A Kac system interacting with two heat reservoirs

Federico Bonetto[1,*], Michael Loss[1] and Matthew Powell[1]
[1] School of Mathematics, Georgia Institute of Technology,Atlanta, GA 30332, USA
[*] corresponding author: bonetto@math.gatech.edu

June 23, 2026

Abstract

We study a system formed by M particles moving in 3 dimension and interacting with 2 heat reservoirs with $N >> M$ particles each. The system and the reservoirs evolve and interact via random collision described by a Kac-type master equation. The initial state of the reservoirs is given by 2 Maxwellian distributions at temperature T_+ and T_-. We show that, for times much shorter than $\sqrt{N}/M$ the interaction with the reservoirs is well approximated by the interaction with 2 Maxwellian thermostats, that is, heat reservoirs with $N = \infty$. As a byproduct, if $T_+ = T_-$ we extend the results in [3] to particles in 3 dimension.

Acknoledgement

FB acknowledge partial support from U.S. National Science Foundation grant DMS-1907643. ML acknowledge partial support from U.S. National Science Foundation grant DMS-2154340.

1 Introduction

In [14] Mark Kac introduced a toy model, the "Kac Evolution", to study kinetic properties of gases. He considered a system of M particles moving in one dimension and interacting via random collisions. After an exponentially distributed time, two particles are randomly and uniformly selected and they suffer a collision; that is, their energy is randomly redistributed. The rate of collision λ_M is chosen in such a way that the average number of collisions a particle suffers in a unit time is independent of M.

Despite its simplicity, the model serves to explain a number of issues. We give a short and informal description before turning to the main point of the paper. The original purpose of Kac was to derive the spatially homogeneous Boltzmann equation. This was achieved through the notion of 'chaos', i.e., states for which position and velocity of the particles are asymptotically independent as the particle number tends to infinity and are, thus, characterized by their single particle marginal. He proved that this property was preserved under the Kac Evolution, a phenomenon known as 'propagation of chaos', and the evolution of the marginal is given by the non-linear Boltzmann-Kac equation. The three dimensional version of this model is discussed in [15] where the analogous results are proved. The main limitation of these papers is that the molecules are Maxwellian: the intensity of the collision process is independent of the velocities. This restriction was removed by Sznitman [18] (see also the seminal paper [17]).

The Kac Evolution is ergodic and hence it makes sense to quantify the rate of approach to equilibrium. One way to achieve this is through the gap of the generator, which Kac [14] conjectured to be bounded away from zero uniformly in the particle number. This was eventually proved in [13]. In [7] and [16] an exact formula for the gap is given. Results for the gap with models with velocity dependent intensity are given in [20], [5], and [6] (in the latter, Kac's conjecture for three dimensional hard spheres is proved).

The notion of a gap means that one views the generator of the evolution on a Hilbert space, but it is well known that probability distributions have L^2 norms that grow exponentially with the particle number. This is remedied by considering the entropy, which, since it is extensive, is a more natural notion of approach to equilibrium. One expectation is that the entropy decreases exponentially in time with a rate that is independent of the particle number. The results, so far, do not lend optimism to this expectation. The best results are those of Villani [20] and Einav [10] where it is shown that the entropy production in the worst case is inversely proportional to the number of particles.

Recall that the entropy production is the ratio of the entropy dissipation and the entropy and hence, strictly speaking, the aforementioned results do not make a statement about the time decay of the entropy of a generic state. If, however, one considers initial states where only a small part of the system is out of equilibrium, then one can show exponential decay for the entropy as shown in [2] (see also [1]).

A rather different interpretation of the setup in [2], which is the central issue for the current work, is that the part of the system initially in equilibrium can be seen as a reservoir and the remaining part as a system that is out of equilibrium. In the simplest case one considers a system of M particles interacting via a Kac Evolution with a reservoir of N particles. Given that the reservoir is initially in equilibrium one expects that for N large the evolution of this state can be approximated by an effective evolution of a system of M particles interacting with a thermostat, i.e., one replaces the reservoir by a process in which one generate a "virtual" particle randomly selected with the same distribution of the initial state of the reservoir and lets it collide with a particle in the system. Thus, the idea is that the distribution of the reservoir, for large N, stays essentially fixed along the evolution. This approximation was carried out in [3] for particles moving in one dimension uniformly in time, using the Gabetta-Toscani-Wennberg metric.

The present paper mainly addresses two issues. Of interest is an extension of the work in [3] to a system of M particles that is coupled to two or more reservoirs of N particles each that are initially in equilibrium at *different* temperatures. In Theorem 2.1 we show that up to a time of order $\sqrt{N}/M$ the evolution of the reservoir system can be approximated uniformly by the corresponding evolution of the thermostats. It is know (see [8, 11] or Corollary A.2 below) that as the time $t \to \infty$ the thermostat system approaches a steady state (which is not explicitly known) at a rate of order one. As a consequence for times $1 << t << \sqrt{N}/M$ the reservoir system stays, approximately, in a non-equilibrium steady state (NESS). One should emphasize that for times $t >> \sqrt{N}/M$ the reservoir system tends to an equilibrium (which in this model is given by the average of the initial condition over all rotations). It should be emphasized that this equilibrium state is not close in any sense to the NESS of the thermostat. The physical distinction between these two states, the equilibrium and the non-equilibrium steady state, is that the latter mediates a heat transfer from the thermostat at higher temperature to the one at lower temperature. This distinction is absent in the one thermostat case where there is no NESS and the thermostat system tends to an equilibrium given by a Maxwellian.

The second issue is how to extend the analysis in [3] to three dimensions. There is an obvious difficulty: any reasonable collision mechanism conserves both energy and momentum; the additional conservation law complicates the analysis. A somewhat deeper issue is the possibility of considering reservoirs that initially have a non zero average momentum. This would produce a shear in the system, but in this situation it is not even clear how to implement the relevant thermostat.

The paper is organized as follows. In Section 2 we present the models and state the results. In Section 3 we state and comment on the key functional inequality. In Section 4 we comment on our results and outline possible extension and ideas for future work. Section 5 contains the proof of Lemmas 3.1, 3.2 and 3.3 while in Section 6 we report the proof of Theorem 2.1. Finally the Appendices contain technical results useful in the proofs.

2 Model and Results

The basic ingredient to build the Kac model evolution is the operator that describes the effect of a collision between two particles. Assume that the particles are spherical and that the unit vector associated with the relative position of their centers is ω. If $\vec{v}_1$ and $\vec{v}_2$, with $\vec{v}_i = (\vec{v}_{i,1}, \vec{v}_{i,2}, \vec{v}_{i,3}) \in \mathbb{R}^3$, are the outgoing velocities just after the collision then the incoming velocities before the collision were

$$\vec{v}_1^*(\omega) = \vec{v}_1 - ((\vec{v}_1 - \vec{v}_2) \cdot \omega)\omega\,, \qquad \vec{v}_2^*(\omega) = \vec{v}_2 - ((\vec{v}_2 - \vec{v}_1) \cdot \omega)\omega\,,$$

where $(\vec{v} \cdot \vec{w})$ denotes the scalar product in $\mathbb{R}^3$. Assume now that before the collision the velocities of the two particles are distributed according to the probability distribution g on $\mathbb{R}^6$. Given ω, the probability of observing the two particles after the collision with velocities $\vec{v}_1$, and $\vec{v}_2$ is $g(\vec{v}_1^*(\omega), \vec{v}_2^*(\omega))$. If we further assume that the unit vector ω is chosen uniformly at random on the unit sphere $\mathbb{S}^2$, that is according to the normalized Haar measure $d\omega$, the effect of the collision on g is given by

$$R[g](\vec{v}_1, \vec{v}_2) = \int_{\mathbb{S}^2} g(\vec{v}_1^*(\omega), \vec{v}_2^*(\omega))d\omega\,. \tag{1}$$

With this collision in hand, we can now start to describe the evolution in the models we will consider. The superscript S will refer to the *system*, as opposed to the reservoirs or the thermostats, for the remainder of this paper. The system contains M particles characterized by their velocities $\underline{v} = (\vec{v}_1, \ldots, \vec{v}_M) \in \mathbb{R}^{3M}$. Since the collision times are randomly chosen, the position of the particles plays no role. Thus, the state of the system is described by a probability distribution $f(\underline{v})$ on $\mathbb{R}^{3M}$. Since the particles are assumed to be identical, we will only consider *permutation invariant* distributions, that is we will assume that, for every permutation π on $\{1, \ldots, N\}$ we have

$$f(\vec{v}_1, \ldots, \vec{v}_M) = f(\vec{v}_{\pi(1)}, \ldots, \vec{v}_{\pi(M)})\,. \tag{2}$$

The effect of a collision between particle i and j on the distribution f is described by the operator R in (1) acting on $\vec{v}_i$ and $\vec{v}_j$, that is

$$R^S_{i,j}[f](\underline{v}) = \int_{\mathbb{S}^2} f(\underline{v}_{i,j}(\omega))d\sigma(\omega) \tag{3}$$

where

$$\begin{aligned} &\underline{v}_{i,j}(\omega) = (\vec{v}_1, \ldots, \vec{v}_i^*(\omega), \ldots, \vec{v}_j^*(\omega), \ldots, \vec{v}_M)\,, \\ &\vec{v}_i^*(\omega) = \vec{v}_i - ((\vec{v}_i - \vec{v}_j)\cdot\omega)\omega\,, \qquad \vec{v}_j^*(\omega) = \vec{v}_j - ((\vec{v}_j - \vec{v}_i)\cdot\omega)\omega\,. \end{aligned} \tag{4}$$

We assume that the collisions take place according to a Poisson process of intensity λ_M and that when a collision takes place, the pair of colliding particles is chosen randomly and uniformly. Choosing $\lambda_M = M\lambda_S/2$ we get that the infinitesimal generator for the evolution of f is

$$\mathcal{L}_S[f] = \frac{\lambda_S}{M-1}\sum_{i<j}(R^S_{i,j}[f] - f). \tag{5}$$

In this way the average number of collisions that a particle suffers in a unit time is equal to λ_S independently of M. λ_S^{-1} is called the *mean free flight* for the system. For simplicity, we will choose a unit of time so that $\lambda_S = 1$.

In [4], in an attempt to introduce more structure in the Kac model, the authors consider the interaction of a Kac system with a *Maxwellian thermostat*. This is thought of as an infinite reservoir of particles in equilibrium at temperature T. When a particle interacts with the thermostat it undergoes a collision with a *virtual* particle randomly selected from a Maxwellian distribution at temperature T. After the collision, the virtual particle disappears. If $g(\vec{v})$ is the probability distribution of the velocity $\vec{v}$ of the particle, then the effect of the interaction with the thermostat is

$$B[g](\vec{v}) = \int_{\mathbb{R}^3} R[g\Gamma_T](\vec{v},\vec{w})d\vec{w} = \int_{\mathbb{R}^3} d\vec{w}\int_{\mathbb{S}^2} g(\vec{v}^*(\omega))\Gamma_T(\vec{w}^*(\omega))d\omega. \tag{6}$$

with

$$\Gamma_T(\vec{w}) = \left(\frac{1}{2\pi T}\right)^{\frac{3}{2}} e^{-\frac{1}{2T}\|\vec{w}\|^2}$$

and

$$\vec{v}^*(\omega) = \vec{v}_i - ((\vec{v}_i - \vec{w})\cdot\omega)\omega \qquad \vec{w}^*(\omega) = \vec{w} - ((\vec{w} - \vec{v}_i)\cdot\omega)\omega\,. \tag{7}$$

At Poisson distributed times, a particle in the system is selected uniformly at random, to interact with the thermostat. We can represent the action of this thermostat on the M-particle distribution f via the infinitesimal generator

$$\mathcal{L}_B[f] = \mu\sum_{i=1}^{M}(B_i[f] - f)\,. \tag{8}$$

where B_i is the operator (6) acting on the velocity $\vec{v}_i$ of the i-th particle. Thus the evolution of the combined system+thermostat can be described via the generator

$$\widetilde{\mathcal{L}}_T = \mathcal{L}_S + \mathcal{L}_B. \tag{9}$$

For the one dimensional version of this evolution, it is possible to show that for every given initial distribution f_0, the evolved distribution $f_t := e^{\widetilde{\mathcal{L}}_T t} f_0$ converges exponentially fast to a Maxwellian distribution both in a suitable L^2 norm and in entropy, see [4]. The L^2 norm is not a very useful measure of approach to equilibrium when M is large, see [4] for more details, while typical alternatives like the entropy are very hard to analyze. For these reasons, in this paper, we will work with the GTW distance d_2 defined as follows (see also [12]). Given two functions $f, g \in L^1(\mathbb{R}^{3M})$ with

$$\int_{\mathbb{R}^{3M}} f(\underline{v})d\underline{v} = \int_{\mathbb{R}^{3M}} g(\underline{v})d\underline{v} \qquad \int_{\mathbb{R}^{3M}} \underline{v}f(\underline{v})d\underline{v} = \int_{\mathbb{R}^{3M}} \underline{v}g(\underline{v})d\underline{v}$$

we set

$$d_2(f,g) = \sup_{\underline{\xi}\neq 0} \frac{|\hat{f}(\underline{\xi}) - \hat{g}(\underline{\xi})|}{|\underline{\xi}|^2} \tag{10}$$

where $\hat{f}(\underline{\xi})$ is the Fourier transform of $f(\underline{v})$ defined as

$$\hat{f}(\underline{\xi}) = \int_{\mathbb{R}^{3M}} e^{2\pi i(\underline{\xi}\cdot\underline{v})} f(\underline{v})d\underline{v}$$

and $(\underline{\xi}\cdot\underline{v}) = \sum_{i=1}^{M}(\vec{v}_i\cdot\vec{\xi}_i)$ is the scalar product in $\mathbb{R}^{3M}$. It is not hard to show that the evolution generated by $\widetilde{\mathcal{L}}_T$ converges exponentially to equilibrium also in the d_2 distance, see [8, 11] and Appendix A below.

It is natural to see the Maxwellian thermostat as an idealization of the interaction of our small system with M particles with a much larger reservoir with $N \gg M$ particles initially in equilibrium at temperature T. We now describe the combined system+reservoir evolution for the three dimensional version of the model we will use in this paper. The state of the combined system is described by a probability distribution $F(\underline{v}, \underline{w})$ on $\mathbb{R}^{3M+3N}$. The generator is given by

$$\mathcal{L}_T = \mathcal{L}_S + \mathcal{L}_R + \mathcal{L}_I \tag{11}$$

where $\mathcal{L}_R$ describes the evolution inside the reservoir. This is modeled as a standard Kac evolution analogous to that of the system with possibly a different mean free flight λ_R^{-1}. This means that

$$\mathcal{L}_R[F] = \frac{\lambda_R}{N-1}\sum_{i<j}(R^R_{i,j}[F] - F) \tag{12}$$

with $R^R_{i,j}$ describing the effect of a collision in the reservoir between particle i with velocity $\vec{w}_i$ and particle j with velocity $\vec{w}_j$ via the operator (1) acting on $\vec{w}_i$ and $\vec{w}_j$. Finally $\mathcal{L}_I$ describes the interaction between the system and the reservoir. It is also modeled via a Kac style collision and we set

$$\mathcal{L}_I[F] = \frac{\mu}{N}\sum_{i=1}^{N}\sum_{j=1}^{M}(R^I_{i,j}[F] - F) \tag{13}$$

where $R^I_{i,j}$ describes the effect of a collision between particle i in the reservoir with velocity $\vec{w}_i$ and particle j in the system with velocity $\vec{v}_j$. It is again given by the operator (1) acting on $\vec{v}_i$ and $\vec{w}_j$.

Observe that the factor μ/N in (13) has the effect that a particle in the system suffers, in average, μ collisions per unit time with particles in the reservoir. At the same time, a particle in the reservoir on average suffers $\mu M/N$ collision per unit time with particles in the system.

Since the reservoir is assumed to be initially in equilibrium, we choose an initial state of the form

$$F_0(\underline{v}, \underline{w}) = f_0(\underline{v})\Gamma_T^N(\underline{w}) \tag{14}$$

where

$$\Gamma_T^N(\underline{w}) = \prod_{i=1}^{N} \Gamma_T(\vec{w}_i)$$

is the Maxwellian distribution at temperature T for N particles.

For the evolution in one dimension, the relation between the finite reservoir and the thermostat was studied in [3]. The authors prove there that, for N large, the combined evolution of system+reservoir is well approximated by the system+thermostat idealization uniformly in time in the d_2 distance. A generalization of this results to the evolution in 3 dimension will be a corollary of our main result that we are going to present now.

As discussed above, to obtain a truly out of equilibrium evolution, we consider the situation when a small system with M particles interacts with two large reservoirs initially at temperature T_+ and T_- with N_+ and N_- particles respectively. The state of the combined system is now given by a distribution $\mathscr{F}(\underline{u}, \underline{v}, \underline{w})$ on $\mathbb{R}^{3(M+N_-+N_+)}$ and, analogously to (14), the initial state is of the form

$$\mathscr{F}_0(\underline{u}, \underline{v}, \underline{w}) = \Gamma_+(\underline{u}) f_0(\underline{v}) \Gamma_-(\underline{w}) \tag{15}$$

where $\Gamma_\sigma = \Gamma_{T_\sigma}^{N_\sigma}$ with $\sigma = \pm$. We also assume that the average velocity of the center of mass of the system vanishes while the initial temperature is finite, that is

$$\int_{R^{3M}} \underline{v} f_0(\underline{v}) d\underline{v} = 0, \qquad \text{and} \qquad T_S := \frac{1}{3M} \int_{R^{3M}} \|\underline{v}\|^2 f_0(\underline{v}) < \infty\,.$$

We can write the generator of the evolution as

$$\mathcal{L} = \mathcal{L}_S + \mathcal{L}_{R_+} + \mathcal{L}_{R_-} + \mathcal{L}_{I_+} + \mathcal{L}_{I_-} \tag{16}$$

where $\mathcal{L}_S$ still describes the internal evolution of the system, see (5), while $\mathcal{L}_{R_\sigma}$ and $\mathcal{L}_{I_\sigma}$, $\sigma = \pm$, describe the internal evolution of the σ reservoir and the interaction of the σ reservoir with the system, respectively, see (12) and (13). Observe that there is no direct interaction between the two reservoirs.

If $T_+ = T_-$ the two reservoir are essentially equivalent to a single reservoir. This situation will be discussed in Corollary 2.2 below. We will thus always assume that $T_+ > T_-$. On the other hand, to simplify notation, we will assume that $N_+ = N_- = N$. Thus we will consider the evolution

$$\mathscr{F}_t := e^{\mathcal{L}t} \mathscr{F}_0 \tag{17}$$

generated by the generator $\mathcal{L}$ in (16) starting from the initial condition $\mathscr{F}_0$ in (15).

It is natural to compare the evolution of the combined system+2 reservoirs described in (17) with the evolution of a small system with M particles interacting with two Maxwellian thermostats at temperature T_+ and T_-. In such a situation the generator of the evolution is given by

$$\widetilde{\mathcal{L}} = \mathcal{L}_S + \mathcal{L}_{B_+} + \mathcal{L}_{B_-} \tag{18}$$

where $\mathcal{L}_{B_\sigma}$ describes the thermostat at temperature T_σ with $\sigma \in \{+,-\}$, see (8), and the initial state is given by f_0. That is we want to find an upper bound for the d_2 distance between $\mathscr{F}_t$ in (15) and

$$\widetilde{\mathscr{F}_t} := \Gamma_+ e^{\widetilde{\mathcal{L}}t} f_0 \Gamma_- =: \Gamma_+ \tilde{f}_t \Gamma_- \tag{19}$$

in term of the temperature difference $T_+ - T_-$, the d_2 distance between f_0 and $\Gamma_\pm^M$, and the moments of f_0. Before stating our result we observe that, since $\mathcal{L}$ and $\widetilde{\mathcal{L}}$ are bounded operators on $L^1(\mathbb{R}^{3(M+2N)})$ and $L^1(\mathbb{R}^{3M})$, repsectively, both $\mathscr{F}_t$ and $\widetilde{\mathscr{F}_t}$ exists for every t. Moreover they are probability distributions with vanishing first moments and finite second moments so that $d_2(\mathscr{F}_t, \widetilde{\mathscr{F}_t})$ is well defined and finite.

To formulate our main Theorem we need a few more definitions. Let $I := \{(i,j) : i = 1, \ldots, M\ ,\ j = 1,2,3\}$ be the set of indices for the components of the vector $\underline{v}$ and, for $k > 0$, consider the set of monomials $\underline{v}_{\mathbf{i}} = \prod_{l=1}^k v_{i_l,j_l}$ where $\mathbf{i} = ((i_1,j_1), \ldots, (i_k,j_k)) \in I^k$. Given a probability distribution f on $\mathbb{R}^{3M}$ we define

$$e_{\mathbf{i}}(f) := \int_{\mathbb{R}^{3M}} |\underline{v}_{\mathbf{i}}|\, f(\underline{v}) d\underline{v} \tag{20}$$

and we set

$$E_4(f) := \max_{\mathbf{i} \in \tilde{I}^4} e_{\mathbf{i}}(f) \tag{21}$$

where $\tilde{I}^4 = \bigcup_{k=0}^4 I^k$.

We are now ready to formulate our results that are contained in the following Theorem.

Theorem 2.1. *Let $f_0(\underline{v})$ be a probability distribution on $\mathbb{R}^{3M}$ with $\int \underline{v} f_0(\underline{v}) d\underline{v} = 0$ and $E_4(f_0) < \infty$ and let $\mathscr{F}_0(\underline{v}, \underline{u}, \underline{w}) = \Gamma_+(\underline{u}) f_0(\underline{v}) \Gamma_-(\underline{w})$ as discussed in* (15)*. Consider the two evolved distributions on $\mathbb{R}^{3(2N+M)}$ given by*

$$\mathscr{F}_t = e^{\mathcal{L}t} \mathscr{F}_0 \qquad\qquad \widetilde{\mathscr{F}_t} = \Gamma_+ e^{\widetilde{\mathcal{L}}_T t} f_0 \Gamma_-$$

with $\mathcal{L}$ and $\widetilde{\mathcal{L}}$ given by (16) *and* (18) *respectively. Then we have*

$$d_2(\mathscr{F}_t, \widetilde{\mathscr{F}_t}) \le C \frac{M}{\sqrt{N}} E_4(f_0)^{\frac{5}{6}} \left((1 - e^{-\frac{\mu}{9} t}) \left(d_2(f_0, \Gamma_+^M)^{\frac{1}{6}} + d_2(f_0, \Gamma_-^M) \right)^{\frac{1}{6}} + (T_+ - T_-)^{\frac{1}{6}} t \right) \tag{22}$$

for a suitable constant C.

Remark 1. Here and in what follows, C will indicate a generic constant, independent of N, M and f_0 whose numerical value is not relevant and may vary even inside the same formula.

Due to the presence of the factor of t in the last term of (22), the above theorem tells us nothing about the difference between $\mathscr{F}_t$ and $\widetilde{\mathscr{F}}_t$ when t is very large. This is not strange since, as opposed to the case of only one reservoir discussed in [3], see also Corollary 2.2 below, we cannot expect $\mathscr{F}_t$ and $\widetilde{\mathscr{F}}_t$ to stay close uniformly in t. To see this it is enough to look at the steady states of the two evolutions considered in Theorem 2.1.

Indeed, on the one hand $\mathscr{F}_\infty := \lim_{t\to\infty} \mathscr{F}_t$ is given by the rotational average at fixed total momentum of $\mathscr{F}_0$ on all its variables. More precisely let $\vec{e}_i$, $i = 1, 2, 3$ be the standard basis in $\mathbb{R}^3$ and, for fixed P consider the vectors $E_i = (\vec{e}_i, \ldots, \vec{e}_i) \in \mathbb{R}^{3P}$. The set $\mathcal{O}$ of unitary transformations with determinant one that preserve E_i, $i = 1, 2, 3$, is a subgroup of the special unitary group $SO(3P)$. Let dO be the normalized Haar measure on $\mathcal{O}$. Given a function $F(\underline{v})$ from $\mathbb{R}^{3P}$ to $\mathbb{R}$ we define its fixed momentum rotational average as

$$\mathcal{R}[F](\underline{v}) = \int_{\mathcal{O}} F(O\underline{v}) d\sigma(O)\,. \tag{23}$$

Clearly $\mathcal{R}[F](\underline{v})$ depends only on $\|\underline{v}\|$ and $\vec{V} = \sum_{i=1}^N \vec{v}_i$. Thus, in particular, $\mathscr{F}_\infty$ is invariant under permutation of all of its arguments. On the other hand, using an argument developed in [8, 11] (see also [9]) we will show in Appendix A that $\tilde{f}_\infty = \lim_{t\to\infty} \tilde{f}_t$ exists and is approached exponentially fast so that $\widetilde{\mathscr{F}}_\infty := \lim_{t\to\infty} \widetilde{\mathscr{F}}_t = \Gamma_+ \tilde{f}_\infty \Gamma_-$. Thus $\widetilde{\mathscr{F}}_\infty$ is far from $\mathscr{F}_\infty$, for any N. See Section 4 for further discussion of this point.

The main ingredient in the proof of Theorem 2.1 is the observation that on a state of the form (15) the action of $\mathcal{L}$ and $\widetilde{\mathcal{L}}$ are very similar. This is due to the product structure of (15). More precisely, this is due to the fact that the velocity of any particle in one of the reservoirs is independent of the velocity of any other particle and has the the same distribution as the virtual particles in the thermostats. The functional inequality discussed in Section 3 makes this heuristic observation quantitative in terms of the d_2 metric. This inequality extends our previous result in [3] where the problem was simpler thanks to particular symmetries of the one dimensional evolution. Nonetheless, the proof presented here simplifies that of [3], makes it easier to understand, and is more general.

It should also be observed that, even if the system+2 reservoirs is initially in a state of the form (15), the product structure is not preserved by the evolution generated by $\mathcal{L}$. Nonetheless, it is possible to interpolate between $e^{\mathcal{L}t}$ and $e^{\widetilde{\mathcal{L}}t}$ in such a way that we mainly have to estimate the d_2 distance between $\mathcal{L}\widetilde{\mathscr{F}}_t$ and $\widetilde{\mathcal{L}}\widetilde{\mathscr{F}}_t$, where $\widetilde{\mathscr{F}}_t$ is given by (19) and thus shares the product structure of (15). In the present work, we achieve this via a Duhamel expansion (see Section 6.2).

We can now look back at the situation where there is only one reservoir. A straightforward adaptation of the proof of Theorem 2.1 yields the following.

Corollary 2.2. *Let $f_0(\underline{v})$ be a probability distribution on $\mathbb{R}^{3M}$ with $E_4(f_0) < \infty$, where E_4 is defined as in (21), and let $F_0(\underline{v}, \underline{w}) = f_0(\underline{v})\Gamma_T^N(\underline{w})$ as discussed in (14). Consider the two evolved distributions on $\mathbb{R}^{3(N+M)}$ given by*

$$F_t = e^{\mathcal{L}_T} F_0 \qquad \text{and} \qquad \widetilde{F}_t = \Gamma_T^N e^{\widetilde{\mathcal{L}}_T} f_0$$

with $\mathcal{L}_T$ and $\widetilde{\mathcal{L}_T}$ given by (11) *and* (9) *respectively. Then we have*

$$d_2(F_t, \widetilde{F}_t) \leq C\frac{M}{\sqrt{N}} E_4(f_0)^{\frac{5}{6}}(1 - e^{-\frac{\mu}{9}t}) d_2(f_0, \Gamma_T^M)^{\frac{1}{6}} . \tag{24}$$

As for the 1 dimensional case, we see that $\widetilde{F}_\infty = \lim_{t\to\infty} \widetilde{F}_t = \Gamma_T^{N+M}$ while $F_\infty = \lim_{t\to\infty} F_t = \mathcal{R}[F_0]$, see (23). In Appendix C we prove that, as for the model in one dimension, if $F_0(\underline{v}, \underline{w}) = f_0(\underline{v})\Gamma_T^N(\underline{w})$ then

$$d_2(F_\infty, \Gamma_{T,N+M}) \leq d_2(f_0, \Gamma_T^M)\frac{M}{N+M}.$$

Thus our result in Corollary 2.2 is surely not optimal, as far as the asymptotic behavior in N and t is concerned. On the other hand, the proof of Lemma 3.3 makes it quite clear that, at least for short time t, the behavior in $1/\sqrt{N}$ is not an artifact of our techniques. Such a behavior is a new feature of the three dimensional evolution, see [3], and it is our opinion that it is linked to the preservation of the total moment by the evolution generated by (16). See Section 4 for further observations.

3 A new inequality

The above results are based on an extension of a functional inequality that first appeared in [3]. The proof contained in that paper is too complex and contains a gap. Thus we first report a more general version of the original statement and its simplified proof and then present the extension we will use here.

Consider a C^3 function $H : \mathbb{R}^3 \mapsto \mathbb{C}$ with $H(0) = 0$ and $\nabla H(0) = 0$ and a C^0 function $g : \mathbb{R}^3 \mapsto \mathbb{C}$ such that

$$|g(\vec{\eta})| \leq \frac{1}{1 + T\|\vec{\eta}\|^2}$$

for some $T > 0$. For $\underline{\eta} = (\vec{\eta}_0, \ldots, \vec{\eta}_N) \in \mathbb{R}^{3N}$ let

$$g^N(\underline{\eta}) = \prod_{i=1}^{N} g(\vec{\eta}_i)$$

and set $\underline{\eta}^i = (\vec{\eta}_0, \ldots, \vec{\eta}_{i-1}, \vec{\eta}_{i+1}, \ldots, \vec{\eta}_N)$. Given $\vec{\xi} \in \mathbb{R}^p$, $p \geq 1$, we want to control the *interlaced sum*

$$\mathcal{D}_N(H, \vec{\xi}) := \sup_{\underline{\eta}\neq 0} \frac{1}{\|\underline{\eta}\|^2 + \|\vec{\xi}\|^2} \left| \sum_{i=1}^{N} g^{N-1}(\underline{\eta}^i) H(\vec{\eta}_i) \right| \tag{25}$$

and show that it can be estimated in terms of

$$\mathcal{D}_1(H, \vec{\xi}) = \sup_{\vec{\eta}\neq 0} \frac{|H(\vec{\eta})|}{\|\vec{\eta}\|^2 + \|\vec{\xi}\|^2}$$

uniformly in N. More precisely we will prove the following.

Lemma 3.1. *Let $H : \mathbb{R}^3 \to \mathbb{C}$ be a C^3 function such that $H(0) = 0$, $\nabla H(0) = 0$ and*

$$\|H\|_3 := \max_{\vec{\beta}\in\mathbb{N}^3, \|\vec{\beta}\|_1\le 3} \left\|\partial^{\vec{\beta}} H\right\|_\infty < \infty\,. \tag{26}$$

and let $g : \mathbb{R}^3 \mapsto \mathbb{C}$ be a C^0 function such that

$$|g(\vec{\eta})| \le \frac{1}{1 + T\|\vec{\eta}\|^2}$$

for some $T > 0$. If we call

$$\mathcal{D}_N(H,\xi) = \sup_{\underline{\eta}\neq 0} \frac{1}{\|\underline{\eta}\|^2 + \|\vec{\xi}\|^2} \left|\sum_{i=1}^N g^{N-1}(\underline{\eta}^i) H(\vec{\eta}_i)\right|$$

then we have

$$\mathcal{D}_N(H,\vec{\xi}) \le C\|H\|_3^{\frac{2}{3}} \mathcal{D}_1(H,\vec{\xi})^{\frac{1}{3}}\,. \tag{27}$$

This Lemma is very similar to Proposition 5 in [3] but it contains a few improvements. For one, it incorporates a more general function g in place of the Fourier transform of the Maxwellian at temperature T. Moreover, it does not require H to be even, which is a necessary improvement for our current setting. This is the reason for the cubic root in the main estimate, as opposed to the square root in the analogous estimate from [3].

In the following we will need to apply Lemma 3.1 to a function $H : \mathbb{R}^3 \mapsto \mathbb{C}$ with $H(0) = 0$, but $\nabla H(0) \neq 0$. We thus first prove the following estimate.

Lemma 3.2. *Let $H : \mathbb{R}^3 \times \mathbb{R}^p \to \mathbb{C}$ be a $C^{3,1}$ function such that $H(0,0) = 0$, $\nabla_{\vec{\eta}} H(0,0) = 0$ and, for every $\vec{\xi} \in \mathbb{R}^p$, $H(0,\vec{\xi}) = 0$. Assume moreover that*

$$\|H\|_{3,0} := \sup_{\vec{\xi}} \|H(\cdot,\vec{\xi})\|_3 < \infty\,. \tag{28}$$

and

$$|H|_{1,1} := \sup_{\vec{\xi}} \left\|\nabla_{\vec{\xi}}\nabla_{\vec{\eta}} H(0,\vec{\xi})\right\| < \infty \tag{29}$$

Then calling

$$\alpha(\vec{\xi}) := \frac{\|\nabla_\eta H(0,\xi)\|}{\|\vec{\xi}\|\sqrt{1+\|\vec{\xi}\|^2}} \tag{30}$$

we get

$$\alpha(\vec{\xi}) \le 4\min\left\{\|H\|_{3,0}^{\frac{1}{2}} \mathcal{D}_1(H,\vec{\xi})^{\frac{1}{2}}, |H|_{1,1}\right\} \tag{31}$$

From Lemmas 3.1 and 3.2, the extension we will need in this paper follows.

Lemma 3.3. *Let $H : \mathbb{R}^3 \times \mathbb{R}^p \to \mathbb{C}$ be be a $C^{3,1}$ function such that $H(0,0) = 0$, $\nabla_{\vec{\eta}} H(0,0) = 0$ and, for every $\vec{\xi} \in \mathbb{R}^p$, $H(0, \vec{\xi}) = 0$. Assume moreover that*

$$\|H\|_{3,1} := \max_{\|\vec{\beta}_\eta\|_1 < 3, \|\vec{\beta}_\xi\| < 1} \left\| \partial_{\vec{\eta}}^{\vec{\beta}_\eta} \partial_{\vec{\xi}}^{\vec{\beta}_\xi} H \right\|_\infty < \infty \,. \tag{32}$$

Let also $g : \mathbb{R}^3 \mapsto \mathbb{C}$ be a C^0 function such that

$$|g(\vec{\eta})| \leq \frac{1}{1 + T\|\vec{\eta}^2\|}$$

for some $T > 0$. Then we have

$$\mathcal{D}_N(H, \vec{\xi}) \leq C\sqrt{N} \|H\|_{3,1}^{\frac{5}{6}} D_1(H, \vec{\xi})^{\frac{1}{6}} \,. \tag{33}$$

4 Outlook

Two limitations of our results appear evident when looking at the statements of Theorem 2.1 and Corollary 2.2.

In the first place, notwithstanding the fact that we consider the system to be 'small', we still want to think of the system as a macroscopic object. That is, we want M to be a fraction of the Avogadro number. In both of our results, the difference between the system+reservoir(s) evolution and system+thermostat(s) evolution behave as $M/\sqrt{N}$. This implies that, for the result to be interesting, we need $N >> M^2$, and thus N has to be un-physically large.

Additionally, as we noted after Corollary 2.2, such behavior, at least for t very large, cannot be optimal. We think that, by adapting the proof of the main theorem in [19] to a three dimensional evolution, one can show that the contribution of the order of $M/\sqrt{N}$ in (24) decay exponentially in time leaving an asymptotic estimate of the order of M^2/N. This would represent just a mild improvement with respect to the present estimate. Moreover the rate of decay, as far as one can get from [19], would be of the order of $1/N$ and thus rather too slow to be relevant.

It is quite clear from the proof of Theorem 2.1 that the main difference between the three-dimensional and the one-dimension Kac evolution is the preservation of the total momentum in addition to the preservation of the total energy. On the other hand, the metric d_2 seems to be well adapted to take advantage of the latter but contains no reference to the former. We think that it should be possible to modify the d_2 metric taking into account the preservation of momentum and obtaining better estimates. As partial validation of this idea, we observe that if $M = 1$ and f_0 is rotationally invariant, then a direct computation shows that we can apply Lemma 3.1 to obtain an estimate of the order of $1/N$.

In the second place, as already observed, Theorem 2.1 is only relevant if $t << \sqrt{N}$. To better understand the long time evolution of the system + 2 reservoir, we can define the temperature, or better average kinetic energy per degree of freedom, in the reservoir, $\tau_+(t)$, $\tau_-(t)$, and in the system,

$\tau_S(t)$, as

$$\tau_S(t) = \frac{1}{3M}\int_{\mathbb{R}^{3(2N+M)}} \|\underline{v}\|^2 \mathscr{F}_t(\underline{u},\underline{v},\underline{w})d\underline{v}d\underline{u}d\underline{w}$$
$$\tau_+(t) = \frac{1}{3N}\int_{\mathbb{R}^{3(2N+M)}} \|\underline{u}\|^2 \mathscr{F}_t(\underline{u},\underline{v},\underline{w})d\underline{v}d\underline{u}d\underline{w}$$

and similarly for $\tau_-(t)$. It is not hard to show that they evolve according to Newton's Law of cooling:

$$\begin{cases} \dot{\tau}_-(t) = \dfrac{\mu M}{3N}(\tau_S(t) - \tau_-(t)) \\ \dot{\tau}_S(t) = \dfrac{\mu}{3}(\tau_-(t) - \tau_S(t)) + \dfrac{\mu}{3}(\tau_+(t) - \tau_S(t)) \\ \dot{\tau}_+(t) = \dfrac{\mu M}{3N}(\tau_S(t) - \tau_+(t)), \end{cases} \tag{34}$$

where $\tau_\sigma(0) = T_\sigma$ for $\sigma \in \{+,-\}$ while we set $\tau_S(0) = T_S$. See Appendix B for a derivation. It is easy to solve (34) and obtain

$$\begin{pmatrix}\tau_+(t)\\ \tau_S(t)\\ \tau_-(t)\end{pmatrix} = T_\infty\begin{pmatrix}1\\1\\1\end{pmatrix} + \frac{T_+ - T_-}{2}\begin{pmatrix}1\\0\\-1\end{pmatrix}e^{-\frac{\mu M}{3N}t} + \frac{T_+ - 2T_S + T_-}{2\overline{N}}\begin{pmatrix}M\\-2N\\M\end{pmatrix}e^{-\frac{\mu}{3}(2+\frac{M}{N})t}. \tag{35}$$

where $\overline{N} = 2N + M$ is the total number of particles while $T_\infty = (NT_+ + MT_S + NT_-)/\overline{N}$ is the average temperature. On a short time scale, $\tau_S(t)$ converges exponentially, with a rate of order 1, to T_∞. This is the time frame over which our Theorem gives interesting information. On a longer time scale, of the order of N, $\tau_+(t) - \tau_-(t)$ tend to vanishes at a rate proportional to N^{-1}. This evolution of $\tau_\pm(t)$ is the main reason why $\mathcal{F}(t)$ and $\widetilde{\mathcal{F}}$ cannot stay close for long time.

On the other hand, due to the slow rate of such evolution and the presence of the Kac collisions in the reservoirs, one can heuristically expect that the state of the reservoirs in uniformly close to $\Gamma^N_{\tau_\pm(t)}$, that is, an equilibrium state at temperature $\tau_\pm(t)$. Thus for t larger than the initial transient, we will have $\mathscr{F}_t \simeq \Gamma^N_{\tau_+(t)}\bar{f}_t\Gamma^N_{\tau_-(t)}$ where $\bar{f}_t$ is the steady state of the system of M particles interacting with two Maxwellian thermostat at temperature $\tau_+(t)$ and $\tau_-(t)$. This suggests replacing the evolution generated by $\widetilde{\mathcal{L}}$ by considering a system interacting with two reservoirs whose temperatures evolve in time.

5 Proof of the key Lemmas

We begin with the proof of Lemma 3.1. The proof reported here is quite similar to the one in [3] but it is simpler in several points and it deals with a more general case. Moreover it fills a gap in the previous proof. We will then show how Lemma 3.3 follows from Lemma 3.1 with the aide of Lemma 3.2.

5.1 Proof of Lemma 3.1

As observed in [3], we can easily show that

$$\mathcal{D}_N(H,\vec{\xi}) \leq \mathcal{D}_1(H,0).$$

On the other hand $\mathcal{D}_1(H,0) \geq \mathcal{D}_1(H,\vec{\xi})$ so that this estimate is not enough for our purpose. What prevent us to get a better estimate this way is the fact that the supremum over $\underline{\eta}$ may be achieved for $\|\underline{\eta}\|$ very large thus making $\|\vec{\xi}\|^2$ essentially negligible in the denominator of (25).

Thus our proof is essentially divided into two parts. In the first part, we show that $\mathcal{D}_N(H,\vec{\xi})$ can be controlled by $\mathcal{D}_1(H,0)/(1+T\|\vec{\xi}\|^2)$ by suitably restricting the set of $\underline{\eta} \neq 0$ we need to consider when taking the supremum in (25) and then applying a convexity argument. Indeed, the presence of the term $g^{N-1}(\underline{\eta}^i)$ essentially implies that we can restrict our attention to those $\underline{\eta}$ which lie in a small box around 0. Convexity of the box and of the function to be estimated inside the box then allows us to explicitly compute the supremum there. In the second part, we show that $\mathcal{D}_1(H,0)/(1+T\|\vec{\xi}\|^2)$ can be controlled in turn by $\mathcal{D}_1(H,\vec{\xi})$ via a Taylor expansion.

First note that $\mathcal{D}_N(H,\vec{\xi})$ only depends on $\xi = \|\vec{\xi}\|$ so that we can assume that $p=1$ and write

$$\mathcal{D}_N(H,\xi) = \sup_{\underline{\eta}\neq 0} \frac{1}{\|\underline{\eta}\|^2+\xi^2} \left| \sum_{i=1}^{N} g^{N-1}(\underline{\eta}^i) H(\vec{\eta}_i) \right|$$

while

$$\mathcal{D}_1(H,\xi) = \sup_{\eta\neq 0} \frac{|H(\eta)|}{\|\vec{\eta}\|^2+\xi^2} \tag{36}$$

It follows that, for all $\vec{\eta} \in \mathbb{R}^3$, we have

$$H(\vec{\eta}) \leq \min\left\{\mathcal{D}_1(H,0)\|\vec{\eta}\|^2, \mathcal{D}_1(H,\vec{\xi})(\|\vec{\eta}\|^2+\xi^2)\right\}.$$

With this in mind, we define $\widetilde{H}(\vec{\eta},\xi) \geq 0$ by

$$\widetilde{H}(\vec{\eta},\xi) := \min\left\{\mathcal{D}_1(H,0)\|\vec{\eta}\|^2, \mathcal{D}_1(H,\vec{\xi})(\|\vec{\eta}\|^2+\xi^2)\right\} \tag{37}$$

so that $H(\vec{\eta}) \leq \widetilde{H}(\vec{\eta},\xi)$ and thus

$$\mathcal{D}_N(H,\xi) \leq \mathcal{D}_N(\widetilde{H},\xi). \tag{38}$$

It will be more convenient to estimate $\mathcal{D}_N(\widetilde{H},\xi)$ instead of $\mathcal{D}_N(H,\xi)$ directly. Observe that for

$$\eta_0^2 = \frac{\mathcal{D}_1(H,\xi)\xi^2}{\mathcal{D}_1(H,0)-\mathcal{D}_1(H,\xi)}.$$

we get

$$\mathcal{D}_1(H,0)\eta_0^2 = \mathcal{D}_1(H,\xi)(\eta_0^2+\xi^2)$$

so that we can write $\widetilde{H}$ as

$$\widetilde{H}(\vec{\eta},\xi) = \begin{cases} \mathcal{D}_1(H,0)\|\vec{\eta}\|^2; & \|\vec{\eta}\| \leq \eta_0 \\ \mathcal{D}_1(H,\xi)(\|\vec{\eta}\|^2+\xi^2); & \|\vec{\eta}\| > \eta_0 \end{cases}. \tag{39}$$

We now call

$$G^N(\underline{\eta}) = \prod_{i=1}^{N} \frac{1}{1+T\|\vec{\eta}_i\|}\,.$$

Since $\widetilde{H}$ is positive we can replace g^N with G^N in the definition of $\mathcal{D}_N(\widetilde{H},\xi)$. Thus from now on we will study

$$\mathcal{D}_N(\widetilde{H},\xi) = \sup_{\underline{\eta}\neq 0} \frac{1}{\|\underline{\eta}\|^2+\xi^2} \sum_{i=1}^{N} G^{N-1}(\underline{\eta}^i)\widetilde{H}(\vec{\eta}_i,\xi)$$

Our next step is to show that we may restrict our attention to the box $\left\{\underline{\eta} : \|\vec{\eta}_j\| \le \eta_0\right\}$. This will essentially follow from the monotonicity of the two functions $G(\vec{\eta})$ and $\frac{ka^2+x^2}{a^2+x^2}$ and the fact that $\frac{|a|+|b|}{|c|+|d|} \le \max\left\{\frac{|a|}{|c|}, \frac{|b|}{|d|}\right\}$.

Claim 1. *For $\widetilde{H}$ as above, we have*

$$\sup_{\underline{\eta}\neq 0} \frac{\sum_{j=1}^N \widetilde{H}(\vec{\eta}_j,\xi)G^{N-1}(\underline{\eta}^j)}{\|\underline{\eta}\|^2+\xi^2} \le 2 \sup_{0<\|\underline{\eta}\|_\infty \le \eta_0} \frac{\sum_{j=1}^N \widetilde{H}(\vec{\eta}_j,\xi)G^{N-1}(\underline{\eta}^j)}{\|\underline{\eta}\|^2+\xi^2}.$$

where $\|\underline{\eta}\|_\infty = \max_k \|\vec{\eta}_k\|$.

Proof. For ease of notation, we will write

$$\mathcal{H}(\underline{\eta},\xi) := \frac{\sum_{j=1}^N \widetilde{H}(\vec{\eta}_j,\xi)G^{N-1}(\underline{\eta}^j)}{\|\underline{\eta}\|^2+\xi^2}.$$

First, suppose, for some $k \ge 1$, $\underline{\eta} = (\vec{\eta}_1, ..., \vec{\eta}_k, 0, ..., 0)$ with $\|\vec{\eta}_j\| \ge \eta_0$ for all $j \le k$, and denote $\underline{\eta}_{0,k} = (\vec{\eta}_0, ..., \vec{\eta}_0, 0, ..., 0)$, for some $\vec{\eta}_0$ such that $\|\vec{\eta}_0\| = \eta_0$, and where there are k non-zero components. Using the monotonicity of G and the fact that the function $\frac{ka^2+x}{a^2+x}$ is decreasing in x, we get

$$\begin{aligned}
\mathcal{H}(\underline{\eta},\xi) &= \frac{\mathcal{D}_1(H,\xi)\sum_{i=1}^k(\xi^2+|\vec{\eta}_i|^2)G^{N-1}(\underline{\eta}^i)}{\xi^2+\|\underline{\eta}\|^2} \\
&\le \frac{\mathcal{D}_1(H,\xi)G(\vec{\eta}_0)^{N-1}(k\xi^2+\|\underline{\eta}\|^2)}{\xi^2+\|\underline{\eta}\|^2} \\
&\le \frac{\mathcal{D}_1(H,\xi)G(\vec{\eta}_0)^{N-1}(k\xi^2+\|\underline{\eta}_{0,k}\|^2)}{\xi^2+\|\underline{\eta}_{0,k}\|^2} \\
&= \frac{\sum_{i=1}^k \widetilde{H}(\vec{\eta}_0,\xi)G^{N-1}(\underline{\eta}_{0,k}^i)}{\xi^2+\|\underline{\eta}_{0,k}\|^2}.
\end{aligned} \tag{40}$$

Hence

$$\mathcal{H}(\underline{\eta},\xi) \le \mathcal{H}(\underline{\eta}_{0,k},\xi).$$

More generally, suppose $\underline{\eta} = (\vec{\eta}_1, ..., \vec{\eta}_N)$ with $\|\vec{\eta}_1\|, ..., \|\vec{\eta}_k\| \geq \eta_0$ and $\|\vec{\eta}_{k+1}\|, ..., \|\vec{\eta}_N\| < \eta_0$, for some k. Set $\underline{\eta}_* = (\vec{\eta}_1, ..., \vec{\eta}_k, 0, ..., 0)$, and $\underline{\eta}_{**} = (0, ..., 0, \vec{\eta}_{k+1}, ..., \vec{\eta}_N)$. Then we have

$$\begin{aligned}\mathcal{H}(\underline{\eta}, \xi) &= \frac{\sum_{i=1}^N \widetilde{H}(\vec{\eta}_i, \xi) G^{N-1}(\underline{\eta}^i)}{\xi^2 + \|\underline{\eta}\|^2} \\ &\leq \frac{\sum_{i=1}^k \mathcal{D}_1(H,\xi)(\xi^2 + \|\vec{\eta}_i^2\|) G^{N-1}(\underline{\eta}^i) + \sum_{i=k+1}^N \mathcal{D}_1(H,\xi)\|\vec{\eta}_i\|^2 G^{N-1}(\underline{\eta}^i)}{\frac{1}{2}\xi^2 + \frac{1}{2}(\|\vec{\eta}_1\|^2 + \cdots + \|\vec{\eta}_k^2\|) + \frac{1}{2}\xi^2 + \frac{1}{2}(\|\vec{\eta}_{k+1}^2\| + \cdots + \|\vec{\eta}_N^2\|)} \\ &\leq 2\max\left\{ \frac{\sum_{i=1}^k \mathcal{D}_1(H,\xi)(|\xi|^2 + \|\vec{\eta}\|^2) G^{N-1}(\underline{\eta}^i)}{\xi^2 + (\|\vec{\eta}_1^2\| + \cdots + \|\vec{\eta}_k^2\|)}, \frac{\sum_{i=k+1}^N \mathcal{D}_1(H,0)\|\vec{\eta}_i\|^2 G^{N-1}(\vec{\eta}^i)}{\xi^2 + (\|\vec{\eta}_{k+1}\|^2 + \cdots + \|\vec{\eta}_N\|^2)} \right\}.\end{aligned} \tag{41}$$

Since η_* is of the form we considered in the first case, for the first term we get

$$\frac{\sum_{i=1}^k \mathcal{D}_1(H,\xi)(\xi^2 + \|\vec{\eta}_i\|^2) G^{N-1}(\underline{\eta}^i)}{\xi^2 + (\|\vec{\eta}_1\|^2 + \cdots + \|\vec{\eta}_k^2)\|} \leq \mathcal{H}(\underline{\eta}_*, \xi) \leq \mathcal{H}(\underline{\eta}_{0,k}, \xi).$$

where we have used again the monotonicity of G for the first inequality. On the other hand, the second term satisfies

$$\frac{\sum_{i=k+1}^N \mathcal{D}_1(H,0)\|\vec{\eta}_i\|^2 G^{N-1}(\underline{\eta}^i)}{\xi^2 + (\|\vec{\eta}_{k+1}\|^2 + \cdots + \|\vec{\eta}_N\|^2)} \leq \mathcal{H}(\underline{\eta}_{**}, \xi).$$

Since $\vec{\eta}_{**}$ has all component of norm less than η_0, combining these two cases yields the claim. □

We are now left with

$$\sup_{\underline{\eta}\neq 0} \frac{\sum_{j=1}^N \widetilde{H}(\vec{\eta}_j, \xi) G^{N-1}(\underline{\eta}^j)}{\|\underline{\eta}\|^2 + \xi^2} \leq 2 \sup_{\vec{\eta}\neq 0, \|\vec{\eta}_j\| \leq \eta_0} \frac{\sum_{j=1}^N \mathcal{D}_1(H,0)\|\vec{\eta}_j\|^2 G^{N-1}(\underline{\eta}^j)}{\|\underline{\eta}\|^2 + \xi^2}. \tag{42}$$

Next, using a convexity argument, we will essentially reduce the supremum over the set $\{\vec{\eta}_j \mid \|\vec{\eta}_j\| \leq \eta_0\}$ to te supremum over an interval.

Claim 2. *The supremum on the right-hand side of* (42) *is achieved when at most one* $\vec{\eta}_i$ *satisfies* $0 < \|\vec{\eta}_i\| < \eta_0$.

Proof. Since the right hand side of (42) depends only on the $\|\vec{\eta}_i\|^2$ we can set $x_i = \|\vec{\eta}_i\|^2$. Thus we call $B_r := \left\{ \underline{x} : \sum_{i=1}^N x_i = r, 0 \leq x_i \leq \eta_0^2 \right\}$ and observe that for $x_i > 0$ we have

$$\prod_{i=1}^N \frac{1}{1 + Tx_i} \leq \frac{1}{1 + T\|\underline{x}\|_1}$$

where $\|\underline{x}||_1 = \sum_{i=1}^N |x_i|$. We thus get

$$\sup_{\vec{\eta}\neq 0, \|\vec{\eta}_j\| \leq \eta_0} \frac{\sum_{j=1}^N \|\vec{\eta}_j\|^2 \hat{G}^{N-1}(\underline{\eta}^j)}{\|\underline{\eta}\|^2 + \xi^2} \leq \sup_{0 < r \leq N\eta_0^2} \sup_{\underline{x}\in B_r} \frac{1}{1 + T\|\underline{x}\|_1} \frac{\sum_{i=1}^N x_i(1 + Tx_i)}{\|\underline{x}\|_1 + \xi^2}$$

Since B_r is a convex set and $\sum_i x_i(1+Tx_i)$ is a convex function we get

$$\sup_{\underline{x}\in B_r} \frac{1}{1+T\|\underline{x}\|} \frac{\sum_{i=1}^N x_i(1+Tx_i)}{\|\underline{x}\|_1+\xi^2} \leq \sup_{\underline{x}\in \partial B_r} \frac{1}{1+T\|\underline{x}\|} \frac{\sum_{i=1}^N x_i(1+Tx_i)}{\|\underline{x}\|_1+\xi^2}$$

Observe now that ∂B_r contains only points for which at most one $0 < x_i < \eta_0^2$. The claim now follows. $\square$

Thus we have

$$\sup_{\underline{\eta}\neq 0, \|\vec{\eta}_j\|\leq \eta_0} \frac{\sum_{j=1}^N \widetilde{H}(\vec{\eta}_j,\xi) G^{N-1}(\underline{\eta}^j)}{|\underline{\eta}|^2+\xi^2}$$
$$\leq \mathcal{D}_1(H,0) \max_{0\leq k\leq N-1} \sup_{0\leq|\eta|\leq\eta_0} \frac{k\eta_0^2(1+T((k-1)\eta_0^2+\eta^2))^{-1}+\eta^2(1+Tk\eta_0^2)^{-1}}{k\eta_0^2+\eta^2+\xi^2}$$

Observe now that

$$\frac{k\eta_0^2(1+T((k-1)\eta_0^2+\eta^2))^{-1}+\eta^2(1+Tk\eta_0^2)^{-1}}{k\eta_0^2+\eta^2+\xi^2} \leq$$
$$\max\left\{\frac{\eta_0^2}{\eta_0^2+\xi^2/2}, \frac{(k-1)\eta_0^2(1+T((k-1)\eta_0^2+\eta^2))^{-1}+\eta^2(1+Tk\eta_0^2)^{-1}}{(k-1)\eta_0^2+\eta^2+\xi^2/2}\right\}.$$

Moreover we have

$$\mathcal{D}_1(H,0)\frac{\eta_0^2}{\eta_0^2+\xi^2/2} \leq 2\mathcal{D}_1(H,0)\frac{\eta_0^2}{\eta_0^2+\xi^2} = 2\mathcal{D}_1(H,\xi)$$

where the last identity follows from the definition of η_0, while

$$\frac{(k-1)\eta_0^2(1+T((k-1)\eta_0^2+\eta^2))^{-1}+\eta^2(1+Tk\eta_0^2)^{-1}}{(k-1)\eta_0^2+\eta^2+\xi^2/2} \leq \frac{((k-1)\eta_0^2+\eta^2)(1+T((k-1)\eta_0^2+\eta^2))^{-1}}{k\eta_0^2+\eta^2+\xi^2/2}$$

Since the right hand side of the last inequality depends only on $(k-1)\eta_0^2+\eta^2$ we have

$$\max_{0\leq k\leq N-1} \sup_{0\leq|\eta|\leq\eta_0} \frac{((k-1)\eta_0^2+\eta^2)(1+T((k-1)\eta_0^2+\eta^2))^{-1}}{k\eta_0^2+\eta^2+\xi^2/2} \leq 2\sup_{y>0}\frac{y(1+Ty)^{-1}}{y+\xi^2} \leq \frac{2}{1+T\xi^2}.$$

Collecting everything we get

$$\mathcal{D}(H,\xi) \leq 4\max\left\{\mathcal{D}_1(H,\xi), \frac{\mathcal{D}_1(H,0)}{1+T\xi^2}\right\}$$

We now need to control $\mathcal{D}_1(H,0)/(1+T|\xi^2|)$ with $\mathcal{D}_1(H,\xi)$. This will be a consequence of the following Claim.

Claim 3. *Under the assumption of Lemma 3.1 we have*

$$\mathcal{D}_1(H,\xi) \geq \frac{1}{57} \frac{\mathcal{D}_1(H,0)^3}{\mathcal{D}_1(H,0)^2 + \|H\|_3^2 \xi^2} \, .$$

Proof. We consider the function $h : \mathbb{R}^+ \mapsto \mathbb{R}$ defined as

$$h(r) = \sup_{\|\omega\|=1} |H(\omega r)|.$$

and observe that

$$\mathcal{D}_1(H,\xi) = \sup_{r>0} \frac{h(r)}{r^2+\xi^2}$$

Since we can write

$$H(\vec{\eta}) = \sum_{i,j=1}^{3} \partial_i \partial_j H(0) \eta_i \eta_j + \sum_{i,j,k=1}^{3} C_{i,j,k}(\vec{\eta}) \eta_i \eta_j \eta_k$$

calling $\bar{H}$ the largest modulus of the eigenvalues of Hessian matrix of H at $\vec{\eta} = 0$, we get

$$|H(\vec{\eta})| \leq \bar{H} \|\vec{\eta}\|^2 + \|H\|_3 \|\vec{\eta}\|^3$$

while

$$\sup_{\omega} |H(r\omega)| \geq r^2 \sup_{\omega} \left| \sum_{i,j=1}^{3} \partial_i \partial_j H(0) \omega_i \omega_j \right| - r^3 \sup_{\omega} \left| \sum_{i,j,k=1}^{3} C_{i,j,k}(r\omega) \omega_i \omega_j \omega_k \right|$$

so that we get

$$\bar{H} r^2 - \|H\|_3 r^3 \leq h(r) \leq \bar{H} r^2 + \|H\|_3 r^3 \, . \tag{43}$$

In particular we obtain that $\mathcal{D}_1(H,0) \geq \bar{H}$. Moreover from the left hand side of (43) we have

$$\frac{h(r)}{r^2+\xi^2} \geq \frac{\bar{H} r^2 - \|H\|_3 r^3}{r^2+\xi^2}.$$

Taking $r = \frac{\mathcal{D}_1(H,0)}{3\|H\|_3}$ we get

$$\mathcal{D}_1(H,\xi) \geq \frac{\mathcal{D}_1(H,0)^2}{27} \frac{3\bar{H} - \mathcal{D}_1(H,0)}{\mathcal{D}_1(H,0)^2 + \|H\|_3^2 \xi^2} \, . \tag{44}$$

On the other hand, the right hand side of (43) gives

$$h(r) - \mathcal{D}_1(H,0) r^2 \leq (\bar{H} - \mathcal{D}_1(H,0)) r^2 + \|H\|_3 r^3 \tag{45}$$

Let r_* be the smallest r such that

$$\mathcal{D}_1(H,0)) = \frac{h(r_*)}{r_*^2} \, .$$

From (45) we see that either $r_* = 0$ or $r_* \geq (\bar{H} - \mathcal{D}_1(H,0))/\|H\|_3$. Observe that, if $r_* = 0$, then $\mathcal{D}_1(H,0) = \bar{H}$ and (44) gives the thesis. Thus we can assume that $r_* \geq (\bar{H} - \mathcal{D}_1(H,0))/\|H\|_3$. Observing that $r^2/(r^2+\xi^2)$ is increasing in r, it follows that

$$\mathcal{D}_1(H,\xi) \geq \frac{h(r_*)}{r_*^2+\xi^2} = \mathcal{D}_1(H,0)\frac{r_*^2}{r_*^2+\xi^2} \geq \frac{\mathcal{D}_1(H,0)(\bar{H}-\mathcal{D}_1(H,0))^2}{(\bar{H}-\mathcal{D}_1(H,0))^2+\|H\|_3^2\xi^2} \geq \frac{\mathcal{D}_1(H,0)^2(\mathcal{D}_1(H,0)-2\bar{H})}{\mathcal{D}_1(H,0)^2+\|H\|_3^2\xi^2}$$

Thus we get

$$\mathcal{D}_1(H,\xi) \geq \frac{\mathcal{D}_1(H,0)^2}{\mathcal{D}_1(H,0)^2+\|H\|_3^2\xi^2}\max\left\{\frac{(3\bar{H}-\mathcal{D}_1(H,0))}{27}, \mathcal{D}_1(H,0)-2\bar{H}\right\}. \tag{46}$$

Observe finally that if $\alpha_1, \alpha_2 > 0$ and $\alpha_1+\alpha_2 = 1$ then $\max\{a_1,a_2\} \geq \alpha_1 a_1 + \alpha_2 a_2$. Taking $\alpha_1 = 18/19$ and $\alpha_2 = 1/19$ in (46) gives

$$\mathcal{D}_1(H,\xi) \geq \frac{1}{57}\frac{\mathcal{D}_1(H,0)^3}{\mathcal{D}_1(H,0)^2+\|H\|_3^2\xi^2}.$$

□

To conclude the proof of Lemma 3.1 we observe that $h(r) \leq \|H\|_3 r^2$ so that $\mathcal{D}_1(H,\xi) \leq \|H\|_3$. In particular, $\mathcal{D}_1(H,0) \leq \|H\|_3$. Combining with Claim 3 it follows that

$$\frac{\mathcal{D}_1(H,0)}{1+\xi^2/3} \leq 4\|H\|_3^{\frac{2}{3}}\mathcal{D}_1(H,\xi)^{\frac{1}{3}}$$

so that

$$\frac{\mathcal{D}_1(H,0)}{1+T\xi^2} \leq 4\max\left\{1,(3T)^{-1}\right\}\|H\|_3^{\frac{2}{3}}\mathcal{D}_1(H,\xi)^{\frac{1}{3}}.$$

Finally we obtain the thesis observing that we can write $\mathcal{D}_1(H,\xi) \leq \|H\|_3^{\frac{2}{3}}\mathcal{D}_1(H,\xi)^{\frac{1}{3}}$.

5.2 Proof of Lemma 3.2 and 3.3

We begin proving Lemma 3.3. We observe that if $\nabla_{\vec{\eta}}H(0,\vec{\xi}) \neq 0$ then, there exists $\vec{\eta}_*$ with $\|\vec{\eta}_*\| = 1$ such that, for ρ small enough we have

$$|H(\rho\vec{\eta}_*,\vec{\xi})G(\rho\vec{\eta}_*)| \geq \frac{1}{2}|\rho|\|\nabla_{\vec{\eta}}H(0,\vec{\xi})\|.$$

For N large enough we can chose $\underline{\eta} = \frac{1}{\sqrt{N}}(\vec{\eta}_*,\ldots,\vec{\eta}_*)$ and obtain

$$\mathcal{D}_N(H,\vec{\xi}) \geq \frac{\sqrt{N}}{2}\frac{\|\nabla_{\vec{\eta}}H(0,\vec{\xi})\|}{1+\|\vec{\xi}\|^2}.$$

This shows that the second term in (33) is not an artifact of our proof.

As for the proof of Lemma 3.1 we can assume that $g(\vec{\eta}) = G(\vec{\eta}) = (1 + T\|\vec{\eta}\|^2)^{-1}$. From (28) and (29), we can write

$$H(\vec{\eta}, \vec{\xi}) = G(\vec{\eta})(\vec{\eta} \cdot \nabla_{\vec{\eta}} H(0, \vec{\xi})) + L(\vec{\eta}, \vec{\xi}). \tag{47}$$

Considering that $G(0) = 1$ and $\nabla G(0) = 0$, we get $L(0, \vec{\xi}) = 0$, $\nabla_{\vec{\eta}} L(0, \vec{\xi}) = 0$ and

$$\|L\|_3 \leq \|H\|_{3,0} + G_3 \|\nabla H\|_\infty \leq 2G_3 \|H\|_{3,0}$$

where

$$G_3 = \max\left\{T^{-1}, T^2\right\} \sum_{i=0}^{3} \sup_x \left| \frac{d^i}{dx^i} \frac{x}{1+x^2} \right|.$$

It follows that

$$\begin{aligned}
\mathcal{D}_N(H, \vec{\xi}) &= \sup_{\underline{\eta} \neq 0} \left| \frac{\sum_{i=1}^N \left(\vec{\eta}_i \cdot \nabla_{\vec{\eta}} H(0, \vec{\xi}) G_N(\underline{\eta}) + L(\vec{\eta}_i, \vec{\xi}) G_{N-1}(\underline{\eta}^i) \right)}{\|\underline{\eta}|^2 + \|\vec{\xi}\|^2} \right| \\
&\leq \sup_{\underline{\eta} \neq 0} \left| \frac{\sum_{i=1}^N \vec{\eta}_i \cdot \nabla_{\vec{\eta}} H(0, \vec{\xi}) G_N(\underline{\eta})}{\|\underline{\eta}\|^2 + \|\vec{\xi}\|^2} \right| + \sup_{\underline{\eta} \neq 0} \left| \frac{\sum_{i=1}^N L(\vec{\eta}_i, \vec{\xi}) G_{N-1}(\underline{\eta}^i)}{\|\underline{\eta}\|^2 + \|\vec{\xi}\|^2} \right| \\
&= \sup_{\underline{\eta} \neq 0} \left| \frac{\sum_{i=1}^N \vec{\eta}_i \cdot \nabla_{\vec{\eta}} H(0, \vec{\xi}) G_N(\underline{\eta})}{\|\underline{\eta}\|^2 + \|\vec{\xi}\|^2} \right| + \mathcal{D}_N(L, \vec{\xi}).
\end{aligned} \tag{48}$$

We can use the Cauchy-Schwartz inequality to obtain

$$\sup_{\underline{\eta} \neq 0} \left| \frac{\sum_{i=1}^N \vec{\eta}_i \cdot \nabla_{\vec{\eta}} H(0, \xi) G_N(\underline{\eta})}{\|\underline{\eta}\|^2 + \|\vec{\xi}\|^2} \right| \leq \sqrt{N} \left| \nabla_{\vec{\eta}} H(0, \xi) \right| \sup_{\vec{x} \neq 0} \frac{|\vec{x}| G(\vec{x})}{\|\vec{x}\|^2 + \|\vec{\xi}\|^2} \tag{49}$$

Observe that we can apply Lemma 3.1 to $\mathcal{D}_N(L, \vec{\xi})$, notwithstanding the dependence of L on $\vec{\xi}$. Using also (47), we obtain

$$\begin{aligned}
\mathcal{D}_N(L, \vec{\xi}) &\leq C \|L\|_3^{\frac{2}{3}} \mathcal{D}_1(L, \vec{\xi})^{\frac{1}{3}} \\
&\leq C \|L\|_3^{\frac{2}{3}} \left(\mathcal{D}_1(H, \vec{\xi})^{\frac{1}{3}} + \mathcal{D}_1(\vec{\eta} \cdot \nabla_{\vec{\eta}} H(0, \vec{\xi}) G(\vec{\eta}), \vec{\xi})^{\frac{1}{3}} \right) \\
&\leq C G_3^{\frac{2}{3}} \|H\|_{3,0}^{\frac{2}{3}} \left(\mathcal{D}_1(H, \vec{\xi})^{\frac{1}{3}} + |\nabla_{\vec{\eta}} H(0, \vec{\xi})|^{\frac{1}{3}} \left(\sup_{\vec{x} \neq 0} \frac{|\vec{x}| G(\vec{x})}{\|\vec{x}\|^2 + \|\vec{\xi}\|^2} \right)^{\frac{1}{3}} \right)
\end{aligned} \tag{50}$$

Since

$$\sup_{\vec{x} \neq 0} \frac{\|\vec{x}\| G(\vec{x})}{\|\vec{x}\|^2 + \|\vec{\xi}\|^2} \leq \frac{1}{\|\vec{\xi}\| \sqrt{1 + \|\vec{\xi}\|^2}},$$

we can combine this with (48), (50), and (49) to obtain

$$\mathcal{D}_N(H,\vec{\xi}) \le \sqrt{N}\frac{\|\nabla_\eta H(0,\vec{\xi})\|}{\|\vec{\xi}\|\sqrt{1+\|\vec{\xi}\|^2}} + CG_3^{\frac{2}{3}}\|H\|_{3,0}^{\frac{2}{3}}\left(\mathcal{D}_1(H,\vec{\xi})^{\frac{1}{3}} + \left(\frac{\|\nabla_{\vec{\eta}} H(0,\vec{\xi})\|}{\|\vec{\xi}\|\sqrt{1+\|\vec{\xi}\|^2}}\right)^{\frac{1}{3}}\right). \tag{51}$$

Proof of Lemma 3.2. Since $H(0,\vec{\xi})=0$ for every $\vec{\xi}$ we see that $\nabla_{\vec{\xi}}H(0,0)=0$ so that

$$\alpha(0) < \infty \tag{52}$$

and $\mathcal{D}(H,0)\le\infty$. Analogously to (47), we can write $H(\vec{\eta},\xi)=\vec{\eta}\cdot\nabla_{\vec{\eta}}H(0,\xi)+\widetilde{L}(\vec{\eta},\xi)$ so that we get

$$\begin{aligned}|H(\vec{\eta},\vec{\xi})| &\ge |\vec{\eta}\cdot\nabla_{\vec{\eta}}H(0,\vec{\xi})| - |\widetilde{L}(\vec{\eta},\vec{\xi})|\\ &\ge |\vec{\eta}\cdot\nabla_{\vec{\eta}}H(0,\vec{\xi})| - \|\vec{\eta}\|^2\,\|H\|_{3,0}\\ &= \|\vec{\xi}\|\frac{|\vec{\eta}\cdot\nabla_{\vec{\eta}}H(0,\vec{\xi})|}{\|\vec{\xi}\|} - \|\vec{\eta}\|^2\,\|H\|_{3,0}\,.\end{aligned}$$

Thus

$$\begin{aligned}\mathcal{D}_1(H,\vec{\xi}) &= \sup_{\vec{\eta}\neq 0}\frac{|H(\vec{\eta},\vec{\xi})|}{\|\vec{\eta}\|^2+\|\vec{\xi}\|^2}\\ &\ge \sup_{\vec{\eta}\neq 0}\frac{\|\vec{\xi}\|\frac{|\vec{\eta}\cdot\nabla_{\vec{\eta}}H(0,\vec{\xi})|}{\|\vec{\xi}\|} - \|\vec{\eta}\|^2\,\|H\|_{3,0}}{\|\vec{\eta}\|^2+\|\vec{\xi}\|^2}\end{aligned}$$

Since the supremum over $\vec{\eta}\neq 0$ is certainly larger than the value at any particular $\vec{\eta}$, setting $\vec{\eta}_0$ to be the vector parallel to $\nabla_{\vec{\eta}}H(0,\xi)$ with norm $\|\vec{\eta}_0\| = \frac{\|\vec{\xi}\|\alpha(\vec{\xi})}{2\|H\|_{3,0}}$ yields

$$\begin{aligned}\mathcal{D}_1(H,\vec{\xi}) &\ge \frac{\|\vec{\xi}\|^2\frac{\alpha(\vec{\xi})^2}{4\|H\|_{3,0}}}{\|\vec{\xi}\|^2\frac{\alpha(\vec{\xi})^2}{4\|H\|_{3,0}^2}+\|\vec{\xi}\|^2}\\ &= \frac{\alpha(\xi)^2\,\|H\|_{3,0}}{\alpha(\xi)^2+4\,\|H\|_{3,0}^2}\end{aligned}$$

It follows that

$$\alpha(\xi)^2 \le \frac{4\|H\|_{3,0}^2\mathcal{D}_1(H,\xi)}{\|H\|_{3,0}-\mathcal{D}_1(H,\xi)} \le 8\|H\|_{3,0}\mathcal{D}_1(H,\xi)$$

where we used that $\|H\|_{3,0}\ge 2\mathcal{D}_1(H,\xi)$. Together with (52), this gives the thesis.

□

Combining (51) and Lemma 3.2, we get

$$\mathcal{D}_N(H,\xi) \le C\sqrt{N}\|H\|_{3,0}^{\frac{5}{6}}\min\left\{D_1(H,\xi)^{\frac{1}{6}}, |H|_{1,1}^{\frac{1}{6}}\right\} + C\|H\|_3^{\frac{2}{3}}D_1(H,\xi)^{\frac{1}{3}}.$$

6 Proof of Theorem 2.1

In this section we prove Theorem 2.1. The proof is based on time interpolation between $\mathscr{F}_t$ and $\widetilde{\mathscr{F}_t}$ based on Duhamel's formula and on Lemma 3.3. But first we start reminding the basic definition from the introduction.

6.1 The setting

We consider the evolution generated by the operator

$$\mathcal{L} = \mathcal{L}_S + \mathcal{L}_{I_+} + \mathcal{L}_{I_-} + \mathcal{L}_{R_+} + \mathcal{L}_{R_-} \tag{53}$$

acting on the probability distributions on $\mathbb{R}^{3(M+2N)}$, where

$$\begin{gathered} \mathcal{L}_S := \frac{1}{M-1} \sum_{1 \le i < j \le M} (R^S_{i,j} - \mathrm{Id}) \\ \mathcal{L}_{R_\sigma} := \frac{\lambda}{N-1} \sum_{1 \le i < j \le N} (R^{R_\sigma}_{i,j} - \mathrm{Id}) \qquad \mathcal{L}_{I_\sigma} := \frac{\mu}{N} \sum_{i=1}^{N} \sum_{j=1}^{M} (R^{I_\sigma}_{i,j} - \mathrm{Id}) \end{gathered} \tag{54}$$

where $\sigma = \pm$ and $R^\alpha_{i,j}$, $\alpha \in \{S, R_+, R_-, I_+, I_-\}$ is the collision operator (1) acting on particles in the system $\alpha = S$, in the reservoirs $\alpha = R_+$, and R_- or one in each as in the interaction terms with $\alpha = I_+$, and I_-.

Moreover we assume that the evolution starts from the initial state

$$\mathscr{F}_0(\underline{u}, \underline{v}, \underline{w}) = \Gamma_+(\underline{u}) f_0(\underline{v}) \Gamma_-(\underline{w})$$

where Γ_σ is the Maxwellian distribution on N particles at temperature T_σ. We thus set

$$\mathscr{F}_t = e^{\mathcal{L}t} \mathscr{F}_0\,.$$

We want to compare $\mathscr{F}_t$ with the distribution

$$\widetilde{\mathscr{F}_t} = \Gamma_+(\underline{u}) \tilde{f}_t(\underline{v}) \Gamma_-(\underline{w}) \tag{55}$$

where

$$\tilde{f}_t = e^{\widetilde{\mathcal{L}}t} f_0$$

and

$$\widetilde{\mathcal{L}} = \mathcal{L}_S + \mathcal{L}_{B^+} + \mathcal{L}_{B^-} \tag{56}$$

acts on $\mathbb{R}^{3M}$, with

$$\mathcal{L}_{B^\sigma} := \mu \sum_{i=1}^{M} (B^\sigma_i - I) \tag{57}$$

where B_i^σ is the operator (6) describing the interaction between particle i and the thermostat at temperature T_σ. Although $\widetilde{\mathcal{L}}$ act on distribution on $\mathbb{R}^{3M}$, we will also consider its extension to the distributions on $\mathbb{R}^{3(M+2N)}$ that acts as the identity on the extra variables. With a slight abuse of notation we will use the same symbol for such an extension thus writing

$$\widetilde{\mathscr{F}_t} = e^{\widetilde{\mathcal{L}}t}\mathscr{F}_0 = \Gamma_+(\underline{u})e^{\widetilde{\mathcal{L}}t}f_0(\underline{v})\Gamma_-(\underline{w})$$

6.2 Estimating $d_2(\widetilde{\mathscr{F}_t}, \mathscr{F}_t)$

Our first step in estimating the difference between $\mathscr{F}_t$ and $\widetilde{\mathscr{F}_t}$ is based on the Duhamel formula

$$e^{\mathcal{L}t} - e^{\widetilde{\mathcal{L}}t} = \int_0^t e^{\mathcal{L}(t-s)}(\mathcal{L} - \widetilde{\mathcal{L}})e^{\widetilde{\mathcal{L}}s}ds.$$

that yields

$$d_2(e^{\mathcal{L}t}\mathscr{F}_0, e^{\widetilde{\mathcal{L}}t}\mathscr{F}_0) \leq \int_0^t d_2(e^{\mathcal{L}(t-s)}\mathcal{L}e^{\widetilde{\mathcal{L}}s}\mathscr{F}_0, e^{\mathcal{L}(t-s)}\widetilde{\mathcal{L}}e^{\widetilde{\mathcal{L}}s}\mathscr{F}_0)ds \leqq \int_0^t d_2(\mathcal{L}\widetilde{\mathscr{F}_s}, \widetilde{\mathcal{L}}\widetilde{\mathscr{F}_s})ds. \tag{58}$$

where $\widetilde{\mathscr{F}_s} = e^{\widetilde{\mathcal{L}}s}\mathscr{F}_0$ and we have used Corollary A.2.

From (55), using the definitions (53) and (56) we get

$$d_2(e^{\mathcal{L}t}\mathscr{F}_0, e^{\widetilde{\mathcal{L}}t}\mathscr{F}_0) \leq \int_0^t d_2(\mathcal{L}_{B_+}\widetilde{\mathscr{F}_s}, \mathcal{L}_{I_+}\widetilde{\mathscr{F}_s})ds + \int_0^t d_2(\mathcal{L}_{B_-}\widetilde{\mathscr{F}_s}, \mathcal{L}_{I_-}\widetilde{\mathscr{F}_s})ds. \tag{59}$$

We will control the first integral on the right-hand side. The second term can be controlled in the same way.

The first step to estimating this expression is understanding the Fourier transform of each term. From now on, we will let $\underline{\xi}, \underline{\eta}$, and $\underline{\zeta}$ denote the dual Fourier variables of $\underline{v}, \underline{u}$, and $\underline{w}$ respectively. We have

$$d_2(\mathcal{L}_{B_+}\widetilde{\mathscr{F}_s}, \mathcal{L}_{I_+}\widetilde{\mathscr{F}_s}) = \sup \frac{\mu}{N}\sum_{i=1}^{N}\sum_{j=1}^{M}\frac{\widehat{R_{i,j}^{I_+}\mathscr{F}_s} - \widehat{B_j^+\mathscr{F}_s}}{|\underline{\xi}|^2 + |\underline{\eta}|^2 + |\underline{\zeta}|^2} \tag{60}$$

Observe that if $f(\vec{v}_1, \vec{v}_2)$ is a distribution on $\mathbb{R}^6$ then

$$\widehat{R[f]}(\vec{\xi}_1, \vec{\xi}_2) = \int_{\mathbb{S}^2} \hat{f}(\vec{\xi}_1^*(\omega), \vec{\xi}_2^*(\omega))d\omega = R[\hat{f}](\vec{\xi}_1, \vec{\xi}_2)$$

while if $g(\vec{v})$ is a distribution on $\mathbb{R}^3$ then

$$\widehat{B[g]}(\vec{\xi}) = \int_{\mathbb{S}^2} \hat{f}(\vec{\xi}^*(\omega))\widehat{\Gamma}_T(0^*(\omega))d\omega = R[\hat{g}\widehat{\Gamma}_T](\vec{\xi}, 0)$$

where $0^*(\omega) = (\xi \cdot \omega)\omega$. Moreover for a general distribution of the form $\mathscr{F}(\vec{u}, \vec{v}, \vec{w}) = \Gamma_-^N(\underline{u}) f(\underline{v}) \Gamma_-^N(\underline{w})$, we have

$$\widehat{R_{i,j}^{I_+}\mathscr{F}}(\underline{\eta}, \underline{\xi}, \underline{\zeta}) = R_{i,j}^{I_+}\left[\hat{f}\widehat{\Gamma}_+^1\right](\vec{\eta}_i, \underline{\xi})\widehat{\Gamma}_+^{N-1}(\underline{\eta}^i)\widehat{\Gamma}_-^N(\underline{\zeta}) \tag{61}$$

$$\widehat{B_j^+\mathscr{F}}(\underline{\eta}, \underline{\xi}, \underline{\zeta}) = \widehat{R_{i,j}^{I_+}F}(\underline{\eta}, \underline{\xi}, \underline{\zeta})\big|_{\vec{\eta}_i=0}\widehat{\Gamma}_+^1(\vec{\eta}_i) = R_{i,j}^{I_+}\left[\hat{f}\widehat{\Gamma}_+^1\right](0, \underline{\xi})\widehat{\Gamma}_+^1(\vec{\eta}_i)\widehat{\Gamma}_+^{N-1}(\underline{\eta}^i)\widehat{\Gamma}_-^N(\underline{\zeta}) \tag{62}$$

For ease of notation, we will not indicate the dependence on $\underline{\zeta}$ if not necessary. For $1 \le i \le N$, define

$$\widehat{G}_s(\vec{\eta}_i, \underline{\xi}) = \sum_{j=1}^{M}\left(R_{i,j}^{I_+}[\hat{f}_s\widehat{\Gamma}_+](\vec{\eta}_i, \underline{\xi}) - R_{i,j}^{I_+}[\hat{f}_s\widehat{\Gamma}_+](0, \underline{\xi})\widehat{\Gamma}_+(\vec{\eta}_i)\right). \tag{63}$$

Now (60) becomes

$$d_2(\mathcal{L}_{B_+}\widetilde{\mathscr{F}}_s, \mathcal{L}_{I_+}\widetilde{\mathscr{F}}_s) = \frac{\mu}{N}\sup_{\underline{\xi},\underline{\eta},\underline{\zeta}\neq 0}\left|\frac{\sum_{i=1}^N \widehat{G}_s(\vec{\eta}_i, \underline{\xi})\widehat{\Gamma}_+^{N-1}(\underline{\eta}^i)\widehat{\Gamma}_+^N(\underline{\zeta})}{|\underline{\xi}|^2 + |\underline{\eta}|^2 + |\underline{\zeta}|^2}\right| \le \frac{\mu}{N}\sup_{\underline{\xi},\underline{\eta}\neq 0}\left|\frac{\sum_{i=1}^N \widehat{G}_s(\vec{\eta}_i, \underline{\xi})\widehat{\Gamma}_+^{N-1}(\underline{\eta}^i)}{|\underline{\xi}|^2 + |\underline{\eta}|^2}\right| \tag{64}$$

With the aim of applying Lemma 3.3 to the above expression we observe that

1. For any $\underline{\xi} \in \mathbb{R}^{3M}$, $\widehat{G}_s$ has a zero of order 1 at $(0, \underline{\xi})$. That is, $\widehat{G}_s(0, \vec{\xi}) = 0$;
2. $\widehat{G}_s$ has a zero of order 2 at $(0, 0)$. That is, $\nabla_{\vec{\eta}_i}\widehat{G}_s(0, 0) = 0$ and clearly $\nabla_{\underline{\xi}}\widehat{G}_s(0, 0) = 0$;
3. In Appendix B we show that there exists a constant C such that for every s we have $\|\widehat{G}_s\|_{3,1} \le MCE_4(f_0)$, see Lemma B.4;
4. It is easy to see that $\widehat{\Gamma}_+(\vec{\eta}) \le \frac{1}{1+T_+\|\vec{\eta}\|^2}$.

Thus, we see that Lemma 3.3 is applicable with $H = \widehat{G}_s$, and we obtain

$$\begin{aligned} d_2(\mathcal{L}_{B_+}\widetilde{\mathscr{F}}_s, \mathcal{L}_{I_+}\widetilde{\mathscr{F}}_s) &\le \frac{\mu}{\sqrt{N}}CM^{\frac{5}{6}}E_4(f_0)^{\frac{5}{6}}d_2(G_s, 0)^{\frac{1}{6}} \\ &\le \frac{\mu M}{\sqrt{N}}CE_4(f_0)^{\frac{5}{6}}d_2\left(R_{i1}^{I_+}[\tilde{f}_s\Gamma_+^1], B_i^+[\tilde{f}_s\Gamma_+^1]\right)^{\frac{1}{6}}. \end{aligned} \tag{65}$$

where we used the monotonicity of Γ_+ and the fact that $\tilde{f}_s$ is permutation invariant so that the M term in (64) gives the same contribution.

We can now apply Lemma A.1 and Corollary A.2 to obtain

$$\begin{aligned} d_2\left(R[\tilde{f}_s\Gamma_+], B^+[\tilde{f}_s]\Gamma_+\right) &\le d_2\left(R[(\tilde{f}_s - \tilde{f}_\infty)\Gamma_+], B^+[\tilde{f}_s - \tilde{f}_\infty]\Gamma_+\right) + d_2\left(R[\tilde{f}_\infty\Gamma_+], B^+[\tilde{f}_\infty]\Gamma_+\right) \\ &\le 2e^{-\frac{2\mu}{3}s}d_2\left(f_0, \tilde{f}_\infty\right) + 2d_2\left(\tilde{f}_\infty, \Gamma_+^M\right). \end{aligned} \tag{66}$$

so that it follows that

$$\begin{aligned}\int_0^t & d_2(\mathcal{L}_{B_+}\widetilde{\mathscr{F}_s},\mathcal{L}_{I_+}\widetilde{\mathscr{F}_s})ds \\ &\le 2\frac{M\mu C}{\sqrt{N}}E_4(f_0)^{\frac{5}{6}}\int_0^t\left(e^{-\frac{2\mu}{3}s}d_2(f_0,\tilde{f}_\infty)+d_2(\tilde{f}_\infty,\Gamma_+^M)\right)^{\frac{1}{6}}\,ds \\ &\le 18\frac{MC}{\sqrt{N}}E_4(f_0)^{\frac{5}{6}}(1-e^{-\frac{\mu}{9}t})d_2(f_0,\tilde{f}_\infty)^{\frac{1}{6}}+2\frac{M\mu C}{\sqrt{N}}E_4(f_0)^{\frac{5}{6}}d_2(\tilde{f}_\infty,\Gamma_+^M)^{\frac{1}{6}}t\end{aligned}$$

We clearly obtain an analogous estimate for the second term in (59) so we conclude that

$$\begin{aligned}d_2(e^{\mathcal{L}t}\mathscr{F}_0,e^{\tilde{\mathcal{L}}t}\mathscr{F}_0)\le{}&\frac{MC}{\sqrt{N}}E_4(f_0)^{\frac{5}{6}}(1-e^{-\frac{\mu}{9}t})d_2(f_0,\tilde{f}_\infty)^{\frac{1}{6}} \\ &+\frac{M\mu C}{\sqrt{N}}E_4(f_0)^{\frac{5}{6}}\left(d_2(\tilde{f}_\infty,\Gamma_+^M)^{\frac{1}{6}}+d_2(\tilde{f}_\infty,\Gamma_-^M)^{\frac{1}{6}}\right)t\end{aligned}\tag{67}$$

We observe that, if $T_+=T_-=T$ then we are essentially in the 1-reservoir situation. In this case $\Gamma_+=\Gamma_-=\Gamma_T$ so that $\tilde{f}_\infty=\Gamma_T^M$ and (67) simplifies to

$$d_2(e^{\mathcal{L}t}\mathscr{F}_0,e^{\tilde{\mathcal{L}}t}\mathscr{F}_0)\le\frac{MC}{\sqrt{N}}E_4(f_0)^{\frac{5}{6}}(1-e^{-\frac{\mu}{9}t})d_2(f_0,\Gamma_T^M)^{\frac{1}{6}}\tag{68}$$

which is the same as the estimate we would obtain by doing the computation directly starting from a 1-reservoir set-up. This proves Corollary 2.2.

In Appendix A we show that

$$d_2(\tilde{f}_\infty,\Gamma_\sigma^M)\le\frac{T_+-T_-}{2}\tag{69}$$

so that the second term in (67) essentially captures the difference between the temperature of the two Maxwellian thermostats. This concludes the proof of Theorem 2.1.

Delarations

Conflict of Interest

The authors have no conflicts of interest to disclose.

Author Contributions

Federico Bonetto: Formal analysis (equal); Writing – original draft (equal); Writing – review & editing (equal). Michael Loss: Formal analysis (equal); Writing – original draft (equal); Writing – review & editing (equal). Matthew Powell: Formal analysis (equal); Writing – original draft (equal); Writing – review & editing (equal).

Data Availability

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

References

[1] Federico Bonetto, Eric A. Carlen, Lukas Hauger, and Michael Loss. Approach to equilibrium for the kac model. In Eric Carlen, Patrícia Gonçalves, and Ana Jacinta Soares, editors, *From Particle Systems to Partial Differential Equations - PSPDE X 2022*, Springer Proceedings in Mathematics and Statistic, pages 187–211. Springer, 2024.

[2] Federico Bonetto, Alissa Geisinger, Michael Loss, and Tobias Ried. Entropy decay for the Kac evolution. *Comm. Math. Phys.*, 363(3):847–875, 2018.

[3] Federico Bonetto, Michael Loss, Hagop Tossounian, and Ranjini Vaidyanathan. Uniform approximation of a Maxwellian thermostat by finite reservoirs. *Comm. Math. Phys.*, 351(1):311–339, 2017.

[4] Federico Bonetto, Michael Loss, and Ranjini Vaidyanathan. The Kac model coupled to a thermostat. *J. Stat. Phys.*, 156(4):647–667, 2014.

[5] Eric A. Carlen, Maria C. Carvalho, and Michael Loss. Determination of the spectral gap for Kac's master equation and related stochastic evolution. *Acta Math.*, 191(1):1–54, 2003.

[6] Eric A. Carlen, Maria C. Carvalho, and Michael Loss. Spectral gaps for reversible markov processes with chaotic invariant measures: The kac process with hard sphere collisions in three dimensions. *Ann. Probab.*, 48:2807–2844, 2020.

[7] Eric Carlen, Maria C. Carvalho, and Michael Loss. Many-body aspects of approach to equilibrium. In *Journées "Équations aux Dérivées Partielles" (La Chapelle sur Erdre, 2000)*, pages Exp. No. XI, 12. Univ. Nantes, Nantes, 2000.

[8] Eric A. Carlen, Joel L. Lebowitz, and Clément Mouhot. Exponential approach to, and properties of, a non-equilibrium steady state in a dilute gas. *Brazilian Journal of Probability and Statistics*, 29(2):372–386, 2015.

[9] Eric A. Carlen, Raffaele Esposito, Joel L. Lebowitz, Rossana Marra, and Clément Mouhot. Approach to the steady state in kinetic models with thermal reservoirs at different temperatures. *Journ, Stat. Phys.*, 172:522–543, 2018.

[10] Amit Einav. On Villani's conjecture concerning entropy production for the Kac master equation. *Kinet. Relat. Models*, 4(2):479–497, 2011.

[11] Josephine Evans. Non-equilibrium steady states in kac's model coupled to a thermostat. *Journ, Stat. Phys.*, 164:1103–1121, 2016.

[12] Irene Gabetta, Giuseppe Toscani, and Bernt Wennberg. Metrics for probability distributions and the trend to equilibrium for solutions of the boltzmann equation. *Journ, Stat. Phys.*, 81:901–934, 1995.

[13] Elise Janvresse. Spectral gap for Kac's model of Boltzmann equation. *Ann. Probab.*, 29(1):288–304, 2001.

[14] Mark Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III*, pages 171–197, Berkeley and Los Angeles, 1956. University of California Press.

[15] Mark Kac. *Probability and related topics in physical sciences*, volume 1957 of *With special lectures by G. E. Uhlenbeck, A. R. Hibbs, and B. van der Pol. Lectures in Applied Mathematics. Proceedings of the Summer Seminar, Boulder, Colo.* Interscience Publishers, London-New York, 1959.

[16] David K. Maslen. The eigenvalues of Kac's master equation. *Math. Z.*, 243(2):291–331, 2003.

[17] Stéphane Mischler and Clément Mouhot. Kac's program in kinetic theory. *Invent. Math.*, 193(1):1–147, 2013.

[18] Alain Solt Sznitman. Équations de type de Boltzmann, spatialement homogènes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 66:559–592, 1984.

[19] Hagop Tossounian. Equilibration in the Kac model using the GTW metric d2. *Journ, Stat. Phys.*, 169:168–186, 2017.

[20] Cédric Villani. Cercignani's conjecture is sometimes true and always almost true. *Comm. Math. Phys.*, 234(3):455–490, 2003.

A Basic Properties of the evolutions

In this appendix we deduce the existence of a steady state for both evolutions and derive (69). In the following, we will restrict our attention to the metric space (X, d_2) where $X = \{f \in L^1 : d_2(f, \Gamma) < \infty\}$.

We start by studying the behavior of the GTW d_2 metric under the action of R and B defined in (1) and (6) respectively.

Lemma A.1. *Suppose $f(\vec{v}_1, \vec{v}_2)$ and $g(\vec{v}_1, \vec{v}_2)$ are distributions on $\mathbb{R}^6$ with zero first moment and finite second moment. Then*

$$d_2(R[f], R[g]) \leq d_2(f, g)$$

while if $f(\vec{v})$ and $g(\vec{v})$ are distribution on $\mathbb{R}^3$ with zero first moment and finite second moment then

$$d_2(B[f], B[g]) \leq \frac{2}{3} d_2(f, g)$$

Proof. Observe that

$$d_2\left(R[f],R[g]\right) = \sup_{\vec{\xi}_1,\vec{\xi}_2}\left|\frac{\int_{\mathbb{S}^2}(\hat{f}(\vec{\xi}_1^*(\omega)\,\vec{\xi}_2^*(\omega)) - \hat{g}(\vec{\xi}_1^*(\omega),\vec{\xi}_2^*(\omega)))\ d\omega}{\|\vec{\xi}_1\|^2+\|\vec{\xi}_2\|^2}\right| \tag{70}$$

$$\leq \sup_{\vec{\xi}_1,\vec{\xi}_2}\int_{\mathbb{S}^2}\frac{\left|\hat{f}(\vec{\xi}_1^*(\omega),\vec{\xi}_2^*(\omega)) - \hat{g}(\vec{\xi}_1^*(\omega),\vec{\xi}_2^*(\omega))\right|}{\|\vec{\xi}_1^*(\omega)\|^2+\|\vec{\xi}_2^*(\omega)\|^2}d\omega \leq d_2\left(f,g\right). \tag{71}$$

Similarly we get

$$\begin{aligned} d_2(B[f],B[g]) =& \sup_{\vec{\xi}}\frac{1}{\|\vec{\xi}\|^2}\left|\int_{\mathbb{S}^2}\left(\hat{f}(\vec{\xi}-(\omega\cdot\vec{\xi})\omega)-\hat{g}(\vec{\xi}-(\omega\cdot\vec{\xi})\omega)\right)\hat{\Gamma}_T((\omega\cdot\vec{\xi})\omega)d\omega\right| \\ &\leq d_2(f,g)\sup_{\vec{\xi}}\left(1-\int\frac{|(\vec{\xi}\cdot\omega)|^2}{\|\vec{\xi}\|^2}d\omega\right) = \frac{2}{3}d_2(f,g) \end{aligned} \tag{72}$$

where we have used if $\sigma\in\mathbb{S}^2$ then

$$n_2 := \int_{\mathbb{S}^2}(\sigma\cdot\omega)^2 d\omega = \frac{1}{2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\sin^2\theta\cos\theta d\theta = \frac{1}{3} \tag{73}$$

□

Observe now that we can write

$$\mathcal{L}_S = \frac{1}{M-1}\sum_{1\leq i<j\leq M}(R_{i,j}^S - \mathrm{Id}) =: Q_S - \Lambda_S\mathrm{Id}$$

and similar decomposition for $\mathcal{L}_{R_\sigma} =: Q_{R_\sigma} - \Lambda_R\mathrm{Id}$, $\mathcal{L}_{I_\sigma} =: Q_{I_\sigma} - \Lambda_I\mathrm{Id}$, and $\mathcal{L}_{B_\sigma} =: Q_{B_\sigma} - \Lambda_B\mathrm{Id}$. We also set

$$\begin{aligned} Q &= Q_S + Q_{I_+} + Q_{I_-} + Q_{R_+} + Q_{R_-} & \Lambda &= \Lambda_S + 2\Lambda_R + 2\Lambda_I \\ \widetilde{Q} &= Q_S + Q_{B^+} + Q_{B^-} & \widetilde{\Lambda} &= \Lambda_S + 2\Lambda_B\,, \end{aligned}$$

so that we get

$$\mathcal{L} = Q - \Lambda\ \mathrm{Id} \qquad \widetilde{\mathcal{L}} = \widetilde{Q} - \widetilde{\Lambda}\ \mathrm{Id}\,.$$

It will also be useful to have the following corollary, which informally states that the Kac evolution is not expanding while the thermostated evolution is a contracting.

Corollary A.2. *Let $\mathscr{F}$ and $\mathscr{G}$ be two distributions on $\mathbb{R}^{3(M+2N)}$ with 0 mean and finite second moment, then*

$$d_2(e^{\mathcal{L}t}\mathscr{F}, e^{\mathcal{L}t}\mathscr{G}) \leq d_2(\mathscr{F},\mathscr{G}). \tag{74}$$

Let now f and g be two distributions on $\mathbb{R}^M$ with 0 mean and finite second moment, then

$$d_2(e^{\widetilde{\mathcal{L}}t}f, e^{\widetilde{\mathcal{L}}t}g) \leq e^{-\frac{2\mu}{3}t}d_2(f,g) \tag{75}$$

Proof. We begin by writing

$$e^{\mathcal{L}t} = e^{-\Lambda t} \sum_{n=0}^{\infty} \frac{t^n}{n!} Q^n, \tag{76}$$

From Lemma A.1, it is easy to see that

$$d_2(Q_\alpha \mathscr{F}, Q_\alpha \mathscr{G}) \leq \Lambda_\alpha d_2(\mathscr{F}, \mathscr{G}); \quad \alpha \in \{S, I_\sigma, R_\sigma\} \ ,$$

so that

$$d_2(Q\mathscr{F}, Q\mathscr{G}) \leq \Lambda d_2(\mathscr{F}, \mathscr{G})\,.$$

Thus

$$d_2(e^{\mathcal{L}t}\mathscr{F}, e^{\mathcal{L}t}\mathscr{G}) \leq e^{-\Lambda t} \sum_{n=0}^{\infty} \frac{t^n}{n!} d_2(Q^n\mathscr{F}, Q^n\mathscr{G}) \leq d_2(\mathscr{F}, \mathscr{G}) e^{-\Lambda t} \sum_{n=0}^{\infty} \frac{t^n}{n!} \Lambda^n$$

where we have used the convexity of d_2.

For (75) we observe that, reasoning like for (72) we get

$$d_2(Q_{B_\sigma} f, Q_{B_\sigma} g) \leq \mu d_2(f,g) \sup_{\underline{\xi} \neq 0} \left(M - \sum_{i=1}^{M} \int \frac{|(\vec{\xi}_i \cdot \omega)|^2}{|\underline{\xi}|^2} d\omega \right) = \left(\Lambda_B - \frac{\mu}{3} \right) d_2(f,g) \tag{77}$$

The second inequality now follows as before. □

An immediate consequence of Corollary A.2 is that the thermostated evolution has a unique steady-state which depends only on the temperatures of the thermostats.

In order to derive (69), it will be helpful to establish the following claim, which follows from a simple computation.

Lemma A.3. *The d_2 distance between two Maxwellian distribution is proportional to their temperature difference, that is*

$$d_2(\Gamma^1_{T_+}, \Gamma^1_{T_-}) = \frac{T_+ - T_-}{2}\,.$$

Proof. Observe that if $T_+ > T_-$, using that $1 - e^{-x} \leq x$, we get

$$d_2(\Gamma^1_{T_+}, \Gamma^1_{T_-}) = \sup_{x>0} \frac{e^{-T_- x/2} - e^{-T_+ x/2}}{x} = \sup_{x>0} \frac{e^{-T_- x/2}(1 - e^{-(T_+ - T_-)x/2})}{x} \leq \sup_{x>0} \frac{|T_+ - T_-|}{2} e^{-|T_+ - T_-|x/2}$$

while the opposite inequality is evident. □

We can now derive (69).

Derivation of (69). We look at the case $\sigma = +$ since the other is very similar. Observe that by definition

$$(Q_S + Q_{B_+} + Q_{B_-})\tilde{f}_\infty = \widetilde{\Lambda} \tilde{f}_\infty$$

while

$$(Q_S + 2Q_{B_+})\Gamma_+^M = \widetilde{\Lambda}\Gamma_+^M.$$

Hence

$$\begin{aligned}
\widetilde{\Lambda} d_2(\tilde{f}_\infty, \Gamma_+^M) &= d_2((Q_S + Q_{B_+} + Q_{B_-})\tilde{f}_\infty, (Q_S + 2Q_{B_+})\Gamma_+^M) \\
&\leq d_2(Q_S\tilde{f}_\infty, Q_S\Gamma_+^M) + d_2((Q_{B_+} + Q_{B_-})\tilde{f}_\infty, 2Q_{B_+}\Gamma_+^M) \\
&\leq \Lambda_S d_2(\tilde{f}_\infty, \Gamma_+^M) + d_2((Q_{B_+} + Q_{B_-})\tilde{f}_\infty, (Q_{B_+} + Q_{B_-})\Gamma_+^M) + d_2(Q_{B_-}\Gamma_+^M, Q_{B_+}\Gamma_+^M)
\end{aligned}$$

Observe now that, reasoning like for (72) and (77) we get

$$d_2(Q_{B_-}\Gamma_+^M, Q_{B_+}\Gamma_+^M) \leq \mu d_2(\Gamma_+, \Gamma_-) \sup_{\vec{\xi}\neq 0} \left(\sum_{i=1}^M \int \frac{|(\vec{\xi}_i \cdot \omega)|^2}{|\underline{\xi}|^2} d\omega \right) \leq \frac{\mu}{3} d_2(\Gamma_+, \Gamma_-)$$

so that, using (77), we obtain

$$\widetilde{\Lambda} d_2(\tilde{f}_\infty, \Gamma_+^M) \leq \Lambda_S d_2(\tilde{f}_\infty, \Gamma_+^M) + 2\left(\Lambda_B - \frac{\mu}{3}\right) d_2(\tilde{f}_\infty, \Gamma_+^M) + \frac{\mu}{3} d_2(\Gamma_+, \Gamma_-)$$

The conclusion thus follows from Lemma A.3 above and recalling that $\Lambda_S + 2\Lambda_B = \widetilde{\Lambda}$. □

B Moments of f_0 and derivatives of $\widehat{G}_s$.

In this appendix we first look at the evolution of the moments of the distribution $\tilde{f}_t$. Although it is quite natural to expect such moments to be uniformly bounded in t, it is less easy to analyze their behavior in M. We will proof here that $E_4(\tilde{f}_t)$ is bounded uniformly in t and M in term of $E_4(f_0)$. The proof is based on a direct computation coupled with a dimensional argument. It extends the analogous proof given in [3], for a much simpler case. Since it turns out to be rather notationally involved, we will not report it in full details. We will then derive the Newton Law of cooling (34).

Continuing the definitions given just before Theorem 2.1, we call V_o^k the space of homogeneous polynomials $p_k(\underline{v}) = \sum_{\mathbf{i}\in I^k} p_{\mathbf{i}}\, \underline{v}_{\mathbf{i}}$ of degree k and V^l the space of polynomials of degree l, that is $p \in V^l$ if $p(\underline{v}) = \sum_{k=0}^l p_k(\underline{v})$ with $p_k \in V_o^k$ where $V_o^0 = \mathbb{R}$. On V^l we consider the scalar product

$$(p, q) = \sum_{\mathbf{i}\in\tilde{I}^l} p_{\mathbf{i}} q_{\mathbf{i}} \tag{78}$$

where $\tilde{I}^l = \bigcup_{k=0}^l I^k$.

To a polynomial p we associate its f_0 expectation

$$\mathbb{E}_0(p) := \int_{\mathbb{R}^{3M}} p(\underline{v}) f_0(\underline{v}) d\underline{v} = \sum_{\mathbf{i}\in\tilde{I}^l} m_{\mathbf{i}}(f_0)\, p_{\mathbf{i}},$$

where $m_{\mathbf{i}}(f) = \int_{\mathbb{R}^{3M}} v_{\mathbf{i}} f(\underline{v}) d\underline{v}$ is the $\mathbf{i}$-th moment of f_0. We can thus write

$$\int_{\mathbb{R}^{3M}} p(\underline{v}) \mathcal{L}_S[f](\underline{v}) d\underline{v} =: \int_{\mathbb{R}^{3M}} L_S[p](\underline{v}) f(\underline{v}) d\underline{v}$$

where L_S is the linear operator acting on V^l associated to $\mathcal{L}_S$. Observe that $L_S V_o^k \subset V_o^k$. We can now write

$$\begin{aligned} \mathbb{E}_t(p) := \int_{\mathbb{R}^{3M}} p(\underline{v}) f_t(\underline{v}) d\underline{v} = \int_{\mathbb{R}^{3M}} p(\underline{v}) e^{\mathcal{L}_S t}[f_0](\underline{v}) d\underline{v} = \\ \int_{\mathbb{R}^{3M}} e^{L_S t}[p](\underline{v}) f_0(\underline{v}) d\underline{v} = \sum_{\mathbf{i} \in \tilde{I}^l} m_{\mathbf{i}} \left(e^{L_S t}[p] \right)_{\mathbf{i}} = (e^{L_S t}[p], \underline{m}) \end{aligned} \tag{79}$$

where $\underline{m} = \{m_{\mathbf{i}} : \mathbf{i} \in \tilde{I}^l\}$ and, for simplicity sake, we set $m_{\mathbf{i}} = m_{\mathbf{i}}(f_0)$.[1]

In a similar way we can define L_B as the linear operator acting on V^l associated to $\mathcal{L}_B$. Observe that $L_B V_o^k \subset V^k$ and that, calling $L_B^{k,l} := P_{V^l} L_B\big|_{V_o^k}$, where P_V is the orthogonal projector on the subspace V, we have $L_B^{k,l} \neq 0$ if and only if $k - l$ is non negative and even, due to the parity of the Maxwellian distribution. The following Lemma collect basic properties of these operators.

Lemma B.1. *L_S is self adjoint, $(L_S[p], p) \le 0$. Moreover $L_B^{k,k}$ is self-adjoint, $(L_B^{k,k}[p], p) \le 0$, $(L_B^{k,k}[p], p) < 0$ if $k > 0$. Finally $\|L_B^{k,l}\| < CM$.*

Proof. Given $i, i' \in \{0, \dots, M\}$ and $\omega \in \mathbb{S}_2$, let $r_{i,i'}(\omega)$ be the linear transformation on $\mathbb{R}^{3M}$ described in (4). A direct computation shows that

$$L_S\big|_{V_0^k} = \frac{\lambda_S}{M-1} \sum_{i<j} \int_{\mathbb{S}^2} (r_{i,j}(\omega)^{\otimes k} - \mathrm{Id}) d\omega \,.$$

The statement follows immediately from the fact that $r_{i,j}(\omega)$ is unitary and self-adjoint.

Let now $\tilde{r}_i(\omega)$ be the linear transformation on $\mathbb{R}^{3M+3}$ described in (7). Then $\tilde{r}_i(\omega)^{\otimes k}$ acts on homogeneous the polynomials of degree k in $3M+3$ variables. Let $\widetilde{P}$ be the projector between the space of homogeneous polynomials in $3M+3$ variables to the space of homogeneous polynomials in $3M$ variables. Then we have

$$L_B^{k,k} = \mu \sum_i \int_{\mathbb{S}^2} (\widetilde{P} \tilde{r}_i(\omega)^{\otimes k} \widetilde{P} - \mathrm{Id}) d\omega$$

and the thesis follows from the fact that $\widetilde{P} \tilde{r}_i(\omega)^{\otimes k} \widetilde{P}$ is a principal minor of a self-adjoint unitary matrix. Finally we observe that also $L_B^{k,l}$ is the sum of M terms that can be written in term of suitable (non-orthogonal) projections of $\tilde{r}_i(\omega)^{\otimes k}$. □

[1] In the following, with a slight abuse of notation, we will use p to indicate both the polynomial $P(\underline{v})$ and the sets of its coefficients $p = \{p_{\mathbf{i}} : \mathbf{i} \in \tilde{I}^l\}$.

Thus Lemma B.1 tells us that $\widetilde{L} = L_S + L_{B_+} + L_{B_-}$ is block triangular with negative definite diagonal blocks but potentially large off-diagonal blocks. It follows that, if $q^2(\underline{v}) := \frac{1}{M}\sum_i \|\vec{v}_i\|^4$ then $\mathbb{E}_t(q^2) \le CM^2\|q^2\|\|E_4(f_t)\|$. Observe moreover that $\|\underline{m}\| = O(M^2)E_4(f_0)$ while $\|q^2\| = M^{-\frac{1}{2}}$. It follows that $\mathbb{E}_t(q^2)$ is bounded uniformly in t but not in M.

To obtain a bound uniform in M we need to look at the structure of L_S and L_B. To avoid overburdening the notation we will only consider the case $l = 4$. Moreover we will restrict our attention on the subspace of permutation invariant polynomials containing only monomials of even degree. They clearly form an L invariant subspace.

We observe that the polynomials

$$p^{4,1}_{\mathbf{j}}(\underline{v}) = \frac{1}{M}\sum_i v_{i,j_1}v_{i,j_2}v_{i,j_3}v_{i,j_4}$$

$$p^{4,2}_{\mathbf{j}}(\underline{v}) = \frac{1}{\binom{M}{2}}\sum_{i_1<i_2} v_{i_1,j_1}v_{i_1,j_2}v_{i_2,j_3}v_{i_2,j_4} \quad p^{4,2'}_{\mathbf{j}}(\underline{v}) = \frac{1}{\binom{M}{2}}\sum_{i_1<i_2} v_{i_1,j_1}v_{i_1,j_2}v_{i_1,j_3}v_{i_2,j_4}$$

$$p^{4,3}_{\mathbf{j}}(\underline{v}) = \frac{1}{\binom{M}{3}}\sum_{i_1<i_2<i_3} v_{i_1,j_1}v_{i_1,j_2}v_{i_2,j_3}v_{i_3,j_4} \quad p^{4,4}_{\mathbf{j}}(\underline{v}) = \frac{1}{\binom{M}{4}}\sum_{i_1<i_2<i_3<i_4} v_{i_1,j_1}v_{i_2,j_2}v_{i_3,j_3}v_{i_4,j_4}$$

with $\mathbf{j} \in J^4 := \{1,2,3\}^4$, form an orthogonal basis in the subspace of permutation invariant polynomials in V^4_o. Let $V^{4,1}_o = \mathrm{span}\left\{p^{4,1}_{\mathbf{j}} : \mathbf{j} \in J^4\right\}$ the subspace of *one particle* polynomials, $V^{4,2}_o = \mathrm{span}\left\{p^{4,2}_{\mathbf{j}}, p^{4,2}_{\mathbf{j}} : \mathbf{j} \in J^4\right\}$ the subspace of *two particles* polynomials, and similarly for $V^{4,3}_o$ and $V^{4,4}_o$. In a similar way we consider the polynomials

$$p^{2,1}_{\mathbf{j}}(\underline{v}) = \frac{1}{Z_{2,1}}\sum_i v_{i,j_1}v_{i,j_2} \qquad p^{2,2}_{\mathbf{j}}(\underline{v}) = \frac{1}{Z_{2,2}}\sum_{i_1<i_2} v_{i_1,j_1}v_{i_2,j_2}$$

with $\mathbf{j} \in \{1,2,3\}^2$, and the subspaces $V^{2,1}_o$ and $V^{2,2}_o$.

Lemma B.2. *For every $\mathbf{j} \in J^4$ we have*

$$(e^{tL}p^{4,k}_{\mathbf{j}}, \underline{m}) \le CE_4(f_0) \tag{80}$$

while for every $\mathbf{j} \in J^2$ we get

$$(e^{tL}p^{2,k}_{\mathbf{j}}, \underline{m}) \le CE_4(f_0)\,. \tag{81}$$

Finally

$$(e^{tL}1, \underline{m}) \le CE_4(f_0)\,. \tag{82}$$

Proof. A direct computation shows that

$$\left\| P_{V^{l,n}_o} L_S\big|_{V^{l,m}_o} \right\| \le CM^{-\frac{|m-n|}{2}}$$

if $|m-n|\leq 1$ and 0 otherwise while $\left\|P_{V_o^{l,n}}L_{B_\pm}\big|_{V_o^{l,m}}\right\|\leq C\delta_{m,n}$. Finally $P_{V_o^{l,n}}L\big|_{V_o^{l,m}}$ is negative definite if $l>0$. It follows from a perturbative argument that

$$\left\|P_{V_o^{l,n}}e^{Lt}\big|_{V_o^{l,m}}\right\| = CM^{-\frac{|m-n|}{2}}e^{-\kappa t}$$

for a suitable $\kappa>0$. Observing that $\|p_{\mathbf{j}}^{4,k}\|\leq CM^{-\frac{k}{2}}$ while $\|P_{V_o^{4,k}}\underline{m}\|\leq CM^{\frac{k}{2}}E_4(f_0)$, we get that

$$(P_{V_o^4}e^{tL}p_{\mathbf{j}}^{4,k},\underline{m})\leq\sum_{l=0}^{4}\left\|P_{V_o^{4,l}}\underline{m}\right\|\left\|P_{V_o^{4,l}}e^{Lt}\big|_{V_o^{4,k}}\right\|\left\|p_{\mathbf{j}}^{4,k}\right\|\leq Ce^{-\kappa t}E_4(f_0)\,. \tag{83}$$

Since $L_{B_\pm}$ are block triangular we can write

$$P_{V_o^2}e^{tL}p_{\mathbf{j}}^{4,k}=\int_0^t P_{V_o^2}e^{sL}(L_{B_+}^{4,2}+L_{B_-}^{4,2})P_{V_o^4}e^{sL}p_{\mathbf{j}}^{4,k}dt$$

Another direct computation show that $\left\|P_{V_o^{2,m}}L_B\big|_{V_o^{4,n}}\right\|\leq CM^{\frac{n-m}{2}}$ for $n-m=0,1$ and 0 otherwise. Repeating the argument in (83) we get that also $(P_{V_o^2}e^{tL}p_{\mathbf{j}}^{4,k},\underline{m})\leq Ce^{-\kappa t}E_4(f_0)$. Finally we write

$$P_{V^0}e^{tL}p_{\mathbf{j}}^{4,k}=\int_0^t P_{V^0}e^{sL}(L_{B_+}^{4,0}+L_{B_-}^{4,0})P_{V_o^4}e^{sL}p_{\mathbf{j}}^{4,k}dt+\int_0^t P_{V^0}e^{sL}(L_{B_+}^{2,0}+L_{2_-}^{2,0})P_{V_o^2}e^{sL}p_{\mathbf{j}}^{4,k}dt$$

that gives $(P_{V^0}e^{tL}p_{\mathbf{j}}^{4,k},\underline{m})\leq C$ since $P_{V^0}L=0$. Summing up we get (80). A similar argument gives (81) and (82).

□

The following Lemma is a direct consequence of the above discussion.

Lemma B.3. *Suppose $f_0(\underline{v})$ is a permutation invariant distribution on $\mathbb{R}^{3M}$ such that $E_4(f_0)<\infty$. Then for every $t>0$ we have*

$$E_4(\tilde f_t)\leq CE_4(f_0)\,.$$

Proof. For every $\mathbf{i}=((i_1,j_1),\dots(i_4,j_4))\in I^k$, using Hölder inequality with suitable exponents we can write

$$e_{\mathbf{i}}(\tilde f_t)\leq\left(\prod_{l=1}^{k}\int_{\mathbb{R}^{3M}}(v_{i_l,j_l})^4\tilde f_t(\underline{v})d\underline{v}\right)^{\frac14}\,.$$

Since $\int_{\mathbb{R}^{3M}}\tilde f_t(\underline{v})d\underline{v}=1$ and thus $E_4(\tilde f_t)\geq 1$, applying Lemma B.2 we get

$$e_{\mathbf{i}}(\tilde f_t)\leq CE_4(f_0)\,.$$

The thesis follows immediately. □

We can now formulate the main result of this Appendix in the following Lemma.

Lemma B.4. *Let $\widehat{G}_s$ be as defined in* (63)*, that is,*

$$\widehat{G}_s(\vec{\eta}_i, \underline{\xi}) = \sum_{j=1}^{M} \Big(R^{I_+}_{i,j}[\hat{f}_s \widehat{\Gamma}_+](\vec{\eta}_i, \underline{\xi}) - R^{I_+}_{i,j}[\hat{f}_s \widehat{\Gamma}_+](0, \underline{\xi}) \widehat{\Gamma}_+(\vec{\eta}_i) \Big) . \tag{84}$$

Then we have

$$\|\widehat{G}_s\|_{3,1} \le CME_4(f_0) .$$

Proof. It is enough to observe that

$$\|R^{I_+}_{i,j}[\hat{f}_s \widehat{\Gamma}_+]\|_{3,1} \le C \max_{\vec{\beta}\in\mathbb{N}^3, \|\vec{\beta}\|_1 \le 4} \left\| \partial^{\vec{\beta}} \hat{f}_s \right\|_\infty \le CE_4(\tilde{f}_s) .$$

□

We can now derive (34).

Derivation of (34)*.* For $f : \mathbb{R}^6 \mapsto \mathbb{R}$ we have

$$\begin{aligned}\int \|\vec{v}\|^2 (R - I)[f](\vec{v}, \vec{w}) d\vec{v} d\vec{w} &= \int \left[\|\underline{v} - ((\underline{v} - \vec{w}) \cdot \omega)\omega\|^2 - \|\vec{v}\|^2 \right] f(\vec{v}, \vec{w}) d\omega d\vec{w} d\underline{v} \\ &= \frac{1}{3} \int (\|\vec{w}\|^2 - \|\vec{v}\|^2) f(\vec{v}, \vec{w}) d\vec{w} d\underline{v} .\end{aligned}$$

we thus get

$$\begin{aligned}&\int \|\underline{v}\|^2 (R^{R_\sigma}_{i,j} - I)[\mathscr{F}_t](\underline{u}, \underline{v}, \underline{w}) d\underline{v} d\underline{u} d\underline{w} = 0 \\ &\int \|\underline{v}\|^2 (R^{S}_{i,j} - I)[\mathscr{F}_t](\underline{u}, \underline{v}, \underline{w}) d\underline{v} d\underline{u} d\underline{w} = 0 \\ &\int \|\underline{v}\|^2 (R^{I_+}_{i,j} - I)[\mathscr{F}_t](\underline{u}, \underline{v}, \underline{w}) d\underline{v} d\underline{u} d\underline{w} = \frac{1}{3} \int (\|\vec{u}_i\|^2 - \|\vec{v}_j\|^2) \mathscr{F}_t(\underline{u}, \underline{v}, \underline{w}) d\underline{v} d\underline{u} d\underline{w}\end{aligned}$$

and similarly for $R^{I_-}_{i,j} - I$. Summing over i and j we get the equation for $e_S(t)$. The others two equations follow in a similar manner. □

C Rotational average at fixed momentum

For N and $k < N$ consider a distribution $f : \mathbb{R}^{3k} \mapsto \mathbb{R}$ such that

$$\int_{\mathbb{R}^{3k}} f(\underline{v}) d\underline{v} = \bar{f} \qquad \text{and} \int_{\mathbb{R}^{3k}} \underline{v} f(\underline{v}) d\underline{v} = 0$$

and its "extension" $f\Gamma^{N-k}_T$ to a distribution on $\mathbb{R}^{3k}$.

Lemma C.1. *We have*
$$d_2\left(\mathcal{R}\big[f\Gamma_T^{N-k}\big],\bar f\Gamma_T^N\right)\le \frac{k}{N}d_2(f,\bar f\Gamma_T^k)$$

Proof. Given $\underline{\eta}\in\mathbb{R}^{3N}$ we write $\underline{\eta}^{\le k}=(\vec\eta_1,\dots,\vec\eta_k)$ and $\underline{\eta}^{>k}=(\vec\eta_{k+1},\dots,\vec\eta_N)$. We have
$$\begin{aligned}
d_2\left(\mathcal{R}\big[f\Gamma_T^{N-k}\big],\bar f\Gamma_T^N\right)=&\sup_{\underline{\eta}\neq 0}\frac{1}{\|\underline{\eta}\|^2}\int_{\mathcal{O}}\Big(f((O\underline{\eta})^{\le k})-\bar f\widehat\Gamma_T^k((O\underline{\eta})^{\le k})\Big)\widehat\Gamma_T^{N-k}((O\underline{\eta})^{>k})d\sigma(O)\\
\le& d_2(f,\bar f\Gamma_T^k)\sup_{\underline{\eta}\neq 0}\frac{1}{\|\underline{\eta}\|^2}\int_{\mathcal{O}}\|(O\underline{\eta})^{\le k}\|^2\widehat\Gamma_T^{N-k}((O\underline{\eta})^{>k})d\sigma(O)\\
\le& d_2(f,\bar f\Gamma_T^k)\sup_{\underline{\eta}\neq 0}\frac{1}{\|\underline{\eta}\|^2}\int_{\mathcal{O}}\|(O\underline{\eta})^{\le k}\|^2d\sigma(O)
\end{aligned}$$

Observe now that for every i,j we have
$$\int_{\mathcal{O}}(O\underline{\eta})_i^2d\sigma(O)=\int_{\mathcal{O}}(O\underline{\eta})_j^2d\sigma(O)$$

so that
$$\int_{\mathcal{O}}\|(O\underline{\eta})^{\le k}\|^2d\sigma(O)=\frac{k}{N}\|\underline{\eta}\|^2\,.$$

□