Journal of Statistical Physics, Vol. 67, Nos. 3/4, 1992

Strict Entropy Production Bounds and Stability of the Rate of Convergence to Equilibrium for the Boltzmann Equation

E. A. Carlen[1] **and M. C. Carvalho**[2]

Received December 18, 1991

We first consider the Boltzmann equation with a collision kernel such that all kinematically possible collisions are run at equal rates. This is the simplest Boltzmann equation having the compressible Euler equations as a scaling limit. For it we prove a stability result for the H-theorem which says that when the entropy production is small, the solution of the spatially homogeneous Boltzmann equation is necessarily close to equilibrium in the entropic sense, and therefore strong L^1 sense. We use this to prove that solutions to the spatially homogeneous Boltzmann equation converge to equilibrium in the entropic sense with a rate of convergence which is uniform in the initial condition for all initial conditions belonging to certain natural regularity classes. Every initial condition with finite entropy and pth velocity moment for some $p > 2$ belongs to such a class. We then extend these results by a simple monotonicity argument to the case where the collision rate is uniformly bounded below, which covers a wide class of slightly modified physical collision kernels. These results are the basis of a study of the relation between scaling limits of solutions of the Boltzmann equation and hydrodynamics which will be developed in subsequent papers; the program is described here.

KEY WORDS: Boltzmann equation; entropy; central limit theorem; hydrodynamics.

1. INTRODUCTION

We begin with the analogy which is the basis of this paper.

Let V_1 and V_2 be independent zero-mean, unit-variance, $\mathbb{R}^3$-valued random variables with the same density $f(v)$. Then clearly $(V_1 + V_2)/\sqrt{2}$

[1] Department of Mathematics, Princeton University, Princeton, New Jersey. On leave from School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia 30332.

[2] Hill Center for Mathematical Sciences, Rutgers University, New Brunswick, New Jersey 08903. On leave from C.F.M.C. and Departamento de Matemática da Faculdade de Ciencias de Lisboa, 1700 Lisboa codex, Portugal.

822/67/3-4-10

0022-4715/92/0500-0575$06.50/0

has the density $\int_{\mathbb{R}^3} f((v+v')/\sqrt{2})\, f((v-v')/\sqrt{2})\, d^3v'$. Now fix a probability density f_0 on $\mathbb{R}^3$ with zero mean and unit variance. Inductively define the sequence of densities $n \mapsto f_n$ by

$$f_{n+1}(v) = \int_{\mathbb{R}^3} f_n\left(\frac{v+v'}{\sqrt{2}}\right) f_n\left(\frac{v-v'}{\sqrt{2}}\right) d^3v' \tag{1.1}$$

Then each f_n is zero mean, unit variance, and $n \mapsto f_n$ uniquely solves the difference equation

$$f_{n+1}(v) - f_n(v) = \int_{\mathbb{R}^3} \left[f_n\left(\frac{v+v'}{\sqrt{2}}\right) f_n\left(\frac{v-v'}{\sqrt{2}}\right) - f_n(v)\, f_n(v') \right] d^3v' \tag{1.2}$$

with the specified initial density. Moreover, the classical central limit theorem says that

$$\lim_{n \to \infty} f_n(v) = \left(\frac{1}{2\pi}\right)^{3/2} e^{-v^2/2}$$

at least weakly.

The Boltzmann collision kernel $\mathcal{Q}(f, f)$ is a quadratic operator on probability densities closely related to the quadratic operator appearing on the right side of (1.2). It exact form is given below, but before concerning ourselves with the details, recall that if we take f_0 as initial data for the spatially homogeneous Boltzmann equation

$$\frac{\partial}{\partial t} f_t(v) = \mathcal{Q}(f_t, f_t)(v) \tag{1.3}$$

then each f_t is a zero-mean, unit-variance probability density and

$$\lim_{t \to \infty} f_t(v) = \left(\frac{1}{2\pi}\right)^{3/2} e^{-v^2/2}$$

at least weakly for a large class of collision kernels $\mathcal{Q}$.[2,14]

There is evidently a very close analogy between the convergence described by the central limit theorem and the convergence to equilibrium described by the Boltzmann equation. This analogy is not new here; aspects of it have been discussed in the work of McKean, Toscani, and others as well. However, we shall develop novel aspects of this analogy here, permitting the application of probabilistic ideas and inequalities to obtain new results on the Boltzmann equation.

In fact, the main technical tool we use is an entropy production inequality adapted from work by Carlen and Soffer[17] on central limit theorems for block sums of dependent random variables. Our entropy production inequality is a strengthening of the "H-theorem" valid at least for a wide class of collision kernels. From it we obtain information on the rate of strong $L^1(\mathbb{R}^3)$ convergence to equilibrium of solutions to the spatially homogeneous Boltzmann equation which is stable in that we have control on how the rate varies when the initial condition is varied. In a forthcoming paper we apply this to control Euler scaling limits of solutions of the full Boltzmann equation and to establish new results concerning the connection between the Boltzmann equation and the Euler equations. But before proceeding further to describe our results and methods, we first establish some notation concerning the Boltzmann equation, and then discuss the problem to which our work is addressed; namely, that of establishing stable rate bounds.

The phase space density $f_t(x, v)$ of a gas of particles moving freely between elastic collisions is supposed to satisfy the Boltzmann equation

$$\frac{\partial}{\partial t} f_t(x, v) = -v \cdot \nabla_x f_t(x, v) + \mathcal{Q}(f_t, f_t)(x, v) \tag{1.4}$$

at least under appropriate assumptions on the mean free path and the nature of the collisions. The first term on the right accounts for the changes in the density due to streaming, i.e., the local average free motion of the particles between collisions. The second accounts for the changes in density due to the collisions themselves. On the basis of mechanical considerations and the "collision number hypothesis," Boltzmann derived

$$\mathcal{Q}(f, f) = \int_{S^2} \int_{\mathbb{R}^3} [f(\tilde{v}) f(\tilde{v}') - f(v) f(v')] \, b(|v - v'|, \vartheta) \, d^3v' \, d\omega \tag{1.5}$$

The notation is this: ω is any unit vector, and

$$\tilde{v} = v + (\omega \cdot (v' - v))\omega, \qquad \tilde{v}' = v' - (\omega \cdot (v' - v))\omega \tag{1.6}$$

denote the resultant velocities in the elastic collision $(v, v') \mapsto (\tilde{v}, \tilde{v}')$. Also, $d\omega$ is the uniform probability measure on S^2. Finally, ϑ is the angle through which the relative velocity is deflected during this collision, and the *microscopic rate function* $b(|v - v'|, \vartheta)$ is the product of the relative speed $|v - v'|$ and the differential cross section $\sigma(|v - v'|, \vartheta)$ for this collision. Under natural symmetry conditions σ depends only on the indicated variables. For more information, see Cercignani's book[20] or, more briefly, Dresden's lectures.[29]

When the density is initially spatially homogeneous, it remains so at later times. There are then no streaming effects, and the density satisfies the spatially homogeneous Boltzmann equation (1.3). We normalize the initial density to be a probability density instead of a mass density. Then $\int_{\mathbb{R}^3} f_t(v)\, d^3v = 1$ for all t, and we define the *bulk velocity* u and the *temperature* θ by

$$u = \int_{\mathbb{R}^3} v f_t(v)\, d^3v, \qquad \theta = \frac{1}{3} \int_{\mathbb{R}^3} |v-u|^2 f_t(v)\, d^3v$$

These quantities are conserved as well, representing conservation of momentum and energy. (Conservation of θ can in general be delicate; see ref. 1.)

Velocity densities of the form

$$M_{u,\theta}(v) = (2\pi\theta)^{-3/2} e^{-|v-u|^2/2\theta}$$

are called *Maxwellian*. They are determined by their bulk velocities and temperatures. If f is any velocity density with finite second moments, we write M^f to denote the Maxwellian with the same bulk velocity and temperature.

The entropy $H(f)$ of a velocity density f with finite second moments is given by

$$H(f) = -\int_{\mathbb{R}^3} f(v) \ln f(v)\, d^3v$$

and is always well defined, though it may take the value $-\infty$. Two classical results concerning entropy and the Boltzmann equation are basic to any study of the subject.

The first result, due to Gibbs,[33] is quite simple; in the notation introduced above

$$H(f) \leqslant H(M^f) \tag{1.7}$$

with equality exactly when $f = M^f$. To see this, introduce the relative entropy $D(f)$ between f and M^f, which is given by

$$D(f) = \int \left(\frac{f}{M^f}\right) \ln \left(\frac{f}{M^f}\right) M^f\, d^3v$$

By Jensen's inequality, $D(f) \geqslant 0$, with equality exactly when $f = M^f$. But $\ln M^f$ is quadratic, and since f and M^f share the same second moments, $D(f) = H(M^f) - H(f)$.

The second result is of a similar nature, and is only slightly more complicated to prove, but it is much more profound. It is Boltzmann's H-theorem, which says that

$$-\int_{\mathbb{R}^3} \ln f(v)\, \mathscr{Q}(f, f)(v)\, d^3v \geqslant 0 \tag{1.8}$$

with equality exactly when $f = M^f$.

The quantity on the left side of (1.8) is called the *entropy production* because for solutions of the spatially homogenous Boltzmann equation, this quantity equals $(d/dt)\, H(f_t)$. Thus, $H(f_t)$ is strictly increasing unless $f = M^f$. This implies that Maxwellian densities are the only equilibria, but it does not suffice to describe the approach to equilibrium or even to establish that

$$\lim_{t \to \infty} f_t = M^{f_0}$$

This was first done by Carleman,[13] who treated a hard-sphere gas and proved a compactness property of $\{f_t \mid t \geqslant 0\}$, and then used the statement concerning cases of equality in the H-theorem to identify any limit as being M^{f_0}. Of course, such an argument cannot give any information on the *rate* at which f_t approaches M^{f_0}.

We emphasize rate information because rate bounds are physically significant. In fact, it is expected that f_t approaches M^{f_0} quite rapidly, i.e., on the scale of the time between collisions. Since streaming changes the density only on a much longer macroscopic time scale, there should then be a separation between the time scale on which collisions and streaming act.

This separation of scales is the basis of a physical picture for the evolution of the phase space density of a spatially inhomogeneous gas.[32,38,58] The physical picture which relates the Boltzmann equation to the equations of hydrodynamics is the following: The collisions going on at each location x rapidly drive the local velocity distribution very close to the *local* Maxwellian

$$\rho(x, t) \left(\frac{1}{2\pi\theta(x, t)} \right)^{3/2} e^{-|v - u(x, t)|^2 / 2\theta(x, t)}$$

conserving the local density, bulk velocity, and temperature before streaming has time to effect any appreciable influence. We think of the set of all phase space densities as a "manifold" and the set of all local Maxwellian densities as a "submanifold" parametrized by the five hydrodynamic moments ρ, u, and θ. Since $\mathcal{Q}(f, f)$ vanishes on this submanifold, the streaming becomes relatively significant near it, and the phase space density then evolves under the combined influence of streaming and collisions, with collisions always keeping it essentially on the "submanifold" of local Maxwellian densities, toward a global Maxwellian equilibrium.

Thus, after a short initial period, the evolution of the phase space density is well described by specifying the evolution of its approximate coordinates on the submanifold of local Maxwellian densities, i.e., the hydrodynamic moments ρ, u, and θ. One can compute the evolution of these in appropriate scaling limits, and one finds that they should satisfy the Euler or Navier–Stokes equations, depending on the scaling employed. See ref. 7 for a particularly helpful heuristic discussion of these scaling limits.

Much work has been done on the connection between the Boltzmann equation and hydrodynamics.[11,12,23,35,37,42,47,51,59] Howevery, many basic issues are still unresolved. In fact, most rigorous work is based on a Hilbert or a Chapman–Enskog–Hilbert expansion for solutions of the Boltzmann equation. While this approach has led to a number of very substantial theorems, its physical basis is somewhat unclear,[29,48] especially as regards the higher order hydrodynamic equations, such as the Burnett equations, which one obtains on continuing the expansion. Moreover, error terms for these expansions have only been controlled for as long as the hydrodynamic equations possess smooth solutions. But Sideris[53] has shown that the compressible Euler equations, for example, develop singularities in arbitrarily short times even for "small" initial data.

Since the physical picture discussed above hinges very much on the separation between the time scales of the collision process and the streaming process, it is expected that strong information on the rate at which f_t approaches M^{f_0} for the spatially homogeneous case would help to resolve a number of questions concerning the relation between the Boltzmann equation and the equations of hydrodynamics. The main result of this paper is a strengthing of the H-theorem which yields such information. It should be stressed that the methods we employ here work *directly* only for special collision kernels. In fact, we will at first restrict our attention to the case where $b(|v|, \vartheta)$ is independent of both $|v|$ and ϑ. We will then extend our entropy production bounds to the class of collision kernels whose rate functions are uniformly bounded below with a simple monotonicity argument. Roughly: the bigger the rate function, the more collisions there are; and the more collisions the are, the more entropy there is produced. Independence of $|v|$ is more crucial to our argument than independence of ϑ, so quite possibly the methods can be adapted without major modification to treat directly Maxwellian molecules as well.

However, the simple Boltzmann equation with uniform collision rate which we do treat directly is of interest in itself. It is the simplest Boltzmann equation which has the compressible Euler equations as a scaling limit. Its nonphysical nature would show up in physical properties depending on the spectrum of the linearized collision operator such as viscosity and heat diffusion constants, but there are none of these on the Euler scale. It is thus natural to study the relation between the Boltzmann equation and the Euler equations in this simplest special case first.

Our method for obtaining rate information is based on stability results for the two strict inequalities (1.7) and (1.8). That is, we use results showing that if f nearly saturates one of these inequalities, then it is necessarily close to M^f.

The first of these stability results is well known:

$$D(f) \geqslant \tfrac{1}{2} \|f - M^f\|^2_{L^1(\mathbb{R}^3)}$$

In the context of information theory, this inequality goes back to Csizlar[21] and Kullback,[36] while in the present context it was later independently discovered by Cercignani.[20]

The second is our main result, a stability result for the H-theorem in the case of the special collision kernels we consider. Before stating it, we introduce two functions associated with any velocity density with zero bulk velocity and unit temperature.

First put

$$\psi_f(R) = \int_{|v| \geqslant R} |v|^2 f(v)\, d^3v$$

Clearly ψ_f decreases to zero, and the rate at which is does so measures the localization of f in velocity space.

Before specifying the next function, define for each $\kappa > 0$

$$\mathscr{P}_\kappa f(v) = \int_{\mathbb{R}^3} [2\pi(1 - r^{-2\kappa})]^{-3/2} \times \{\exp[-|v - v'|^2/2(1 - e^{-2\kappa})^{1/2}]\}\, e^{3\kappa} f(e^\kappa v')\, d^3v'$$

which is a sort of Maxwellian regularization of f. The operators $\mathscr{P}_\kappa$ constitute the adjoint Ornstein–Uhlenbeck semigroup, which plays an important role in our analysis.

The second function mentioned above is given by

$$\chi_f(\kappa) = H(\mathscr{P}_\kappa f) - H(f)$$

We will show in Section 3 that if $H(f)$ is finite, then χ_f is a positive, continuous, strictly increasing (unless $f = M^f$) function of κ. Also, for each $\kappa > 0$, $f \mapsto \chi_f(\kappa)$ is strictly convex. Like D, it is a measure of the deviation of f from M^f, and it is a mild measure of the smoothness of f.

We can now roughly formulate our stability result; a careful formulation is given in Section 4.

For any pair of continuous functions ψ and χ which are, respectively, decreasing to zero and increasing from zero, there is a function $\Phi_{\psi,\chi}$ which is *strictly* increasing from zero, and which depends only on ψ and χ so that for all velocity densities f with zero bulk velocity, unit temperature, and finite entropy satisfying

$$\psi_f \leqslant \psi, \qquad \chi_f \leqslant \chi \tag{1.9}$$

it holds that

$$-\int_{\mathbb{R}^3} \ln f \mathscr{Q}(f, f)\, d^3v \geqslant \Phi_{\psi,\chi}(D(f)) \tag{1.10}$$

We can apply this to the Boltzmann equation because of two further results. First, it is possible to show that for any f_0 with zero bulk velocity, unit temperature, and finite entropy, there is a continuous function χ increasing from zero and *depending only on* χ_{f_0} so that

$$\chi_{f_t} \leqslant \chi$$

for all t, where f_t is the solution of (1.3) starting from f_0.

It would be very useful, here and elsewhere, to known the analogous statement for ψ. We can prove it fairly easily in the context of the central limit theorem,[16] but the problem is more difficult for the Boltzmann equation and remains open.

Therefore, we must now assume that $\int_{\mathbb{R}^3} |v|^p f_0(v)\, d^3v$ is finite for some $p > 2$. A result of Elmroth[30] then implies that for all t this moment is bounded *uniformly* in t. This readily implies the existence of a ψ depending only on $\int_{\mathbb{R}^3} |v|^p f(v)\, d^3v$ and decreasing to zero so that

$$\psi_{f_t} \leqslant \psi$$

for all t. Then, since $(d/dt)\, D(f_t) = -(d/dt)\, H(f_t)$,

$$\frac{d}{dt} D(f_t) \leqslant -\Phi_{\psi,\chi}(D(f_t)) \tag{1.11}$$

Clearly this provides information on the rate of decrease of $D(f_t)$, and hence $\|f_t - M^{f_0}\|_{L^1}$ as well by the first stability result. In particular, it implies that as long as the initial condition is varied within the class

$$\left\{f_0 \,\middle|\, \int_{\mathbb{R}^3} |v|^p f_0(v)\, d^3v \leqslant B,\, \chi_{f_0} \leqslant \chi\right\}$$

for some fixed $p > 2$, $B < \infty$, and χ increasing continuously from zero, the strong rate of convergence to equilibrium is uniformly controlled.

In fact, it is possible to proceed further with the methods introduced here and to calculate the dependence of $\Phi_{\psi,\chi}$ on ψ and χ; i.e., to give a formula for it. This will be done in the second paper of this series, and again the argument is adapted from the used to prove a similar result on the central limit theorem.[18] This allows one to read off an explicit rate for the convergence to equilibrium, but the method is complicated and therefore extravagant, so that the rate is only sharp in rather special circumstances. Moreover, it is the above explicit condition for *uniformity* of the rate—not the rate itself—which plays the crucial role in our forthcoming study of the relation between the Boltzmann equation and the compressible Euler equations.

Results asserting strong L^1 convergence to equilibrium have been obtained in a very general setting by Gustafsson[34] and Arkeryd.[2,3] However, their results do not give any indication as to how the rate might vary with the initial condition. There is some other rate information available; Truesdell[57] has shown that all finite moments converge exponentially fast to their equilibrium values in the Maxwellian case.

Stability results for both (1.7) and (1.8) were conjectured by Cercignani,[19] whose paper can be consulted for other ways in which they might be useful. In fact, Cercignani conjectured somewhat more than we have proved here; namely, that (1.10) holds for quite general collision kernels with $\Phi_{\psi,\chi}(D(f))$ being a positive multiple of $D(f)$. In general, depending on the assumed regularity, our function Φ will vanish faster than linearly at the origin. This may well be spurious, but as explained above, it is no obstacle to our intended applications.

McKean[44] has studied the rate of convergence to equilibrium in the Kac model, which is a one-dimensional model Boltzmann equation. In this context he conjectures (1.10) without specifying the nature of the positive term on the right side. The methods we use in this paper all apply to the Kac model, and as we will show elsewhere, allow us to answer several questions raised by McKean in his paper. As mentioned earlier, his paper is also informed by an analogy with the central limit theorem—to which we will soon return. We have drawn a number of technical and general ideas from McKean's paper, and in fact it was the starting point of our investigation.

The only other result expressing a stability property of the entropy production by the Boltzmann equation is due to Desvillettes,[24] who shows that for any $R>0$ there is a constant C_R so that

$$-\int_{\mathbb{R}^3} \ln f\mathcal{Q}(f,f)\,d^3v \geqslant C_R \inf\left\{\int_{|v|\leqslant R} |\ln f - \ln M|^2\,d^3v \,\middle|\, M\right\}$$

where the infimum is taken over all Maxwellian densities M. The proof gives no information on the way C_R depends on R, so the right side is not comparable to $D(f)$, and one does not obtain a differential inequality such as (1.11) from Desvillettes' results. However, he is able to derive his bounds for a very wide class of collision kernels.

Finally, we mention the work of Toscani,[55,56] whose preprints were brought to our attention by Arkeryd after this work was publicly presented. Toscani's papers do not contain any of our main results, but he is motivated by the same analogy with the central limit theorem. For two-dimensional Maxwellian molecules, he proves entropic convergence to equilibrium using methods Barron[8] developed for the central limit theorem to prove a strong compactness result. Then he uses Carleman's classical argument to identify any limit as Maxwellian. This gives no rate information, and he does not discuss stability for the H-theorem. However, any reader who has read this far into our paper will find Toscani's papers very interesting.

We now return to the analogy with the central limit theorem, and explain its relevance to the program we have been discussing.

Let us rewrite (1.1) as $f_{n+1}=f_n \star f_n$; evidently $\star$ is a rescaled convolution. A special case of the Shannon–Stam inequality[52,54] says that

$$H(f\star f)\geqslant H(f)$$

which equality exactly when f is normal, i.e., Gaussian. Thus the Shannon–Stam inequality plays the role of the H-theorem for the difference equation (1.2).

However, as with the Boltzmann equation, it is not immediately clear that

$$\lim_{n\to\infty} H(f_n) = H((2\pi)^{-3/2}e^{-|\cdot|^2/2}) \tag{1.12}$$

even *already knowing the central limit theorem*, since H is only weakly upper semicontinuous. This was first established in general by Barron,[8] building on previous work of Brown[10] and Linnik.[40,41] Barron's compactness argument, however, gives no rate information.

A stability result for the Shannon–Stam inequality was proved in ref. 17. As a special case, it says that for f satisfying (1.9)

$$H(f\star f)-H(f)\geqslant \Phi_{\psi,\chi}(D(f)) \tag{1.13}$$

for some $\Phi_{\psi,\chi}$ as before.

It is easy to show that for any f_0 with zero mean, unit variance, and finite entropy, one can find continuous functions ψ and χ, respectively decreasing to zero and increasing from zero, so that

$$\chi_{f_n}\leqslant\chi, \qquad \psi_{f_n}\leqslant\psi$$

for all n, where f_n is the solution of (1.2) starting from f_0. But then we may apply (1.13) uniformly at each n and conclude that (1.12) holds. In particular, we will obtain a uniformity result on the rate of entropic convergence to the normal distribution.

The stability inequality (1.13) has other implications; in ref. 17 it was used to control the effects of dependence and prove central limit theorems for sums of dependent random variables.

Returning to the Boltzmann equation, it has been observed a number of times that the *gain term* in the collision kernel

$$\int_{S^2}\int_{\mathbb{R}^3} f(\tilde v)\, f(\tilde v')\, b(|v-v'|,\vartheta)\, d^3v'\, d\omega$$

is especially close to being a sort of rescaled convolution, as in (1.1), when b is independent of $|v-v'|$. This was done by Wild[60] in 1951, whose ideas are discussed more carefully in the next section. Briefly, however, with ν denoting the spherical average of b, i.e., the average collision rate, Wild defined an operation $f\mapsto f\circ f$ on densities by

$$f\circ f(v)=\frac{1}{\nu}\int_{S^2}\int_{\mathbb{R}^3} f(\tilde v)\, f(\tilde v')\, b(\vartheta)\, d^3v'\, d\omega$$

The homogeneous Boltzmann equation can then be written as

$$\frac{\partial}{\partial t}f_t(v)=\nu[f_t\circ f_t(v)-f_t(v)]$$

This enables us to prove an intermediate bound for the entropy production; namely

$$-\int_{\mathbb{R}^3} \ln f_t(v)\, \mathcal{Q}(f_t, f_t)(v)\, d^3v \geqslant \nu[H(f_t \circ f_t) - H(f_t)] \tag{1.14}$$

Formally the proof is very simple. Let dt be infinitesimal so that

$$f_{t+dt} = f_t + \nu[f_t \circ f_t - f_t]\, dt = \nu\, dt\, f_t \circ f_t + (1 - \nu\, dt) f_t$$

Then since H is a concave function of the entropy,

$$\begin{aligned} H(f_{t+dt}) - H(f_t) &= H(\nu\, dt\, f_t \circ f_t + (1 - \nu\, dt) f_t) - H(f_t) \\ &\geqslant \nu\, dt\, H(f_t \circ f_t) - \nu\, dt\, H(f_t) \end{aligned}$$

Now formal division by dt yields (1.14). We make this argument rigorous in the next section.

Moreover, we will show by another simple convexity argument that the right side of (1.14) is strictly positive unless $f_t = M^{f_0}$. The analogy with the Shannon–Stam inequality is now quite clear. And fortunately it turns out to be close enough that we can adapt the techniques used to pove stability for it to prove stability for the H-theorem here.

Actually, we will only proceed *directly* along this route when b is independent of ϑ as well. But then the results transfer to all collision kernels with the rate b bounded below by monotonicity in b of the entropy production, as we will show in the final section.

Section 2 contains background on the Wild form of the homogeneous Boltzmann equation in our setting and contains preliminary entropy estimates. Section 3 contains some *Fisher information* estimates which we use to obtain our final entropy estimates. As we will explain there, many entropy inequalities are most easily proved via a Fisher information inequality. This strategy was developed by Stam[54] in his proof of the Shannon–Stam inequality, which had been conjectured by Shannon.[52] A very different proof is given in Lieb's paper,[39] which was one of our original motivations for contemplating the Shannon–Stam inequality in the context of statistical mechanics. Section 4 contains the precise statements and the proofs of our main results.

2. PRELIMINARY ENTROPY BOUNDS

We now restrict our attention, until the end of Section 4, to the Boltzmann equation in which the microscopic rate function is a constant ν. This means that all kinematically possible collisions are assigned equal probability. It is possible that the methods we employ here can be extended to cover directly the case where the microscopic rate function b depends only on the scattering angle ϑ and not the relative velocity, and moreover

$$\int_{S^2} b(\vartheta)\, d\omega = \tfrac{1}{2} \int_0^\pi b(\vartheta) \sin \vartheta \, d\vartheta = v < \infty$$

This includes the case of cutoff Maxwellian molecules, i.e., molecules interacting through a $1/r^5$ force law, but with small-angle collisions suppressed—the class of generalized Maxwellian collision kernels. However, since we will be able to make an easy extension to many such cases and more by an additional but simple monotonicity argument, we have not delved into the matter.

The main goal of this section is to establish a preliminary lower bound for the entropy production in our setting.

Wild[60] observed that for generalized Maxwellian collision kernels, and therefore in our case, the homogeneous Boltzmann equation can be rewritten as

$$\frac{\partial}{\partial t} f_t(v) = v[f_t \circ f_t(v) - f_t(v)] \tag{2.1}$$

where is our case

$$f \circ g(v) = \frac{1}{v} \int_{S^2} \int_{\mathbb{R}^3} f(v + [(v' - v) \cdot \omega] \omega)\, g(v' - [(v' - v) \cdot \omega] \omega) v \, dv' \, d\omega$$

denotes the *Wild convolution* of the densities f and g.

Wild rewrote (2.1) as an integral equation and solved it by iteration under some assumptions. Morgenstern[49,50] developed Wild's method further and showed that (2.1) is uniquely solvable in $L^1(\mathbb{R}^3, (1 + |v|^2)\, d^3v)$ for all nonnegative initial data in this space. Moreover, DiBlasio[36,37] showed that $\int_{\mathbb{R}^3} f_t(v)\, d^3v$, $\int_{\mathbb{R}^3} v f_t(v)\, d^3v$, and $\int_{\mathbb{R}^3} |v|^2 f_t(v)\, d^3v$ are all conserved, and $t \mapsto f_t$ is strongly differentiable in $L^1(\mathbb{R}^3, (1 + |v|^2)\, d^3v)$. Other useful information on Wild's method can be found in refs. 44 and 46.

The main result of this section is the following theorem:

Theorem 2.1. For all velocity densities f_0 in $L^1(\mathbb{R}^3, (1 + |v|^2)\, d^3v)$ with $H(f_0) > -\infty$, $t \mapsto H(f_t)$ is continuous on $[0, \infty)$ and differentiable on $(0, \infty)$ with

$$\frac{d}{dt} H(f_t) \geqslant v[H(f_t \circ f_t) - H(f_t)] \geqslant 0 \tag{2.2}$$

There is equality on the right exactly when $f_0 = M^{f_0}$.

Before establishing this result, we collect some useful properties of entropy, relative entropy, and a regularization procedure which is very convenient in our setting. These results will be used in later sections as well.

For any two probability densities η and ρ on $\mathbb{R}^n$, $D(\eta \mid \rho)$ denotes the relative entropy of η with respect to ρ, and is given by

$$D(\eta \mid \rho) = \int_{\mathbb{R}^n} \left(\frac{\eta}{\rho}\right) \ln \left(\frac{\eta}{\rho}\right) \rho(y)\, d^n y \tag{2.3}$$

Since $t \mapsto t \ln t$ is convex and bounded below, $D(\eta \mid \rho)$ is always well defined and nonnegative. In fact, by the results quoted before,[31,36]

$$D(\eta \mid \rho) \geqslant \tfrac{1}{2} \|\eta - \rho\|^2_{L^1(\mathbb{R}^n)} \tag{2.4}$$

and since $(s, t) \mapsto s \ln s - s \ln t$ is jointly convex, it follows easily that $(\eta, \rho) \mapsto D(\eta \mid \rho)$ is jointly convex.

The entropy $H(\eta)$ is given by

$$H(\eta) = -\int_{\mathbb{R}^n} \eta(y) \ln \eta(y)\, d^n y \tag{2.5}$$

which is always well defined when η has finite second moments. Actually, to *define* $H(\eta)$, we assume finite second moments and let G^η denote the normal, i.e., Gaussian, density with the same mean and variance as η. We then put

$$H(\eta) = H(G^\eta) - D(\eta \mid G^\eta) \tag{2.6}$$

If $D(\eta \mid G^\eta)$ is finite, then $|\ln \eta - \ln G^\eta| \eta$ is integrable (see ref. 8 for a simple proof), and hence $|\ln \eta| \eta$ is integrable. So H is given by (2.5). Whenever we refer to the entropy of a density, it is implicitly assumed that the density has finite second moments. In the context of the Boltzmann equation, this convention is entirely natural.

Entropy is well known to be a strictly concave function of the density and upper semicontinuous in the weak $L^1(\mathbb{R}^n)$ topology. There is a proof in ref. 17. Here we shall use only the concavity of H restricted to densities of zero mean and unit variance. This follows directly from (2.6) and what we have already said about the convexity of D.

The strict positivity of the relative entropy immediately implies the *strict subadditivity* of the entropy: Let $\eta(y, z)$ be a density on $\mathbb{R}^n \times \mathbb{R}^n$, and let $\eta_1(y) = \int_{\mathbb{R}^n} \eta(y, z)\, d^n z$ and $\eta_2(z) = \int_{\mathbb{R}^m} \eta(y, z)\, d^m y$ denote the marginals densities of η. Then

$$H(\eta) \leqslant H(\eta_1) + H(\eta_2) \tag{2.7}$$

and provided $H(\eta)$ is finite, there is equality exactly when $\eta(y, z) = \eta_1(y)\,\eta_2(z)$ almost everywhere. To see this, suppose $H(\eta)$ is finite, or there is nothing to prove. Then computing the relative entropy of η with respect to the product of its marginals, one finds

$$\tfrac{1}{2}\|\eta - \eta_1\eta_2\|^2_{L^1(\mathbb{R}^n)} \leqslant D(\eta \mid \eta_1\eta_2) = -H(\eta) + H(\eta_1) + H(\eta_2)$$

(Note that η_1 and η_2 have finite second moments whenever η does.) Since $H(\eta)$ is finite, it can be added to both sides.

This has the following consequence:

Lemma 2.2. For all densities f on $\mathbb{R}^3$ with $H(f) > -\infty$,

$$H(f \circ f) \geqslant H(f) \tag{2.8}$$

with equality exactly when $f = M^f$.

Proof. Fix any unit vector ω, and put

$$\eta(v, v') = f(v + [(v'-v)\cdot\omega]\,\omega)\, f(v' - [(v'-v)\cdot\omega]\,\omega) = f(\tilde{v})\, f(\tilde{v}')$$

Put

$$f \circ_\omega f(v) = \int_{\mathbb{R}^3} f(v + [(v'-v)\cdot\omega]\,\omega)\, f(v' - [(v'-v)\cdot\omega]\,\omega)\, d^3v' \tag{2.9}$$

Then

$$\int_{\mathbb{R}^3} \eta(v, v')\, d^3v' = f \circ_\omega f(v), \qquad \int_{\mathbb{R}^3} \eta(v, v')\, d^3v = f \circ_\omega f(v') \tag{2.10}$$

Also, since $d^3v\, d^3v' = d^3\tilde{v}\, d^3\tilde{v}'$, $H(\eta) = 2H(f)$. Subadditivity of the entropy (2.7) and (2.10) then yield $2H(\eta) \leqslant 2H(f \circ_\omega f)$. Averaging both sides over ω and appealing to the concavity of the entropy yields (2.8). If there is equality, then $\eta(v, v')$ must be the product of its marginals for almost every ω, and this is only the case when $f = M^f$. (See ref. 15 for a method of proof which applies even without moment assumptions. The basic observation, however, goes back to Maxwell.[43]) ∎

The relative entropy is affine invariant. Let $\mathscr{A}: \mathbb{R}^n \to \mathbb{R}^n$ be the affine transformation given by $\mathscr{A}: y \mapsto Ay - y_0$ with a nonsingular, and for any density η define $\eta^{\mathscr{A}}$ by $\eta^{\mathscr{A}}(y) = |\det(A)|\, \eta(\mathscr{A}(y))$. Then clearly

$$D(\eta_1^{\mathscr{A}} \mid \eta_2^{\mathscr{A}}) = D(\eta_1, \eta_2) \tag{2.11}$$

The Wild convolution commutes with the action of a special class of affine transformations:

Lemma 2.3. Let $\mathscr{A}\colon \mathbb{R}^3 \to \mathbb{R}^3$ be given by $\mathscr{A}\colon v \mapsto av - u$ with $a > 0$ and $u \in \mathbb{R}^3$. Then for any probability density f on $\mathbb{R}^3$,

$$f^{\mathscr{A}} \circ f^{\mathscr{A}} = (f \circ f)^{\mathscr{A}}$$

Proof.

$$\begin{aligned} f^{\mathscr{A}} \circ f^{\mathscr{A}} &= a^6 \int_{S^2} \int_{\mathbb{R}^3} f(a\{v + [(v'-v)\cdot\omega]\omega\} - u) \\ &\qquad \times f(a\{v'-v)\cdot\omega]\omega\} - u)\, d^3v'\, d\omega \\ &= a_6 \int_{S^2} \int_{\mathbb{R}^3} f(\mathscr{A}(v) + \{[\mathscr{A}(v') - \mathscr{A}(v)]\cdot\omega\}\omega) \\ &\qquad \times f(\mathscr{A}(v') - \{[\mathscr{A}(v') - \mathscr{A}(v)]\cdot\omega\}\omega)\, d^3v'\, d\omega \\ &= a^3 f \circ f(\mathscr{A}(v)) \quad \blacksquare \end{aligned}$$

Next we introduce the heat semigroup $\mathscr{G}_\lambda = e^{\lambda\Delta/2}$, $\lambda > 0$, which acts by convolution with $G_\lambda(y) = (2\pi\lambda)^{-n/2} e^{-|y|^2/2\lambda}$. Since G_λ is a probability density, $\mathscr{G}_\lambda \eta$ is a probability density whenever η is; in fact $\mathscr{G}_\lambda \eta$ is just an average of translates of η.

Lemma 2.4. For all probability densities η_1, η_2 and all $\lambda > 0$,

$$D(\mathscr{G}_\lambda \eta_1 \mid \mathscr{G}_\lambda \eta_2) \leqslant D(\eta_1 \mid \eta_2)$$

Proof. For $y' \in \mathbb{R}^n$, let $T_{y'}\eta(y) = \eta(y - y')$. Then

$$\begin{aligned} D(\mathscr{G}_\lambda \eta_1 \mid \mathscr{G}_\lambda \eta_2) &= D\left(\int G_\lambda(y')\, T_{y'}\eta_1\, d^n y' \,\middle|\, \int G_\lambda(y')\, T_{y'}\eta_2\, d^n y' \right) \\ &\leqslant \int G_\lambda(y')\, D(T_{y'}\eta_1 \mid T_{y'}\eta_2)\, d^n y' = D(\eta_1 \mid \eta_2) \end{aligned}$$

using the joint convexity and translation invariance of the relative entropy. ∎

The action of the heat semigroup commutes with our Wild convolution. This result is due to Morgenstern,[50] who also establishes if for planar Maxwellian molecules. The following proof is patterned after that used by McKean to establish the analogous result for the Kac model.[44]

Lemma 2.5. For all probability densities f on $\mathbb{R}^3$ and all $\lambda > 0$

$$\mathscr{G}_\lambda(f \circ f) = (\mathscr{G}_\lambda f) \circ (\mathscr{G}_\lambda f)$$

Proof. For any fixed ω, let $(v, v') \mapsto (\tilde{v}, \tilde{v}')$ be the transformation given in (1.6). The proof follows very easily from two simple invariances, namely $d^3v\, d^3v' = d^3\tilde{v}\, d^3\tilde{v}'$ and $G_\lambda(v)\, G_\lambda(v') = G_\lambda(\tilde{v})\, G_\lambda(\tilde{v}')$, and Fubini's theorem. We have

$$\begin{aligned}
&(\mathcal{G}_\lambda f) \circ (\mathcal{G}_\lambda f)(v) \\
&= \int_{S^2} \int_{\mathbb{R}^3} \left\{ \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} G_\lambda(w)\, G_\lambda(w')\, f(v + [(v'-v)\cdot\omega]\omega - w) \right. \\
&\qquad \left. \times f(v' - [(v'-v)\cdot\omega]\omega - w')\, d^3w\, d^3w' \right\} d^3v'\, d\omega \\
&= \int_{S^2} \int_{\mathbb{R}^3} \left[\int_{\mathbb{R}^3} \int_{\mathbb{R}^3} G_\lambda(\tilde{w})\, G_\lambda(\tilde{w}') \right. \\
&\qquad \times f((v-w) + \{[v'-w') - (v-w)]\cdot\omega\}\omega) \\
&\qquad \left. \times f((v'-w') - \{[(v'-w') - (v-w)]\cdot\omega\}\omega)\, d^3w\, d^3w' \right] d^3v'\, d\omega \\
&= \int_{\mathbb{R}^3} G_\lambda(w) \left\{ \int_{S^2} \int_{\mathbb{R}^3} \left[\int_{\mathbb{R}^3} G_\lambda(w') \right.\right. \\
&\qquad \times f((v-w) + \{[(v'-w') - (v-w)]\cdot\omega\}\omega) \\
&\qquad \left.\left. \times f((v'-w') - \{[v'-w') - (v-w)]\cdot\omega\}\omega)\, d^3v' \right] d^3w'\, d\omega \right\} d^3w \\
&= \mathcal{G}_\lambda(f \circ f)(v) \quad \blacksquare
\end{aligned}$$

The heat semigroup clearly regularizes probability densities in several ways, but for our purposes the most useful regularizing semigroup is the adjoint Ornstein–Uhlenbeck semigroup. Its operators $\mathcal{P}_\lambda$ are given by

$$\mathcal{P}_\lambda \eta(y) = \int_{\mathbb{R}^n} G_{(1-e^{-2\lambda})^{1/2}}(y')\, e^{-n\lambda} \eta(e^{-\lambda}(y - y'))\, d^ny' \tag{2.12}$$

For background on this semigroup see ref. 17. Note that $\mathcal{P}_\lambda$ is the composition of a scale change and an action of the heat semigroup. Also, since the variance of the convolution of two densities is the sum of their variances, if f is a velocity distribution with temperature θ, $\mathcal{P}_\lambda f$ has temperature $(1 - e^{-2\lambda}) + e^{-2\lambda}\theta$. So it is easy to see that if f is a velocity density with zero bulk velocity and unit temperature, then so is each $\mathcal{P}_\lambda f$ and moreover

$$\lim_{\lambda \to 0} \mathcal{P}_\lambda f = f, \qquad \lim_{\lambda \to \infty} \mathcal{P}_\lambda f = M^f \tag{2.13}$$

Thus, the adjoint Ornstein–Uhlenbeck semigroup provides a natural interpolation between f and M^f which preserves the class of zero-bulk-velocity, unit-temperature distributions. Its real advantage over the heat semigroup will only be explained in the next section, however. The smoothing properties of $\mathscr{P}_\lambda$ will be used in the next section; here the only regularizing properties we use are given by the following lemmas. See Barron[8] for a lemma related to our next lemma.

Lemma 2.6. For any density f with zero bulk velocity and unit temperature

$$|\ln \mathscr{P}_\lambda f(v)| \leqslant A_\lambda(1+|v|^2)$$

for some constant A_λ depending only on λ.

Proof. Clearly from (2.12), $\mathscr{P}_\lambda f(v) \leqslant G_{(1-e^{-2\lambda})^{1/2}}(0)$. For a lower bound,

$$\begin{aligned}
\mathscr{P}_\lambda f(v) &= \int_{\mathbb{R}^3} G_{(1-e^{-2\lambda})^{1/2}}(v-v')e^{-3\lambda}f(e^{-\lambda}(v'))\, d^3v' \\
&\geqslant \int_{|v'|\leqslant R} G_{(1-e^{-2\lambda})^{1/2}}(v-v')e^{-3\lambda}f(e^{-\lambda}(v'))\, d^3v' \\
&\geqslant \left(\frac{1}{2\pi(1-e^{-2\lambda})^{1/2}}\right)^{3/2} \exp[-(|v|^2+R^2)(1-e^{-2\lambda})^{1/2} \\
&\quad \times \int_{|v'|\leqslant e^\lambda R} f(v')\, d^3v' \\
&\geqslant \left(\frac{1}{2\pi(1-e^{-2\lambda})^{1/2}}\right)^{3/2} \exp[-(|v|^2+R^2)(1-e^{-2\lambda})^{-1/2}] \\
&\quad \times \left(1-\frac{e^{2\lambda}}{R^2}\int_{|v'|\geqslant e^\lambda R} |v'|^2 f(v')\, d^3v'\right)
\end{aligned}$$

Choosing $R = e^{2\lambda}$ now gives the desired result. ∎

Lemma 2.7. For any two probability densities η_1 and η_2, and all $\lambda > 0$,

$$D(\mathscr{P}_\lambda \eta_1 \,|\, \mathscr{P}_\lambda \eta_2) \leqslant D(\eta_1 \,|\, \eta_2)$$

In particular, for any velocity density with zero bulk velocity and unit temperature,

$$H(\mathscr{P}_\lambda f) \geqslant H(f) \tag{2.14}$$

Proof. Since $\mathscr{P}_\lambda$ is the composition of a scale change and an action of the heat semigroup, the first inequality follows from Lemmas 2.3 and 2.4. Next, if f has zero bulk velocity and unit temperature, so does M^f. Therefore in this case $\mathscr{P}_\lambda M^f = M^f$, so that

$$D(\mathscr{P}_\lambda f \mid M^f) = D(\mathscr{P}_\lambda f \mid \mathscr{P}_\lambda M^f) \leqslant D(f \mid M^f)$$

which is equivalent to (2.14) by (2.6). ∎

Lemma 2.8. For any density f, and all $\lambda > 0$,

$$\mathscr{P}_\lambda(f \circ f) = (\mathscr{P}_\lambda f) \circ (\mathscr{P}_\lambda f)$$

Proof. This follows directly from Lemmas 2.3 and 2.5.

Proof of Theorem 2.1. Since the relative entropy is affine invariant, so is $(d/dt)\, H(f_t) = -(d/dt)\, D(f_t)$. Clearly $H(f \circ f) - H(f)$ is affine invariant, so by Lemma 2.3, it suffices to assume f_0 has zero bulk velocity and unit temperature. By Lemma 2.7, if f_t solves (2.1) with initial condition f_0, then $\tilde{f}_t = \mathscr{P}_\lambda f_t$ solves (2.1) with initial condition $\tilde{f}_0 = \mathscr{P}_\lambda f_0$. Then, since by Lemma 2.6, $|\ln \tilde{f}_t(v)| \leqslant A_\lambda(1 + |v|_2)$, $\ln\ \tilde{f}_t[\tilde{f}_t \circ \tilde{f}_t - \tilde{f}_t] \in L^1(\mathbb{R}^3)$, and

$$\frac{d}{dt} H(\tilde{f}_t) = -\int \ln \tilde{f}_t(v)\, \frac{d}{dt} \tilde{f}_t(v)\, d^3v$$

To better exploit the concavity of H, define $g(s) = e^{-\nu s}\tilde{f}_t + (1 - e^{-\nu s})\tilde{f}_t \circ \tilde{f}_t$. Then

$$g(0) = \tilde{f}_t, \qquad \frac{d}{ds} g(s)\bigg|_{s=0} = \frac{d}{dt} \tilde{f}_t$$

so that

$$\frac{d}{ds} H(g(s))\bigg|_{s=0} = \frac{d}{dt} H(\tilde{f}_t) \tag{2.15}$$

But

$$\begin{aligned} &\frac{d}{dt} H(g(s))\bigg|_{s=0} \\ &\quad = \lim_{s \to 0} \frac{1}{s} \left[H(e^{-\nu s}\tilde{f}_t + (1 - e^{-\nu s})\tilde{f}_t \circ \tilde{f}_t) - H(\tilde{f}_t) \right] \\ &\quad \geqslant \lim_{s \to 0} \frac{1}{s} \left[e^{-\nu s} H(\tilde{f}_t) + (1 - e^{-\nu s})\, H(\tilde{f}_t \circ \tilde{f}_t) - H(\tilde{f}_t) \right] \\ &\quad = \nu [H(\tilde{f}_t \circ \tilde{f}_t) - H(\tilde{f}_t)] \end{aligned}$$

In particular, by Lemma 2.2, $H(\tilde{f}_t)$ is increasing. Now recalling that $\tilde{f}_t = \mathscr{P}_\lambda f_t$, we have for any $t_2 > t_1$

$$H(\mathscr{P}_\lambda f_{t_2}) - H(\mathscr{P}_\lambda f_{t_1}) \geqslant v \int_{t_1}^{t_2} [H(\mathscr{P}_\lambda(f_t \circ f_t)) - H(\mathscr{P}_\lambda f_t)]\, dt \tag{2.16}$$

where we have used Lemma 2.8 on the right side. Then by Lemma 2.7 we have

$$H(f) \leqslant H(\mathscr{P}_\lambda(f_t \circ f_t)), \quad H(\mathscr{P}_\lambda f_t) \leqslant H(M^f)$$

Now the dominated convergence theorem allows us to take λ to 0 in (2.16), which proves the theorem. ∎

Remark. While the heuristic argument for Theorem 2.1 given at the end of the Introduction applies to the whole class of generalized Maxwelian collision kernels, the rigorous proof we have just given makes use of Lemma 2.5, which is not true is such generality. Probably the use of this lemma can be avoided. However, since we can extend the applicability of our main results to an even more general class of collision kernels by the monotonicity argument described in the Introduction, we have not investigated the matter too closely.

3. FISHER INFORMATION BOUNDS

The fisher information $I(f)$ of a velocity density is given by

$$I(f) = 4 \int_{\mathbb{R}^3} |\nabla f^{1/2}(v)|^2\, d^3v = \int_{\mathbb{R}^3} |\nabla \ln f(v)|^2 f(v)\, d^3v \tag{3.1}$$

for all f such that $f^{1/2}$ has a square-integrable distributional gradient. Otherwise it is defined to be ∞. Under the same conditions, there relative Fisher information $J(f)$ between f and M^f is defined by

$$\begin{aligned} J(f) &= \int_{\mathbb{R}^3} \left| \nabla \left(\frac{f}{M^f} \right)^{1/2} \right|^2 M^f(v)\, d^3v \\ &= \int_{\mathbb{R}^3} |\nabla \ln f(v) - \ln M^f(v)|^2 f(v)\, d^3v \end{aligned} \tag{3.2}$$

Then an integration by parts gives

$$J(f) = I(f) - \frac{3}{\theta(f)} \tag{3.3}$$

where $\theta(f)$ is the temperature of f. Since $J(f)$ clearly vanishes only when $f = M^f$,

$$I(f) \geqslant I(M^f) \tag{3.4}$$

which equality exactly when $f = M^f$, which is to be compared with (1.7).

Other useful properties which I and $-H$ have in common are strict convexity, weak upper semicontinuity, and strict superadditivity.[25,28] The convexity is easy to establish directly. Fisher information and entropy are related in many ways; e.g., both arise as rate functions in the theory of large deviations, as explained in refs. 25 and 28. The connection which is useful to us here is given by the following result due to Bakry and Emery[6] and in an equivalent form to Barron.[8]

Lemma 3.1. For any velocity density with zero bulk velocity and unit temperature and with $H(f) > -\infty$, $\lambda \mapsto H(\mathscr{P}_\lambda f)$ is continuous on $[0, \infty)$ and is continuously differentiable on $(0, \infty)$. Moreover,

$$\frac{d}{d\lambda} H(\mathscr{P}_\lambda f) = J(\mathscr{P}_\lambda f) \tag{3.5}$$

so that in particular

$$D(f) = \int_0^\infty J(\mathscr{P}_\lambda f)\, d\lambda \tag{3.6}$$

Proof. See, for example, Lemma 1.2 of ref. 17. Roughly, one computes the generator of $\mathscr{P}_\lambda$ by differentiating and finds

$$\mathscr{P}_\lambda = e^{\lambda \nabla \cdot (\nabla + v)}$$

Then formal differentiation in λ and integration by parts in v leads to (3.5). Abundant convexity and the regularizing properties of $\mathscr{P}_\lambda$ make it fairly easy to justify the formal computations and to prove the remaining assertions. ∎

Many entropy inequalities can most easily be proved by first proving an inequality for J and then using (3.6) to deduce an entropy inequality as a consequence. In our remarks at the end of this section we shall have more to say on why this is the case, but one advantage of working with J is its essentially quadratic nature.

The next result is an information counterpart of Lemma 2.2. It is also an analog of the Blachman–Stam convolution inequality for information.[9,54] However, its proof will only be given we have proved another technical lemma.

Lemma 3.2. For any velocity density f,

$$J(f \circ f) \leqslant J(f) \tag{3.7}$$

and when $J(f) \leqslant \infty$, there is equality exactly when $f = M^f$.

Remark. Just as Lemma 2.2 followed easily from the strict subadditivity of the entropy, Theorem 3.2 follows from the strict superadditivity of the Fisher information, which is a close relation of Lemma 3.3 below; see ref. 15. The proof of Lemma 3.3 is essentially the same as the proof of the $p = 2$ case of Theorem 2 in ref. 15, but there the remainder term on the right side of (3.9) was not obtained. It will be used in the next paper of this series.

Lemma 3.3. For all square-integrable functions G: $\mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}_+$ with square-integrable distribution gradient, let $g(y) = [\int_{\mathbb{R}^n} G^2(y, z)\, d^n z]^{1/2}$. Then g has a square-integrable distribution gradient and with

$$\begin{aligned} \|\nabla_y G\|^2 &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} |\nabla_y G(y, z)|^2\, d^m y\, d^n z \\ \|\nabla g\|^2 &= \int_{\mathbb{R}^m} |\nabla g(y)|\, d^m y \end{aligned} \tag{3.8}$$

we have

$$\begin{aligned} &\|\nabla_y G\|^2 - \|\nabla g\|^2 \\ &\quad \geqslant \frac{\|\nabla g\|}{\|\nabla_y G\|} \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} \left| \frac{\|\nabla_y G\|^2}{\|\nabla g\|} \nabla \ln g(y) - \nabla_y \ln G(y, z) \right|^2 \\ &\quad\quad \times G^2(y, z)\, d^m y\, d^n z \end{aligned} \tag{3.9}$$

In particular, the right side is strictly positive unless $G(y, z) = G_1(y)\, G_2(z)$ a.e. for some pair of functions G_1, G_2.

Proof. We will formally manipulate distribution derivatives, since the arguments required to justify this can be found in full detail in the proof of Theorem 2 in ref. 15. We have

$$\begin{aligned} \|\nabla g\|^2 &= \int_{\mathbb{R}^m} \nabla g(y) \cdot \frac{1}{g(y)} \left[\int_{\mathbb{R}^n} G(y, z)\, \nabla_y G(y, z)\, d^n z \right] d^m y \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} [\nabla \ln g(y) \cdot \nabla_y \ln G(y, z)]\, G(y, z)^2\, d^m y\, d^n z \end{aligned}$$

For any positive α, this is the same as

$$\frac{1}{2}\int_{\mathbb{R}^n}\int_{\mathbb{R}^m}\Bigg[\alpha^2|\nabla \ln g(y)|^2+\frac{1}{\alpha^2}|\nabla_y \ln G(y,z)|^2 \\ -\left|\alpha\nabla \ln g(y)-\frac{1}{\alpha}\nabla_y \ln G(y,z)\right|^2\Bigg]G^2(y,z)\,d^my\,d^nz$$

Choosing $\alpha=\|\nabla_y G\|/\|\nabla g\|$ leads to the result after simple calculation. ∎

Proof of Lemma 3.2. Fix any unit vector ω and put

$$G(v,v')=f^{1/2}(v+[(v'-v)\cdot\omega]\omega)\,f^{1/2}(v'-[(v'-v)\cdot\omega]\omega) \\ =f^{1/2}(\tilde{v})\,f^{1/2}(\tilde{v}')$$

As before, put

$$f\circ_\omega f(v)=\int_{\mathbb{R}^3}f(v+[(v'-v)\cdot\omega]\omega)\,f(v'-[(v'-v)\cdot\omega]\omega)\,d^3v'$$

Then

$$\nabla_v G(v,v')=\{[\nabla-(\omega\cdot\nabla)\omega]f^{1/2}\}(\tilde{v})\,f^{1/2}(\tilde{v}') \\ +f^{1/2}(\tilde{v})[(\omega\cdot\nabla)\,\omega f^{1/2}](\tilde{v}')$$

so that in the notation of Lemma 3.3 and again using $d^3v\,d^3v'=d^3\tilde{v}\,d^3\tilde{v}'$, $\|\nabla_v G\|^2=4I(f)$. Plainly, $\|\nabla g\|=4I(f\circ_\omega f)$, so

$$I(f\circ_\omega f)\leqslant I(f) \tag{3.10}$$

with equality just when $G(v,v')$ is a product of functions of v and v' alone. Since I is a convex function of the density,

$$I(f\circ f)=I\left(\int_{S^2}(f\circ_\omega f)\,d\omega\right)\leqslant\int_{S^2}I(f\circ_\omega f)\,d\omega \tag{3.11}$$

Combining (3.10) and (3.11) yields (3.7) with equality exactly when there is equality in (3.10) for almost every ω, and this is only the case when $f=M^f$. ∎

Lemma 3.4. For all velocity densities f with zero bulk velocity and unit temperature,

$$\lambda\mapsto J(\mathscr{P}_\lambda f) \tag{3.12}$$

is a continuous, monotone-decreasing function on $(0, \infty)$. The decrease is strict unless $f = M^f$ and

$$\lim_{\lambda \to 0} J(\mathscr{P}_\lambda f) = J(f) \tag{3.13}$$

whether $J(f)$ is finite or not.

Proof. First suppose that $J(f)$ is finite. Let g be any other density with zero bulk velocity, unit temperature, and finite J. For any $\lambda > 0$ put

$$f^{(\lambda)}(v) = e^{3\lambda} f(e^{\lambda} v)$$
$$g_{(\lambda)}(v) = (1 - e^{-2\lambda})^{-3/2}\, g((1 - e^{-2\lambda})^{-1/2} v)$$

Then if $*$ denotes convolution, $g_{(\lambda)} * f^{(\lambda)}$ is another zero-bulk-velocity, unit-temperature distribution. The Blachman–Stam inequality[9,54] says that

$$J(g_{(\lambda)} * f^{(\lambda)}) \leqslant (1 - e^{-2\lambda})\, J(g) + e^{-2\lambda} J(f) \tag{3.14}$$

and there is equality exactly when both f and g are Maxwellian. [The similarity between (3.7) and (3.14) should be noted; (3.14) can be given a simple proof starting from Lemma 3.3 very much along the lines of the proof of Lemma 3.2. This is done in ref. 15.]

It follows from (2.12) that when $g = M^f$, then $g_{(\lambda)} * f^{(\lambda)} = \mathscr{P}_\lambda f$. Since $J(M^f) = 0$, (3.14) implies the monotonic decrease.

Next, it is easy to see that for any $\lambda > 0$, $\mathscr{P}_\lambda f$ is finite for all f. McKean proves (3.12) in ref. 44, where J is replaced by I and $\mathscr{P}_\lambda$ is replaced by $\mathscr{G}_\lambda$. His argument is easily adapted to our case; or, just as easily, one can use his result and the relation between $\mathscr{P}_\lambda$ and $\mathscr{G}_\lambda$ expressed by (2.12) to prove the claim. By the semigroup property of $\mathscr{P}_\lambda$, continuity at $\lambda = 0$ implies continuity at all λ. ▌

To state the main result of this section, we introduce some useful notation. For all velocity density f_0 with zero bulk velocity and unit temperature, defined

$$\chi_f(\kappa) = H(\mathscr{P}_\kappa f) - H(f) \tag{3.15}$$

which is always well defined since $\mathscr{P}_\lambda f$ is bounded (as seen in the proof of Lemma 2.6), which implies that $H(\mathscr{P}_\lambda f)$ is finite.

Lemma 3.5. On the set of all velocity denities f_0 with zero bulk velocity and unit temperature,

$$f \mapsto \chi_f(\kappa)$$

is strictly convex. Moreover,

$$\chi_f(\kappa) \geqslant 0 \tag{3.16}$$

with equality exactly when $f = M^f$.

Proof. First note that $\mathscr{P}_\lambda f = M^f$ exactly when $f = M^f$, as can be seen by Fourier transforming (2.12), for example. By Lemma 3.1,

$$H(\mathscr{P}_\kappa f) - H(f) = \int_0^\kappa J(\mathscr{P}_\lambda f)\, d\lambda \tag{3.17}$$

and since $\mathscr{P}_\lambda$ is linear and J is strictly convex (on densities with unit temperature), it follows that $f \mapsto \chi_f(\kappa)$ is strictly convex for every $\kappa > 0$. ∎

Theorem 3.6. For all velocity densities f_0 with zero bulk velocity and unit temperature, let f_t be the corresponding solution of (2.1). Then for every $\kappa > 0$, $t \mapsto \chi_{f_t}(\kappa)$ is a monotone-decreaing function which is continuous on $[0, \infty)$ and is differentiable on $(0, \infty)$ and satifies

$$\frac{d}{dt} \chi_{f_t}(\kappa) \leqslant \nu[\chi_{f_t \circ f_t}(\kappa) - \chi_{f_t}(\kappa)] \leqslant 0$$

There is equality on the right side exactly when $f_t = M^{f_0}$.

When in addition $J(f_0)$ is finite, $t \mapsto J(f_t)$ is a monotone-decreasing function which is continuous on $[0, \infty)$ and is differentiable on $(0, \infty)$ and satisfies

$$\frac{d}{dt} J(f_t) \leqslant \nu[J(f_t \circ f_t) - J(f_t)] \leqslant 0$$

with equality on the right side exactly when $f_t = M^{f_0}$.

Proof. First note that because of (3.17) and (3.13), $\lim_{\kappa \to 0} (1/\kappa)\, \chi_f(\kappa) = J(f)$, so that the statements about $\chi_f(\kappa)$ imply those about $J(f)$, and only the former need be proved.

As in the proof of Theorem 2.1, we put $\bar{f}_t = \mathscr{P}_\lambda f_t$, so that solves (2.1) with initial condition $\bar{f}_0 = \mathscr{P}_\lambda f_0$.

The idea of the proof is really the same as in Theorem 2.1 with the convexity of $f \mapsto \chi_f(\kappa)$ in place of the concavity H. Formally the same argument can be made with J, which is also convex. However, it is difficult to work directly with $J(\bar{f}_t)$ in a rigorous manner because the formal time derivative—even taking the regularization into account—involves some delicate terms. By working with $\chi_f(\kappa)$, we avoid this, and in fact it turns out that $\chi_f(\kappa)$ is the more useful convex functional of f for our applications in the next section.

As before, put $g(s) = e^{-\nu s}\tilde{f}_t + (1 - e^{-\nu s})\tilde{f}_t \circ \tilde{f}_t$. Then, by the argument which established (2.15),

$$\frac{d}{ds}\chi_{g(s)}(\kappa)\bigg|_{s=0} = \frac{d}{dt}\chi_{\tilde{f}_t}(\kappa)$$

But

$$\begin{aligned}
&\frac{d}{dt}\chi_{g(s)}(\kappa)\bigg|_{s=0} \\
&\quad = \lim_{s \to 0}\frac{1}{s}\left[\chi_{[e^{-\nu s}\tilde{f}_t + (1-e^{-\nu s})\tilde{f}_t \circ \tilde{f}_t]}(\kappa) - \chi_{\tilde{f}_t}(\kappa)\right] \\
&\quad \leqslant \lim_{s \to 0}\frac{1}{s}\left[e^{-\nu s}\chi_{\tilde{f}_t}(\kappa) + (1-e^{-\nu s})\,\chi_{\tilde{f}_t \circ \tilde{f}_t}(\kappa) - \chi_{\tilde{f}_t}(\kappa)\right] \\
&\quad = \nu\left[\chi_{\tilde{f}_t \circ \tilde{f}_t}(\kappa) - \chi_{\tilde{f}_t}(\kappa)\right]
\end{aligned}$$

By Lemmas 3.1 and 3.2, this last quantity is strictly negative unless $f_t = M^{f_0}$.

This proves the result for the regularized solutions $\tilde{f}_t$ of the Boltzmann equation. Using the dominated convergence theorem to remove the regularization on the solutions exactly as in the proof of Theorem 2.1, we obtain the result. ▮

4. STABILITY FOR THE *H*-THEOREM

Before stating our results, we introduce some useful notation. We say that a function ψ is *decreasing to zero* if ψ: $\mathbb{R}_+ \to \mathbb{R}_+$ is continuous and monotone decreasing, and $\lim_{R \to \infty} \psi(R) = 0$. We say it is *decreasing strictly to zero* if in addition $\psi(R) > 0$ for all $R > 0$.

We say a function χ is *increasing from zero* if $\chi: \mathbb{R}_+ \to \mathbb{R}_+$ is continuous and monotone increasing and $\chi(0) = 0$. We say χ is *increasing strictly from zero* if in addition $\chi(R) > 0$ for all $R > 0$.

For any velocity density f with $\theta(f)$ finite, define

$$\psi_f(R) = \int_{|v| \geqslant R} |v|^2 f(v)\, d^3v \tag{4.1}$$

By the dominated convergence theorem, ψ_f is decreasing to zero.

By Lemma 3.2, for every f with zero bulk velocity and unit temperature, $\kappa \mapsto \chi_f(\kappa)$ is increasing from zero.

Our main result is the following:

Theorem 4.1. For any pair of functions ψ and χ which are, respectively, decreasing to zero and increasing from zero, there is a function $\Phi_{\psi,\chi}$ which is *strictly* increasing from zero so that for all velocity densities f with zero bulk velocity, unit temperature, and finite entropy, the inequalities

$$\psi_f \leqslant \psi, \qquad \chi_f \leqslant \chi \tag{4.2}$$

imply the inequality

$$[H(f \circ f) - H(f)] = [D(f) - D(f \circ f)] \geqslant \Phi_{\psi,\chi}(D(f)) \tag{4.3}$$

Proof. The argument we give closely follows the proof of Theorem 1.2 of ref. 17.

It suffices to show that given any $\varepsilon > 0$, there is a $\delta > 0$ so that

$$\psi_f \leqslant \psi, \qquad \chi_f \leqslant \chi, \qquad D(f) > \varepsilon \tag{4.4}$$

together imply

$$D(f) - D(f \circ f) \geqslant \delta$$

By Lemma 3.1,

$$\begin{aligned}[][D(f) - D(f \circ f)] &- [D(\mathscr{P}_\kappa f) - D(\mathscr{P}_\kappa f \circ \mathscr{P}_\kappa f)] \\ &= \int_0^\kappa [J(\mathscr{P}_\lambda f) - J(\mathscr{P}_\lambda f \circ \mathscr{P}_\lambda f)]\, d\lambda \geqslant 0 \end{aligned} \tag{4.5}$$

With $*$ denoting convolution, the fact that $D(f) - D(f * f) \geqslant D(\mathscr{P}_\kappa f) - D(\mathscr{P}_\kappa f * \mathscr{P}_\kappa f)$ was discovered by Dembo[22] independently of ref. 17; he gave no application.

Next, since

$$D(\mathscr{P}_\kappa f) = D(f) - \chi_f(\kappa) \geqslant D(f) - \chi(\kappa)$$

we can choose a $\kappa > 0$ small enough that

$$D(\mathscr{P}_\kappa f) \geqslant \frac{\varepsilon}{2}$$

for all densities f satisfying (4.4).

Furthermore, it is easy to see that if $\psi_f \leqslant \psi$, then

$$\psi_{\mathscr{P}_\kappa f}(R) \leqslant 2[\psi(R/2) + \psi_M(R/2)]$$

for all $R > 0$, $\kappa > 0$, where M is the Maxwellian with zero bulk velocity and unit temperature.

Now a standard argument shows that the closure of

$$\{\mathscr{P}_\kappa f \mid \psi_f \leqslant \psi\} = \mathscr{C}_{\kappa,\psi}$$

is strongly compact in $L^1(\mathbb{R}^3, (1+|v|^2)\, d^3v)$. For an explicit proof, see ref. 17.

By Lemma 2.6, there is constant A_κ so that the $L^1(\mathbb{R}^3, (1+|v|^2)\, d^3v)$ closure of $\mathscr{C}_{\kappa,\psi}$ is contained in

$$\{g \in L^1(\mathbb{R}^3, (1+|v|^2)\, d^3v) \mid |\ln g(v)| \leqslant A_\kappa(1+|v|^2)\} = \mathscr{B}_\kappa$$

which is also a closed set. We equip it with the $L^1(\mathbb{R}^3, (1+|v|^2)\, d^3v)$ topology. It is then easy to see that

$$g \mapsto H(g)$$

is continuous on each $\mathscr{B}_\kappa$. Then, since $g \mapsto g \circ g$ is continuous on $L^1(\mathbb{R}^3, (1+|v|^2)\, d^3v)$,

$$g \mapsto [D(g) - D(g \circ g)]$$

is a continunous function on $\mathscr{C}_{\kappa,\psi} \subset \mathscr{B}_\kappa$. Thus

$$\inf\{D(\mathscr{P}_\kappa f) - D(\mathscr{P}_\kappa f \circ \mathscr{P}_\kappa f) \mid \psi_f \leqslant \psi,\ \chi_f \leqslant \chi,\ D(f) \geqslant \varepsilon\}$$

is a minimum attained at some g_0 contained in the closure of $\mathscr{C}_{\kappa,\psi}$. Moreover, by the continuity of D on $\mathscr{B}_\kappa$, $D(g_0) \geqslant \varepsilon/2$. In particular, g is not M^f, so that $D(g_0) - D(g_0 \circ g_0) > 0$. Put

$$\delta = D(g_0) - D(g_0 \circ g_0)$$

Then by (4.5), for any f satisfying (4.4),

$$[D(f) - D(f \circ f)] \geqslant [D(\mathscr{P}_\kappa f) - D(\mathscr{P}_\kappa f \circ \mathscr{P}_\kappa f)] \geqslant \delta > 0 \quad \blacksquare$$

In order to apply successfully this to the Boltzmann equation, we must be able to apply it uniformly in time. That is, we need to be able to find a pair of functions ψ and χ, respectively, decreasing to zero and increasing from zero so that

$$\psi_{f_t} \leqslant \psi, \qquad \chi_{f_t} \leqslant \chi$$

for all $t \geqslant 0$.

The second condition by now poses no problem. Theorem 3.6 shows that as long as $H(f_0)$ is finite,

$$\chi_{f_t} \leqslant \chi_{f_0}$$

so χ_{f_0} is a suitable choice for χ.

The first condition is slightly more problematic, and we must rely on global moment bounds of Elmroth[30] to proceed. His method for obtaining global moment bounds applies in our setting without modification (and unfortunately without the possibility for significant simplification) and yields the result that if for some $p > 2$

$$\int_{\mathbb{R}^3} |v|^p f_0(v)\, d^3v = B < \infty \tag{4.6}$$

then there is a constant C_B depending only on B so that

$$\int_{\mathbb{R}^3} |v|^p f_t(v)\, d^3v \leqslant C_B$$

for all $t > 0$.

Clearly, when (4.6) holds

$$\psi_{f_t}(R) \leqslant \int_{|v| \geqslant R} R^{2-p} |v|^p f_t(v)\, d^3v \leqslant C_B R^{2-p}$$

Thus, when (4.6) is satisfied, $\psi(R) = C_B R^{2-p}$ is a suitable choice for ψ. It is now easy to prove the following convergence theorem for the Boltzmann equation.

Theorem 4.2. Let f_0 be a velocity density with $H(f_0) > -\infty$ and $\int_{\mathbb{R}^3} |v|^p f_0(v)\, d^3v < \infty$. Let f_t be the unique solution to (2.1) with the initial density f_0. Then there is a function Ψ decreasing to zero and depending only on χ_{f_0} and $\int_{\mathbb{R}^3} |v|^p f_0(v)\, d^3v$ such that

$$D(f_t) \leqslant \Psi(t)$$

In particular, f_t tends to M^{f_0} strongly in $L^1(\mathbb{R}^3)$ at a depending only on χ_{f_0} and $\int_{\mathbb{R}^3} |v|^p f_0(v)\, d^3v$.

Proof. This is now an easy consequence of Theorem 2.1, Theorem 4.1, and the remarks made just above. ∎

Remark. Since local L^1 stability is now well known,[27] one can combine it with the L^1 convergence to equilibrium implies by Theorem 4.2

to obtain a global L^1 statilibity result as in refs. 3 and 4. Furthermore, one can use Theorem 4.2 to control the time it takes for f_t to reach a given small neighborhood of M^{f_0} in $L^1(\mathbb{R}^3, (1+|v|^2)^{p'/2}\, d^3v)$ for any $2 < p' < p$. Then with the neighborhood chosen small enough, the linearized Boltzmann equation takes over an describes the final approach to equilibrium. From a spectral analysis of the linearized collision operator like that in refs. 3 and 4 one should obtain exponential bounds on the approach to equilibrium with constants which are uniform in the classes of initial conditions to which Theorem 4.2 refers.

Now for the first time since the beginning of Section 2 we turn to a wide class of collision kernels. The collision kernels we consider are those with rate functions which are uniformly bounded below. Making modifications on small subsets of phase space, this includes many important physical cases. For example, in the case of hard spheres, the microscopic rate function expressed as a function of v, v', and ω is a constant multiple of $|(v-v')\cdot\omega|$. This vanishes when ω is orthogonal to the relative velocity. Since such collisions have no effect, the effect of boosting the rate at which they occur should be negligible. Thus, while the hard-sphere collision kernel does not fall into the class we consider as it stands, it joins the class we consider after modification of its rate function on a *small set of phase which anyway corresponds to null collisions*. This is an indication that the next result can be extended to include the hard-sphere collision kernel even without modification, as will be seen in our next paper.

We give an analog only of Theorem 4.1. Then an analog of Theorem 4.2 again follows whenever the usual H-theorem can be established for the collision kernel under consideration. This can be done in great generality.[1,34] Of course, Theorem 3.6 is no longer available to control χ_{f_t} in terms of χ_{f_0}. The simplest way to proceed is to use a convexity property of the entropy production which implies that the entropy production of $\mathscr{P}_\kappa f_t$ is less than that of f_t itself. This leads to estimates on the rate at which $\|\mathscr{P}_\kappa f_t - M^{f_0}\|_{L^1(\mathbb{R}^3)}$ converges to zero. A better way to proceed is to prove propagation of smoothness results for the spatially homogeneous Boltzmann equation. In fact, Gustafsson's results can be used in this way. We will elaborate on these remarks, as well as the other remarks following Theorem 4.2, in the next paper of this series.

Theorem 4.3. Let $\mathscr{Q}$ be a collision kernel in which the rate function is bounded below by a positive constant ν:

$$\inf\{b(t, \vartheta) \mid t, 0, 0 \leqslant \vartheta \leqslant \pi\} = \nu > 0$$

Then for any pair of functions ψ and χ which are, respectively, decreasing to zero and increasing from zero, there is a function $\Phi_{\psi,\chi}$ which is *strictly* increasing from zero so that for all velocity densities f with zero bulk velocity, unit temperature, and finite entropy, the inequalities

$$\psi_f \leqslant \psi, \qquad \chi_f \leqslant \chi$$

imply the inequality

$$-\int_{\mathbb{R}^3} \ln f\, \mathcal{Q}(f, f)\, d^3v \geqslant \Phi_{\psi,\chi}(D(f))$$

Proof. Put $b_{(1)} = b - v$, $b_{(2)} = v$. Both are positive. Let $\mathcal{Q}_{(1)}$ and $\mathcal{Q}_{(2)}$ be the corresponding collision kernels. Then

$$-\int_{\mathbb{R}^3} \ln f\, \mathcal{Q}(f, f)\, d^3v$$
$$= \left[-\int_{\mathbb{R}^3} \ln f\, \mathcal{Q}_{(1)}(f, f)\, d^3v \right] + \left[-\int_{\mathbb{R}^3} \ln f\, \mathcal{Q}_{(2)}(f, f)\, d^3v \right] \qquad (4.7)$$

But the first term in (4.7) is positive by the easy part of the usual H-theorem, and Theorem 4.1 applies to the second term. ∎

ACKNOWLEDGMENTS

We thank Henry McKean for discussions on his work and for encouragement to undertake this work. We thank Rafaello Esposito, Joel Lebowitz, and Rossana Marra for many discussions on the Boltzmann equation and hydrodynamics, and much encouragement. We thank Carlo Cercignani, Carlo Scgara, and Diego DeFalco for discussions on the Boltzmann equation. Finally, we thank Avy Soffer for many discussions on Lyapunov functions in mathematical physics, and Elliott Lieb and Michael Loss for many discussions on the role of convexity in mathematical physics. M.C.C. was partially supported by Fundacão Luso Americana and JNICT.

REFERENCES

1. L. Arkeryd, Intermolecular forces of infinite range and the Boltzmann equation, *Arch. Rat. Mech. Anal.* **77**:11–21 (1981).
2. L. Arkeryd, Asymptotic behavior of the Boltzmann equation with infinite range forces, *Commun. Math. Phys.* **86**:475–484 (1982).
3. L. Arkeryd, Existence theorems for certain kinetic equations and large data, *Arch. Rat. Mech. Anal.* **103**:139–150 (1988).
4. L. Arkeryd, Stability in L^1 for the spatially homogeneous Boltzmann equation, *Arch. Rat. Mech. Anal.* **103**:151–168 (1988).

5. L. Arkeryd, R. Esposito, and M. Pulvirenti, The Boltzmann equation for weakly inhomogeneous data, *Commun. Math. Phys.* **11**:393–407 (1987).
6. D. Bakry and M. Emery, Diffusions hypercontractives, in *Séminaire de Probabilités XIX* (Lecture Notes in Mathematics 1123) (Springer, New York, 1985), pp. 179–206.
7. C. Bardos, F. Golse, and D. Levermore, Fluid dynamical limits of kinetic equations I: formal derivations, preprint.
8. A. R. Barron, Entropy and the central limit theorem, *Ann. Prob.* **14**:366–342 (1986).
9. N. M. Blachman, The convolution inequality for entropy powers, *IEEE Trans. Inform Theory* **2**:267–271 (1965).
10. L. D. Brown, A proof of the central limit theorem motivated byu the Cramer–Rao inequality, in *Statistics and Probability: Essays in Honor of C. R. Rao*, Kallianpur *et al.*, eds. (North-Holland, Amsterdam, 1982), pp. 314–328.
11. R. Caflisch, The fluid dynamical limit of the nonlinear Boltzmann lquation, *Commun. Pure Appl.* **33**:651–666 (1980).
12. R. Caflisch and G. Papanicolau, The fluid dynamical limit of a nonlinear Boltzmann equation, *Commun. Pure Appl. Math.* **32**:589–616 (1979).
13. T. Carleman, Sur la solution de l'équation intégrodifférential de Boltzmann, *Acta Math.* **60**:91–146 (1933).
14. T. Carleman, *Problèmes mathématique dans la théorie cinétique des gaz* Almqvist Wiksells, Uppsala, 1957).
15. E. A. Carlen, Superadditivity of Fisher's information and logarithmic Sobolev inequalities, *J. Funct. Anal.* **101**:194–211 (1991).
16. E. A. Carlen and M. C. Carvalho, The central limit theorem, entropy production, and stability of the rate of convergence to equilibrium for the Boltzmann equation, in *Proceedings of the 1991 Sienna on Probabilistic Methods in Mathematical Physics*, to appear.
17. E. A. Carlen and Soffer, Entropy production by block variable summation central limit theorems, *Commun. Math. Phys.* **140**:339–371 (1991).
18. E. A. Carlen and A. Soffer, Explicit entropy production bounds for convolution and central limit theorems with strong rate information, in preparation.
19. C. Cercignani, *H*-theorem and trend to equilibrium in the kinetic theory of gases, *Arch. Mech.* **34**:231–241 (1982).
20. C. Cercignani, *The Boltzmann Equation and Its Applications* (Springer-Verlag, New York, 1988).
21. I. Csizlar, Informationstheoretische Konvergenenzbegriffe im Raum der Wahrscheinlichkeitsverteilungen, *Pub. Math. Inst. Hung. Acad. Sci. VII Ser. A* **1962**:137–157.
22. A. Dembo, Information Inequalities and Uncertainty Principles, Standford University Technical Report No. 75 (1990).
23. A. DeMasi, R. Esposito, and J. Lebowitz, Incompressible Navier–Stokes and Euler limits of the Boltzmann equation, *Commun. Pure Appl. Math.* **42**:1189–1214 (1989).
24. L. Desvillettes, Entropy dissipation rate and convergence in kinetic equations, *Commun. Math. Phys.* **123**:687–702 (1989).
25. J. D. Deuschel and D. W. Stroock, *Large Deviations* (Academic Press, Boston, 1989).
26. G. DiBlasio, Strong solution for Boltzmann equation in the spatially homogeneous case, *Boll. Unione Mat. It.* **8**:127–136 (1973).
27. G. DiBlasio, Differentiability of spatially homogeneous solutions of the Boltzmann equation in the non-Maxwellian case, *Commun. Math. Phys.* **38**:331–352 (1974).
28. M. D. Donsker and S. R. S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time, I, *Commun. Pure Appl. Math.* **28**:1–47 (1975).

29. M. Dresden, *Kinetic Theory Applied to Hydrodynamics* (Field Research Laboratory, Dallas, Texas, 1956).
30. T. Elmroth, Global boundedness of moments of solutions of the Boltzmann equation for forces of infinite range, *Arch. Rat. Mech. Anal.* **82**:1–12 (1983).
31. T. Elmroth, On the *H*-function and convergence towards equilibrium for a space-homogeneous molecular density, *SIAM J. Appl. Math.* **44**:150–159 (1984).
32. G. Ford and G. Uhlenbeck, *Lectures in Statistical Mechanics* (American Mathematical Society, Providence, Rhode Island, 1963).
33. J. W. Gibbs, *Elementary principles in statistical mechanics, developed with especial reference to the foundations of thermodynamics* (Scribner, New York, 1902), esp. Theorem II, Chapter XI.
34. T. Gustafsson, Global L^p properties for the spatially homogeneous Boltzmann equation, *Arch. Rat. Mech. Anal.* **103**:1–38 (1988).
35. S. Kawashima, A. Matsumura, and T. Nishida, On the fluid-dynamical approximation to the Boltzmann equation at the level of the Navier–Stokes equation, *Commun. Math. Phys.* **70**:97–124 (1979).
36. S. Kullback, A lower bound for discrimination information in terms of variation, *IEEE Trans. Inform. Theory* **4**:126–127 (1967).
37. M. Lachowicz, On the initial layer and tyhe existence theorem for the nonlinear Boltzmann equation, *Math. Meth. Appl. Sci.* **9**:342–366 (1987).
38. J. Lebowitz and E. Montroll, *Nonequilibrium Phenomena I: The Boltzmann Equation* (North-Holland, Amsterdam, 1984).
39. E. H. Lieb, Proof of an entropy conjecture of Wehrl, *Commun. Math. Phys.* **62**:35–41 (1978).
40. Ju. V. Linnik, An information theoretic proof of the central limit theorem with Lindeberg conditions, *Theory Prob. Appl.* **4**:288–299 (1959).
41. Ju. V. Linnik, On certain connections of the information theory of C. Shannon and R. Fisher with the theory of symmetrization of random vectors, in *Transactions of the Second Prague Conference on Information Theory* (Academic Press, New York, 1960), pp. 313–327.
42. N. Maslova and Yu. Romanovskii, Justification of the Hilbert method in the theory of Kinetic equations, *Zh. Vychisl. Mat. Mat. Fiz.* **27**:1680–1689 (1987).
43. J. C. Maxwell, On the dynamical theory of gases, *Phil. Trans. R. Soc. Lond.* **157**:49–88 (1867), esp. p. 63.
44. H. P. McKean, Speed of approach to equilibrium for Kac's caricature a Maxwellian gas, *Arch. Rat. Mech. Anal.* **21**:343–367 (1966).
45. H. P. McKean, An exponential formula for solving Boltzmann's equation for a Maxellian gas, *J. Comb. Theory* **2**:358–382 (1967).
46. H. P. McKean, Chapman–Enskog–Hilbert expansion for a class of solutions of the telegraph equation, *J. Math. Phys.* **8**:547–552 (1967).
47. H. P. McKean, A simple model of the derivation of fluid mechanics from the Boltzmann equation, *Bull. Am. Math. Soc.* **75**:1–10 (1969).
48. H. P. McKean, The central limit theorem for Carleman's equation, *Isr. J. Math.* **21**:54–92 (1975).
49. D. Morgenstern, General existence and uniqueness proof for spatially homogeneous solutions of the Maxwell–Boltzmann equation in the case of Maxwellian molecules, *Proc. Natl. Acad. Sci. USA* **40**:719–721 (1954).
50. D. Morgenstern, Analytical studies related to the Maxwell–Boltzmann equation, *J. Rat. Mech. Anal.* **4**:154–183 (1955).

51. T. Nishida, Fluid dynamical limit of the nonlinear Boltzmann equation to the level of the incompressible Euler equation, *Commun. Math. Phys.* **61**:119–148 (1978).
52. C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, Illinois, 1949).
53. T. Sideris, Formation of singularities in three-dimensional compressible fluids, *Commun. Math. Phys.* **101**:475–485 (1985).
54. A. Stam, Some inequalities satisfied by the quantities of information of Fisher and Shannon, *Inform. Control* **2**:101–112 (1959).
55. G. Toscani, New *a priori* estimates for the spatially homogeneous Boltzmann equation, Ferrara preprint (1991).
56. G. Toscani, Strong convergence towards equilibrium for a gas of Maxwellian pseudomolecules, Ferrara preprint (1991).
57. C. Truesdell, On the pressures and flux of energy in a gas according to Maxwell's kinetic theory, Part II, *J. Rat. Mech. Anal.* **5**:55–128 (1956).
58. G. Uhlenbeck, The statistical mechanics of non-equilibrium phenomena, 1954 Higgins Lectures, Mimeographed notes, Fine Hall Library, Princeton University, Princeton, New Jersey.
59. S. Ukai, On the existence of global solutions of mixed problem for nonlinear Boltzmann equation, *Proc. Jpn. Acad.* **50**:179–184 (1974).
60. E. Wild, On Boltzmann's equation in the kinetic theory of gases, *Proc. Camb. Phil. Soc.* **47**:602–609 (1951).