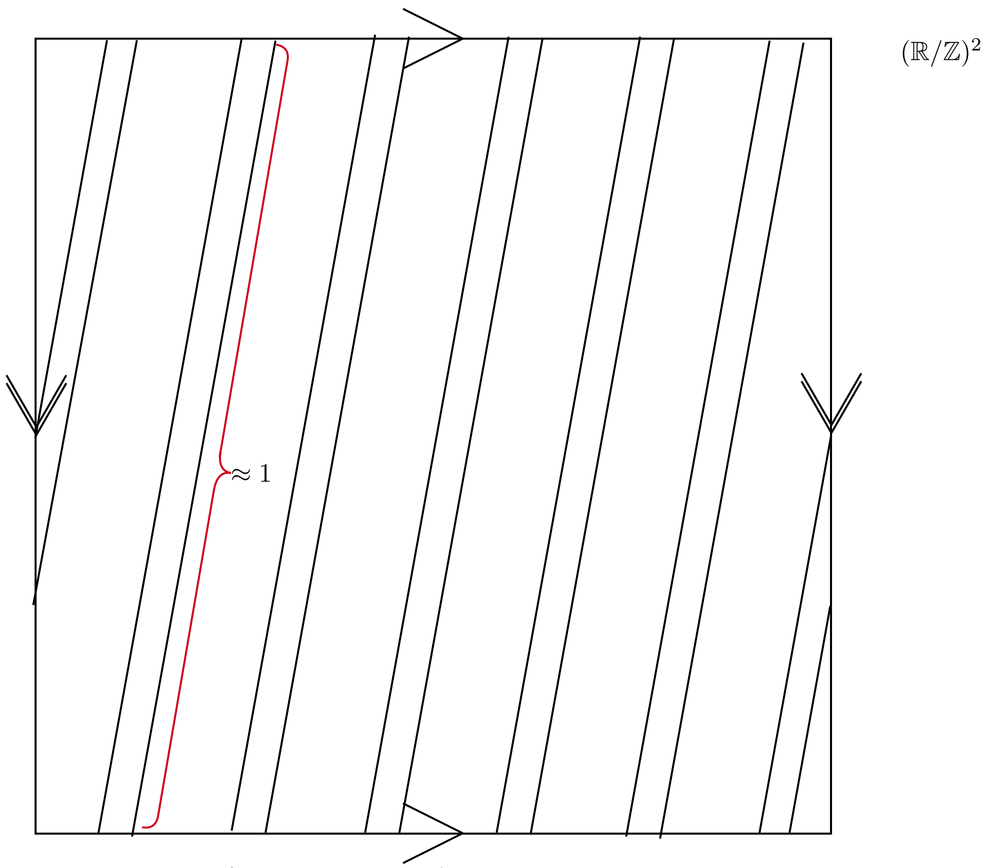



Fig: An exaggeration of a slightly tilted tube.

Looking at the tube on the torus we will see it is now split up into a number of segments (a segment being a part of the original tube that wraps around the torus) each of which have length close to 1 (provided the tilt angle is small) but also each segment is strictly greater in length than 1. Thus one might hazard a guess the most number of segments one can is simply given by the maximum number of segments is given by $(V(a_0) + \overline{V})t$. However, this is wrong as if we look at the figure above there is one segment at the far left and far right of the torus that is actually split into two segment each of which are not necessarily close in length to 1, but this can only happen once. Thus the maximum number of segments is given by $(V(a_0) + \overline{V})t + 1$. Finally, consider a tube of length roughly 1 that is perpendicular to the tubes in the above figure but that is also not long enough to wrap around the torus. Then this tube can intersect the tube in the figure a maximum of $(V(a_0) + \overline{V})t + 1$

times. Now consider another tube of length $(V(a_0) + \overline{V})t$ that is perpendicular to the above tube in the figure then by the same reasoning each segment of the new tube can intersect the tube in the figure a maximum of $(V(a_0)+\overline{V})t+1$ times, but there are $(V(a_0)+\overline{V})t+1$ such segments and so the maximum number of intersections between any two tubes is $((V(a_0) + \overline{V})t + 1)^2$.

Now we want to bound the volume of the intersection of two cylinders. When two cylinders intersect you get an angled Steinmetz. When you look down on such a shape (i.e. project it on the plane) you get a rhombus. We wish to bound the intersection by above and we do the by fitting the Steinmetz into a tube.

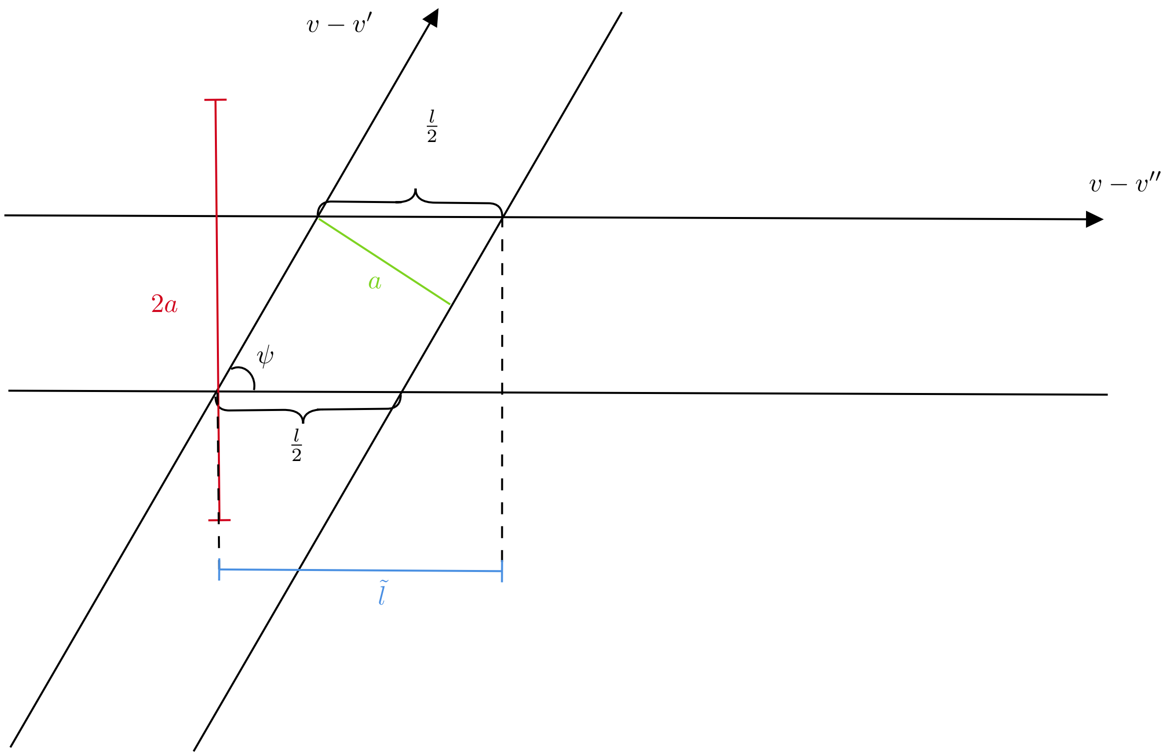



Fig: The intersection of two tubes viewed from above.

Now place a $(d-1)$-dimensional circle or diameter $2a$ whose face is facing the same way as the corresponding tube and stretch it by a length of $\tilde{l}$ which is bounded above by $l$. Thus the intersection of two tubes can fit inside a tube, $T$ of length $l$ and radius $a$. Thus it follows that $\mathcal{H}^d(T) = (2a)^{d-1}l$ (note Hausdorff measure of a sphere is equal to Lebesgue measure of a cube of length equal to diameter of the sphere). Thus the volume of the intersection of any two tubes is bounded above by $(2a)^{d-1}l$. Note further that $l = \frac{2a}{\sin(\psi)}$.

Next we introduce a new angle $\psi_0$ (see figure below) and use our conditions. By the sine rule we have that,

$$\sin(\psi) = \frac{|v' - v''|}{|v - v'|}\sin(\psi_0).$$

Then recall our assumptions that,

$$\left| \frac{v - v'}{|v - v'|} \cdot \frac{v'' - v'}{|v'' - v'|} \right| \leq 1 - \varepsilon(a_0),$$

which is equivalent to,

$$\cos(\psi_0) \leq 1 - \varepsilon(a_0).$$

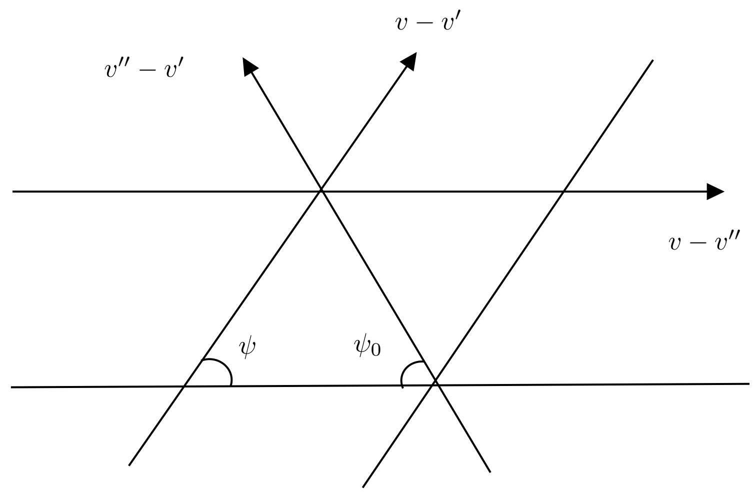



Fig: A figure for calculations

Furthermore, with the assumption that $|v' - v''| \geq \varepsilon$ we have from the following short calculation that,

$$\begin{aligned}
|\sin(\psi)| &= \frac{|v' - v''|}{|v - v'|}\sin(\psi_0) \\
&\geq \frac{\varepsilon}{|v - v'|}\sin(\psi_0) \\
&\geq \frac{\varepsilon}{|v - v'|}(1 - (1 - \varepsilon(a_0))^2)^{\frac{1}{2}} \\
&= \frac{\varepsilon}{|v - v'|}(2\varepsilon - \varepsilon^2)^{\frac{1}{2}}
\end{aligned}$$

and since $\sqrt{2x - x^2} \geq \sqrt{x}$ on [0,1], we have,

$$\geq \frac{\varepsilon^{\frac{3}{2}}}{|v - v'|}.$$

We finally now can see that,

$$\begin{aligned}
\mathcal{H}^d((C_{\bar{l}} \cap C_{l'}) \cap (\mathbb{T}^d \cap \{v'\}) &\leq (\text{max number of intersections}) \times (\text{max size of intersections}) \\
&\leq (2a)^{d-1} l((V(a_0) + \overline{V})t + 1)^2 \\
&= \frac{2^d a^{d-1} a}{\sin(\psi)}((V(a_0) + \overline{V})t + 1)^2
\end{aligned}$$

and since $|v - v'| \leq \overline{V} + V(a_0)$

$$\leq 2^d a^{d-1} a \varepsilon(a_0)^{-\frac{3}{2}}(V(a_0) + \overline{V})((V(a_0) + \overline{V})t + 1)^2$$

Then recall that,

$$\begin{aligned}
e_1 &= n \int_{\mathbb{R}^d} \mathcal{H}^d((\mathcal{C}_j(\phi) \cap \bar{\mathcal{C}}_j(\phi)) \cap (\mathbb{T}^d \times \{v'\}) \, \mathrm{d}f_0(v') \\
&= n \int_{\mathbb{R}^d \setminus B_{\overline{V}}(0)} \mathcal{H}^d((\mathcal{C}_j(\phi) \cap \bar{\mathcal{C}}_j(\phi)) \cap (\mathbb{T}^d \times \{v'\}) \, \mathrm{d}f_0(v') + n \int_{B_{\overline{V}}(0)} \mathcal{H}^d((\mathcal{C}_j(\phi) \cap \bar{\mathcal{C}}_j(\phi)) \cap (\mathbb{T}^d \times \{v'\}) \, \mathrm{d}f_0(v').
\end{aligned}$$

Observing that if $\#m \leq r$ then for any given set $(\mathcal{C}_j(\phi) \cap \bar{\mathcal{C}}_j(\phi))$ there are at most $\binom{r}{2} \leq r^2$ sets of the form $C_{\bar{l}} \cap C_{l'}$ (i.e. choose $l$ and $l'$). Then using the Boltzmann-Grad scaling limit (2) we get,

$$\leq n \int_{\mathbb{R}^d \setminus B_{\overline{V}}(0)} \mathcal{H}^d((\mathcal{C}_j(\phi) \cap \bar{\mathcal{C}}_j(\phi)) \cap (\mathbb{T}^d \times \{v'\}) \,\mathrm{d}f_0(v') + 2^d r^2 a\varepsilon(a_0)^{-\frac{3}{2}}(V(a_0) + \overline{V})((V(a_0) + \overline{V})t + 1)^2$$

To deal with the second term we see that,

$$\begin{aligned}
\mathcal{H}^d((\mathcal{C}_j(\phi) \cap \bar{\mathcal{C}}_j(\phi)) \cap (\mathbb{T}^d \times \{v'\})) &\leq \mathcal{H}^d((\mathcal{C}_j(\phi) \cap (\mathbb{T}^d \times \{v'\})) \\
&\leq \sum_{|l|=j} \mathcal{H}^d((C_l(\phi) \cap (\mathbb{T}^d \times \{v'\})) \\
&\leq \sum_{|l|=j} \frac{\kappa_d}{n} |v_l - v'| t,
\end{aligned}$$

where the last inequality comes from (47). Thus,

$$\begin{aligned}
n \int_{\mathbb{R}^d \setminus B_{\overline{V}}(0)} \mathcal{H}^d((\mathcal{C}_j(\phi) \cap \bar{\mathcal{C}}_j(\phi)) \cap (\mathbb{T}^d \times \{v'\}) \,\mathrm{d}f_0(v') &\leq r \int_{\mathbb{R}^d \setminus B_{\overline{V}}(0)} \kappa_d |v_l - v'| t \,\mathrm{d}f_0(v') \\
&\leq r \int_{\mathbb{R}^d \setminus B_{\overline{V}}(0)} \kappa_d (|v'| + V(a_0)) t \,\mathrm{d}f_0(v').
\end{aligned}$$

But from the very initial assumption (5) we have that $\int_{\mathbb{R}^d} (|v'| + V(a_0)) t \,\mathrm{d}f_0(v') < \infty$, and so for sufficiently large $\overline{V}$ we obtain that,

$$r\kappa_d t \int_{\mathbb{R}^d \setminus B_{\overline{V}}(0)} (|v'| + V(a_0)) \,\mathrm{d}f_0(v') < \eta.$$

Thus we finish this part of the proof by noting that,

$$e_1 \leq \eta + 2^d r^2 a\varepsilon(a_0)^{-\frac{3}{2}}(V(a_0) + \overline{V})((V(a_0) + \overline{V})t + 1)^2$$

which yields the result that $\lim_{a\to 0} e_1 = 0$ for any fixed $a_0$ and in turn we have that, $\lim_{a\to 0} K(a_0, a) = 0$ and so finally we can say, without losing much sleep that, $\lim_{a\to 0} I_1 = 0$. We have now proved (66) in the sense when $a_0$ is fixed and we bound the tree size by $r$ that is to say we have shown,

$$\lim_{a\to 0} \sup_{\substack{\Omega \in \mathcal{MT}, \mathrm{Borel}, \\ k \geq 1}} \left| \widehat{P}_{t,k}(\Omega \cap \widehat{\mathcal{G}}(a_0) \cap \bigcup_{\substack{m \in \mathcal{T} \\ \#m \leq r}} \mathcal{E}(m)) - P_{t,k}(\Omega \cap \widehat{\mathcal{G}}(a_0) \cap \bigcup_{\substack{m \in \mathcal{T} \\ \#m \leq r}} \mathcal{E}(m))) \right| = 0. \tag{73}$$

We finish the proof by showing that,

$$\lim_{\delta \to 0} \lim_{a_0 \to 0} \lim_{a \to 0} (I_2 + I_3).$$

From equation (68) applied to the definition of $I_3$ we have that for sufficiently large $r$,k

$$I_3 = \sup_k \sum_{\substack{m \in \mathcal{T} \\ \#m > r}} P_{t,k}(\widehat{\mathcal{G}}(a_0) \cap \mathcal{E}(m)) \leq \frac{1}{r-1} K_{\mathrm{ini}} e^{\kappa_d K_{\mathrm{ini}}} < \delta. \tag{74}$$

Then also,

$$\begin{aligned}
\lim_{a_0 \to 0} \lim_{a \to 0} I_2 &= \lim_{a_0 \to 0} \lim_{a \to 0} \sup_k \sum_{\substack{m \in \mathcal{T} \\ \#m > r}} \widehat{P}_{t,k}(\widehat{\mathcal{G}}(a_0) \cap \mathcal{E}(m)) \\
&\leq \lim_{a_0 \to 0} \lim_{a \to 0} \sup_k \widehat{P}_{t,k}(\widehat{\mathcal{G}}(a_0)) - \lim_{a_0 \to 0} \lim_{a \to 0} \inf_k \sum_{\substack{m \in \mathcal{T} \\ \#m \leq r}} \widehat{P}_{t,k}(\widehat{\mathcal{G}}(a_0) \cap \mathcal{E}(m))
\end{aligned}$$

But then by (73) we have that,

$$\lim_{a\to 0}\sup_k \widehat{P}_{t,k}(\widehat{\mathcal{G}}(a_0)) = \sup_k P_{t,k}(\widehat{\mathcal{G}}(a_0))$$

and by Lemma 2.47

$$\lim_{a_0\to 0}\sup_k P_{t,k}(\widehat{\mathcal{G}}(a_0)) = 1.$$

Again by equation (73) we have that,

$$\lim_{a\to 0}\inf_k \sum_{\substack{m\in\mathcal{T}\\ \#m\leq r}} \widehat{P}_{t,k}(\widehat{\mathcal{G}}(a_0)\cap\mathcal{E}(m)) = \inf_k \sum_{\substack{m\in\mathcal{T}\\ \#m\leq r}} P_{t,k}(\widehat{\mathcal{G}}(a_0)\cap\mathcal{E}(m)). \tag{75}$$

Now we are near to the end just quickly note from the above that,

$$\lim_{a_0\to 0}\lim_{a\to 0} I_2 = 1 - \lim_{a_0\to 0}\inf_k \sum_{\substack{m\in\mathcal{T}\\ \#m\leq r}} P_{t,k}(\widehat{\mathcal{G}}(a_0)\cap\mathcal{E}(m)) \leq \delta$$

where the last equality follows from (70). Thus putting everything together we have that $\lim_{a_0\to 0}\lim_{a\to 0}(I_1 + I_2 + I_3) = 0$ and hence believe it or not we have finished the proof and shown (66) to be true. ■

**Proposition 2.50** (Tightness)**.**
Let $\mathcal{G}(a_0)$ be defined as in Definition 2.36 and let $\Omega\subseteq\mathcal{G}(a_0)$ be a Borel set. Then,

$$\lim_{a_0\to 0}\sup_k P_{t,k}(\mathcal{G}(a_0)) = 1 \tag{76}$$

$$\lim_{a\to 0}\left(\sup\left\{|\widehat{P}_{t,k}(\Omega) - P_{t,k}(\Omega)| \,\middle|\, k>0,\ \Omega\subseteq\mathcal{G}(a_0)\right\}\right) = 0 \tag{77}$$

if $a_0$ is fixed. ▲

*Proof.* For (76) clearly $\widehat{\mathcal{G}}(a_0)\subseteq\mathcal{G}(a_0)$ for all $a_0$ and then Lemma 2.47 yields the result.
For (77), recall $P_{t,k}$ and $\widehat{P}_{t,k}$ are probability measures then with Lemma 2.48, one gets,

$$\lim_{a_0\to 0}\lim_{a\to 0}\sup_{k\geq 1}\left|\widehat{P}_{t,k}(\mathcal{MT}\setminus\widehat{\mathcal{G}}(a_0)) - P_{t,k}(\mathcal{MT}\setminus\widehat{\mathcal{G}}(a_0))\right| = 0. \tag{78}$$

Now let $\Omega\subseteq\widehat{\mathcal{G}}(a_0)$ for some $a_0>0$ and fix $\varepsilon>0$. Then,

$$\begin{aligned}
\lim_{a_0\to 0}\lim_{a\to 0}\sup_{\substack{\Omega\in\mathcal{MT},\text{Borel},\\ k\geq 1}}\left|\widehat{P}_{t,k}(\Omega) - P_{t,k}(\Omega)\right| &\leq \lim_{a_0\to 0}\lim_{a\to 0}\sup_{\substack{\Omega\in\mathcal{MT},\text{Borel},\\ k\geq 1}}\left|\widehat{P}_{t,k}(\Omega\cap\widehat{\mathcal{G}}(a_0)) - P_{t,k}(\Omega\cap\widehat{\mathcal{G}}(a_0))\right|\\
&\quad + \lim_{a_0\to 0}\lim_{a\to 0}\left(\sup_{\substack{\Omega\in\mathcal{MT},\text{Borel},\\ k\geq 1}}\widehat{P}_{t,k}(\Omega\setminus\widehat{\mathcal{G}}(a_0))\right.\\
&\quad \left. + \sup_{\substack{\Omega\in\mathcal{MT},\text{Borel},\\ k\geq 1}} P_{t,k}(\Omega\setminus\widehat{\mathcal{G}}(a_0))\right)
\end{aligned}$$

Which by Lemma 2.48,

$$= \lim_{a_0\to 0}\lim_{a\to 0}\left(\sup_{\substack{\Omega\in\mathcal{MT},\text{Borel},\\ k\geq 1}}\widehat{P}_{t,k}(\Omega\setminus\widehat{\mathcal{G}}(a_0)) + \sup_{\substack{\Omega\in\mathcal{MT},\text{Borel},\\ k\geq 1}} P_{t,k}(\Omega\setminus\widehat{\mathcal{G}}(a_0))\right)$$

But also recall that $P_{t,k}$ is the macroscopic description of the dynamics and so independent of $a$ which in turn implies the above expression is equal to

$$\lim_{a_0\to 0}\lim_{a\to 0}\sup_{\substack{\Omega\in\mathcal{MT},\text{Borel},\\ k\geq 1}}\widehat{P}_{t,k}(\Omega\setminus\widehat{\mathcal{G}}(a_0))+\lim_{a_0\to 0}\sup_{\substack{\Omega\in\mathcal{MT},\text{Borel},\\ k\geq 1}}P_{t,k}(\Omega\setminus\widehat{\mathcal{G}}(a_0))$$

Which in turn due to (78) we have,

$$\leq 2\lim_{a_0\to 0}\sup_{\substack{\Omega\in\mathcal{MT},\text{Borel},\\ k\geq 1}}P_{t,k}(\mathcal{MT}\setminus\widehat{\mathcal{G}}(a_0))=0.$$

where the final equality comes from Lemma 2.47. ■

### 2.3.7 Proof of the main Result

We now approach the end of this body of work, thankfully most of the hard work has been done and all that remains is to use the proven results to prove the main theorems. For ease of reading I have re-written the main results again before proving them.

**Theorem 2.51** (Main theorem c.f. Theorem2.6)**.** Suppose that $f_0\in PM(\mathbb{R}^d)$, then further suppose that, $f_0$ has finite total mass and Kinetic energy

$$\int_{\mathbb{R}^d}(1+|v|)^2\,\mathrm{d}f_0(v)=K_{\text{ini}}<\infty$$

and that $f_0$ does not concentrate any mass on lines,

$$\int_{\rho(v,\nu)}\mathrm{d}f_0\left(v'\right)=0\text{ for all }v\in\mathbb{R}^d,\nu\in S^{d-1},$$

where $\rho(v,\nu)=v+\nu\mathbb{R}^d$. Let $\omega\subset\mathbb{T}^d\times\mathbb{R}^d$ be a realisation of the *tagged Poisson point process* with intensity $\mu=n(\mathbb{1}_{\mathbb{T}^d}\otimes f_0)$ where $\mathbb{1}_{\mathbb{T}^d}$ is the Lebesgue measure on the unit torus. Define $N:=\#\omega$ and further suppose we have $N$ particles adhering to

$$\frac{\partial}{\partial t}u(i,t)=v(i,t),\qquad\frac{\partial}{\partial t}v(i,t)=0.$$

with initial positions and velocities given by $\omega$. Then for all $t\in[0,\infty)$,

$$\lim_{a\to 0}\sup_{A\subset\mathbb{T}^d\times\mathbb{R}^d\text{ Borel}}\left|\text{Prob}_{\text{tppp}}\left(z(1,t)\in A\text{ and }\beta^{(a)}(1,t)=1\right)-\int_A\mathrm{d}u\;\mathrm{d}f_t(v)\right|=0,$$

where $f:[0,\infty)\to PM\left(\mathbb{R}^d\right)$ is the unique solution of the homogeneous, gainless Boltzmann equation

$$\dot{f}=Q_-[f,f],\quad f_{t=0}=f_0$$

with $Q_-[f,f](v)=-\int_{\mathbb{R}^d}\kappa_d\,|v-v'|\,f(v)\;\mathrm{d}f\left(v'\right)$, and $\kappa_d$ the volume of $d-1$ dimensional unit-ball. ▲

*Proof.* Let $A\subseteq\mathbb{T}^d\times\mathbb{R}^d$ be a Borel set and define $\Omega(A)\subseteq\mathcal{MT}$ to be given by,

$$\Omega(A):=\{\Phi\in\mathcal{MT}|\beta_1(m)=1\text{ and }Z_1\in A\}$$

which is a Borel set in $\mathcal{MT}$ since $\beta_1(m)$ is a property of $m$ alone. Thus for all $a_0>0$ we have that,

$$\begin{aligned}\left|\lim_{a\to 0}\lim_{k\to\infty}\widehat{P}_{t,k}(\Omega)-\int_A\mathrm{d}u\,\mathrm{d}f_t(v)\right|&=\lim_{a\to 0}\lim_{k\to\infty}\left|\widehat{P}_{t,k}(\Omega)-\int_A\mathrm{d}u\,\mathrm{d}f_{t,k-1}(v)\right|\\&=\lim_{a\to 0}\lim_{k\to\infty}\left|\widehat{P}_{t,k}(\Omega)-P_{t,k}(\Omega)\right|\\&=\lim_{a\to 0}\lim_{k\to\infty}\left|\widehat{P}_{t,k}(\Omega\cap\mathcal{G}(a_0))-P_{t,k}(\Omega\cap\mathcal{G}(a_0))-\widehat{P}_{t,k}(\Omega\setminus\mathcal{G}(a_0))+P_{t,k}(\Omega\setminus\mathcal{G}(a_0))\right|\end{aligned}$$

Where the first equality comes from Lemma 2.11 and the second equality comes from Proposition 2.28. Next we will use the Tightness proposition, Proposition 2.50, in particular line (77) to conclude that,

$$\leq \lim_{a\to 0}\lim_{k\to\infty}\left[P_{t,k}(\mathcal{MT}\setminus\mathcal{G}(a_0)) + \widehat{P}_{t,k}(\mathcal{MT}\setminus\mathcal{G}(a_0))\right]$$

Next since $P_{t,k}$ and $\widehat{P}_{t,k}$ are probability measures and using (77) again from the Tightness proposition and noting that the supremum is over $k$ we have that since,

$$\lim_{a\to 0}\widehat{P}_{t,k}(\mathcal{MT}\setminus\mathcal{G}(a_0)) = 1 - \lim_{a\to 0}\widehat{P}_{t,k}(\mathcal{G}(a_0)) = 1 - P_{t,k}(\mathcal{G}(a_0)) = P_{t,k}(\mathcal{MT}\setminus\mathcal{G}(a_0)),$$

that our initial limit is bounded above by

$$\left|\lim_{a\to 0}\lim_{k\to\infty}\widehat{P}_{t,k}(\Omega) - \int_A \mathrm{d}u\,\mathrm{d}f_t(v)\right| \leq 2\lim_{k\to\infty}P_{t,k}(\mathcal{MT}\setminus\mathcal{G}(a_0))$$

for all $a_0 > 0$. Finally by the Tightness proposition and line (76) we establish that,

$$\lim_{a\to 0}\lim_{k\to\infty}\widehat{P}_{t,k}(\Omega) = \int_A \mathrm{d}u\,\mathrm{d}f_t(v).$$

This completes the proof. ■

Recall Corollary 2.8

**Corollary 2.52.** The measures $\mathbf{1}_{\mathbb{T}^d}\otimes f_t$

$$\mathrm{d}\hat{f}_t^{(a)}(u,v) = \mathrm{Prob}_{\mathrm{tppp}}\Big(z(1,t)\in[u,u+\mathrm{d}v)\times[v,v+\mathrm{d}v) \text{ and } \beta^{(a)}(1,t)=1\Big)$$

are both absolutely continuous with respect to $\mathbf{1}_{\mathbb{T}^d}\otimes f_0$. Furthermore,

$$\lim_{a\to 0}\hat{f}_t^{(a)} = \mathbf{1}_{\mathbb{T}^d}\otimes f_t \tag{79}$$

in the $L^1(\mathbf{1}_{\mathbb{T}^d}\otimes f_0)$ norm. ▲

*Proof.* Since $f_{t,k} = e^{-\int_0^t L[f_{s,k-1}]\,\mathrm{d}s}f_0$ clearly $f_{t,k} << f_0$. Then since $\lim_{k\to\infty}f_{t,k} = f_t$ implies that $f_t << f_0$.

To show $\mathrm{d}\hat{f}_{t,k}$ is absolutely continuous with respect to $\mathbf{1}_{\mathbb{T}^d}\otimes f_t$ follows from equation (40). This is because,

$$\widehat{P}_{t,k}(\Omega) = \int_{\mathcal{MT}_1}\int_{\mathcal{MT}_2}\cdots\int_{\mathcal{MT}_{k-1}}\widehat{P}_{t,k}(\Omega\,|\Phi')\,\mathrm{d}\widehat{P}t,k-1(\Phi_{k-1}\mid\Phi_{k-2})\ldots\mathrm{d}\widehat{P}t,2(\Phi_2\mid\Phi_1)\,\mathrm{d}\widehat{P}_{t,1}(\Phi_1)$$

where $\widehat{P}_{t,1}(\Phi_1) = \big(\mathbf{1}_{\mathbb{T}^d}\otimes f_t\big)(z_1)$ since $z_1 = \Phi_1$. So by a cheeky abuse of notation,

$$\mathrm{d}\hat{f}_{t,k} \leq \int_{[u,u+\mathrm{d}u)\times[v,v+\mathrm{d}v)}\mathrm{d}\widehat{P}_{t,1}(\Phi_1).$$

For my sake only by the above we mean to say, given any Borel set $A\subseteq\mathcal{MT}_1$ we have,

$$\int_A \mathrm{d}\hat{f}_{t,k}^{(a)}(u,v) \leq \int_A \mathrm{d}\widehat{P}_{t,1}(\Phi_1). \tag{80}$$

and so the measure on the left hand integral is clearly absolutely continuous with respect to $\mathbf{1}_{\mathbb{T}^d}\otimes f_t$. Finally, the fact that

$$\lim_{a\to 0}\hat{f}_t^{(a)} = \mathbf{1}_{\mathbb{T}^d}\otimes f_t \tag{81}$$

in the $L^1(\mathbf{1}_{\mathbb{T}^d}\otimes f_0)$ norm follows from the calculations made in the proof of the main theorem. ■

We are now done with the theory.

## 2.4 The Effect of Concentrations

Equation (9) is the assumption that says we can't have concentrations on lines and that convergence fails in this case. In this chapter we will assume $f_0$ has strong concentrations and hence shown that convergence fails. To make things a little less painful we will work in $d = 2$.

**Theorem 2.53.** Let $v \in \mathbb{R}^2$ be such that $|v| = \dfrac{1}{2}$ and set $f_0 := \frac{1}{2}(\delta_{-v} + \delta v)$. If,

$$\widehat{Q}(t) = \lim_{a\to 0}\lim_{k\to\infty} \widehat{P}_{t,k}(\beta_1 = 1),$$

denotes the empirical probability that a tagged particle does not collide, then,

$$\lim_{t\to 0}\frac{1}{t^3}\left(\widehat{Q}(t) - \int_{\mathbb{R}^2} \mathrm{d}f_t(v)\right) = \frac{1}{9}, \tag{82}$$

where $f_t = \dfrac{\kappa_2}{1+t} f_0$ is the unique solution of the Boltzmann equation (11) which satisfies the initial condition $f_{t=0} = f_0$. ▲

**Remark 2.54.** It is easy to see that $f_t$ is indeed a solution. Clearly,

$$\frac{\partial f_t}{\partial t} = \frac{-\kappa_2 f_0}{(1+t)^2}.$$

On the other hand notice that,

$$\begin{aligned}
Q[f_t, f_t](v) &= \frac{-\kappa_2 f_0}{(1+t)^2}\int_{\mathbb{R}^2} |v - v'|\,\mathrm{d}\left(\frac{1}{2}(\delta_{-v} + \delta v)\right)\\
&= \frac{-\kappa_2 f_0}{(1+t)^2}\left[\frac{|v-v|}{2} + \frac{|v+v|}{2}\right]\\
&= \frac{-\kappa_2 f_0}{(1+t)^2}(v).
\end{aligned}$$

similarly, $\dfrac{\partial f_t}{\partial t}(-v) = Q[f_t, f_t](-v)$. So indeed $f_t$ is a solution. ▲

Next without loss of generality assume the tagged particle has initial data $(0, v)$. Define,

$$M_\lambda = \left\{u \in \mathbb{T}^2 \;\middle|\; \min_{s\in[0,t]} |2sv - u| \le \rho\right\}.$$

Then $M_\lambda$ looks like,

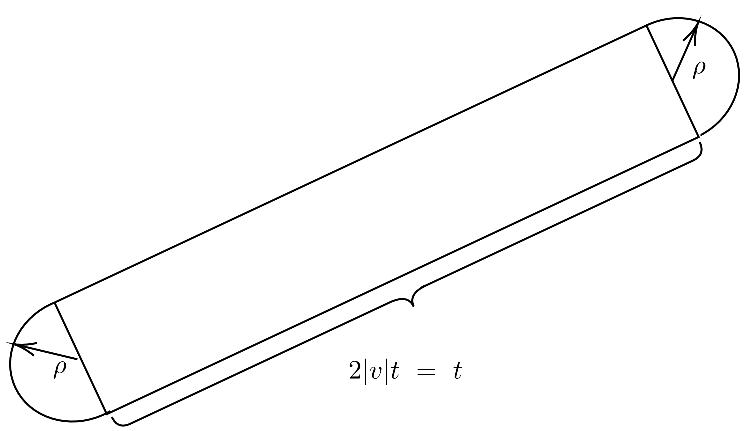



Fig: Illustration of $M_\lambda$

which are the initial positions of particles such that if they have velocity $-v$ they will collide $(0, v)$ at some point in time in $[0, t]$ when the diameter of the particles is $\rho$. It is worth noting particle with velocity $v$ can never collide with the tagged particle $(0, v)$. Where by colliding we mean the particles touch but have at some point prior not been touching. This excludes the case when two particles have the same velocity and have initial positions very close to each other, which means they will be touching for all time but never actually collide. Thus,

$$Vol(M_\lambda) = \pi\rho^2 + 2\rho t, \tag{83}$$

and if we require $Vol(M_\lambda) = at\lambda$, then $\pi\rho^2 + 2\rho t - at\lambda = 0$.

**Lemma 2.55.** For all $j \in \mathbb{N}$ we have that,

$$\lim_{t\to 0} \frac{1}{t^j}\left|\widehat{Q}(t) - e^{-\lambda t}\sum_{k=0}^{j} \frac{(\lambda t)^k}{k!} p_k(\lambda)\right| = 0.$$

where $p_k(\lambda, t)$ is the probability that the tagged particle does not collide before time t if there exists precisely $k$ particles in $M_\lambda$. ▲

We will show that $p_k(\lambda, t)$ is independent of $t$ in the limit as $a \to 0$. Furthermore, recall that $\mu = n(\mathbb{1}_{\mathbb{T}^d} \otimes f_0)$ and that in any given volume there are a Poisson number of particles further more in $d = 2$ we have that $na = 1$ and so it follows that,

$$\text{Prob}(\#(\omega \cap M_\lambda) = k) = e^{-\mu(M_\lambda)}\frac{\mu(M_\lambda)^k}{k!} = e^{-nat\lambda}\frac{(nat\lambda)^k}{k!} = e^{-t\lambda}\frac{(t\lambda)^k}{k!},$$

and so the probability that there exists between 0 and $j$ particles in $M_\lambda$ and that in each case none of the particles collide with the tagged particle, is given by,

$$e^{-t\lambda}\sum_{k=0}^{j}\frac{(t\lambda)^k}{k!}p_k(\lambda, t).$$

Thus one could view $\widehat{Q}(t)$ as being given by,

$$\widehat{Q}(t) = \lim_{a\to 0} e^{-t\lambda}\sum_{k=0}^{\infty}\frac{(t\lambda)^k}{k!}p_k(\lambda, t) = e^{-t\lambda}\sum_{k=0}^{\infty}\frac{(t\lambda)^k}{k!}p_k(\lambda) \tag{84}$$

We now pursue the proof of Lemma 2.55.

*Proof.*

$$\begin{aligned}
\left|\widehat{Q}(t) - e^{-t\lambda}\sum_{k=0}^{j}\frac{(t\lambda)^k}{k!}p_k(\lambda, t)\right| &= \left|\widehat{Q}(t) - e^{-t\lambda}\sum_{k=j+1}^{\infty}\frac{(t\lambda)^k}{k!}p_k(\lambda, t)\right| \\
&= \text{Prob(At least } j+1 \text{ particles and in each case the tagged particle is unscattered)} \\
&\leq \text{Prob}(\#(\omega \cap M_\lambda) > j) \\
&= e^{-\lambda t}\left(e^{\lambda t} - \sum_{k=0}^{j}\frac{(t\lambda)^k}{k!}\right)
\end{aligned}$$

Then recall Taylor's theorem which states $e^{\lambda t} = \sum_{k=0}^{j}\frac{(t\lambda)^k}{k!} + R_k(x)$ where $|R_k(x)| \leq C\frac{t^{j+1}}{(j+1)!}$ and $\frac{d^{j+1}}{dt^{j+1}}e^{\lambda t} \leq C$. However,

$$\frac{d^{j+1}}{dt^{j+1}}e^{\lambda t} = \lambda^{j+1}e^{\lambda t} \leq \sup_{s\in[0,t]}\frac{\lambda^{j+1}}{(j+1)!}e^{\lambda s} = \frac{\lambda^{j+1}}{(j+1)!}e^{\lambda t} = O(t^{j+1}).$$

Consequently,

$$\lim_{t\to 0}\frac{1}{t^j}\left|\widehat{Q}(t)-e^{-t\lambda}\sum_{k=j+1}^{\infty}\frac{(t\lambda)^k}{k!}p_k(\lambda,t)\right|=0$$

■

We now study the idealised behaviour and try to express $f_t$ in a different way that will be more useful. Firstly define $Q(t)=\frac{1}{1+t}=\sum_{k=0}^{\infty}(-t)^k$. However, we'd somehow like to insert the idealised probabilities, $p_k^{(id)}(\lambda)$ so that we can write,

$$Q(t)=\sum_{k=0}^{\infty}e^{-\lambda t}\frac{(\lambda t)^k}{k!}p_k^{(id)}(\lambda). \tag{85}$$

Expanding the exponential term by its power series in (81) one gets,

$$\sum_{l,m=0}^{\infty}\frac{(-\lambda t)^l}{l!}\frac{(\lambda t)^m}{m!}p_m^{(id)}(\lambda)=\sum_{k=0}^{\infty}(-t)^k$$

equating coefficients yields that

$$\sum_{l=0}^{k}\frac{(-1)^l}{l!(k-l)!}p_{k-l}^{(id)}(\lambda)=\Big(-\frac{1}{\lambda}\Big)^k,$$

which recursively yields,

$$p_k^{id}(\lambda)=(-1)^k\frac{k!}{\lambda^k}-\sum_{l=1}^{k}(-1)^l\binom{k}{l}p_{k-l}^{(id)}(\lambda)=\sum_{l=1}^{k}\frac{k!}{(k-l)!}\Big(-\frac{1}{\lambda}\Big)^l \tag{86}$$

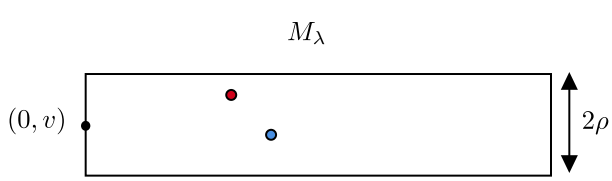
Thus if we can find a $k$ such that $p_k\neq p_k^{(id)}$ then we can easily show by Lemma 2.55 that $\widehat{Q}(t)\neq Q(t)$. It turns out that $k=3$ is sufficient.

The only thing that we have left to do is explicitly compute $p_k$ which can be done explicitly when $\lambda\geq 4$. We will show that,

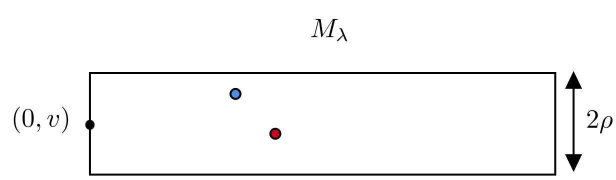
$$p_0=1,\quad p_1=1-\frac{1}{\lambda},\quad p_2=1-\frac{2}{\lambda}+\frac{2}{\lambda^2},\quad p_3=1-\frac{3}{\lambda}+\frac{6}{\lambda^2}-\frac{6}{\lambda^3}+\frac{\alpha}{\lambda^3},$$

where $\alpha=\frac{2}{3}$ which in turn implies, Theorem 2.53. For ease of computation the tagged particle is an interval of length $a$ perpendicular to $v$. There will be an error depending on $a$ which is not present for $p_1$ and $p_2$.

Let $k\in\{0,1,2,3\}$ be the number of particles in $M_\lambda$. For simplicity particles with velocity $v$ are *red* and particles with velocity $-v$ are *blue*. One obtains $2^k$ different colour combinations, each of those cases have the same probability of occurring. Note also that all particles are intervals of length $a$ and in this case $M_\lambda$ is a tube with flat ends and $p_{\cdots}^{\cdots}$ has subscript representing the number of particles in $M_\lambda$ and the superscript represents the order of particles appearing, see figure below for more clarity on this last point.




Corresponds to $p_2^{rb}$




Corresponds to $p_2^{br}$

Computation of $p_0$

$p_0 = 1$ is clear no obstacle in $M_\lambda$

Computation of $p_1$

$p_1^r = 1$ since $(0, v)$ can't collide with the red particle as they are travelling with same velocity. We move to the other case, that is $p_1^b$.

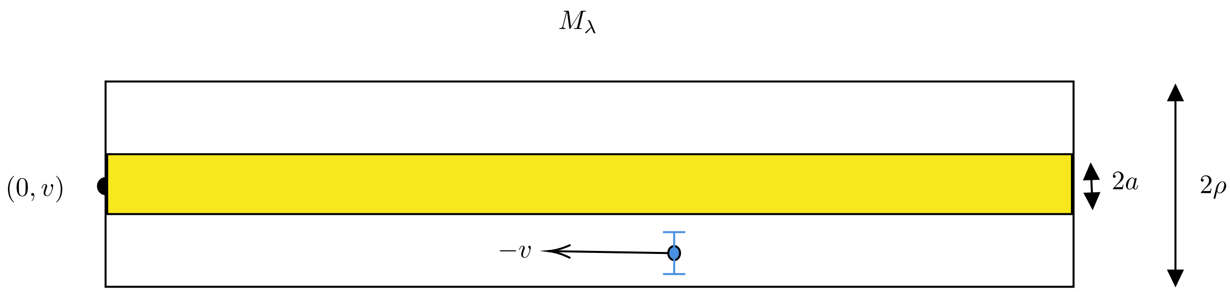



Fig: The case $p_1^b$

What we show in this figure is if the initial position of in *not* in the yellow zone the the event described by $p_1^b$ will occur, as when the two particles inevitably pass each other they will be a distance greater than $a$ apart. Next the number of particles of a particles generated is a Poisson amount and since we are conditioning on there being one particle the position of the blue particle is a uniform distribution, thus we have that,

$$p_1^b = \frac{2\rho t - 2at}{2\rho t} = 1 - \frac{a}{\rho},$$

and so by the equation $\pi\rho^2 + 2\rho t - at\lambda = 0$. We can conclude that $p_1^b = 1 - \frac{2}{\lambda}$. For the sake of clarity I'll spell out the final step in only this case.

$$\begin{aligned}
p_1 &= \text{Prob(tagged particle alive at time } t \mid 1 \text{ other particle exists in } M_\lambda) \\
&= \frac{1}{2}\text{Prob(tagged particle alive at time } t \mid 1 \text{ other particle exists in } M_\lambda \text{ and it is red)} \\
&+ \frac{1}{2}\text{Prob(tagged particle alive at time } t \mid 1 \text{ other particle exists in } M_\lambda \text{ and it is blue)} \\
&= \frac{1}{2}(p_1^b + p_1^r)
\end{aligned}$$

Thus we have $p_1 = 1 - \frac{1}{\lambda}$.

Computation of $p_2$

It is clear that $p_2^{rr} = 1$ as no collision can occur almost surely. Additionally, $p_2^{bb} = (1 - \frac{2}{\lambda})^2$, this is because each blue particle is independently distributed and they are almost surely never going to collide and hence for each particle on its own has a probability of $p_1^b = 1 - \frac{2}{\lambda}$ of hitting the red particle and by independence we get our claimed result. In the case of $p_2^{br}$ we have the situation where the blue particle can hit the tagged particle but the red one can't. Moreover, the the red and blue particle can never collide as the red particle is *after* the blue particle. Thus $p_1^b = p_2^{br}$.

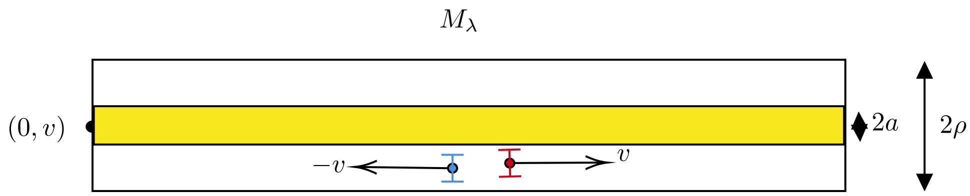



Fig: The case $p_2^{br}$

The only case that requires some thinking is the case $p_2^{rb}$. It needs to be split into two further cases. The first case being when the red particle misses the blue particle which then might hit the tagged particle. By independence it is quite simple to see that the probability of the event described by this case is given to be $(1 - \frac{2}{\lambda})^2$. The second case is that the red and blue particle collide and hence leave the root particle unscattered. If one imagines that the red particle is just for this instance the root particle the the chance of the red and blue colliding is given by $1 - p_1^b$

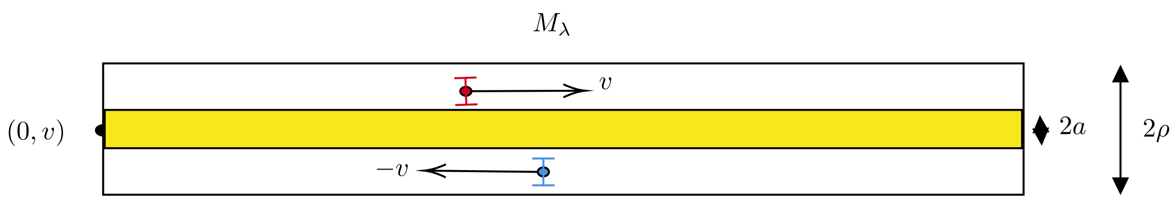



Fig: The case $p_2^{rb}$ case 1

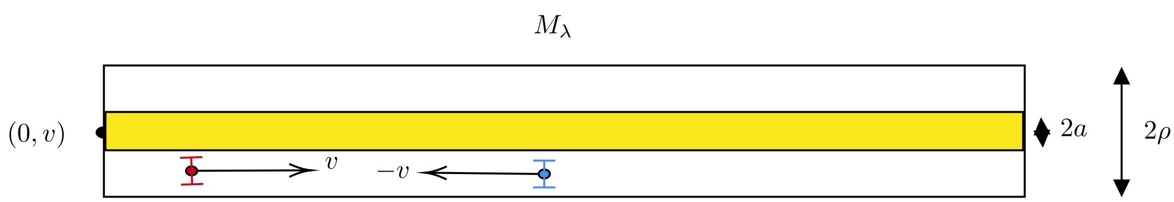



Fig: The case $p_2^{rb}$ case 2

Clearly we have,

$$p_2 = \frac{1}{4}(p_2^{rr} + p_2^{bb} + p_2^{br} + p_2^{rb}) = 1 - \frac{2}{\lambda} + \frac{2}{(\lambda)^2}$$

<u>Computation of $p_3$</u>

Just like before we have $p_3^{rrr} = 1$ and that $p_3^{bbb} = (1 - \frac{2}{\lambda})^3$. In an identical trail of thought to the case of $p_2^{br}$ we have that $p_3^{brr} = (1 - \frac{2}{\lambda})$. Using the same process as in the ace of $p_2^{br}$ one see that $p_3^{rbr} = p_2^{rb}$. The case $p_3^{bbr}$ is simply the same as $p_2^{bb}$ and so $p_3^{bbr} = (1 - \frac{2}{\lambda})^2$. The case $p_3^{rrb}$ has 3 sub-cases. The first sub-case is blue particle misses the second red but hits the first red. This event has probability given by $\frac{2}{\lambda}(1 - \frac{2}{\lambda})$. The second sub-case is the event where the blue particle hits the last red particle, which has probability $\frac{2}{\lambda}$. The last case is the event that the blue particle misses the first and last red particle in addition to missing the tagged particle. By Independence this last case has probability, $(1 - \frac{2}{\lambda})^3$. $p_3^{rrb}$ is simply the sum of the above 3 sub-cases, thus, $p_3^{rrb} = 1 - \frac{2}{\lambda}(1 - \frac{2}{\lambda})^2$.

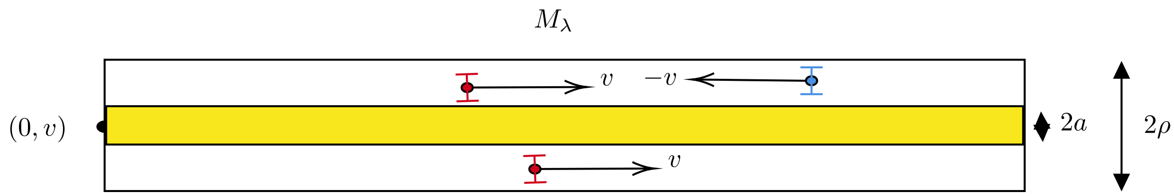



Fig: The case $p_3^{rrb}$ sub-case 1

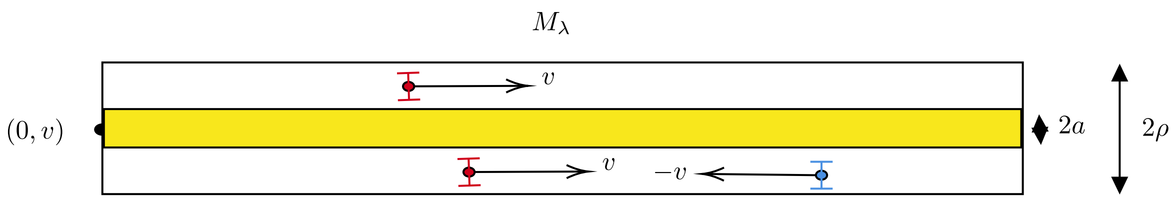



Fig: The case $p_3^{rrb}$ sub-case 2

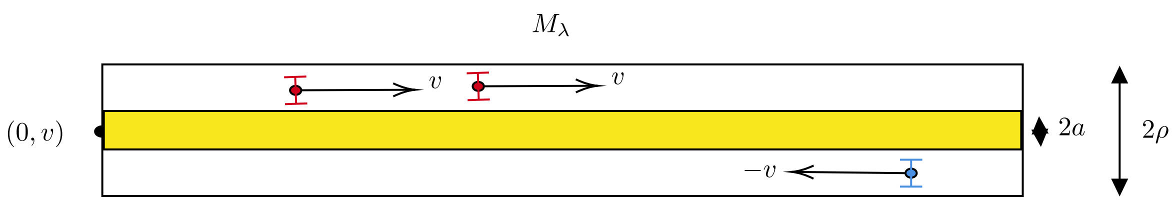



Fig: The case $p_3^{rrb}$ sub-case 3

The case $p_3^{brb}$ has two sub-cases. The first one being that the first blue particle misses the tagged particle and the second blue particle hits the red particle. By looking at the figure below it should be easy by now to see that this second case has probability $(1 - \frac{2}{\lambda})(\frac{2}{\lambda})$. The second sub-case is the event that both blue particles miss the tagged particle and the second blue particle misses the red particle. This has probability $(1 - \frac{2}{\lambda})^3$. This results us in realising that $p_3^{brb} = (1 - \frac{2}{\lambda})((1 - \frac{2}{\lambda}(1 - \frac{2}{\lambda}))$

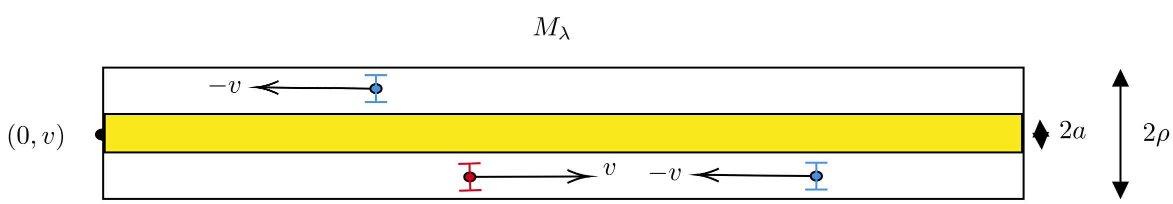



Fig: The case $p_3^{brb}$ sub-case 1

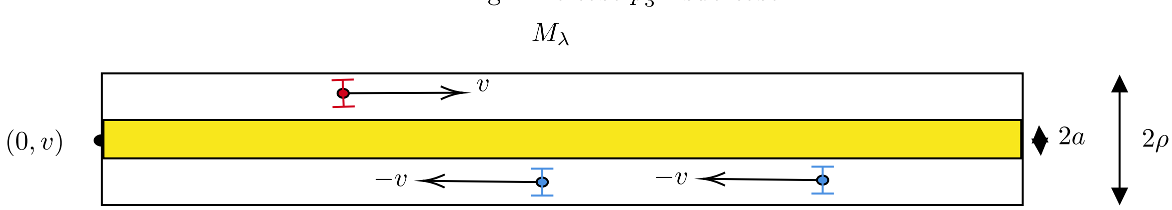



Fig: The case $p_3^{brb}$ sub-case 2

We now proceed with the final case. This case is a bit complicated, well at least it was for me. So it helps to this case properly without the aid of pictures and just do a brute force calculation. The case $p_3^{rbb}$ has three sub-cases of which the third sub-case has two sub-sub-cases. We introduce 3 coordinate parameters that are perpendicular to $v$, call them, $u_1, u_2, u_3 \in \mathbb{R}$ and consider in the following cases that $a = 1$, just to make our lives a little less miserable. The first case is that both blue particles are a distance bigger than 1 away from the tagged particle, in this case the red particle can be ignored. It is simple to see that in this case,

$$\text{Prob}(|u_2| > 1,\ |u_2| > 1) = \left(1 - \frac{2}{\lambda}\right)^2.$$

The second sub-case is where the first blue particle hits the red particle and hence doesn't collide with the tagged particle and the second blue particle misses the tagged particle. That is to say,

$$\begin{aligned}\text{Prob}(|u_2| \le 1,\ |u_2 - u_1| \le 1,\ |u_3| > 1) &= \text{Prob}(|u_2| \le 1,\ |u_2 - u_1| \le 1)\text{Prob}(|u_3| > 1)\\ &= \frac{4}{\lambda^2}\left(1 - \frac{2}{\lambda}\right)\end{aligned}$$

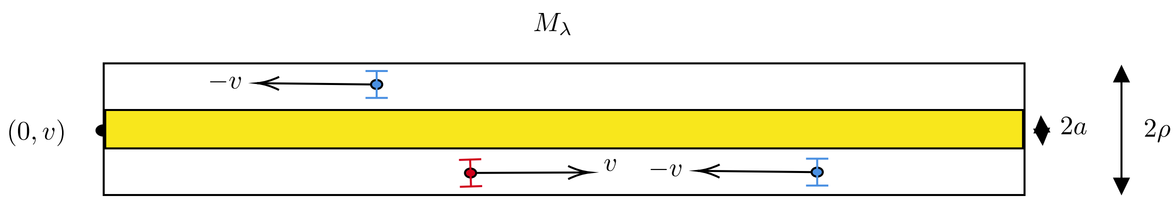



Fig: The case $p_3^{rbb}$ sub-case 1

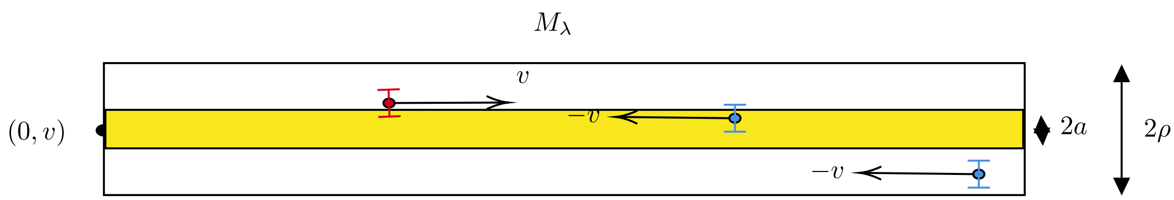



Fig: The case $p_3^{rbb}$ sub-case 2

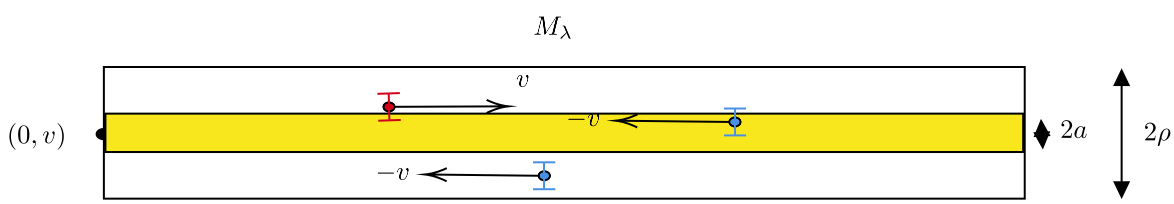



Fig: The case $p_3^{rbb}$ sub-case 3

The third and final sub-case being the one where the first blue particle misses both the tagged particle and the incoming red one and the second blue particle collides with the red particle. That is to say we wish to compute the following probability,

$$\text{Prob}(|u_3| \le 1,\ |u_2| > 1,\ |u_3 - u_1| \le 1,\ |u_2 - u_1| > 1)$$

Now consider the first sub-sub-case, we wish to compute,

$$\text{Prob}(|u_3| \le 3,\ |u_2| > 1,\ |u_3 - u_1| \le 1),$$

but not in this case the given event automatically implies that $|u_2 - u_1| > 1$. It is not too hard to see that,

$$\text{Prob}(|u_3| \le 1,\ |u_2| > 3,\ |u_3 - u_1| \le 1) = \left(1 - \frac{6}{\lambda}\right)\left(\frac{2}{\lambda^2}\right).$$

Finally, we proceed with the final sub-sub-case of the final sub-case of the final case. We wish to compute the following probability,

$$\text{Prob}(|u_3| \leq 1,\ |u_2| \in [1,3),\ |u_3 - u_1| \leq 1,\ |u_1 - u_2| > 1) = \frac{I_2}{\lambda^3}.$$

For this computation we need to explicitly compute the integral $I_2$.

$$\begin{aligned}
I_2 &= \int_{-1}^{1}\int_{u_3-1}^{u_3+1}\int_{1}^{3}(1 - \mathbb{1}_{[-1,+1]}(u_1 - u_2))\,\mathrm{d}u_2\,\mathrm{d}u_1\,\mathrm{d}u_3 + \int_{-1}^{1}\int_{u_3-1}^{u_3+1}\int_{-3}^{-1}(1 - \mathbb{1}_{[-1,+1]}(u_1 - u_2))\,\mathrm{d}u_2\,\mathrm{d}u_1\,\mathrm{d}u_3 \\
&= 2\int_{-1}^{1}\int_{u_3-1}^{u_3+1}\int_{1}^{3}(1 - \mathbb{1}_{[-1,+1]}(u_1 - u_2))\,\mathrm{d}u_2\,\mathrm{d}u_1\,\mathrm{d}u_3,
\end{aligned}$$

where we use the transformation $(u_1.u_2, u_3) \mapsto (-u_1, -u_2, u_3)$.

$$= 2\int_{-1}^{1}\int_{0}^{u_3+1}\int_{1}^{3}(1 - \mathbb{1}_{[-1,+1]}(u_1 - u_2))\,\mathrm{d}u_2\,\mathrm{d}u_1\,\mathrm{d}u_3 + 2\int_{-1}^{1}\int_{u_3-1}^{0}\int_{1}^{3}(1 - \mathbb{1}_{[-1,+1]}(u_1 - u_2))\,\mathrm{d}u_2\,\mathrm{d}u_1\,\mathrm{d}u_3,$$

where the second integrand is equal to 1 on the whole domain. Yielding,

$$\begin{aligned}
&= 4\int_{-1}^{1}(1 - u_3)\,\mathrm{d}u_3 + 2\int_{-1}^{1}\int_{0}^{u_3+1}\int_{1}^{3}\mathrm{d}u_2\,\mathrm{d}u_1\,\mathrm{d}u_3 - 2\int_{-1}^{1}\int_{0}^{u_3+1}\int_{1}^{3}\mathbb{1}_{[-1,+1]}(u_1 - u_2)\,\mathrm{d}u_2\,\mathrm{d}u_1\,\mathrm{d}u_3 \\
&= 8 + 8 - 2\int_{-1}^{1}\int_{0}^{u_3+1}\int_{1}^{3}\mathbb{1}_{[-1,+1]}(u_1 - u_2)\,\mathrm{d}u_2\,\mathrm{d}u_1\,\mathrm{d}u_3 \\
&= 16 - 2\int_{-1}^{1}\int_{0}^{u_3+1}\int_{\max(1,u_1-1)}^{1+u_1}\mathrm{d}u_2\,\mathrm{d}u_1\,\mathrm{d}u_3.
\end{aligned}$$

But since $u_1 \in [0,2]$ and we need $|u_1 - u_2| \leq 1$ which is equivalent to $u_2 \in [u_1 - 1, u_1 + 1]$ but since $u_3 \in [1,3]$ and $u_1 - 1 \in [-1,1]$ we can conclude that $u_2 \in u_1 + 1$ giving us,

$$\begin{aligned}
&= 16 - 2\int_{-1}^{1}\int_{0}^{u_3+1}\int_{1}^{1+u_1}\mathrm{d}u_2\,\mathrm{d}u_1\,\mathrm{d}u_3 \\
&= \frac{40}{3}
\end{aligned}$$

Piecing everything together we have that,

$$p_3^{rbb} = \left(1 - \frac{2}{\lambda}\right)^2 + \frac{4}{\lambda^2}\left(1 - \frac{2}{\lambda}\right) + \frac{4}{\lambda^2}\left(1 - \frac{6}{\lambda}\right) + \frac{I_2}{\lambda^3},$$

which in turn implies that,

$$\begin{aligned}
p_3 &= \frac{1}{8}(p_3^{rrr} + p_3^{rrb} + p_3^{rbr} + p_3^{brr} + p_3^{rbb} + p_3^{brb} + p_3^{bbr} + p_3^{bbb}) \\
&= 1 - \frac{3}{\lambda} + \frac{6}{\lambda^2} + \frac{6}{\lambda^3} + \frac{\alpha}{\lambda^3}
\end{aligned}$$

where $\alpha = \frac{I_2 - 8}{8} = \frac{2}{3}$. Now recall (86) and notice that the coefficient of $\lambda^{-3}$ is equal to $-6$ and since $\alpha$ is non-zero we have that $p_3$ and $p_3^{(\mathrm{id})}$ disagree proving Theorem 2.53.

# References


[1] K.Matthies and F.Theil. *Validity and Failure of the Boltzmann approximation of Kinetic Annihilation* J. Nonlinear Sci., 20 (2010), pp. 1–46.

[2] K.Matthies and F.Theil. *A Semigroup Approach to the Justification of Kinetic Theory* SIAM J. MATH. ANAL. Vol. 44, No. 6, pp. 4345–4379

[3] Isabelle Gallagher, Laure Saint-Raymond and Benjamin Texier. *From Newton To Boltzmann: Hard Sphere and Short-Range Potentials*

[4] R.Durrett, *Probability: Theory and Example.* 3rd edition. Duxbury (2004).