Inconsistency in theories of violent relaxation

I. Arad* and D. Lynden-Bell

Institute of Astronomy, Madingley Road, Cambridge CB3 OHA

ABSTRACT

We examine an inconsistency in theories of violent relaxation by Lynden-Bell and Nakamura. The inconsistency arises from the non-transitive nature of these theories: a system that undergoes a violent relaxation, relaxes and then, upon an addition of energy, undergoes violent relaxation once again would settle in an equilibrium state that is different from the one that is predicted had the system gone directly from the initial to the final state. We conclude that a proper description of the violent relaxation process cannot be achieved by an equilibrium statistical mechanics approach, but instead a dynamical theory for the coarse-grained phase-space density is needed.

Key words: methods: analytical – galaxies: evolution – galaxies: kinematics and dynamics – galaxies: statistics.

1 INTRODUCTION

The main purpose of this paper is to demonstrate that theories predicting a definite statistical equilibrium, dependent only on the energy and the initial volume $\eta$ at phase-space densities in the range $\eta \rightarrow \eta + d\eta$, do not predict the same final state when the system undergoes two violent relaxation sessions separated in time, as they do when the two sessions are treated as one. This may be called an inconsistency or at best a lack of transitivity.

Consider an $N$-body gravitating system that starts in some state far from equilibrium, vibrates violently and settles to a dynamically steady state $s_1$ that lasts long enough that it may be considered the final product of violent relaxation with energy $E_1$. Now suppose that this system suffers a significant tidal disturbance from a passing object that causes violent vibrations and leaves our system to relax again but now with energy $E_2$. There are now two ways of predicting the outcome. Either we take the function $\tau_0(\eta)$, giving the volume at each phase-space density from the initial state $s_0$, or we use $\tau_1(\eta)$, the predicted outcome for that function after the first relaxation process. Of course, if we used the fine-grained phase-space density, both would be the same, but by hypothesis the system lasted so long in state $s_1$ that is at 'equilibrium', and only the coarse-grained density can any longer be relevant to the dynamics of the final relaxation. We show that the outcomes predicted with energy $E_2$ and volume functions $\tau_0(\eta)$ and $\tau_1(\eta)$ are certainly different in the Lynden-Bell theory of violent relaxation (Lynden-Bell 1967), as well as in a more recent theory by Nakamura (Nakamura 2000).

This paper is organized as follows. In Section 2 we give a short overview of some of the difficulties in the theory of violent relaxation, which we feel complement the main subject of this paper. Then in Section 3 we demonstrate the non-transitivity of the Lynden-Bell theory of violent relaxation, using the thought experiment that was presented above. In Section 4 we give a brief description of Nakamura’s theory, which is based on the information-theory approach. We rederive his theory using a combinatorial approach that enables us to compare it to Lynden-Bell’s theory. Then in Section 5 we demonstrate that also Nakamura’s theory is non-transitive, using the thought experiment once again. In Section 6 we present our conclusions.

2 OVERVIEW OF DIFFICULTIES IN THEORY OF VIOLENT RELAXATION

In this section we offer a short discussion of other problems connected with the process of violent relaxation and ‘theories’ that aim to predict its outcome. Let us start from first principles. Most large $N$-body systems governed by long-range forces that are not initially in balance will oscillate with decreasing amplitude before they settle into a state in which the potential of the long-range force becomes almost steady. Such violent relaxation processes are known to occur in gravitational $N$-body systems. Thereafter, evolution may continue as a result of the shorter-range interaction, in which the graininess of the individual particles is of importance, but there is a large class of systems in which this secondary evolution is on a much longer time-scale. Violent relaxation under gravity does not last long. After a few oscillations on the time-scale $(G\bar{\rho})^{-1/2}$ it is over. Thus the whole idea that the interaction of the particles with the mean field will lead to some unique detailed statistical equilibrium state, only dependent on the initial conditions via the dynamically conserved quantities, is more of a vain hope and a confession of ignorance than an established fact. Nevertheless, Hénon (1968) gave some evidence in favour of its prediction, and for cosmological initial conditions Navarro, Frenk & White (1995, 1996, 1997) show a considerable...
universality in their results. Binney (2004) gives a lovely toy model that he uses to criticize deductions from N-body simulations. Even in the initial discussion (Lynden-Bell 1967), it was admitted that there were many stable steady states into which a gravitating system could settle and that violent relaxation would be incomplete, so that the system would not inevitably attain a state close to the more probable one.

A second worrying aspect of any equilibrium theory is the apparent lack of an analogue to the law of detailed balance. At real thermodynamic equilibrium there are no cyclic processes going around and around, but each individual emission process is exactly balanced by the corresponding absorption process. In radiation theory, Einstein introduced his stimulated emission process just to ensure that this would be so. In the process of violent relaxation, each element is interacting with the potential of the whole system, and one might expect some to be highly accelerated, as Fermi argued to get his cosmic ray acceleration process. Conservation of energy must, however, lead to some dynamical friction term, but this cannot be mass-related, as that is at odds with our earlier arguments that energy gain is independent of mass.

Finally, not all systems have violent relaxation. In some early experiments with pulsating concentric spherical shells, Hénon (1968) found a few examples of gravitating systems with persistent oscillation that defied the general decay. Newton in Principia showed that systems with a force law between particles proportional to separation (rather than inverse square) oscillated forever. Indeed, Newton solved that N-body problem completely, showing that each particle moved on a central ellipse centred on the barycentre and that all those orbits had the same period. Lynden-Bell introduced his cosmic ray acceleration process. Conservation of energy must, one might expect some to be highly accelerated, as Fermi argued to get his cosmic ray acceleration process. Conservation of energy must, however, lead to some dynamical friction term, but this cannot be mass-related, as that is at odds with our earlier arguments that energy gain is independent of mass.

3 NON-TRANSITIVITY IN LYNDEN-BELL THEORY OF VIOLENT RELAXATION

3.1 Main results of LB67 theory

In LB67 theory we deal with a system that is initially out of equilibrium. Its initial state is specified by its total energy $E$ and the phase-space volumes $V_1, V_2, \ldots$ of the initial phase-space density levels $\eta_1, \eta_2, \ldots$. Phase space is then divided into micro-cells of fixed volume $\omega$, which can be either empty or hold a phase-space element of one of the prescribed levels. The state of all these micro-cells defines a micro-state.

Next, we let the system redistribute its phase-space elements as it approaches an equilibrium. The micro-cells are then grouped into macro-cells, each macro-cell containing $v$ micro-cells. A macro-state of the system is defined by the matrix $\{\eta_j\}$, which specifies how many phase-space elements of type $J$ ended up in the macro-cell $i$. The coarse-grained phase-space density function (DF) at the macro-cell $i$ is therefore given by

$$f_i = \frac{1}{v} \sum_j \eta_j n_{ij}.$$  

Then, in the spirit of ordinary statistical mechanics, one assumes that the system has an equal a priori probability of being in each one of the micro-states. To find the equilibrium state, one maximizes the function $W(\{\eta_j\})$, which counts the number of micro-states that correspond to the macro-state $\{\eta_j\}$, hence obtaining the most probable state.

When maximizing $W$, one has to consider only those macro-states for which the total energy is $E$ and the overall volume in each of the initial phase-space levels is $V_1, V_2, \ldots$. This can be done in the usual way with Lagrange multipliers. After passing to a continuous description of the macro-cells, the resultant DF is

$$f_{LB}(r, \nu) = f_{LB}[\{r, \nu\}] = A(\epsilon) \sum \eta_j e^{-\beta \eta_j (\epsilon - \mu)},$$

with

$$A(\epsilon) = \frac{1}{1 + \sum_j e^{-\beta \epsilon (\epsilon - \mu)}}.$$  

Here $\epsilon(r, \nu) = \nu^2/2 + \Phi_{LB}(r)$ is the energy per unit mass of the $(r, \nu)$ phase-space cell, and $\Phi_{LB}(r)$ is the gravitational potential, calculated self-consistently from the Poisson equation

$$V^2 \Phi(r) = 4\pi G \int d^3 v f_{LB}(r, \nu).$$

The dimensional constants $\beta, \mu_1, \mu_2, \ldots$ are Lagrange multipliers, which are calculated from the energy-conservation constraint

$$\int d^3 \tau f_{LB}(r, \nu) [\frac{1}{2} \nu^2 + \frac{1}{2} \Phi_{LB}(r)] = E,$$  

and the initial-conditions constraint

$$\int d^3 \tau A(\epsilon) e^{-\beta \epsilon (\epsilon - \mu)} V_J = \eta_J, \quad J = 1, 2, \ldots,$$  

where we have used the notation $d^3 \tau \equiv d^3 r d^3 v$.

Equations (2)–(6) are the main results of the LB67 theory. For our needs, however, two small modifications are needed. First, by introducing the energy density function

$$g(\epsilon_0) = \int d^3 \tau \delta[\epsilon(r, \nu) - \epsilon_0],$$  

the integral on the left-hand side (LHS) of equation (6) can be written as

$$\int g(\epsilon) A(\epsilon) e^{-\beta \epsilon (\epsilon - \mu)} d\epsilon.$$  

1 In the original formulation of LB67, Lynden-Bell used the masses of the different levels instead of their phase-space volumes. However, these are trivially related to each other by $M_J = \eta_J V_J$.  

Secondly, we pass to a continuous description of the initial density levels using the phase-space volume distribution function $\tau(\eta)$, which is defined so that $\tau(\eta)\,d\eta$ is the phase-space volume initially occupied by phase-space densities in the range $\eta \to \eta + d\eta$. Formally, if $f_1(r, v)$ is the initial DF, then $\tau(\eta)$ is given by

$$
\tau(\eta) = \int d^6 \tau \, \delta[f_1(r, v) - \eta].
$$

By letting each density level $\eta_j$ have a small width $\Delta \eta$, the sums in equations (2) and (3) can be changed to integrals by

$$
\sum_j \rightarrow \frac{1}{\Delta \eta} \int d\eta,
$$

and hence equations (2)–(6) are now

$$
f_{LB}(\epsilon) = \frac{1}{\Delta \eta} A(\epsilon) \int \eta \, e^{-\beta(\epsilon - \mu \eta)} \, d\eta,
$$

$$
A(\epsilon) = \frac{1}{1 + (1/\Delta \eta) \int e^{-\beta(\epsilon - \mu \eta)} \, d\eta},
$$

$$
\nabla^2 \Phi(r) = 4\pi G \int d^3 v \, f_{LB}(r, v),
$$

$$
E = \int d^6 \tau \, f_{LB}(r, v) \left[ \frac{1}{2} v^2 + \frac{1}{2} \Phi(r) \right],
$$

$$
\tau(\eta) \Delta \eta = \int g(\epsilon) A(\epsilon) e^{-\beta(\epsilon - \mu \eta)} \, d\epsilon.
$$

Notice that the amplitude of $\Delta \eta$ is unimportant, as it can always be absorbed into the Lagrange multipliers $\mu_{\eta}$.

Finally we recall that in a spatially infinite domain equations (10)–(14) have no solution, since the density can spread indefinitely while conserving its energy and increasing its entropy. A common way to overcome this problem, which will be adopted here, is to work within a rigid sphere of radius $R$, and to assume that the resulting equilibrium configuration is spherical. Such a model, although not realistic, is an easy way to obtain a finite solution.

### 3.2 Double-relaxation experiment

To test the transitivity of the LB67 theory, we propose the following four-step thought experiment, which was mentioned in the Introduction.

(i) We prepare a system with one density level (the water-bag configuration) and a total energy of $E$. In accordance with the above section, we put the system inside a sphere of radius $R$. We denote the initial state of the system by $\eta_0$.

(ii) We let the system go through a violent relaxation process to an equilibrium that is denoted by $s_1$.

(iii) We add an amount $\Delta E$ of energy to the system by, for example, a strong impulse of an external gravitational field. The system then goes once again through a violent relaxation process and settles down in a new equilibrium state, denoted by $s_2$.

(iv) We prepare a new system with the same parameters as $s_0$ except for the energy, which is set to $E + \Delta E$. We let it go through a violent relaxation process to an equilibrium that is denoted by $s_3$.

Now if the theory is transitive, then necessarily $s_2 = s_1$.

To see if this is really the case in the LB67 theory, we begin by calculating the $s_1$ and $s_3$ states, which are relatively simple to calculate, being the outcome of a water-bag configuration. This calculation has been fully done in Chavanis & Sommeria (1998), and here we use some of their results.

The $s_0$ step was prepared with total mass $M = 1$, $G = 1$, $R = 1$ and an initial phase-space density level

$$
\eta_0 = \frac{1}{\sqrt{512 \pi^2 G^2 MR^3}},
$$

which according to Chavanis & Sommeria (1998) guarantees that for each energy there would be only one equilibrium state.

From equation (2) we find that the DF of $s_1$ and $s_3$ is the well-known Fermi–Dirac distribution

$$
f_1(\epsilon) = \frac{\eta_0}{1 + e^{(\epsilon - \mu_1)G}},
$$

$$
f_3(\epsilon) = \frac{\eta_0}{1 + e^{(\epsilon - \mu_3)G}},
$$

with $\beta_1, \beta_3, \mu_1, \mu_3$ Lagrange multipliers to be fixed from the energy constraint and the initial conditions. As there is only one density level, the initial-conditions constraint can be replaced with the conservation-of-mass constraint.

In Chavanis & Sommeria (1998) it is shown how Lagrange multipliers can be found for a given mass and energy, and we therefore do not repeat these steps here but instead give the values of these parameters in Table 1. All numerical calculations were done using the GNU Scientific Library 1.5 (GSL 1.5) software. The differential equation for $\Phi(r)$ was solved using an embedded Runge–Kutta Prince–Dormand (8,9) method, whereas integration was done using a 51-point Gauss–Kronrod rule. In all calculations a relative error of less than $10^{-3}$ was maintained.

Fig. 1 shows the radial density profiles of $s_1$ and $s_3$. As the $s_3$ configuration is more energetic, its mean kinetic energy is higher, and as a result the distribution is less concentrated than the $s_1$ distribution, and has a lower-density core. Fig. 2 shows the DFs of $s_1$ and $s_3$. Notice how both distributions have a substantial degenerate part, as for both configurations the Fermi energies are $\epsilon_{\text{F}} \simeq \Phi_i(0)/2$.

Finally, to verify that the $s_3$ configuration is indeed more mixed than the $s_1$ configuration and therefore a transition $s_1 \rightarrow s_3$ is permitted by the mixing theorem (Tremaine, Hénon & Lynden-Bell 1986) – we have calculated the $M(\nu)$ functions of $s_1$ and $s_3$. The $M(\nu)$ function is defined in the following way. We first define $M(\eta)$ and $V(\eta)$ as the cumulative mass and phase-space volumes above the phase-space density $\eta$.

$$
V(\eta) = \int_{\eta}^\infty \frac{\tau(\eta')}{\Delta \eta} \, d\eta',
$$

and

Table 1. The numerical parameters that specify the $s_1$ and $s_3$ states. Here $E$ is the total energy of the state, $\eta_0$ is the initial density level given by equation (15), $\beta$ and $\mu$ are the Lagrange multipliers in equations (16) and (17), and $\Phi(0)$ is the gravitational potential at $r = 0$.

<table>
<thead>
<tr>
<th>State</th>
<th>$E$</th>
<th>$\eta_0\beta$</th>
<th>$\mu$</th>
<th>$\Phi(0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>-6.157</td>
<td>1.2</td>
<td>-14.87</td>
<td>-33.96</td>
</tr>
<tr>
<td>$s_2$</td>
<td>-1.589</td>
<td>1.0</td>
<td>-8.322</td>
<td>-15.12</td>
</tr>
</tbody>
</table>

2 Free from http://www.gnu.org/software/gsl/
energy, which makes it less concentrated.

The phase-space densities of the states in the double-relaxation experiment of the LB67 theory. Here $\bar{\rho}$ is the average density, given by $\bar{\rho} = 3M/(4\pi R^3)$. The state corresponds to a hotter system with higher energy, which makes it less concentrated.

Figure 1. Density profiles of the $s_1$ and $s_3$ states in the double-relaxation experiment of the LB67 theory. Here $\bar{\rho}$ is the average density, given by $\bar{\rho} = 3M/(4\pi R^3)$. The state corresponds to a hotter system with higher energy, which makes it less concentrated.

Figure 2. The phase-space densities of the $s_1$ and $s_3$ states as functions of the energy per unit mass $\epsilon$. The functional form of these densities is the well-known Fermi–Dirac distribution, given by equations (16) and (17) and Table 1.

$$M(\eta) = \int_{r<R} d^6r \, f(r, v) H[f(r, v) - \eta]$$

$$= \int_0^\infty \eta^* \tau(\eta^*) d\eta^*,$$

with $H(\ldots)$ being the Heaviside step function. As $V(\eta)$ is a monotonically decreasing function, we can invert it and express $\eta$ in terms of the cumulative volume $V$. Plugging this into $M(\eta)$, we get the $M(V)$ function. According to the mixing theorem, the distribution $s_3$ is more mixed than the distribution $s_1$ if and only if $M_3(V) \leq M_1(V)$ for every $V$, as shown in Fig. 3.

### 3.3 Analysing the $s_2$ Configuration

Let us now turn our attention to the $s_2$ configuration. Seemingly, we need to calculate the $\tau_1(\eta)$ function of the $s_1$ configuration, and, together with an energy of $E + \Delta E$, solve the equations

$$f_2(\epsilon) = \frac{1}{\Delta \eta} A(\epsilon) \int_0^\infty \eta \, e^{-\beta_2(\epsilon - \mu_1)} d\eta,$$

$$A(\epsilon) = \frac{1}{1 + (1/\Delta \eta) \int_0^\infty \eta \, e^{-\beta_2(\epsilon - \mu_1)} d\eta},$$

$$\tau_1(\eta) = \frac{1}{\Delta \eta} \int_{\Phi(0)}^\infty g_2(\epsilon) A(\epsilon) e^{-\beta_2(\epsilon - \mu_1)} d\epsilon.$$  \hspace{2cm} (22)

The function $g_2(\epsilon)$ needs to be calculated from equation (7) using the gravitational potential $\Phi_2(r)$, which has to be recovered from $f_2(\epsilon)$ using the Poisson equation (12). Finally, the resultant $f_2(\epsilon)$ would be compared to $f_3(\epsilon)$ to see if the two configurations are equal.

There is, however, a much simpler way to see if $f_2(\epsilon) = f_3(\epsilon)$. Let us assume that this is indeed the case, and that consequently also $\Phi_2(\epsilon) = \Phi_3(\epsilon)$ and $g_2(\epsilon) = g_3(\epsilon)$. In such case it is possible to recover the full expression

$$\frac{1}{\Delta \eta} A(\epsilon) e^{-\beta_2(\epsilon - \mu_1)}$$

in terms of the known functions $f_3(\epsilon)$ and $g_3(\epsilon)$.

We start by replacing $f_2(\epsilon) \leftrightarrow f_3(\epsilon)$ and $g_2(\epsilon) \leftrightarrow g_3(\epsilon)$ in equations (20)–(22). Differentiating equation (21) with respect to $\epsilon$ and substituting $f_3(\epsilon)$ from equation (20), we obtain

$$A'_2(\epsilon) = \beta_2 A(\epsilon) f_3(\epsilon),$$

which yields

$$A(\epsilon) = a_0 \exp \left[ -\beta_2 \int_0^\infty f_3(\epsilon') d\epsilon' \right] = a_0 A_0(\epsilon).$$

Here $a_0$ is an unknown integration constant. The integral in $A_0(\epsilon)$ can be easily done analytically if we recall the definition of $f_3(\epsilon)$, which is given in equation (17), yielding

$$A_0(\epsilon) = \left[ 1 + e^{-\beta_2(\epsilon - \mu_3)} \right]^{-\beta_2/\beta_3}. \hspace{2cm} (25)$$

Having found $A(\epsilon)$, we use equation (22) to find the Lagrange multipliers $\mu_\alpha$;

$$e^{-\beta_2 \mu_\alpha} = \left[ \tau_1(\eta) \right]^{-1} a_0 \frac{1}{\Delta \eta} \frac{1}{\Phi_0(0)} \int_0^\infty g_3(\epsilon) A_0(\epsilon) e^{-\beta_2 \mu_\alpha} d\epsilon.$$  \hspace{2cm} (26)

Therefore

$$A(\epsilon) \frac{1}{\Delta \eta} e^{-\beta_2(\epsilon - \mu_\alpha)} = \frac{\tau_1(\eta)}{B_0(\eta)} A_0(\epsilon) e^{-\beta_2 \mu_\alpha},$$

with

$$B_0(\eta) = \int_0^\infty g_3(\epsilon) A_0(\epsilon) e^{-\beta_2 \mu_\alpha} d\epsilon'.$$  \hspace{2cm} (28)
This finally gives us

$$f_2(\epsilon) = A_0(\epsilon) \int_0^{\eta} \frac{\tau_i(\eta)}{B_0(\eta)} e^{-\beta_2 \epsilon} \, d\eta.$$  \hspace{1cm} (29)

Note that the unknown integration constant $A_0$ has been cancelled out. The only remaining unknown is $\beta_2$, which can be fixed by requiring that the energy of $s_2$ will be equal to $E + \Delta E$. Once this is done, we have an expression for $f_2(\epsilon)$, which is equal to $f_3(\epsilon)$ if and only if $s_2$ is identical to $s_3$.

The procedure above is mathematically straightforward. However, numerically, it is slightly more complicated, as $\tau_i(\eta)$ has a very strong peak near $\eta = \eta_0$ as a result of the degeneracy. It is therefore preferable to perform the calculation using the cumulative version of $\tau_i(\eta)$, which was defined in equation (18). For a spherical, isotropic system in a sphere of radius $R$ with a DF $f(\epsilon)$ and a gravitational potential $\Phi(r)$, it is easy to verify that $V(\eta)$ is given by

$$V(\eta) = \frac{2^{3/2}(4\pi)^3}{3} \int_0^{r(\eta)} s^2(\epsilon(\eta) - \Phi(s))^{3/2} \, ds,$$  \hspace{1cm} (30)

with $\epsilon(\eta)$ being the inverse function of $f(\epsilon)$, and $r(\eta)$ is

$$r(\eta) = \frac{\Phi^{-1}(\epsilon(\eta))}{R}, \quad \epsilon(\eta) < -GM/R,$$  \hspace{1cm} (31)

$$r(\eta) = 0, \quad \epsilon(\eta) \geq -GM/R.$$  \hspace{1cm}

Once $V_s(\eta)$ is calculated from the formula above [using $f_1(\epsilon)$ and $\Phi(\epsilon)$], we can calculate $f_2(\epsilon)$ from equation (29) using integration by parts:

$$f_2(\epsilon) = A_0(\epsilon) \int_0^{\eta} V_s(\eta) \frac{e^{-\beta_2 \epsilon}}{B_0(\eta)} \left[ 1 - \beta_2 \epsilon \eta - \eta \frac{B_0(\eta)}{B_0(\eta)} \right] \, d\eta.$$  \hspace{1cm} (32)

### 3.4 Results

To satisfy the energy constraint $E_2 = E_3 = -1.589$, we calculated $E_2$ for various values of $\beta_2$, and chose $\beta_2 = 0.37071$, which gives the correct energy as shown in Fig. 4. We did not find any other solution in the range $0.1 < \beta < 100$ and therefore we believe that $\beta_2 = 0.37071$ is the only relevant solution.

Fig. 5 shows the graphs of $f_2(\epsilon)$ and $f_3(\epsilon)$ once $\beta_2$ was fixed. Clearly the two graphs strongly disagree, in some places by more than one order of magnitude – much more than the numerical error in our calculations. The conclusion is therefore that the LB67 theory is not transitive.

![Figure 4](https://example.com/figure4.png)

**Figure 4.** The energy $E_2$ of different values of $\beta_2$, together with $E_3 = -1.589$. The value $\beta_2 = 0.3707$ is the value that gives $E_2 = E_3$. For higher values up to $\beta_2 = 100$ we did not find any other solution, and therefore we believe that the above solution is the only physical solution.

### 4 INFORMATION-THEORY APPROACH TO VIOLENT RELAXATION AND ITS RELATION TO LB67 THEORY

Recently, a new approach to violent relaxation was proposed in an interesting paper by Nakamura (2000, hereafter NK00). In that paper, Nakamura uses an information-theory approach (Jaynes 1957a,b) to define the entropy of a collisionless system and thereby find its equilibrium state. NK00’s theory predicts a different equilibrium state than LB67, and it is therefore interesting to check whether his theory is transitive or not. We will not, however, try to answer the question of which one of these theories is more correct, as, in our opinion, this is still an open question. Instead, we shall first give a brief description of NK00 and its main results, and then rederive the theory using a combinatorial approach, which would enable us to compare it with the LB67 theory, and point to the reasons why they differ. Finally we will analyse the two-level configuration that goes into the water-bag configuration in the limiting case. The result of this analysis will be used in the next section when we examine the transitivity of the NK00 theory in a double-relaxation experiment.

#### 4.1 Outline of NK00 theory

In the NK00 theory we adopt the probabilistic description of the phase-space density $f(r, v, t)$. Let $f_0(r, v)$ be the initial phase-space density of the system, and define the initial probability distribution $p_0(r, v)$ of finding a (single) test point at $t = 0$ by

$$p_0(r, v) = \frac{1}{M} f_0(r, v).$$  \hspace{1cm} (33)

Then we let the test point move under gravity just like any phase-space element, and we define the probability distribution $p(r, v, t)$ as the probability distribution of finding the test point at time $t > 0$. The conservation of phase-space volume guarantees that $p(r, v, t) = f(r, v, t)/M$ for all $t > 0$.

Next, we divide phase space into macro-cells $i = 1, 2, 3, \ldots$ of volume $\bar{V}_i$ and define the coarse-grained probability $\bar{p}_i$ as the probability of finding the point in the $i$th macro-cell when the system reaches an equilibrium. From the above discussion, it is clear that $\bar{p}_i$ is equal to $f_i/M$ with $f_i$ being the coarse-grained DF in the macro-cell $i$ at equilibrium.

To calculate $\bar{p}_i$, using the information-theory approach, we define the joint probability distribution $p_i(r, v, t)$, which measures the probability of initially finding the test point at $(r, v)$ and later at the
macro-cell $i$. Then we maximize the Shannon entropy

$$S = -\sum_i \int d^6\tau \, p_i(\mathbf{r}, \mathbf{v}) \log p_i(\mathbf{r}, \mathbf{v}),$$

subject to constraints of energy conservation, phase-space volume conservation and initial conditions. The resultant distribution can best be written in terms of the conditional probability $K_i(\mathbf{r}, \mathbf{v}) = p_i(\mathbf{r}, \mathbf{v})/p_i(\mathbf{r}, \mathbf{v})$:

$$K_i(\mathbf{r}, \mathbf{v}) = e^{-\beta \epsilon_i - \delta(x, \mathbf{v}) - \lambda_i}$$

with $\epsilon_i$ being the energy per unit mass of the macro-cell $i$, and $\beta, \delta(x, \mathbf{v}), \lambda_i$ are Lagrange multipliers, to be found from the energy-conservation constraint, from the phase-space volume-conservation constraint

$$\int K_i(\mathbf{r}, \mathbf{v}) d^6\tau = \bar{\omega},$$

and from the phase-space volume-conservation constraint

$$\int K_i(\mathbf{r}, \mathbf{v}) d^6\tau = \bar{\omega}.\tag{37}$$

From equations (35)–(37) it is evident that the dependence of $K_i(\mathbf{r}, \mathbf{v})$ on the indices $i, \mathbf{r}, \mathbf{v}$ is only via $\epsilon_i$ and $p_i(\mathbf{r}, \mathbf{v})$; the latter can trivially be replaced by $f_0(\mathbf{r}, \mathbf{v})$. Therefore, the above equations can be rewritten using the $K(\epsilon, \eta)$ function, together with the $\tau(\eta)$ function, and $g(\epsilon)$ functions that were defined in Section 3.1:

$$K(\epsilon, \eta) = \bar{\omega}^{-1} e^{-\beta \epsilon - \delta(\epsilon) - \lambda(\epsilon)/\eta},$$

$$\int_0^{\infty} g(\epsilon) K(\epsilon, \eta) d\epsilon = 1,$$

$$\int_0^{\infty} \tau(\eta) K(\epsilon, \eta) d\eta = 1.$$

The coarse-grained equilibrium DF is then given by

$$f(\epsilon) = \int_0^{\infty} \tau(\eta) K(\epsilon, \eta) d\eta.$$

As noted by NK00, a prominent difference between his result and the LB67 results is that in the non-degenerate limit his expression reduces to a single Maxwellian distribution, whereas the LB67 expression is a superposition of Maxwellian distributions with different dispersions. This difference can be attributed to the fact that in LB67 we discretize phase space using phase-space elements of equal volume and different masses, while, as we shall see below, the NK00 theory can be derived by discretizing phase space using elements of equal mass, which are associated with different phase-space volumes.

Another evident difference comes from the phase-space volume-conservation constraint (36). This constraint guarantees that the total phase-space volume of all phase-space patches that ended up in macro-cell $i$ will be equal to the macro-cell volume. Consequently, the total phase-space volume of the initial system $\int_0^{\infty} \tau(\eta) d\eta$ must be equal to the total phase-space volume of the non-vanishing phase-space density in the equilibrium configuration. This constraint does not exist in the LB67 theory, where the macro-cells can be only partly full – as is the case, for example, in the non-degenerate equilibrium of a system that is initially in the water-bag configuration. Furthermore, this cannot be trivially changed by adding a volume of zero phase-space density to the initial condition, because setting $p(\mathbf{r}, \mathbf{v}) = 0$ would lead to divergences in equation (35). In Section 4.3 we shall see how this problem can be overcome by using a limiting procedure.

### 4.2 Deriving NK00 theory in a combinatorial approach

To derive the NK00 theory in a combinatorial approach, we realize the phase-space density distribution using $N \gg 1$ elements of equal mass $m$. As in Section 3.1, we assume that initially the system is made of a discrete set of density levels $\eta_1, \eta_2, \ldots$ occupying phase-space volumes $V_1, V_2, \ldots$. Then the overall number of elements that realize a phase-space density $\eta_j$ is $N_j = V_j \eta_j/m$.

Next, we let the system reach an equilibrium through the process of violent relaxation, and divide phase space into macro-cells of equal volume $\bar{\omega}$, which are labelled by the index $i = 1, 2, \ldots$. We define a micro-state by specifying the macro-cell in which every element ended up. A macro-state is then defined by the matrix $\{n_{ij}\}$, which counts how many elements that initially realized the density level $\eta_j$ ended up in the macro-cell $i$. Using $\{n_{ij}\}$, the coarse-grained DF at macro-cell $i$ is given by

$$\bar{f}_i = \frac{m}{\bar{\omega}} \sum_j n_{ij}.$$

Finally, we define the function $W(\{n_{ij}\})$, which counts how many micro-states give the macro-state $\{n_{ij}\}$. It is then simple combinatorics to show that

$$W(\{n_{ij}\}) = \prod_j N_j! / \prod_i n_{ij}!\tag{43}.$$
The last constraint is the energy constraint
\[ \sum_i \omega f_i \left( \frac{1}{2} v_i^2 - \frac{1}{2} G \sum_j \frac{\omega f_j}{|r_i - r_j|} \right) = E, \] (47)
where \( r_i \) and \( v_i \) are the mean position and velocity of the \( i \)th macrocell.

To maximize the entropy under the above constraint, we use Lagrange multipliers. The function that we wish to maximize with respect to \( n_{i,j} \) is therefore
\[ I = S - \sum_j \lambda_i n_{i,j} - \sum_i \delta_j n_{i,j} - \beta E. \] (48)

Differentiating \( I \) with respect to \( n_{i,j} \) and equating it to zero, we get
\[ \frac{dI}{dn_{i,j}} = -\log n_{i,j} - \delta_j - \frac{\lambda_i}{\eta_j} - \beta m \epsilon_i = 0, \] (49)

with \( \epsilon_i = v_i^2/2 + \Phi(r_i) \) as usual, and therefore
\[ n_{i,j} = e^{-\mu m (\epsilon_i - k_i \eta_j)}. \] (50)

Finally, we pass to a continuous description by giving every initial phase-space density level a small width \( \Delta \eta \). Then using the \( \tau(\eta) \) and \( g(\epsilon) \) function we replace
\[ \sum_j \rightarrow \frac{1}{\Delta \eta} \int d\eta, \] (51)
\[ \sum_i \rightarrow \frac{1}{\omega} \int d\epsilon g(\epsilon), \] (52)
\[ V_J \rightarrow \Delta \eta \tau(\eta_J), \] (53)
\[ n_{i,j} \rightarrow \frac{\Delta \eta}{m} \tau(\eta_J) \eta_j K(\epsilon_i, \eta_j). \] (54)

Plugging these replacements into equations (50), (45) and (46) and redefining \( m \beta \rightarrow \beta \), we recover the NK00 equations (38), (39) and (40).

This combinatorial formulation of the NK00 theory is very much along the lines of ordinary statistical mechanics of a classical Boltzmann gas. Indeed, if we replace the notion of phase elements with particles of equal mass and discard the constraint of conservation of phase-space volume (equation 40), we have a textbook derivation of the Boltzmann gas statistics. It is therefore not surprising that NK00 found that his equilibrium DF reduces to the well-known Maxwell–Boltzmann distribution in such a case. We do not agree, however, with Nakamura’s claim that this property is a proof for its correctness over the LB67 theory. This is because a collisionless relaxation is essentially a very different process from the collision-full relaxation that occurs in a Boltzmann gas, driven by different physical processes over different time-scales. However, as previously mentioned, deciding which theory is more correct is not the goal of this paper.

4.3 Analysing the two-level configuration

As was noted at the end of Section 4.1, the NK00 theory cannot handle a zero phase-space density directly. Therefore, it is not straightforward to analyse the equilibrium state that results from an initial water-bag configuration, as in this configuration there is one patch of phase-space density \( \eta_0 \) surrounded by an infinite volume of zero phase-space density. The way this can be done is to consider an initial state with two density levels \( \eta_0 \) and \( \eta_1 \) with corresponding volumes \( V_0 \) and \( V_1 \). The water-bag configuration is then recovered by taking the limits \( \eta_1 \rightarrow 0, V_1 \rightarrow \infty \) and \( \eta_1 V_1 \rightarrow 0 \).

To derive the equilibrium configuration of the two-level system, we use the fact that in this particular case the matrix \( \{n_{i,j}\} \) can be expressed in terms of \( f_i \), thereby greatly simplifying the end result. Let us then redefine the equilibrium equation for this particular case instead of using equations (38)–(41). Denoting by \( n_{i,0} \) and \( n_{i,1} \) the total number of elements of \( \eta_0 \) and \( \eta_1 \) that end up in the \( i \)th macrocell, the coarse-grained DF is given by
\[ f_i = \frac{m}{\omega} (n_{i,0} + n_{i,1}). \] (55)

Then using the conservation of phase-space volume constraint (46),
\[ n_{i,0} \frac{V_0}{N_0} + n_{i,1} \frac{V_1}{N_1} = \bar{\omega}, \] (56)
with \( N_0, N_1 \) being the total number of elements with densities \( \eta_0, \eta_1 \), given by equation (45), we express \( n_{i,0} \) and \( n_{i,1} \) in terms of \( f_i \):
\[ n_{i,0} = \frac{\bar{\omega}}{m} f_i - \frac{\bar{\omega}}{m} \eta_1, \] (57)
\[ n_{i,1} = \frac{\bar{\omega}}{m} \eta_0 - \frac{\bar{\omega}}{m} f_i - \frac{\bar{\omega}}{m} \eta_1. \] (58)

The energy constraint is given by equation (47), and the initial-conditions constraint is
\[ \sum_i n_{i,0} = N_0 = \frac{\eta_0 V_0}{m}. \] (59)

Notice that we need only the \( N_0 \) constraint since the \( N_1 \) constraint follows directly from requiring that the total phase-space volume occupied by the equilibrium system would be equal to \( V = V_0 + V_1 \). In fact, instead of equation (59), we can use an alternative total mass constraint, provided that the overall phase-space volume is conserved. This is done as follows. Expressing \( n_{i,0} \) in terms of \( f_i \) in equation (59) we get
\[ \omega \sum_i \frac{f_i}{\eta_0} - \frac{\eta_0}{\eta_1} = V_0, \] (60)
which gives us
\[ \sum_i \omega f_i/n \eta_0 - \eta_1 = V_0 \eta_0 - V_1 \eta_1. \] (61)

However, \( \sum_i \omega = V_0 + V_1 \) (conservation of total phase-space volume) and therefore we find
\[ \sum_i \omega f_i = V_0 \eta_0 + V_1 \eta_1 = M. \] (62)

Adding these constraints together with the appropriate Lagrange multipliers to the entropy, the expression that we need to maximize is
\[ I \propto \text{constant} - \sum_i n_{i,0} (\log n_{i,0} - 1) \]
\[ - \sum_i n_{i,1} (\log n_{i,1} - 1) + \bar{\beta} E + \bar{\mu} \sum_i \bar{\omega} f_i. \] (63)

Differentiating \( I \) with respect to \( f_i \) and equating to 0, we find
\[ - \frac{\eta_0}{\eta_0 - \eta_1} \log \left( \frac{\bar{\omega} f_i - \eta_1}{m \eta_0 - \eta_1} \right) \]
\[ + \frac{\eta_1}{\eta_0 - \eta_1} \log \left( \frac{\bar{\omega} \eta_0 - \bar{\omega} f_i}{m \eta_0 - \eta_1} \right) + \bar{\beta} \nu - \bar{\mu} \bar{\omega} = 0. \] (66)
After some trivial algebra and redefinition of the Lagrange multipliers $\beta$ and $\mu$ to $\tilde{\beta}$ and $\mu$, we obtain

$$
\frac{(f_i - \eta_1)/(\eta_0 - \eta_1)}{([\eta_0 - f_i]/(\eta_0 - \eta_1)\eta_1/\eta_0)} = e^{-\beta(\epsilon - \mu)}.
$$

(67)

Notice how the denominator provides an upper cut-off for $f_i$, as it forbids it from exceeding $\eta_0$.

Consider now the $\eta_1 \rightarrow 0$ limit. Seemingly, it would go into an isothermal sphere $f_i = \eta_0 e^{-\beta(\epsilon - \mu)}$, but this is not the case, as for every finite $\eta_1$, $f_i$ cannot exceed $\eta_0$. It is easy to see that the right limit is therefore

$$
\tilde{f}_i = \eta_0 e^{-\beta(\epsilon - \mu)},
$$

(68)

but this is not the case, as for every finite $\eta_1$, $f_i$ cannot exceed $\eta_0$. It is easy to see that the right limit is therefore

$$
\tilde{f}_{NK} = \begin{cases} 
\eta_0, & \epsilon < \mu, \\
\eta_0 e^{-\beta(\epsilon - \mu)}, & \epsilon \geq \mu.
\end{cases}
$$

(69)

This distribution is not the LB67 Fermi–Dirac distribution given by equation (16) or equation (17), but is what corresponds to that distribution in the NK00 theory. It is not smooth, and is exactly isothermal for energies $\epsilon > \mu$.

For the water-bag model in LB67, the condition that no two elements of phase density can overlap leads to statistics with exclusion, equivalent to the Fermi–Dirac problem. It is not clear to us how Nakamura’s formulation could obtain the Fermi–Dirac statistics.

### 5 Non-Transitivity in NK00 Theory of Violent Relaxation

Having found the equilibrium configuration of the water-bag initial configuration in the NK00 theory, we are in a position to test the theory’s transitivity. The procedure for that is identical to the one that was used in the LB67 case, in Sections 3.2 and 3.3, and therefore will not be repeated. Instead, we shall first describe how the $s_1$ and $s_3$ configurations are found and then how $s_2$ is compared to the $s_3$ configuration.

To find the $s_1$ and $s_3$ configurations, we must first find the gravitational potential of the DF in equation (69) in a sphere of radius $R$, and then fix $\beta$ and $\mu$ so that the overall energy and mass will be equal to $E$ and $M$. Additionally, just as in the LB67 case, we assume that the final equilibrium state is spherical and therefore the Poisson equation for $\Phi(r)$ is

$$
\frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{d\Phi}{dr} \right) = (4\pi\gamma)^2 G \int_0^\infty v^2 f_{NK}[v^2/2 + \Phi(r)] \, dv.
$$

(70)

Passing from $\Phi(r)$ to the dimensionless $\psi(r)$ by using

$$
\psi(r) \equiv \beta[\mu - \Phi(r)],
$$

(71)

equation (70) simplifies to

$$
-\frac{1}{\beta r^2} \frac{d}{dr} \left( r^2 \frac{d\psi}{dr} \right) = (4\pi\gamma)^2 G \eta_0
$$

(72)

$$
\times \int_0^\infty v^2 \left\{ 1, \frac{\beta v^2/2 < \psi(r)}{\beta v^2/2 \geq \psi(r)} \right\}
$$

(73)

Finally, changing variables $r \rightarrow x$,

$$
x \equiv \left( \frac{16\pi^2 \sqrt{2} G \eta_0}{\beta^{1/2}} \right)^{1/2} r,
$$

(74)

and using the error function

$$
\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \, dt,
$$

the ordinary differential equation for $\psi(x)$ is

$$
-\frac{1}{x^2} \frac{d}{dx} \left( x^2 \frac{d\psi}{dx} \right) = I(\psi),
$$

(75)

with

$$
I(\psi) = \begin{cases} 
\frac{1}{2} \sqrt{\pi} e^\psi, & \psi < 0, \\
\frac{1}{2} \psi^{3/2} + \sqrt{\psi} + \frac{1}{2} \sqrt{\pi} e^\psi(1 - \text{erf}(\sqrt{\psi})), & \psi \geq 0.
\end{cases}
$$

(76)

To integrate this equation, we must first set its initial condition. We let $\psi(0) = \psi_0$ be a free parameter, and $\psi'(0) = 0$, since in a spherical system the gravitational force vanishes at the centre. Then once $\psi(x)$ is (numerically) found, we fix $\mu$ by requiring that $\Phi(R) = -GM/R$. In this way we can find the gravitational potential, and thereafter the total mass and energy for any given $\beta$ and $\psi_0$. The last step is to find the right $\beta$ and $\psi_0$ that would give us $M$ and $E$.

Practically, instead of looking for $\beta$ and $\psi_0$ for a specific choice of $E$ and $M$ in configurations $s_1$ and $s_3$, we have picked two different values of $\beta$ and two corresponding values of $\psi_0$ to satisfy the total mass constraint. Once $\psi_0$ is fixed it also defines a total energy. We chose $\beta_1 > \beta_3$ in order to obtain $E_1 < E_2$.

Let us now see how the $s_2 \leftrightarrow s_3$ comparison can be done. According to Section 4, $f_{s2}(\epsilon)$ is determined by the following set of equations:

$$
K(\epsilon, \eta) = \tilde{\omega}^{-1} e^{-\beta(\epsilon - \delta) - \lambda(\epsilon)/\sqrt{\eta}},
$$

(77)

$$
1 = \int_{\Phi(0)}^\infty g_{s2}(\epsilon) K(\epsilon, \eta) \, d\epsilon,
$$

(78)

$$
1 = \int_0^\infty \tau_s(\eta) K(\epsilon, \eta) \, d\eta,
$$

(79)

$$
f_{s2}(\epsilon) = \int_0^\infty \tau_s(\eta) K(\epsilon, \eta) \, d\eta.
$$

(80)

Additionally, we know that $f_{s3}(\epsilon)$ is given by

$$
f_{s3}(\epsilon) = \begin{cases} 
\eta_0, & \epsilon < \mu_3, \\
\eta_0 e^{-\beta_3(\epsilon - \mu_3)}, & \epsilon \geq \mu_3.
\end{cases}
$$

(81)

and $g_{s1}(\epsilon)$ and $\Phi_s(\epsilon)$ have been found as described above. Assuming that $s_2 = s_3$, we replace these functions with $f_{s2}(\epsilon)$, $g_{s2}(\epsilon)$ and $\Phi_{s2}(\epsilon)$ in equations (78)–(80), and differentiate equation (80) with respect to $\epsilon$. Using equation (79) we get

$$
f_{s2}'(\epsilon) = -\beta_2 f_{s2}(\epsilon) - \lambda'(\epsilon),
$$

(82)

and therefore

$$
\lambda(\epsilon) = C + \beta_2 \int f_{s2}(\epsilon') \, d\epsilon' - f_{s2}(\epsilon) \equiv C + \lambda_0(\epsilon),
$$

(83)

with $C$ some unknown integration constant. The integral in $\lambda_0(\epsilon)$ can be done analytically, yielding

$$
\lambda_0(\epsilon) = \begin{cases} 
\frac{\eta_0[(\beta_2/\beta_3) - 1 + \beta_2(\mu_3 - \epsilon)]}, & \epsilon < \mu_3, \\
\frac{\eta_0[(\beta_2/\beta_3) - 1] e^{-\beta_2(\epsilon - \mu_3)}}, & \epsilon \geq \mu_3.
\end{cases}
$$

(84)
Then from equation (78) we find that
\[
e^\delta_{\eta} = e^{\beta \delta_{\eta}} = \int_{\Phi_0(0)}^{\infty} g_\delta(\epsilon) e^{-\beta \epsilon \lambda_\eta(\epsilon)/\eta} \, d\epsilon
\]
\[
\equiv e^{-\beta \delta_{\eta}/\eta} D_\delta(\eta), \tag{85}
\]
and therefore
\[
K(\epsilon, \eta) = \tilde{\omega}^{-1} e^{-\beta \epsilon \lambda_\eta(\epsilon)/\eta} D_\delta(\eta).	ag{86}
\]
\[
= e^{-\beta \epsilon \lambda_\eta(\epsilon)/\eta} D_\delta(\eta).	ag{87}
\]
Notice how the unknown integration constant \(C\) and the dimensional constant \(\tilde{\omega}\) are cancelled out.

Next we calculate \(f_3(\epsilon)\) using equation (80) and fix \(\beta_2\) such that the energy of the system is \(E + \Delta E\). Once \(\beta_1\) is fixed, \(K(\epsilon, \eta)\) is completely resolved in terms of \(s_1\) and \(s_3\) functions, and we can check whether it solves the maximum-entropy equations by plugging it into equation (79).

Finally, we should note that, as in the LB67 case, the \(\tau_1(\eta)\) function has a strong peak at \(\eta = \eta_0\) as a result of the degeneracy. Here, however, this peak is proportional to \(\delta(\eta - \eta_0)\) as \(f_3(\epsilon) = \eta_0\) for every \(\Phi_3(0) < \epsilon < \mu_1\). The prefactor in front of this delta function is \(V_{\text{deg}}\) – the volume of phase space for which \(\Phi_3(0) < \epsilon < \mu_1\), which can be easily calculated from equation (30). Therefore, to perform the integration over \(\tau_1(\eta)\) in equations (79) and (80) numerically, we first calculate the smooth contribution from the \(\tau_1(\eta)\) with \(\eta < \eta_0\), and then add the delta-function contribution by evaluating the integrands at \(\eta = \eta_0\) and multiplying them by \(V_{\text{deg}}\).

5.1 Numerical results

As in the LB67 case, the \(s_0\) state was constructed as a water-bag configuration with \(\eta\) given by equation (15), \(M = 1\), \(R = 1\) and \(G = 1\). The \(s_1\) and \(s_3\) configurations were then chosen as described in the previous subsection, by fixing \(\beta\) and varying \(\psi_3\) until the total mass constraint was satisfied. The main numerical parameters of these configurations are summarized in Table 2.

Fig. 6 shows the density profiles of the \(s_1\) and \(s_3\) configurations. As expected, the hotter system, \(s_3\), has a core with a lower density than the \(s_1\) system. Fig. 7 shows the DF of these systems, and Fig. 8 compares the \(M(V)\) functions of these two states, showing that \(M_3(V) \leq M_1(V)\) for every \(V\) – and therefore the transition \(s_1 \rightarrow s_3\) is allowed.

Finally, to compare the \(s_2\) configuration to the \(s_3\) configuration, we have varied \(\beta_2\) in the range \(0.1 < \beta_2 < 100\) until we obtained \(E_2 = E_3 = -4.306\) with \(\beta_2 = 0.6944\). This was the only solution in that range, and we believe that it is the only physical solution in general.

Table 2. The numerical parameters that specify the \(s_1\) and \(s_3\) states in the NK00 double-relaxation experiment. Here \(E\) is the total energy, \(\beta\) and \(\mu\) are the Lagrange multipliers in equation (69), and \(\Phi(0)\) is the gravitational potential at \(r = 0\).

<table>
<thead>
<tr>
<th>State</th>
<th>(E)</th>
<th>(\beta)</th>
<th>(\mu)</th>
<th>(\Phi(0))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(s_1)</td>
<td>-6.405</td>
<td>3.0</td>
<td>-15.30</td>
<td>-35.08</td>
</tr>
<tr>
<td>(s_3)</td>
<td>-4.306</td>
<td>0.68</td>
<td>-14.08</td>
<td>-27.80</td>
</tr>
</tbody>
</table>
The energy $E_2$ of $k_2$, together with the energy $E_3 = -4.306$. The value of $\beta_2$ for which the two energies agree is $\beta_2 = 0.6944$. No other solutions were found for $\beta_2$ up to 100, and we therefore believe that this is the only physical solution.

Once $\beta_2$ was found, we used the expression of $K(\epsilon, \mu)$ in equation (85), to calculate the right-hand side (RHS) of equation (79) and compare it to 1. Fig. 10 shows this comparison. The disagreement between the RHS and 1 is sometimes as high as three orders of magnitude – much higher than any possible numerical error. We therefore conclude that also the NK00 theory is non-transitive.

### 6 CONCLUSIONS

In this paper we have demonstrated that the statistical-mechanical theories of violent relaxation by Lynden-Bell and Nakamura are both non-transitive. This non-transitivity is a result of the phase mixing that occurs when the system relaxes; as the fine-grained phase-space density filaments become thinner and thinner, the system is better described in terms of the coarse-grained phase-space density. Any further relaxation of the system should be therefore considered in terms of the coarse-grained phase-space density – which as we have seen would yield different results from a prediction that is based on the initial fine-grained phase-space density. This is a worrying aspect of these theories, as it is easy to imagine a scenario where part of the system mixes, then fluctuates, and then mixes once again. The predictions of the theory, based on the fine-grained density, will then give us a wrong result.

In some sense we have been breaking into an open door. Even without considering the non-transitivity of the theories, they are plagued by severe problems. There exist two equally plausible ways of discretizing phase space, one with equal volume elements and one with equal mass elements, which yield two different results. More importantly, the ability of the theories to predict the final outcome of a violent relaxation process is very limited. Indeed, as was mentioned in Section 2, the most important reason for this is that violent relaxation is almost never complete; the fluctuations of the gravitational potential die much faster for the system to settle in the most probable state.

Nevertheless, we believe that these difficulties and ambiguities in exactly how to do the statistical mechanics of the collisionless Boltzmann equation teach us an important lesson. The non-transitivity that we have shown is a sign that a kinetic description of violent relaxation is probably incomplete, as the equilibrium is dependent on the evolutionary path of the system. Instead, what is probably needed is a dynamical approach to the problem. Indeed, most of the above difficulties are circumvented if, instead of aiming to derive a universal most probable state, we reduce our aim to that of finding an appropriate and useful evolution equation for the coarse-grained $\bar{f}$.

An interesting attempt to find such an equation was taken by Chavanis (1998), who used the maximal entropy-production principle (MEPP) to obtain a closed equation for $\bar{f}$. His analysis, however, uses the initial fine-grained $\tau(\eta)$ to define the (Lynden-Bell) entropy rather than the instantaneous, coarse-grained $\tau(\eta)$, whichaccording to the above discussion is more correct. Indeed, in a more recent paper, Chavanis & Bouchet (2005) use an alternative ‘deterministic approach’ to describe the evolution of $\bar{f}$. They define the coarse-grained distribution function (DF) by a convolution of the fine-grained DF with a Gaussian window and obtain a closed equation for $\bar{f}$ that gives different predictions from the MEPP model. Their study, as well as this work, show that the notion of coarse-graining is a complicated concept. Derivation of a useful dynamical equation for $\bar{f}$ thus remains a challenging open problem.

### ACKNOWLEDGMENTS

We thank Peter Johansson for his help with the manuscript. This work was supported by a Marie-Curie Individual Fellowship of the European Community No. HPMF-CT-2002-01997.

### REFERENCES


This paper has been typeset from a \TeX\ file prepared by the author.