On the well-posedness of equations for smoothed phase space distribution functions and irreversibility in classical statistical mechanics



2001 J. Phys. A: Math. Gen. 34 4641

(http://iopscience.iop.org/0305-4470/34/22/305)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **34** (2001) 4641–4651 www.iop.org/Journals/ja PII: S0305-4470(01)20209-1

# On the well-posedness of equations for smoothed phase space distribution functions and irreversibility in classical statistical mechanics

**C B Muratov**

New Jersey Institute of Technology, Department of Mathematical Sciences, University Heights, Newark, NJ 07102, USA

Received 7 December 2000, in final form 9 April 2001

**Abstract**

We obtain an exact equation for a smoothed phase space distribution function for a system of $N$ non-relativistic particles of unit mass obeying Hamiltonian dynamics. We show that this equation is well posed in only one time direction in the sense of the continuous dependence of the solution on the initial data. We interpret the ill-posedness of this equation in the backward time direction as the manifestation of irreversibility of the observable statistical quantities for systems obeying time-reversible dynamical laws.

PACS numbers: 0520, 4520

## 1. Introduction

Perhaps one of the most fundamental problems in physics is understanding the nature of irreversibility of physical processes. After more than 100 years since its formulation it remains a mystery how in the world where all the dynamical laws are time reversible we have overwhelming evidence of irreversible phenomena. Since the works of Maxwell and Boltzmann [1], this problem has been the subject of the debate which goes on to the present day (for a recent discussion, see [2, 3] and references therein).

It should be pointed out that from the physical point of view the contradiction between the reversibility of the dynamical laws of classical mechanics and the irreversibility of the observed physical phenomena has essentially been resolved already in the works of Maxwell and Boltzmann. It arises due to one's essential inability to control the initial conditions for the dynamical equations exactly. On the timescale of the observations a small uncertainty in the initial conditions may result in large deviations in the observable quantities, so one has to perform averaging to extract information which does not appear to be strongly affected by this uncertainty. Therefore, one has to go to the probabilistic (statistical) description which predicts the dynamics of only certain macroscopic quantities (for a thorough discussion on this subject, see [2]).

0305-4470/01/224641+11$30.00  

The foundation for the link between the dynamics and the statistical mechanics was laid in the mid-1940s in the works of Bogoliubov, Born, Green and Kirkwood [4]. They recognized the importance of introducing random initial conditions in passing to the statistical description in a general dynamical setting. This allowed the derivation of the irreversible kinetic equations from the original dynamical equations in certain limits [5]. At roughly the same time Landau derived the irreversible time-dependent density correlation function and the famous Landau damping for the ideal gas and plasma [6]. However, the transition to the statistical description introduced by Bogoliubov *et al* turned out to lead once again to the time-reversible equations. This creates an apparent mathematical paradox, since from the mathematical point of view reversible equations fundamentally *cannot* lead to irreversible behaviours.

The reason for the time reversibility in the statistical description introduced by Bogoliubov *et al* is that it is in one-to-one correspondence with the original system's dynamics. In other words, this statistical description does not involve any loss of the dynamical information, so it is still a microscopic description and does not exhibit macroscopic irreversible behaviour. On the other hand, if the system's dynamics is mixing, the above-mentioned statistical description becomes more and more complicated with time. So, in the physical context one has to perform some kind of regularization or smoothing in order to extract quantities which correspond to those measured in the experiments. This smoothing introduces loss of information about small regions in the phase space and therefore should lead to the breakdown of time reversibility for the smoothed quantities.

In the present paper we show that a particular kind of smoothing which formally does not produce any loss of information leads to a closed equation for the smoothed probability density, which is irreversible.

## 2. Smoothed distribution function

We begin with a dynamical system of $N$ particles of unit mass with coordinates $x_n$ and momenta $p_n$ obeying Hamilton equations

$$\dot{x}_n = \frac{\partial H}{\partial p_n} \qquad \dot{p}_n = -\frac{\partial H}{\partial x_n} \tag{1}$$

with the Hamiltonian

$$H = \sum_n \frac{p_n^2}{2} + U(x_1, x_2, \ldots, x_N). \tag{2}$$

A statistical description of such a system is obtained by introducing a random distribution of initial data in the phase space. Then, from the Liouville theorem one obtains that the phase space distribution function $f(x, p, t)$, where $x$ and $p$ denote the $N$-component vectors, evolves according to the equation

$$\frac{\partial f}{\partial t} + \{f, H\} = 0 \tag{3}$$

where $\{f, H\}$ is the Poisson bracket. Equation (3) is time reversible. This immediately follows from the symmetry of this equation with respect to the transformation

$$x \to x \qquad p \to -p \qquad t \to -t. \tag{4}$$

This means that $f$ must retain the memory of the initial conditions for all times.

To get to the source of the distinction between the reversible microscopic behaviour and the irreversible macroscopic one, we need to go back to the derivation of equation (3) in the physical context. Two issues have to be considered here: preparation of the initial conditions and measurement of the state of the system at a given moment. Let us address the issue of *measurement* first. It is clear that in any experiment the state of the system, that is, the coordinate of a point in the phase space, cannot be measured exactly, but will have some uncertainty. In a measurement, one is certain to lose information about points that are very close in the phase space. Therefore, if at a moment $t$ an experimentalist is given a distribution $f(x, p, t)$ of the points in the phase space, he will actually measure a function $\bar{f}$, where

$$\bar{f}(x, p, t) = \int J_\epsilon(x - x', p - p') f(x', p', t)\, \mathrm{d}x'\, \mathrm{d}p' \tag{5}$$

where $J_\epsilon$ is the probability density of the error in the measurement and $\epsilon$ is the measure of the error, with $J_\epsilon$ approaching a $\delta$-function as $\epsilon \to 0$. This formula implies averaging over the realizations of the phase space points and repeated measurements. In other words, $f$ is not a physically *observable* quantity.

Let us now turn to the issue of the *preparation* of the initial conditions. In order to get the system into a given point of the phase space, an experimentalist starts from an arbitrary state of the system, that is, an arbitrary phase space point. He then measures the state of the system and applies time-dependent *control* to move the phase space point towards the target point. Once the desired degree of closeness between the position of the phase space point and the target is achieved, the control is removed and the system is allowed to evolve according to equations (1). It is clear that because of the uncertainty in the measurements, one can never achieve such a control that will move the point *exactly* to the target position. Instead, it is natural to expect a Gaussian distribution of phase space points in the $\epsilon$-neighbourhood of the target point. Then, on the scale of $\epsilon$ or greater one can approximate an arbitrary smooth distribution function by sampling different points of the phase space. In other words, the initial condition $f_0(x, p) = f(x, p, 0)$ for equation (3) should have the form

$$f_0(x, p) = \sum_i \frac{w_i}{(2\pi c\epsilon^2)^N} \exp\left(-\frac{(p - p_i)^2}{2c\epsilon^2} - \frac{(x - x_i)^2}{2c\epsilon^2}\right) \tag{6}$$

where $w_i$ is the weight of an $i$th sample point, $x_i$, $p_i$ are its phase space coordinates and $c$ is a positive constant.

Given the initial conditions of the form of equation (6), we wish to study the observed (macroscopic) distribution function $\bar{f}$ generated by the unobservable (microscopic) distribution function $f$ which is the solution of equation (3) with these initial conditions. One issue here is whether one can obtain a closed description for the dynamics of $\bar{f}$. If $\epsilon$ is small, one can use equation (3) to write down an infinite hierarchy of equations for $\bar{f}$ and the higher moments. One can further introduce a closure to this hierarchy assuming that $f$ is a sufficiently smooth function and thus obtain a closed equation for $\bar{f}$ and arbitrary $J_\epsilon$.

Another issue is whether the obtained equation for $\bar{f}$ is well posed and whether it is valid for all times. The latter is especially important, since the irreversible behaviours should be manifested at long times. It is clear that the closure approximations discussed above should in general break down at long times for mixing systems since the original distribution function $f$ becomes highly oscillatory at long times.

We circumvent all these difficulties by considering a special form of $J_\epsilon$:

$$J_\epsilon(x - x', p - p') = (2\pi\epsilon^2)^{-N/2}\delta(x - x')\mathrm{e}^{-(p-p')^2/2\epsilon^2}. \tag{7}$$

This special choice of $J_\epsilon$ has an important property that will be used below.

**Lemma 2.1.** *Let $J_\epsilon$ be given by equation (7), $f(x, p, t)$ bounded, and $\bar{f}$ given by equation (5). Then*

$$(2\pi\epsilon^2)^{-N/2} \int (p_n - p_n')\mathrm{e}^{-(p-p')^2/2\epsilon^2} f(x, p', t)\, \mathrm{d}p' = -\epsilon^2 \frac{\partial \bar{f}}{\partial p_n}. \tag{8}$$

**Proof.** The proof is obtained by differentiating $\bar{f}$ in equation (5) with $J_\epsilon$ from equation (7) with respect to $p_n$. □

It turns out that for this particular form of $J_\epsilon$ one can obtain a closed equation for $\bar{f}$ exactly which is thus valid for all times. This result is formulated in the following:

**Theorem 2.2.** *Let $f(x, p, t) \in C^1$ be a solution of equation (3) with the Hamiltonian from equation (2), vanishing outside a compact subset of the phase space. Then the averaged distribution function $\bar{f}(x, p, t)$ obtained from equation (5) with $J_\epsilon$ given by equation (7) satisfies*

$$\frac{\partial \bar{f}}{\partial t} + \{\bar{f}, H\} = -\epsilon^2 \sum_n \frac{\partial^2 \bar{f}}{\partial x_n \partial p_n}. \tag{9}$$

**Proof.** Equation (3) with the Hamiltonian from equation (2) is explicitly written as follows:

$$\frac{\partial f}{\partial t} + \sum_n p_n \frac{\partial f}{\partial x_n} - \sum_n \frac{\partial U}{\partial x_n}\frac{\partial f}{\partial p_n} = 0. \tag{10}$$

Let us multiply this equation by $(2\pi\epsilon^2)^{-N/2} \mathrm{e}^{-(p-p')^2/2\epsilon^2}$ and integrate over $p$. If we then add and subtract a term $-(2\pi\epsilon^2)^{-N/2} \int p'_n \mathrm{e}^{-(p-p')^2/2\epsilon^2} f(x, p, t)\,\mathrm{d}p$ and use equation (5), we obtain an equation that relates $\bar{f}(x, p', t)$ and $f(x, p, t)$

$$\begin{aligned}\frac{\partial \bar{f}}{\partial t} &+ \sum_n p'_n \frac{\partial \bar{f}}{\partial x_n} - (2\pi\epsilon^2)^{-N/2} \sum_n \frac{\partial U}{\partial x_n} \int \mathrm{e}^{-(p-p')^2/2\epsilon^2} \frac{\partial f}{\partial p_n}\,\mathrm{d}p \\ &= -(2\pi\epsilon^2)^{-N/2} \sum_n \int (p_n - p'_n) \mathrm{e}^{-(p-p')^2/2\epsilon^2} \frac{\partial f}{\partial x_n}\,\mathrm{d}p.\end{aligned} \tag{11}$$

Let us now integrate the third term in the left-hand side of this equation by parts, taking into account that the surface integral vanishes, and swap $p$ and $p'$. We obtain for $\bar{f}(x, p, t)$

$$\begin{aligned}\frac{\partial \bar{f}}{\partial t} &+ \sum_n p_n \frac{\partial \bar{f}}{\partial x_n} + (2\pi\epsilon^2)^{-N/2}\epsilon^{-2} \sum_n \frac{\partial U}{\partial x_n} \int (p_n - p'_n) \mathrm{e}^{-(p-p')^2/2\epsilon^2} f(x, p', t)\,\mathrm{d}p' \\ &= (2\pi\epsilon^2)^{-N/2} \sum_n \int (p_n - p'_n) \mathrm{e}^{-(p-p')^2/2\epsilon^2} \frac{\partial f(x, p', t)}{\partial x_n}\,\mathrm{d}p'.\end{aligned} \tag{12}$$

Now, applying the result of lemma 2.1 to the integrals, we arrive at equation (9). □

**Remark 2.1.** Equation (9) is valid for all $\epsilon$, so one does not have to restrict oneself to the case of $\epsilon$ small.

**Remark 2.2.** Note that a similar kind of smoothing in a different context was introduced by Klimas as a method for solving numerically the Vlasov equation [7].

**Remark 2.3.** The derivation of equation (9) did not rely on the fact that the system is conservative, so it can be extended to the case when the interaction potential $U$ in the Hamiltonian is time dependent.

Observe that for $J_\epsilon$ from equation (7) the correspondence between $f$ and $\bar{f}$ is one-to-one since the Fourier transform of $J_\epsilon$ never vanishes, so formally there is no information loss when going from $f$ to $\bar{f}$ and the solutions of equation (9) are in *exact* correspondence with those of equation (3). What happens is that the information about the shorter and shorter distances gets progressively more hidden in $\bar{f}$, so one can talk about the effective loss of information about *arbitrarily* short distances. Note, however, that this is different from coarse-graining, where the analogue of $\bar{f}$ would be obtained by averaging over finite regions of the phase space, thus leading to the actual information loss.

## 3. Well-posedness of the initial-value problem

Equation (9) still formally possesses symmetry with respect to the transformation in equation (4). It has to be solved as an initial-value problem. Here, however, we come to the difficulty that this problem is actually ill posed. This is due to the fact that now we have a second-order hyperbolic differential operator in the right-hand side of equation (9), which in general leads to the blowup of the short-scale modes of $\bar{f}$ in arbitrarily short times. Indeed, let us introduce the Fourier transform

$$\bar{f}_{kl}(t) = \int \mathrm{e}^{\mathrm{i}k\cdot x+\mathrm{i}l\cdot p}\bar{f}(x, p, t)\,\mathrm{d}x\,\mathrm{d}p \tag{13}$$

where $k$ and $l$ are $N$-component vectors. For large $|l|$ and $|k|$ the right-hand side of equation (9) will dominate, so we will have approximately $\bar{f}_{kl}(t) \simeq \bar{f}_{kl}(0)\mathrm{e}^{\epsilon^2 k\cdot lt}$. For $k = l$ we will have $\bar{f}_{kl} \sim \mathrm{e}^{\epsilon^2 k^2 t}$. The difficulty here is similar to the one that appears in the problem of negative diffusion. So, we have to make more precise what we actually mean by equation (9).

To do this, we should recall that the physically relevant initial conditions for equation (3) and, therefore, equation (9) must have the form given by equation (6). These initial conditions possess a very high degree of regularity. It is expressed in the very fast decay of $\bar{f}$ in the Fourier space: $|\bar{f}_{kl}(0)|^2 \sim \mathrm{e}^{-c\epsilon^2(k^2+l^2)}$. In this situation the growth of the short-scale modes can actually be controlled for finite times. Indeed, with these initial conditions we will have $|\bar{f}_{kl}(t)|^2 \sim \mathrm{e}^{2\epsilon^2 k\cdot lt-c\epsilon^2(k^2+l^2)} \leqslant \mathrm{e}^{-\epsilon^2(c-t)(k^2+l^2)}$.

More formally, we require that the initial condition $\bar{f}_0$ be in the Banach space $L^2_c$:

**Definition 3.1.** *Let $L^2_c$ denote the Banach space of functions with the norm*

$$||\bar{f}||^2_{L^2_c} = \int \frac{\mathrm{d}k}{(2\pi)^N}\frac{\mathrm{d}l}{(2\pi)^N}\mathrm{e}^{c\epsilon^2(k^2+l^2)}|\bar{f}_{kl}|^2 \tag{14}$$

*where $\bar{f}_{kl}$ are defined as in equation (13), and $c > 0$ is a constant.*

The weighted $L^2$-norm ensures sufficiently rapid decay of $|\bar{f}_{kl}|$ for large $|k|$, $|l|$.

**Lemma 3.2.** *The relationships*

$$L^2_c \subseteq L^2_{c'} \subseteq L^2 \tag{15}$$

*hold for $c \geqslant c' \geqslant 0$.*

**Proof.** The proof follows from the obvious estimates

$$\begin{aligned}\int \frac{\mathrm{d}k}{(2\pi)^N}\frac{\mathrm{d}l}{(2\pi)^N}|\bar{f}_{kl}|^2 &\leqslant \int \frac{\mathrm{d}k}{(2\pi)^N}\frac{\mathrm{d}l}{(2\pi)^N}\mathrm{e}^{c'\epsilon^2(k^2+l^2)}|\bar{f}_{kl}|^2\\ &\leqslant \int \frac{\mathrm{d}k}{(2\pi)^N}\frac{\mathrm{d}l}{(2\pi)^N}\mathrm{e}^{c\epsilon^2(k^2+l^2)}|\bar{f}_{kl}|^2\end{aligned} \tag{16}$$

for $c \geqslant c' \geqslant 0$. □

Let us introduce the following notation:

$$\mathcal{L} = -\epsilon^2\sum_n \frac{\partial^2}{\partial x_n \partial p_n}. \tag{17}$$

The operator $\mathcal{L}$ has the following property:

**Lemma 3.3.** *$\mathcal{L}$ is a bounded operator from $L^2_c$ to $L^2_{c'}$ with $c > c'$. Moreover, the following estimate holds:*

$$||\mathcal{L}\bar{f}||_{L^2_{c'}} \leqslant \frac{1}{c-c'}||\bar{f}||_{L^2_c}. \tag{18}$$

**Proof.** This can be seen from the following inequalities:

$$
\begin{aligned}
||\mathcal{L}\bar{f}||^2_{L^2_{c'}} &\leqslant \int \frac{\mathrm{d}k}{(2\pi)^N}\frac{\mathrm{d}p}{(2\pi)^N}\epsilon^4 k^2 l^2 \mathrm{e}^{c'\epsilon^2(k^2+l^2)}|\bar{f}_{kl}|^2 \\
&= \int \frac{\mathrm{d}k}{(2\pi)^N}\frac{\mathrm{d}p}{(2\pi)^N}\epsilon^4 k^2 l^2 \mathrm{e}^{-\epsilon^2(c-c')(k^2+l^2)}|\bar{f}_{kl}|^2 \mathrm{e}^{c\epsilon^2(k^2+l^2)} \\
&\leqslant \frac{1}{(c-c')^2}\int \frac{\mathrm{d}k}{(2\pi)^N}\frac{\mathrm{d}p}{(2\pi)^N}\mathrm{e}^{c\epsilon^2(k^2+l^2)}|\bar{f}_{kl}|^2 \\
&= \frac{1}{(c-c')^2}||\bar{f}||^2_{L^2_c}.
\end{aligned}
\tag{19}
$$

□

Let us also introduce the fundamental solution $\mathcal{E}(t)$ of equation (3):

$$
\mathcal{E}(t) = \theta(t)\delta\big(x-\bar{x}(t)\big)\delta\big(p-\bar{p}(t)\big)
\tag{20}
$$

where $\bar{x}(t)$, $\bar{p}(t)$ satisfy equations (1) with the initial conditions $\bar{x}(0) = x'$, $\bar{p}(0) = p'$; $\theta(t)$ is the Heaviside step. Thus, the solution of equation (3) with the initial data $f(x, p, 0) = f_0(x, p)$ is $f(x, p, t) = \mathcal{E}(t) * f_0(x, p)$, where '$*$' denotes the convolution in the phase space. What we are going to show below is that under certain assumptions on $\mathcal{E}(t)$ and $\bar{f}_0$, the initial-value problem for equation (9) is well posed in a certain sense for finite times.

**Theorem 3.4.** *Let $c > 0$, $\bar{f}_0 \in L^2_c \cap C^\infty(\mathbb{R}^{2N})$, $\mathcal{E}(t) * \bar{f} \in C^\infty(\mathbb{R}^{2N}) \times C^1[0, T]$ for $\bar{f} \in C^\infty(\mathbb{R}^{2N})$, and for any $0 < r \leqslant c$*

$$
||\mathcal{E}(t) * \bar{f}||_{L^2_{\alpha(t)r}} \leqslant M||f||_{L^2_r} \qquad 0 \leqslant t \leqslant T
\tag{21}
$$

*where $M$ and $T$ are constants independent of $r$, and $0 < \alpha(t) < 1$ is a function also independent of $r$. Also, let $\alpha(t)$ be Lipschitz continuous on the interval $0 \leqslant t \leqslant T$. Then for any $0 < c' < c$ there exists a unique solution $\bar{f}(x, p, t)$ of equation (9) with the initial data $\bar{f}(x, p, 0) = \bar{f}_0(x, p)$ that is uniformly bounded in $L^2_{c'}$ on a sufficiently small interval $0 \leqslant t \leqslant T'$.*

**Proof.** We use the standard perturbation argument to prove the existence and uniqueness of the solution of equation (9). We first write the equivalent integral equation

$$
\bar{f}(x, p, t) = \mathcal{E}(t) * \bar{f}_0(x, p) + \int_0^t \mathcal{E}(t-t') * \mathcal{L}\bar{f}(x, p, t')\,\mathrm{d}t'.
\tag{22}
$$

Any solution of equation (22) that lies in $C^2(\mathbb{R}^{2N}) \times C^1[0, T']$ is a classical solution of equation (9). Since $L^2_{c'} \subset H^\infty(\mathbb{R}^{2N})$, for each function $\bar{f} \in L^2_{c'}$ there is a unique function $\tilde{f} \in C^\infty(\mathbb{R}^{2N})$, such that $\tilde{f} = \bar{f}$ almost everywhere. By this and the assumption of smoothness of $\mathcal{E}(t)*$, for any solution $\bar{f}(\cdot, t) \in L^2_{c'}$ of equation (22) there is a unique classical solution of equation (9).

The formal solution of equation (22) has the form

$$
\begin{aligned}
\bar{f}(x, p, t) = \mathcal{E}(t) * \bar{f}_0 + \sum_{n=1}^{\infty}\int_0^t \mathrm{d}t_1 \int_0^{t_1}\mathrm{d}t_2 \cdots \int_0^{t_{n-1}}\mathrm{d}t_n \\
\times\mathcal{E}(t-t_1) * \mathcal{L}\mathcal{E}(t_1-t_2) * \mathcal{L}\cdots\mathcal{E}(t_n) * \bar{f}_0(x, p).
\end{aligned}
\tag{23}
$$

Let us show that under the assumptions of the theorem this series converges in $L^2_{c'}$ on a sufficiently small time interval. Consider the $n$th term $\bar{f}_n$ in the series in equation (23).

According to equation (21) and lemma 3.3, the operator $\mathcal{E}(t-t_1)*\mathcal{L}\mathcal{E}(t_1-t_2)*\mathcal{L}\cdots\mathcal{E}(t_n)*$ is a bounded operator from $L^2_c$ to $L^2_{c_n}$, where

$$c_n = c\left(1-\frac{at}{n}\right)^n \alpha(t-t_1)\alpha(t_1-t_2)\cdots\alpha(t_n) \tag{24}$$

and $a > 0$ is a fixed constant. Here we treated the operator $\mathcal{L}$ as an operator from $L^2_r$ to $L^2_{r(1-at/n)}$ for $0 < r \leqslant c$.

For any value of $a$ it is possible to choose a sufficiently small interval $0 \leqslant t \leqslant T'$ such that $c_n \geqslant c'$ for any $n$. Indeed, let us denote the Lipschitz constant for $\alpha(t)$ as $K$. Obviously, $\alpha(0) = 1$. Then, for sufficiently small $T' \leqslant T$ we will have $\alpha(t) \geqslant 1 - Kt \geqslant \frac{1}{2}$ for all $t$ in the interval. Also, $T'$ can be chosen so small that $1 - \frac{at}{n} \geqslant \frac{1}{2}$ for all $n$ as well. Let us take the natural logarithm of both sides of equation (24). It is not difficult to verify that for any $\frac{1}{2} \leqslant x \leqslant 1$ we have $\ln x \geqslant -2(1-x)$ for all $x$ in this interval. Then, taking into account that $0 \leqslant t_n \leqslant t$ for all $t_n$, we can write

$$\ln c_n \geqslant \ln c - 2at - 2K(t - t_1 + t_1 - t_2 + \cdots + t_n) \tag{25}$$

so we have

$$c_n \geqslant c\mathrm{e}^{-2(a+K)t}. \tag{26}$$

From this formula it is clear that for any $0 < c' < c$ and any $a$ it is possible to choose the value of $T'$ such that $c_n > c'$ for all $0 \leqslant t \leqslant T'$ and any $n$. This and the fact that $\mathcal{E}(t) * \bar{f}_0 \in L^2_{c\alpha(t)} \subseteq L^2_{c(1-Kt)}$ means that $\bar{f}$ from equation (23) is in $L^2_{c'}$ for sufficiently small times, if the series converges.

Let us now show that the series in equation (23) indeed converges. From lemma 3.3 we have

$$||\mathcal{L}\bar{f}||_{L^2_{r(1-\frac{at}{n})}} \leqslant \frac{n}{atr}||\bar{f}||_{L^2_r} \leqslant \frac{n}{atc'}||\bar{f}||_{L^2_r} \tag{27}$$

for any $c' < r < c$. Then, we have the following estimate for $||\bar{f}_n||_{L^2_{c'}}$:

$$||\bar{f}_n||_{L^2_{c'}} \leqslant ||\bar{f}_0||_{L^2_c} \frac{M^{n+1}n^n}{a^n t^n c'^n} \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \cdots \int_0^{t_{n-1}} \mathrm{d}t_n \leqslant ||\bar{f}_n||_{L^2_c} \left(\frac{6M}{ac'}\right)^n \tag{28}$$

where we used the Stirling formula for large enough $n$. Since $a$ is arbitrary, we can choose it so large that the bracket in the right-hand side of equation (28) is less than one, and the series in equation (23) converges in $L^2_{c'}$. By construction, the operator that maps $\bar{f}_0(x,p)$ into $\bar{f}(x,p,t)$ is a uniformly bounded linear operator from $L^2_c$ to $L^2_{c'}$ for all $0 \leqslant t \leqslant T'$.

Let us now show that the solution of equation (9) is unique. If the opposite is true, there must exist a solution $\bar{f} \in L^2_{c'}$ of equation (22) with $\bar{f}_0 = 0$ on the interval $0 \leqslant t \leqslant T'$. Then, from equation (22) we have the following estimate:

$$||\bar{f}(x,p,t)||_{L^2_{\alpha(t)c'/2}} \leqslant \left(\frac{2Mt}{c'}\right) \max_{0\leqslant t\leqslant T'} ||\bar{f}(x,p,t)||_{L^2_{c'}} \tag{29}$$

where we teated the operator $\mathcal{L}$ as an operator from $L^2_{c'}$ to $L^2_{c'/2}$ and used lemma 3.3. Continuing in this fashion, we get

$$||\bar{f}(x,p,t)||_{L^2} \leqslant ||\bar{f}(x,p,t)||_{L^2_{c'\alpha^n(t)/2^n}} \leqslant \max_{0\leqslant t\leqslant T'} ||\bar{f}(x,p,t)||_{L^2_{c'}} \frac{2^n M^n t^n}{c'^n n!} \to 0 \tag{30}$$

as $n \to \infty$. For the classical solution of equation (9), this implies that $\bar{f} = 0$. □

**Corollary 3.5.** *The solution continuously depends on the initial data in the sense that small errors in the initial data that lie in $L^2_c$ produce small errors in the solution that lie in $L^2_{c'}$.*

The assumption of boundedness of the operator $\mathcal{E}(t)*$ in equation (21) and the Lipschitz property of $\alpha(t)$ are the non-trivial assumptions on the dynamics. Note that it is not difficult to show that these assumptions are satisfied for the system of $N$ non-interacting particles ($U = 0$).

## 4. Breakdown of time-reversibility

We have proved that under certain assumptions the initial value problem for equation (9) is well posed for finite times. This well-posedness, however, is time *irreversible*. Indeed, the obtained solution $\bar{f}(t)$ lies in the larger space $L^2_{c'}$ and not necessarily in the space $L^2_c$ in which the initial condition lies. Remarkably, this statement remains true even in the limit $\epsilon \to 0$, for which the solution of equation (9) should go to the solution of equation (3)! In other words, the right-hand side of equation (9) is a *singular* perturbation to the Liouville equation.

The reason for the breakdown of time reversibility is that although equation (9) is formally time reversible, this does *not* extend to the continuous dependence of the solution on the initial data. Corollary 3.5 guarantees such a dependence in a sense that small perturbations of the initial conditions that lie in $L^2_c$ will transform into small perturbations of the solution in the larger space $L^2_{c'}$, and not necessarily in $L^2_c$. This is the key point of the whole analysis above.

Let us point out that the assumptions of theorem 3.4 already have an irreversible statement that $\mathcal{E}(t)*$ maps $L^2_c$ into a larger space $L^2_{c'}$ with $c' = \alpha(t)c < c$. This exemplifies the view of irreversibility as progressive loss of regularity of solutions of equation (3) with time adopted in this paper. Note, however, that while the unobservable distribution function $f$ becomes more and more irregular in the limit $t \to \infty$, the observable function $\bar{f}$ should retain certain degree of regularity for all times.

Let us demonstrate these points explicitly in the case of non-interacting particles, that is, when $U = 0$. When $U = 0$, equation (9) can be solved exactly. Let us introduce the new variable $x' = x - pt$. Then, equation (9) can be rewritten as

$$\frac{\partial \bar{f}}{\partial t} = -\epsilon^2 \sum_n \frac{\partial^2 \bar{f}}{\partial x_n \partial p_n} + t\epsilon^2 \sum_n \frac{\partial^2 \bar{f}}{\partial x_n^2} \tag{31}$$

where the primes were dropped. Introducing the Fourier transform, after simple calculations we obtain that the solution of equation (31) is

$$\bar{f}_{kl}(t) = \bar{f}_{kl}(0)\mathrm{e}^{\frac{1}{2}\epsilon^2 l^2 - \frac{1}{2}\epsilon^2 (l-kt)^2}. \tag{32}$$

Thus, at long times $\bar{f}_{kl}(t) \to 0$ for $k \neq 0$, and $\bar{f}_{kl}(t) = \bar{f}_{kl}(0)$ for $k = 0$, that is, any initial condition $\bar{f}_0$ *relaxes* to a uniform distribution in $x$. Thus, in contrast to equation (3), our equation (9) is *dissipative*.

Let us see how the solution of equation (9) with $U = 0$ and the initial condition in the form of a single Gaussian peak (equation (6)) in the infinite domain evolves with time relative to the spaces $L^2_c$. A straightforward calculation shows that the solution will lie in the space $L^2_{c'}$ with $c' < \bar{c}(t)$, where

$$\bar{c} = \tfrac{1}{2}\left(1 + 2c + ct^2 - \sqrt{1 - 2ct^2 + 4c^2t^2 + c^2t^4}\right). \tag{33}$$

For $c > 1$ the function $\bar{c}(t)$ is decreasing with time, so the solution indeed moves from the smaller space $L^2_c$ to a larger space $L^2_{c'}$ for $t > 0$. Also, observe that according to equation (33), no matter what the constant $c$ we start with, we will have $\bar{c} \to 1$ when $t \to \infty$, so $L^2_1$ is an *attracting* space in this example. This also means that the observable $\bar{f}$ will not lose its regularity at long times, in contrast to the unobservable $f$, which will leave all the spaces $L^2_c$ as $t \to \infty$.

Let us now go back to the case of general $U$. The analysis above dealt with the well-posedness of equation (9) taken in its own right. Recall, however, that the physically relevant solutions of this equation are those distributions $\bar{f}$ which are related to $f$ by equations (5) and (7). So, they *a priori* possess a high degree of regularity in $p$. For those physically relevant distributions one can construct the solutions of equation (9) by first applying the inverse of the convolution, then finding $f(t) = \mathcal{E}(t) * f_0$, and then applying the convolution to find $\bar{f}(t)$. This can be done for arbitrary $t$ assuming that the solutions of equations (1) extend to infinite times. So, in fact, for physically relevant initial conditions the solution of equation (9) should in fact exist for all times.

It is not difficult to see that in the case of an anharmonic oscillator with one degree of freedom the solutions of equation (9) will spread uniformly over the energy surfaces at long times. Indeed, consider an initial condition for $f$ to be a function which is different from zero only in a small rectangle close to an energy surface $E = \text{const}$. As time passes, this rectangle will become distorted and progressively more skewed, until at very late times it will start to wind around the energy surface. Since the solutions of equation (9) are obtained from $f$ by a convolution (equation (5)), at long times the values of $\bar{f}$ in the $\epsilon$ neighbourhood of the energy surface will be essentially an integral over the entire region where $f$ is nonzero (for small $\epsilon$). So, the distribution of $\bar{f}$ will in fact be spreading over the energy surface. Thus, we can talk about the onset of the microcanonical distribution in this situation.

An interesting property of the smoothing operation given by equations (5) and (7) is that it changes the average energy of the system by a constant independent of the distribution $f$. Indeed, calculating the averages of $H$ with respect to $f$ and $\bar{f}$, we get

$$E = E_\epsilon - \frac{N}{2}\epsilon^2 \tag{34}$$

where $E$ and $E_\epsilon$ are the respective averages, and we used equations (5) and (7). Of course, the smoothing operation preserves the normalization of $\bar{f}$. This allows us to introduce the Boltzmann entropy $S_\epsilon$:

$$S_\epsilon = -\int \bar{f} \ln \bar{f} \,\mathrm{d}x\,\mathrm{d}p \tag{35}$$

which is bounded for the normalized distributions $\bar{f}$ having a finite average energy $E$. Substituting this definition into equation (9), we obtain an equation for the rate of change of this entropy

$$\frac{\mathrm{d}S_\epsilon}{\mathrm{d}t} = -\epsilon^2 \sum_n \int \frac{1}{\bar{f}} \frac{\partial \bar{f}}{\partial x_n} \frac{\partial \bar{f}}{\partial p_n} \,\mathrm{d}x\,\mathrm{d}p. \tag{36}$$

One can see from this equation that $\mathrm{d}S_\epsilon/\mathrm{d}t$ can have both signs and it is a question whether $S_\epsilon$ can play the role of a Lyapunov functional in the system. This is also a consequence of the formal reversibility of equation (9). Note, however, that $S_\epsilon$ may in fact be a monotonically increasing function of time for some sets of initial conditions. For example, if one takes the initial condition to be a single Gaussian peak in equation (6), in the case of the non-interacting system in an infinite domain one can explicitly calculate $S_\epsilon$ as a function of time: $S_\epsilon(t) = N \ln \sqrt{1 + c + t^2}$ up to a constant. This function is monotonically increasing with time. It is interesting to note that the growth rate of this function is independent of $\epsilon$. Also, it is not difficult to see from the exact solution (equation (32)) that for the system of non-interacting particles in a finite domain $\mathrm{d}S_\epsilon/\mathrm{d}t > 0$ for sufficiently long times.

The Boltzmann entropy defined in equation (35) allows us to reformulate the question of irreversibility in the dynamics given by equations (1). Let us introduce the spaces $\mathcal{H}_+$, $\mathcal{H}_-$ and $\mathcal{H}_0$ of all functions $\bar{f}$ for which the right-hand side of equation (36) is positive, negative or zero, respectively. In that sense, the space $\mathcal{H}_+$ corresponds to the initial conditions that are 'oriented' forward in time, the space $\mathcal{H}_-$ is for backward and the space $\mathcal{H}_0$ is for the initial conditions for which both directions of time are equivalent. Then, the dynamics can be considered irreversible if $\mathcal{H}_+$ is an attracting set. Alternatively, for $\bar{f}_0 \in \mathcal{H}_+$ the dynamics will be irreversible if $\mathcal{H}_+$ is an invariant space of equation (9). In general, the question of irreversibility may be studied from the point of view of the dynamics generated by equation (9) on $\mathcal{H}_0$.

## 5. Discussion

In conclusion, we have derived an equation for a smoothed phase space distribution function $\bar{f}$ from equation (3) and shown that with the appropriate initial conditions and certain assumptions on the dynamics the obtained equation is well posed. However, this well-posedness is shown to exist only in the forward time direction. We interpret this result as a manifestation of irreversibility of the dynamics of the physically observable quantities. Of course, this does not solve the big problem here, which is what this irreversibility is coming from. The latter is an essential property of the system's dynamics and has to do with mixing. However, the question of well-posedness of the 'macroscopic' equation for the smoothed phase space distribution function $\bar{f}$ seems to give a nice characterization of breaking of the time-reversal symmetry. This is in contrast with the coarse-graining approach, in which the irreversibility in the coarse-grained quantities is the consequence of the loss of information during coarse-graining. Note that the direction of time in our analysis is determined by causality. The latter is incorporated into the fundamental solution of equation (3) (function $\theta(t)$ in equation (20)).

One of the nice features of equation (9) is that because this equation is already irreversible, it may be used as a more suitable starting point for deriving the kinetic equations as opposed to the Liouville equation [5]. In particular, one can use equation (9) as a starting point for deriving the Boltzmann equation. In a rarefied gas the smoothed distribution function $\bar{f}$ will obey equation (9) with $U = 0$ everywhere in the phase space except when the particles are within the small-interaction region. One can therefore choose $\epsilon$ to be sufficiently small but yet big enough that the timescale $\epsilon^{-1}$ associated with the decay of $\bar{f}$ in a non-interacting system of size of order 1 (equation (32)) is much smaller than the collision timescale. This means that on this timescale the smoothed distribution function $\bar{f}$ will become homogeneous in space. Yet, when the particles are within the interaction range, the right-hand side of equation (9) will be a small perturbation, so in these regions the solution of equation (9) will be close to that of equation (3). One can then retrace the arguments of Bogoliubov [5] to obtain the Boltzmann equation for the distribution function in the momentum space. This derivation, however, will not have the weakness that it breaks down after a finite time [5, 8].

## Acknowledgments

The author would like to acknowledge valuable discussions with W E, J Goodman, D V Ingerman, J L Lebowitz and V V Osipov.

## References

[1] Maxwell J C 1866 *Phil. Trans. R. Soc.* **157** 26
Boltzmann L 1872 *Wein. Ber.* **66** 275
Boltzmann L 1964 *Lectures on Gas Theory* translated by S G Brush (Berkeley, CA: University of California Press)
[2] Lebowitz J L 1993 *Physica* A **194** 1
Lebowitz J L 1999 *Rev. Mod. Phys.* **71** S346
[3] Prigogine I 1999 *Physica* A **263** 528
[4] Bogoliubov N N 1946 *J. Phys. USSR* **10** 265
Born M and Green H S 1946 *Proc. R. Soc.* A **188** 10
Kirkwood J G 1946 *J. Chem. Phys.* **14** 180
Isihara A 1971 *Statistical Physics* (London: Academic)
[5] Bogoliubov N N 1965 *The Dynamical Theory in Statistical Physics* (Dehli: Hindustan)
[6] Landau L D 1946 *J. Phys. USSR* **10** 25
Lifshitz E M and Pitaevskii L P 1981 *Physical Kinetics* (Oxford: Pergamon)

[7] Klimas A J 1986 *J. Comput. Phys.* **68** 202

[8] Lanford O E 1981 *Physica* A **106** 70
Lanford O E 1975 *Dynamical Systems, Theory and Applications* ed J Moser (New York: Springer)