Journal of Statistical Physics, Vol. 74, Nos. 3/4, 1994

Entropy Production Estimates for Boltzmann Equations with Physically Realistic Collision Kernels

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

Received June 10, 1993

We establish strict entropy production bounds for the Boltzmann equation with the hard-sphere collision kernel. Using these entropy production bounds, we prove results asserting that the rate at which strong L^1 convergence to equilibrium occurs is uniform in wide classes of initial data. This extends our previous results in this direction, which applied only to a very special collision kernel. Moreover, the present results provide computable lower bounds; compactness arguments are entirely avoided. The uniformity is an important ingredient in our study of scaling limits of solutions of the non-spatially homogeneous Boltzmann equation, and is the main focus of this paper. However, the results obtained here provide the only framework known to us in which one can obtain computable estimates on the time it takes a solution of the spatially homogeneous Boltzmann equation with initial data far from equilibrium to reach any given small strong L^1 neighborhood of equilibrium.

KEY WORDS: Entropy production; Boltzman Equations.

1. INTRODUCTION

In this paper we bound the rate of approach to equilibrium in the strong L^1 sense by solutions to the spatially homogeneous Boltzmann equation with physically realistic collision kernels. Our main tool is a new quantitative entropy production inequality. Using it, we develop a method for computing a bound on the time it takes a solution which starts arbitrarily far from equilibrium to reach any given neighborhood of equi-

[1] School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia 30332.
[2] On leave from C.F.M.C. and Departamento de Matemática da Faculdade de Ciencias de Lisboa, 1700 Lisboa codex, Portugal.

0022-4715/94/0200-0743$07.00/0

librium. Once an appropriately small neighborhood of equilibrium is attained, methods based on the spectral theory of the linearized collision operator are applicable, and provide control over the rate of the final approach to equilibrium.[4] Our result enables us to estimate how long it takes a solution with initial data far from equilibrium to reach any desired neighborhood of equilibrium. *In particular, we can control the way in which this time varies with the initial data.*

In a previous paper on the subject, we presented a qualitative version of these results for a special collision kernel, and discussed its application to the study of the relation between hydrodynamics and scaling limits of the Boltzmann equation. This paper builds on the results obtained in that paper. Familiarity with the first paper will surely facilitate the reading of this paper. We shall try, however, to make our exposition readable on its own.

In the rest of this introduction, we first describe our results. We then explain the methods by which we obtain them, and finally, we discuss the relations between our results and methods and those of others.

We begin by introducing some notation and terminology. The spatially homogeneous Boltzmann equation, which, under certain conditions,[15,18,26] describes the time evolution of the velocity distribution of a gas of molecules undergoing binary collisions, is

$$\frac{\partial}{\partial t} f_t(v) = \mathcal{Q}(f_t, f_t)(v) \tag{1.1}$$

where

$$\mathcal{Q}(f, f)(v) = \int_{S^2} \int_{\mathbb{R}^3} [f(\tilde{v}) f(\tilde{v}') - f(v) f(v')]\, b(v, v', \omega)\, d^3v'\, d\omega \tag{1.2}$$

The velocities $\tilde{v}$ and $\tilde{v}'$ are related to the velocities v and v' and the unit vector ω by

$$\tilde{v} = v + (\omega \cdot (v' - v))\omega \qquad \text{and} \qquad \tilde{v}' = v' - (\omega \cdot (v' - v))\omega \tag{1.3}$$

Here, ω is a unit vector, and throughout the paper, $d\omega$ denotes the uniform probability measure on S^2. Finally, the *microscopic rate function* $b(v, v', \omega)$ is the product of the relative speed $|v - v'|$ and the differential cross section for the elastic collision taking (v, v') into $(\tilde{v}, \tilde{v}')$. For example, in the case of so-called hard-sphere collisions, the microscopic rate function is $b(v, v', \omega) = |(v - v') \cdot \omega|$. Our method applies to this case and others to be described below.

Since f in (1.1) represents a density, we look for solutions of (1.1) in the space of nonnegative integrable functions. It is convenient to normalize

f to be a probability density instead of a mass density; this normalization is conserved in time by solutions to (1.1).

In addition to $\int_{\mathbb{R}^3} f_t(v)\, d^3v$, there are two more physically significant functionals of f which are conserved under the time evolution described by (1.1), and one which is monotone. The two conserved functionals are the *bulk velocity* $u(f)$ and the *temperature* $\theta(f)$, which are defined by

$$u(f) = \int_{\mathbb{R}^3} v f_t(v)\, d^3v, \qquad \theta(f) = \tfrac{1}{3} \int_{\mathbb{R}^3} |v-u|^2 f_t(v)\, d^3v \tag{1.4}$$

The monotone functional of f is the *entropy* $H(f)$, which is defined by

$$H(f) = -\int_{\mathbb{R}^3} \ln f(v)\, f(v)\, d^3v \tag{1.5}$$

This always well defined, admitting $-\infty$ as a possible value, for velocity distributions with finite temperature. The Boltzmann H-theorem says that for solutions of (1.1), $H(f_t)$ is monotone increasing.

Throughout this paper we shall consider solutions of (1.1) with initial data that has finite temperature and finite entropy. For all the collision kernels we consider here, global existence and uniqueness of such solutions for quite general initial data have been established, along with the fact that for such solutions, the mass, bulk velocity, and temperature are conserved, and the entropy is monotone increasing (see, e.g., refs. 1 and 29). That is, the conservation laws and the H-theorem, whose classical derivations are somewhat formal, do actually hold in our setting.

The equilibrium solutions of (1.1) are the Maxwellian densities; i.e., those of the form

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

We shall denote by M^f the Maxwellian density with the same bulk velocity u and temperature θ as f.

The *entropy production* at f is defined by

$$-\int_{\mathbb{R}^3} \ln f(v)\, \mathcal{Q}(f, f)(v)\, d^3v \tag{1.6}$$

since at f_t this equals $(d/dt)\, H(f_t)$ when f_t is a solution of (1.1) and the integrand in (1.6) is integrable at f_t.

A more precise statement of the Boltzmann H-theorem is that

$$-\int_{\mathbb{R}^3} \ln f(v)\, \mathcal{Q}(f, f)(v)\, d^3v \geqslant 0 \tag{1.7}$$

with equality exactly when $f = M^f$. While this does show that Maxwellian densities are the only equilibrium solutions to the Boltzmann equation, it does not by itself show that all solutions f_t converge to the corresponding equilibrium M^{f_0}, and it certainly does not by itself give us information on the rate at which such convergence would occur.

To see how the H-theorem might be strengthened to yield rate information, introduce the relative entropy $D(f)$ of f, always with respect to M^f, by

$$D(f) = \int_{\mathbb{R}^3} \left(\frac{f}{M^f}\right) \ln\left(\frac{f}{M^f}\right) M^f \, d^3v \tag{1.8}$$

Because $\ln M^f$ is quadratic,

$$D(f) = H(M^f) - H(f) \tag{1.9}$$

It is clear from Jensen's inequality that $D(f) \geqslant 0$, with equality exactly when $f = M^f$; this is known as Gibbs' lemma. More is true; when $D(f)$ is close to zero, then f is close to M^f. This stability result for Gibbs' lemma, which is precisely expressed by the Csizlar–Kullback inequality,[16]

$$D(f) \geqslant \tfrac{1}{2} \|f - M^f\|^2_{L^1(\mathbb{R}^3)} \tag{1.10}$$

is one of the reasons we can use entropy production to control the strong $L^1(\mathbb{R}^3)$ convergence to equilibrium.

Our main result is a quantitative entropic stability result for the H-theorem. Precise statements are made in Theorems 1.1 and 3.1 below, but what we shall establish is a bound of the form

$$-\int_{\mathbb{R}^3} \ln f(v) \, \mathcal{Q}(f, f)(v) \, d^3v \geqslant \Phi(D(f)) \tag{1.11}$$

where $\Phi: \mathbb{R}_+ \to \mathbb{R}_+$ is a strictly increasing function that naturally depends on the microscopic rate function b, and that depends on f in a simple, explicitly computable way. We shall soon explain what this simple dependence is, but for now we note that what is important is that Φ depends on f only through characteristics *which are preserved by the time evolution*. This means that given initial data f_0, we can choose a single function Φ so that (1.11) holds for all f_t. Then since $M^{f_t} = M^{f_0}$,

$$\frac{d}{dt} H(f_t) = -\frac{d}{dt} D(f_t)$$

and we have

$$\frac{d}{dt} D(f_t) \leqslant -\Phi(D(f_t)) \tag{1.12}$$

Since Φ is computable in terms of f_0, a standard comparison theorem for differential inequalities now gives us the desired quantitative rate information.

The fact that Φ *must* depend on the initial data f_0 is made clear by a result of Bobylev,[7] who has shown that for Maxwellian molecules, there exist for each $\lambda > 0$ physical initial data for which

$$\|f_t - M^{f_0}\|_{L^1(\mathbb{R}^3)} \geqslant Ce^{-\lambda t}$$

Roughly speaking, the characteristics of the initial data that determine Φ measure how far "within" the class of finite-temperature, finite-entropy densities it is.

The temperature part is easy. Let f be any finite-temperature density. For each $R > 0$, define

$$\psi_f(R) = \frac{1}{\theta(f)} \int_{|v - u(f)| \geqslant R} |v - u(f)|^2 f(v)\, d^3v \qquad (1.13)$$

Evidently, $\psi_f(R)$ decreases to zero as R increases. The rate at which this occurs is a measure of the *concentration* of the velocity distribution f. Since $\psi_f(R)$ can decrease arbitrarily slowly, a bound on its rate of decrease is a measure of how far "within" the class of finite-temperatue densities our data is.

The entropy part requires an introductory discussion, since it is based on properties of the adjoint Ornstein–Uhlenbeck semigroup. This is the semigroup $\{\mathscr{P}_\lambda : \lambda \geqslant 0\}$ of operators on $L^1(\mathbb{R}^n, d^nv)$ whose action is defined by

$$\mathscr{P}_\lambda \eta(y) = \int_{\mathbb{R}^n} M_{(1 - e^{-2\lambda})}(y')\, e^{n\lambda} \eta(e^\lambda (y - y'))\, d^n y' \qquad (1.14)$$

where M_α denotes the Maxwellian with zero bulk velocity and temperature α. This can be expressed most clearly in probabilistic terms: if X is an $\mathbb{R}^3$-valued random variable with density f, and Z is an $\mathbb{R}^3$-valued random variable with the density M^f, *independent* of X, then $\mathscr{P}_\lambda f$ is the density of $e^{-\lambda} X + (1 - e^{-2\lambda})^{1/2} Z$. Evidently $\mathscr{P}_\lambda$ effects a sort of Maxwellian regularization, and it is clear that

$$\lim_{\lambda \to 0} \mathscr{P}_\lambda f = f \qquad \text{and} \qquad \lim_{\lambda \to \infty} \mathscr{P}_\lambda f = M^f$$

for all densities f with zero bulk velocity and unit temperature. It is not difficult to show (ref. 12, Lemma 2.7) that $H(\mathscr{P}_\lambda f) \geqslant H(f)$.

For all finite-temperature, finite-entropy densities f, we define the positive function χ_f on $\mathbb{R}_+$ by

$$\chi_f(\lambda) = H(\mathscr{P}_\lambda f) - H(f) \tag{1.15}$$

We have shown that $\lambda \mapsto \chi_f(\lambda)$ increases continuously from zero, and that, furthermore, for fixed $\lambda > 0$, $f \mapsto \chi_f(\lambda)$ is strictly convex and vanishes just when $f = M^f$ (ref. 12, Lemma 3.5). We shall explain later how these statements are proved, and why the regularization provided by $\mathscr{P}_\lambda$ is a natural tool to use in investigating the Boltzmann equation, but for now we simply note that an upper bound on the rate at which χ_f increases from zero is a measure of how far "within" the class of finite-entropy densities a given finite-temperature, finite-entropy density f lies.

The function Φ in our main estimate depends on f only through the rate of decrease of ψ_f and the rate of increase of χ_f.

To obtain our lower bounds on the entropy production, we first consider the physically artificial Boltzmann equation arising when all kinematically allowed collisions are run at the same rate; i.e., in which the microscopic rate function b equals constant ν. Because of its simplicity, and in particular the extra symmetries that it posseses, we are able to control solutions of this Boltzmann equation more closely than is *directly* possible with a physical collision law.

It is worthwhile to obtain this control because entropy production estimates for the simple Boltzmann equation can be transfered to physical Boltzmann equations essentially on account of the simple fact that the *entropy production is monotone in the microscopic rate function.* Observe that while the hard-sphere rate function is not bounded below, it vanishes only when the relative velocity $v - v'$ is orthogonal to ω. As is evident in (1.3), these collisions have no effect. Thus, there will be only a small price to pay for modifying the hard-sphere rate function so that it is uniformly positive, and this small price can be safely absorbed into the entropy production inherited by monotonicity from the corresponding collision kernel with constant rate.

The passage from the constant rate function to the hard-sphere rate function that we have just described is explained in full detail in Section 3. We show in fact that classical estimates of a type first considered by Carleman[9,10] provide the means to carry out the monotonicity argument described above. Carleman's estimates, which we recall in Lemma 3.1, have been greatly generalized by recent authors, notably Elmroth[11] and Gustafsson.[23] Their results could be used to prove a direct analog of Theorem 3.1, our theorem on the strong rate of convergence to equilibrium for solutions to the hard-sphere Boltzmann equation, for all initial data f_0 satisfying

$$\int_{\mathbb{R}^3} |f_0(v)|^p \, d^3v < \infty \qquad \text{and} \qquad \int_{\mathbb{R}^3} |v|^{2p} f_0(v) \, d^3v < \infty$$

for some $p > 1$. However, the analysis would be more involved and less explicit. For this reason, we instead work, following Carleman, with Lipschitz initial data possessing many moments. And even in the number of moments we have been extravagent when convenience could be bought. Still, the class of initial data with which we work is broad enough to include much that is of physical interest, and so we have sacrificed nothing essential in our efforts to keep things simple.

With this much said about the hard-sphere case, we now turn to the matter that is the subject of Section 2: the entropy production estimates for Boltzmann equation with constant microscopic rate function $b = \nu$ that are the basis of our monotonicity arguments.

First, it is advantageous to rewrite this equation as follows: We introduce a bilinear operation $f, g \mapsto f \circ g$ on densities, called the Wild convolution[39] of f and g, by

$$f \circ g(v) = \int_{S^2} \int_{\mathbb{R}^3} f(v + ((v' - v) \cdot \omega)\omega) \, g(v' - ((v' - v) \cdot \omega)\omega) \, dv' \, d\omega \qquad (1.16)$$

The Boltzmann equation with constant rate function $b = \nu$ can then be written as

$$\frac{\partial}{\partial t} f_t(v) = \nu[f_t \circ f_t(v) - f_t(v)] \qquad (1.17)$$

[Note that we are using f_t to denote $f(t, \cdot)$.]

The use we make of the special symmetries of this equation stems from the fact that on account of those symmetries, its time evolution commutes with the action of the adjoint Ornstein–Uhlenbeck semigroup. Equivalently put, the Wild convolution commutes with the action of this semigroup (ref. 12, Lemma 2.8):

$$\mathscr{P}_\lambda(f \circ f) = (\mathscr{P}_\lambda f) \circ (\mathscr{P}_\lambda f) \qquad (1.18)$$

This result is actually a simple variant of Morgenstern's result[30] that

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

and $\mathscr{G}_\lambda$ is an operator from the heat semigroup.

Next note that formally, for any solution f_t of (1.17),

$$f_{t+\Delta t} = (1 - \nu\, \Delta t)\, f_t + (\nu\, \Delta t)\, f_t \circ f_t + o(\Delta t)$$

i.e., $f_{t+\Delta t}$ is a convex combination of f_t and $f_t \circ f_t$. Then, since the entropy functional is concave, we formally have

$$\frac{H(f_{t+\Delta t}) - H(f_t)}{\Delta t} \geqslant [H(f_t \circ f_t) - H(f_t)] + o(1) \tag{1.19}$$

Indeed, using a Maxwellian regularization argument made available by (1.18), we have proved in our previous paper (ref. 12, Theorem 2.1) that 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

$$\frac{d}{dt} H(f_t) \geqslant \nu [H(f_t \circ f_t) - H(f_t)] \tag{1.20}$$

We have also established there that (ref. 12, Lemma 2.2)

$$H(f \circ f) - H(f) \geqslant 0 \tag{1.21}$$

and moreover the inequality is strict unless $f = M^f$.

Thus, (1.21) implies the Boltzmann H-theorem for the special Boltzmann equation (1.17), and our problem of obtaining lower bounds for the entropy production is reduced to that of obtaining lower bounds on $H(f \circ f) - H(f)$.

If the Wild convolution $\circ$ in (1.21) were replaced by the ordinary convolution $*$, (1.21) would become a special case of the Shannon–Stam inequality[32] for the entropy of convolutions. It is fortunately the case that the Wild convolution has enough in common with the usual convolution that the analogy with the Shannon–Stam inequality can be made quite precise. In particular, we are able to adapt, and further develop, techniques introduced by Carlen and Soffer in their proof[13] of a stability result for the Shannon–Stam inequality.

The main reason that (1.18) is so important to us is that, as we shall explain below, it allows us to rewrite advantaneously the entropy difference $H(f \circ f) - H(f)$ in terms of a quantity called the Fisher information.

The *Fisher information* $I(f)$ is defined by

$$I(f) := 4 \int_{\mathbb{R}^3} |\nabla [f(v)]^{1/2}|^2\, d^3v = \int_{\mathbb{R}^3} |\nabla \ln f(v)|^2\, d^3v\, f(v) \tag{1.22}$$

The *relative Fisher information* $J(f)$ of f with respect to M^f is defined by

$$J(f) := 4 \int_{\mathbb{R}^3} \left| \nabla \left(\frac{f}{M^f} \right)^{1/2} \right|^2 M^f(v)\, d^3v$$
$$= \int_{\mathbb{R}^3} |\nabla \ln f(v) - \nabla \ln M^f(v)|^2 f(v)\, d^3v \tag{1.23}$$

Easy computations reveal that

$$J(f) = I(f) - I(M^f) \tag{1.24}$$

and that $I(M^f) = 3/\theta(f)$, and hence that

$$I(f) = J(f) + 3/\theta(f) \tag{1.25}$$

The Fisher information functional is a convex functional closely related to the entropy functional. The analog of Gibb's lemma, for example, is the fact that $I(f) \geqslant I(M^f)$ with equality exactly when $f = M^f$; this fact is easily deduced from (1.24) and (1.23). Further background on the Fisher information functional and its relation to entropy functional can be found in ref. 13.

The crucial connection between the entropy and the Fisher information is mediated by the adjoint Ornstein–Uhlenbeck semigroup. As demonstrated in ref. 13, Lemma 1.2, $\lambda \mapsto H(\mathscr{P}_\lambda f)$ is continuous on $[0, \infty)$ and continuously differentiable on $(0, \infty)$ for all finite-temperature, finite-entropy distributions f. Moreover, for such densities f we have

$$\frac{d}{d\lambda} H(\mathscr{P}_\lambda f) = J(\mathscr{P}_\lambda f) \tag{1.26}$$

A consequence of this is that

$$D(f) = \int_0^\infty J(\mathscr{P}_\lambda f)\, d\lambda \tag{1.27}$$

Then, since $M^f = M^{f \circ f}$, $D(f) - D(f \circ f) = H(f \circ f) - H(f)$ and $J(f) - J(f \circ f) = I(f) - I(f \circ f)$. Therefore, using (1.18) and (1.26), we have

$$H(f \circ f) - H(f) = \int_0^\infty [I(\mathscr{P}_\lambda f) - I(\mathscr{P}_\lambda f \circ \mathscr{P}_\lambda f)]\, d\lambda$$
$$\geqslant \int_a^b [I(\mathscr{P}_\lambda f) - I(\mathscr{P}_\lambda f \circ \mathscr{P}_\lambda f)]\, d\lambda \tag{1.28}$$

for all $0 < a < b < \infty$.

The problem of obtaining lower bounds on $H(f \circ f) - H(f)$ is now reduced to the problem of obtaining a lower bound on $[I(\mathscr{P}_\lambda f) - I(\mathscr{P}_\lambda f \circ \mathscr{P}_\lambda f)]$ in terms of $D(f)$ for values of λ that are neither too large nor too small. Since the microscopic rate function is fixed, the bound is only allowed to depend on the rate of increase of χ_f and on the rate of decrease of ψ_f.

We now explain our solution of this problem. We shall assume throughout our discussion that $\chi: \mathbb{R}_+ \mapsto \mathbb{R}_+$ is a *fixed* continuous increasing function with the property that $\chi(0) = 0$, and that $\psi: \mathbb{R}_+ \mapsto \mathbb{R}_+$ is a *fixed* continuous decreasing function with the property that $\lim_{R \to \infty} \psi(R) = 0$. Suppose also that f is a density with zero bulk velocity, unit temperature, and $D(f) \geqslant \varepsilon$ for some *fixed* $\varepsilon > 0$, and suppose that $\chi_f \leqslant \chi$ and $\psi_f \leqslant \psi$.

The function χ_f enters our analysis as follows: we base our selection of a and b in (1.28) on the rate of increase of χ_f. We show in Lemma 2.1 that there is a number $\Lambda(\chi, \varepsilon) > 0$, explicitly computable in terms of the indicated arguments, so that

$$J(\mathscr{P}_\lambda f) \geqslant D(f) \qquad \text{whenever} \quad \lambda \leqslant \Lambda(\chi, D(f)) \tag{1.29}$$

Since $J(\mathscr{P}_\lambda f) \geqslant D(f)$ throughout the region of integration on the right side of (1.28), our problem is now reduced to that of obtaining a lower bound on

$$\inf\{I(g) - I(g \circ g) \mid J(g) \geqslant \varepsilon, g = \mathscr{P}_\lambda f \text{ and } \Lambda(\lambda, \varepsilon)/2 \leqslant \lambda \leqslant \Lambda(\chi, \varepsilon)\} \tag{1.30}$$

Since $\mathscr{P}_\lambda$ has a number of evident regularizing properties, it is easier to estimate the infimum in (1.30) than it is to estimate directly $H(f \circ f) - H(f)$. The regularizing properties that are crucial here are that $\mathscr{P}_\lambda$ is *smoothing* and that it *produces pointwise Maxwellian lower bounds.* In fact, we show in Lemma 2.2 that for all g of the form $g = \mathscr{P}_\lambda g_0$, where g_0 is any density with $u(g_0) = 0$ and $\theta(g_0) = 1$,

$$|\nabla \ln g(v)|^2 \leqslant K_1(\lambda)\, \rho(v) \tag{1.31}$$

where ρ is a density with ψ_ρ comparable to ψ_g, and

$$g(v) \geqslant K_2(\lambda)\, M_{(1 - e^{-2\lambda})/2}(v) \tag{1.32}$$

where the constants are explicitly computed functions of λ alone. Recall that M_α denotes the Maxwellian with zero bulk velocity and temperature α.

Maxwellian lower bounds, such as we have in (1.32), seem to be essential in estimating the entropy production, and the fact that $\mathscr{P}_\lambda$ produces them automatically saves us from imposing any artificial assumptions about the existence of such bounds.

In addition to the help provided by the regularization, we are helped by the fact that $I(g)$ is quadratic in $g^{1/2}$. Using all of this together, we show in Lemma 2.4 that the infimum in (1.30) is bounded below by

$$K_3(\lambda) \inf \left\{ \int_{\mathbb{R}^3} |cv + d + \nabla \ln g(v)|^2 \, M_{(1-e^{-2\lambda})/2}(v) \, d^3v \,\middle|\, c \in \mathbb{R}, d \in \mathbb{R}^3 \right\} \tag{1.33}$$

where, again, K_3 is an explicitly computed function of λ. This is the most involved part of our argument. In an intermediate step, we need bounds on the eigenvalues of an operator closely related to the quadratic form on the right side of (1.33); these are worked out in Lemma 2.4.

The right side of (1.33) bears a close resemblance to a multiple of $J(g)$. We now make use, for the first time, of the fact that $\psi_f \leqslant \psi$ to turn this resemblance into a bound. Because of (1.31) and our assumption on ψ_g, we can choose an R large enough that at the minimizing c and d,

$$\int_{|v| \geqslant R} |\nabla \ln g(v) + cv + d|^2 \, g(v) \, d^3v \leqslant J(g)/2 \tag{1.34}$$

For this value of R, consider

$$\int_{|v| \geqslant R} |cv + d + \nabla \ln g(v)|^2 \, M_{(1-e^{-2\lambda})/2}(v) \, d^3v \tag{1.35}$$

which is obviously less than the quantity in (1.33). Call this value of R, which depends on ψ and $\varepsilon := J(g)$, $R(\psi, \lambda, \varepsilon)$.

Now clearly, we can replace $M_{(1-e^{-2\lambda})/2}$ in (1.35) by g at the expense of a multiplicative constant depending on λ and $R(\psi, \lambda, \varepsilon)$, i.e., on λ, ψ, and ε. What we are left with is a multiple of $J(g)$ which exceeds $D(f)$ for the values of λ under consideration. This gives us our bound on $H(f \circ f) - H(f)$.

Now simply keeping track of the explicit form of the constants in the cited lemmas, we prove:

Theorem 1.1 (Entropy production bounds for the constant rate collision kernel). Let f be a density with zero bulk velocity and unit temperature. Let χ be a function that increases continuously from zero, and let ψ be a function that decreases monotonically to zero. Suppose that

$$\chi_f \leqslant \chi \qquad \text{and} \qquad \psi_f \leqslant \psi \tag{1.36}$$

Then,

$$H(f \circ f) - H(f) \geqslant \Phi_{\chi,\psi}(D(f)) \tag{1.37}$$

where

$$\Phi_{\chi,\psi}(\varepsilon) := (\Lambda(\chi,\varepsilon)\varepsilon/4)\,\Gamma_\psi(\Lambda(\chi,\varepsilon)/2,\varepsilon)$$

$$\Gamma_\psi(\lambda,\varepsilon) := [10^{-1}e^{-4/(5\lambda)}]^2\,e^{-R_\psi(\varepsilon,\lambda)^2/\lambda}\left(\frac{\varepsilon}{6+2\varepsilon}\right) \tag{1.38}$$

$$R_\psi(\varepsilon,\lambda) := \inf\left\{R \mid [2\psi_f(R/2)+\psi_{M_{3\lambda}}(R/2)+\psi_{M_{(1-\varepsilon-2\lambda)}}(R/2)] \leqslant \left(\frac{\varepsilon}{3+\varepsilon}\right)10^{-3}\lambda[10^{-1}e^{-4/(5\lambda)}]\right\} \tag{1.39}$$

and

$$\Lambda(\chi,\varepsilon) := \sup\{\lambda \mid \chi(\lambda)\leqslant\varepsilon/2\} \tag{1.40}$$

Remarks. The recipe for computing $\Phi_{\chi,\psi}$ is somewhat complicated. While it clearly produces a strictly increasing function, this function in general increases very slowly.

As described in our first paper,[12] our goal is to obtain bounds on the time it takes to reach a given small neighborhood of equilibrium that are *independent of the initial data for all initial data within a given large class*. This we obtain for the classes determined by χ and ψ. For the applications described in ref. 12 it is the uniformity, as opposed to the times themselves, that matters.

However, the fact that our results provide a framework within which bounds on the times of approach to equilibrium can be computed is not without physical interest. Indeed, by the spectral analysis in Arkeryd[4] and Wennberg,[38] one knows that after one reaches such a neighborhood, one has uniform exponential control on the final approach to equilibrium. However, published results give no clue as to how this first arrival time might vary with the initial condition; this time is always shown to be finite by a compactness argument.

The problem is that the finite times provided by a compactness argument may well be so large as to have nothing to do with physical applications: consider Poincaré recurrence times, for example. (We thank Ed Nelson for suggesting this example in this context.)

Thus, it is natural to enquire whether or not our bounds lead to physically realistic estimates on the time of approach to a sufficiently small neighborhood of equilibrium that the exponential convergence driven by the linearized Boltzmann equation must set in. The answer is that we obtain physically realistic times *only* under the assumption of substantial additional regularity of the initial data with *presently available* results on the regularity of solutions of the Boltzmann equation. It may be possible to prove that such regularity as we need is *quickly produced* for quite general initial data, and we are studying the question. [A relevant example of this "automatic regularization" is given in (3.10) of Lemma 3.1.] Thus, it is worthwhile to discuss briefly the nature of this additional regularity.

To get a feeling for the situation, let us consider a concrete example in which the various quantities in Theorem 1.1 can be worked out explicitly.

Suppose that

$$I(f_0) < \infty \qquad \text{and} \qquad C(f_0) := \int_{\mathbb{R}^3} |v|^{10} f_0(v)\, d^3v < \infty \tag{1.41}$$

While these conditions are not terribly restrictive, they permit very explicit computations.

As shown in ref. 12, $\chi_g(\lambda) \leqslant I(g)\lambda$ for all g, and $\chi_{f_t}(\lambda)$ is monotonically decreasing in t. Thus, $\chi_{f_t}(\lambda) \leqslant \chi(\lambda) := I(f_0)\lambda$ for all t and λ.

One now easily works out from the definitions that

$$\Lambda(\chi, \varepsilon) = \min(\varepsilon/(2I(f_0)), 1/10)$$
$$R_\psi^2(\varepsilon, \lambda) \leqslant 10^4 C(f_0)\, \varepsilon\lambda^{-2}[10^{-1}e^{-4/(5\lambda)}]^{-1}$$

The problem is that if we have, say, $I(f_0) = 10^2$ and we want to bound the time it takes to drive $D(f_t)$ below, say, 10^{-2}, we have to work with $\Lambda(\chi, \varepsilon) \approx 10^{-4}$. Because of the factor $[10^{-1}e^{-4/(5\lambda)}]$ in (1.39), this leads to alarming numbers. To arrive at physically realistic numbers, we must find a way to work with larger values of λ or to eliminate the term $[10^{-1}e^{-4/(5\lambda)}]$ from our analysis, or both.

These things can be done under additional hypotheses on the initial data f_0 that we now explain. Without additional hypotheses, however, we must content ourselves with the uniformity in the initial data.

We first briefly explain how the term $[10^{-1}e^{-4/(5\lambda)}]$ enters our analysis and what can be done to avoid it. In all estimates of entropy production, e.g., those of Desvillettes[17] and Wennberg,[38] pointwise Maxwellian lower bounds of the form

$$f(v) \geqslant AM_\alpha(v) \tag{1.42}$$

for some $A > 0$ and some $\alpha > 0$ play a crucial role. Unfortunately, there are no proven *a priori* bounds of this type for the Boltzmann equation (uniformly in time) even when they are assumed for the initial data. A crucial role of the Ornstein–Uhlenbeck semigroup in our analysis is that it produces such bounds, as we have explained regarding Eq. (1.32).

While (1.32) gives us the lower bounds we need, so that we need make no artificial *assumption* concerning (1.42), it is from the constant in (1.32) that the term $[10^{-1}e^{-4/(5\lambda)}]$ enters.

Under the additional hypothesis that

$$f(v) \geqslant A \qquad \text{for} \quad |v| \leqslant 1 \tag{1.43}$$

we can replace $[10^{-1}e^{-4/(5\lambda)}]$ by $A\lambda^{-3/2}$ throughout our analysis. Fortunately, by slightly extending an argument of Carleman,[9] it is possble to establish (1.43) as an *a priori* bound at least in the case of hard spheres. If one, moreover, had a bound of the form (1.42), one could, as we shall see, eliminate the inverse power of λ from the exponent in (1.38). One would then be able to obtain realistic times.

Without such an estimate, one must find a way to work with larger values of λ. The problem is, as should be clear from the description of our method, that large values of λ tend to almost entirely wash away $D(\mathscr{P}_\lambda g)$ and $J(\mathscr{P}_\lambda g)$. The simple Lemma 2.1 which we use to control this "washing away" seems to be an obvious place to look for room for improvement. This can be made, but—again—with additional strong regularity assumptions on the initial data.

It is clear that if $D(g) \neq 0$, then $D(\mathscr{P}_\lambda g)$ cannot vanish for any value of λ. Thus one expects that it should be possible to prove lower bounds on $D(\mathscr{P}_\lambda g)$ even for $\lambda \approx 1$, in terms of $D(g)$, λ, bounds on the moments of g (or, what would be the same, bounds on ψ_g), and bounds on the derivatives of g. If one replaces the condition on $I(f)$ in (1.41) by a condition on, say, $\int_{\mathbb{R}^3} |(-\Delta)^5 f(v)|\, d^3v$, one can easily derive a lower bound on $D(\mathscr{P}_\lambda f)$ suitable for the derivation of realistic times. Such bounds are propagated uniformly in time by the constant-rate-function Boltzmann equation, but we do not know if this is the case for more physical collision kernels (though this seems likely).

We shall not develop this part of the subject further in this paper because it is somewhat incidental to the main purpose of our investigation, and we just do not know what combination of improved pointwise lower bounds and higher-order smoothness estimates provide the best route to realistic estimates on the times of approach to equilibrium. We do intend for it to be clear that the results of this paper do provide a framework within which such bounds can be established. For this reason, we carefully track all constants throughout Section 2. In Section 3, our emphasis is different. We wish to show that the methods introduced in ref. 12 and further developed here do apply to physically realistic collision kernels—such as that for hard spheres. Our aim in this section is to show how the size of which constants in which inequalities affect the rate of approach to equilibrium in which quantitative way, but we shall not actually try to compute the rates themselves.

We finally turn to the matter of related investigations, of which there are many. The use of Fisher information inequalities in proving entropy inequalities goes back to Stam.[32] Fisher information was used by Linnik[25] in his proof of the central limit theorem by means of entropy production. McKean[28] was the first to bring such ideas to bear on the Boltzmann equation in his study of the Kac model, a one-dimensional caricature of Maxwellian molecules. Later, Toscani[33,34] developed some of these ideas for actual Maxwellian molecules.

Another paper of Toscani[35] is of particular interest here: He shows that for solutions of the Boltzmann equation for Maxwellian molecules, $(d/dt)\, I(f_t) \leqslant 0$: this controls χ_{f_t}. Since the rate function for Maxwellian molecules is bounded below, the bounds of Section 2 are directly applicable to Maxwellian molecules by monotonicity of the entropy production in the rate function.

The method of getting lower bounds on the entropy production in terms of the relative entropy itself was introduced for the central limit theorem by Carlen and Soffer.[13] The method was developed further and extended to the Boltzmann equation by Carlen and Carvalho,[12] but there only for rather special collision kernels, and only qualitatively. We show here, as promised there, that the methods apply for physically realistic collision kernels as well.

Much more detailed discussion of the vast literature on entropy production can be found in our first paper[12]; we have repeated only the essential points here.

2. ENTROPY PRODUCTION ESTIMATES FOR THE BOLTZMANN EQUATION WITH CONSTANT MICROSCOPIC RATE FUNCTION

In this section, we prove the lemmas concerning the Boltzmann equation with a constant microscopic rate function that were described in the introduction. Here, we state them in complete detail, and keep explicit track of all of the constants.

We begin with a simple consequence of Gross' logarithmic Sobolev inequality.[22]

Lemma 2.1 (Lower bound on $J(\mathscr{P}_\lambda f)$ in terms of $D(f)$, χ_f, and λ). Let f be any density with zero bulk velocity, unit temperature, and finite entropy. For any function χ increasing continuously from zero, define

$$\Lambda(\chi, \varepsilon) := \sup\{\lambda \mid \chi(\lambda) \leqslant \varepsilon/2\} \tag{2.1}$$

Then, whenever $\lambda \leqslant \Lambda(\chi_f, D(f))$,

$$J(\mathscr{P}_\lambda f) \geqslant D(f) \tag{2.2}$$

Proof. By the definition (1.15) of χ_f, $D(\mathscr{P}_\lambda f) \geqslant D(f)/2$ for all λ such that $\chi_f(\lambda) \leqslant D(f)/2$. Gross' logarithmic Sobolev inequality[22] states that $J(f) \geqslant 2D(f)$. (See ref. 11 for a statement in this form as well as several proofs.) It is thus clear from Gross' inequality that $J(\mathscr{P}_\lambda f) \geqslant 2D(\mathscr{P}_\lambda f) \geqslant 2(D(f)/2)$ for all λ such that $\chi_f(\lambda) \leqslant D(f)/2$. ∎

We next summarize the relevant regularizing properties of $\mathscr{P}_\lambda$. All of the statements in the following theorem are simple variants of known estimates for the heat semigroup.[5]

Lemma 2.2 (Regularity estimates on the range of $\mathscr{P}_\lambda$). Let f be any velocity density with zero bulk velocity and unit temperature. Suppose that $0 < \lambda \leqslant 1/10$. Then the following hold:

(i) *Pointwise upper and lower bounds:*

$$10^{-1} e^{-5/(4\lambda)} M_{(1-e^{-2\lambda})/2}(v) \leqslant \mathscr{P}_\lambda f(v) \leqslant (4\pi)^{-3/2} \lambda^{-3/2} \tag{2.3}$$

(ii) *Smoothness bounds:*

$$|\nabla \ln \mathscr{P}_\lambda f(v)|^2 \mathscr{P}_\lambda f(v) \leqslant 10\lambda^{-1} \, (M_{3\lambda} * f^{(\lambda)}(v))' \tag{2.4}$$

where $*$ denotes convolution, and $f^{(\lambda)}(v)$ denotes $e^{3\lambda} f(e^\lambda v)$.

(iii) *Concentration bounds:* For all $R > 1$

$$\psi_{\mathscr{P}_\lambda f}(R) \leqslant 50[\psi_{f^{(\lambda)}}(R/2) + \psi_{M_{(1-e^{-2\lambda})}}(R/2)] \tag{2.5}$$

Proof. The condition $0 < \lambda \leqslant 1/10$ is imposed for the sake of convenience (since anyway we will mainly be concerned with small values of λ). In this range we have

$$(8/5)\lambda \leqslant (1 - e^{-2\lambda}) \leqslant 2\lambda \tag{2.6}$$

The upper bound in (2.3) is elementary. The lower bound is established as follows: Let $g(v)$ be any density with zero bulk velocity and temperature no greater than one. Then for any $\alpha > 0$, making use of the fact that

$$\int_{|v| \leqslant \sqrt{2}} g(v) \, d^3v \geqslant 1/2$$

by Chebychev's inequality,

$$\int_{\mathbb{R}^3} M_\alpha(v-v')\, g(v')\, d^3v' \geqslant \int_{|v'| \leqslant \sqrt{2}} M_\alpha(v-v')\, g(v')\, d^3v'$$
$$\geqslant (2)^{-5/2}\, e^{-2/\alpha} M_{\alpha/2}(v) \tag{2.7}$$

Choosing $\alpha = (1 - e^{-2\lambda})$ and g as in the definition (1.14) of $\mathscr{P}_\lambda f$, we obtain the lower bound.

The inequality (2.3) is proved by adapting a heat semigroup argument of Barron.[5] This is easy to do, but because we need explicit constants, we carry out the estimate here. Let $g = \mathscr{P}_\lambda f$. Then by (1.14)

$$|\nabla g(v)| = \left| \frac{1}{1-e^{-2\lambda}} \int_{\mathbb{R}^3} (v'-v)\, M_{(1-e^{-2\lambda})}(v-v')\, e^{3\lambda} f(e^\lambda(v))\, d^3v' \right|$$
$$\leqslant \frac{1}{1-e^{-2\lambda}} \left[\int_{\mathbb{R}^3} |v'-v|^2\, M_{(1-e^{-2\lambda})}(v-v')\, e^{3\lambda} f(e^\lambda(v))\, d^3v' \right]^{1/2}$$
$$\times \left[\int_{\mathbb{R}^3} M_{(1-e^{-2\lambda})}(v-v')\, e^{3\lambda} f(e^\lambda(v))\, d^3v' \right]^{1/2}$$

But the last factor on the right is $g^{1/2}$, and we have

$$|\nabla \ln g(v)|^2\, g(v) \leqslant (1-e^{-2\lambda})^{-2} \int_{\mathbb{R}^3} |v'-v|^2\, M_{(1-e^{-2\lambda})}(v-v')\, e^{3\lambda} f(e^\lambda(v))\, d^3v'$$

Now note that for any $\beta > \alpha$

$$w^2 M_\alpha(w)/M_\beta(w) = (\beta/\alpha)^{3/2}\, w^2 e^{-(\beta-\alpha)\, w^2/2\alpha\beta} \leqslant (\beta/\alpha)^{3/2}\, [2\alpha\beta/(\beta-\alpha)]$$

With $\alpha = (1-e^{-2\lambda})$ and $\beta = 3\lambda$, we can simplify this expression using (2.6) and bound it above by $100/\lambda$ uniformly in w. We use this uniform bound to absorb the factor of $|v'-v|^2$ into $M_{(1-e^{-2\lambda})}(v'-v)$, and (2.4) is now established by elementary estimates.

Finally, let V_1 and V_2 be independent $\mathbb{R}^3$-valued random variables with zero means and variances no greater than 3 (i.e., temperatures no greater than unity). Suppose that they have densities g_1 and g_2, respectively. Then $V_1 + V_2$ has the density $g_1 * g_2$. Therefore, using E to denote the expectation,

$$\psi_{g_1 * g_2}(R) = E 1_{\{|V_1+V_2| \geqslant R\}}\, |V_1 + V_2|^2$$
$$\leqslant 2E 1_{\{|V_1| \geqslant R/2\}}(|V_1|^2 + |V_2|^2) + 2E 1_{\{|V_2| \geqslant R/2\}}(|V_1|^2 + |V_2|^2)$$
$$\leqslant 2\psi_{g_1}(R/2) + (48/R^2)\, \psi_{g_2}(R/2) + 2\psi_{g_2}(R/2) + (48/R^2)\, \psi_{g_1}(R/2)$$

This clearly establishes (2.13), and we record the bound

$$\psi_{g_1 * g_2}(R) \leqslant 50[\psi_{g_1}(R/2) + \psi_{g_2}(R/2)] \tag{2.8}$$

for all $R > 1$. ▌

Remark. The unfortunate factor of $e^{-5/(4\lambda)}$ in the lower bound of (2.3) arises because the fact that $\theta(f) = 1$ only gives us a lower bound on

$$\int_{|v'| \leqslant R} f(v')\, d^3v' \qquad \text{for} \quad R > 1$$

This prevents us from replacing the $\sqrt{2}$ in (2.7) by a variable R to be optimally chosen depending on λ. If, however, we posses *a priori* lower bound of the form

$$f(v) \geqslant A \qquad \text{for} \quad |v| \leqslant 1 \tag{2.9}$$

we can replace (2.7) by

$$\begin{aligned}\int_{\mathbb{R}^3} M_\alpha(v - v')\, g(v')\, d^3v' &\geqslant \int_{|v'| \leqslant R} M_\alpha(v - v')\, g(v')\, d^3v' \\ &\geqslant A(4\pi/3)\, R^3 e^{-R^2/\alpha} M_{\alpha/2}(v) \end{aligned} \tag{2.10}$$

We then choose $R^2 = \alpha$ and $\alpha = (1 - e^{-2\lambda})$ and obtain the following estimate:

Let f and λ satisfy the hypotheses of Lemma 2.2, and suppose in addition that (2.9) holds. Then

$$\mathscr{P}_\lambda f(v) \geqslant A\lambda^{3/2} M_{(1 - e^{-2\lambda})/2}(v) \tag{2.11}$$

We shall see in the next section that for hard-sphere collisions and uniformly continuous initial data, we have bounds of the form (2.9) uniformly in times $t \geqslant 1$ for solutions of the Boltzmann equation with uniformly continuous initial data. In fact, by slightly extending a result of Carleman,[9] we have the *a priori* bound

$$|\ln f(v)| \leqslant A'(1 + |v|^3)$$

for all v in the setting described above. The constant A' depends only on the modulus of continuity of the initial data. The exponent 3 can be brought arbitrarily close to 2 at the expense of increasing A', but it cannot be brought all the way to 2 with available estimates.

Thus, in the hard-sphere setting that we discuss in the next section, the estimates (2.9) and (2.11) will be available. For general initial data, however, (2.4) is the best estimate we have.

We now prove the main estimate of this section. It is a descendant, considerably more involved, of a convolution estimate of Brown.[8]

Lemma 2.3 (Variational bound for the information dissipation). Let f be a velocity density with zero bulk velocity and unit temperature. Suppose further that $f = \mathscr{P}_\lambda g$ for some $0 < \lambda \leqslant 10^{-1}$ and some other such density g. Then

$$\begin{aligned} &I(f) - I(f \circ f) \\ &\quad \geqslant I(f)^{-1} \lambda [10^{-1} e^{-4/(5\lambda)}]^2 \\ &\qquad \times \left[\inf \left\{ \int_{\mathbb{R}^3} |av + b + \nabla \ln f(v)|^2 \, M_{(1-e^{-2\lambda})/2}(v) \, d^3v \,\middle|\, a, b \right\} \right] \end{aligned} \tag{2.12}$$

where a and b range over $\mathbb{R}$ and $\mathbb{R}^3$, respectively. Furthermore,

$$I(f) \leqslant 10\lambda^{-1} \tag{2.13}$$

and the infimum in (2.12) is attained at some pair a, b with

$$a^2 + |b|^2 \leqslant [10^{-1} e^{-4/(5\lambda)}] \, I(f) \tag{2.14}$$

Remark. We have kept the term $[10^{-1} e^{-4/(5\lambda)}]$ separate and in square brackets since this terribly small term may be replaced by the much larger (for small λ) term $[A\lambda^{3/2}]$ under the condition (2.9) so that (2.11) holds. To make clear that this is so, we shall avoid simplifying constants containing the term $[10^{-1} e^{-4/(5\lambda)}]$ throughout the proof. We shall thus have that: *Under the further condition (2.9), (2.12) and (2.14) hold with* $[10^{-1} e^{-4/(5\lambda)}]$ *replaced by* $[A\lambda^{3/2}]$.

The most involved part of the proof is an estimate on the spectrum of a quadratic form closely related to the integral on the right side of (2.12). We now state and prove this estimate as Lemma 2.4, and then return to the proof of Lemma 2.3.

Let M_α denote the Maxwellian with zero bulk velocity and temperature α, and let $\mathscr{H}_\alpha$ denote the Hilbert space of nonconstant functions ϕ with square-integrable distributional gradient equipped with the norm

$$\|\phi\|^2_{\mathscr{H}_\alpha} = \int_{\mathbb{R}^3} |\nabla \phi(v)|^2 \, M_\alpha(v) \, d^3v$$

in the usual way. For each $\omega \in S^2$, let

$$P_\omega^{\|} = \omega \otimes \omega \qquad \text{and} \qquad P_\omega^{\perp} = I - \omega \otimes \omega \tag{2.15}$$

be the corresponding orthogonal projections on $\mathbb{R}^3$. Define the quadratic forms $\mathscr{E}_\alpha^{\|}$ and $\mathscr{E}_\alpha^{\perp}$ by

$$\mathscr{E}_\alpha^{\|}(\phi, \phi) := \int_{S^2} \int_{\mathbb{R}^3} \left| \int_{\mathbb{R}^3} P_\omega^{\|} \nabla\phi(\tilde{v}') \, M_\alpha(v') \, d^3v' \right|^2 M_\alpha(v) \, d^3v \, d\omega$$

$$\mathscr{E}_\alpha^{\perp}(\phi, \phi) := \int_{S^2} \int_{\mathbb{R}^3} \left| \int_{\mathbb{R}^3} P_\omega^{\perp} \nabla\phi(\tilde{v}) \, M_\alpha(v') \, d^3v' \right|^2 M_\alpha(v) \, d^3v \, d\omega$$

Then define

$$\mathscr{E}_\alpha(\phi, \phi) = \mathscr{E}_\alpha^{\|}(\phi, \phi) + \mathscr{E}_\alpha^{\perp}(\phi, \phi) \tag{2.16}$$

Define positive operators C_α and D_α on $L^2(\mathbb{R}^3, M_\alpha(v)\, d^3v)$ by

$$\langle \phi, C_\alpha \phi \rangle_{L^2(M_\alpha(v)\, d^3v)} = \mathscr{E}_\alpha(\phi, \phi) \tag{2.17}$$

and

$$\langle \phi, D_\alpha \phi \rangle_{L^2(M_\alpha(v)\, d^3v)} = \int_{\mathbb{R}^3} |\nabla\phi(v)|^2 \, M_\alpha(v) \, d^3v \tag{2.18}$$

It is easy to see that if $\phi(v)$ is a polynolmial of degree m in the components of v, then so are $C_\alpha \phi$ and $D_\alpha \phi$. As Grad[21] observed in connection with the linearized collision operator, this implies that the eigenfunctions of C_α and D_α are Hermite polynomials. Moreover, since both operators are rotationally invariant, the particular Hermite polynomials that arise as their eigenfunctions are exactly those that can be written as products of Laguerre polynomials in $|v|^2$ and solid spherical harmonics. Thus, C_α and D_α can be diagonalized together.

In fact, D_α is a well-known operator—it is the quantum mechanical harmonic oscillator Hamiltonian (with restoring force depending on α) in the ground-state representation. Its spectrum is therefore well understood, and easy to work out. We need certain information concerning the relative sizes of corresponding eigenvalues of C_α and D_α in what follows. First, it is clear from Jensen's inequality and the identity $M_\alpha(v)\, M_\alpha(v') = M_\alpha(\tilde{v})\, M_\alpha(\tilde{v}')$ that

$$\mathscr{E}_\alpha(\phi, \phi) \leqslant \int_{\mathbb{R}^3} |\nabla\phi(v)|^2 \, M_\alpha(v) \, d^3v$$

In other words, $C_\alpha \leqslant D_\alpha$. We shall need a somewhat stronger statement, which we give in the next lemma.

Lemma 2.4 (Eigenvalue estimates). Let $\mathscr{K}_\alpha$ denote the subspace of $\mathscr{H}_\alpha$ spanned by

$$v_1, \quad v_2, \quad v_3, \quad \text{and} \quad |v|^2 - 3\alpha \tag{2.19}$$

Then

$$C_\alpha \phi = D_\alpha \phi \qquad \text{if and only if} \quad \phi \in \mathscr{K}_\alpha \tag{2.20}$$

Furthermore,

$$\mathscr{E}_\alpha(\phi, \phi) \leqslant 13/15 \int_{\mathbb{R}^3} |\nabla\phi(v)|^2 \, R^{(l)}_{0,0}(\omega) \qquad \text{for all} \quad \phi \in \mathscr{K}^\perp_\alpha \tag{2.21}$$

Proof. We first note that since scaling relates the spectrum of C_α to that of C_β in the same way it relates the spectrum of D_α to that of D_β, it suffices to prove the result for any fixed value of α. For convenience, we choose $\alpha = 1/2$, and write M for $M_{1/2}$ in the rest of the proof. We shall likewise drop the α subscript from all operators and quadratic forms during the proof.

Let L^γ_n denote the nth Laguerre polynomialof order γ; i.e.,

$$\sum_{n=0}^{\infty} L^\gamma_n(y)\, t^n = (1-t)^{-(\gamma+1)}\, e^{-yt/(1-t)}, \qquad |t| < 1 \tag{2.22}$$

Let $\mathscr{Y}_{l,m}(v)$ denote the solid spherical harmonic of order l. By what we have said above, the eigenfunctions of both C and D are the functions $\phi_{l,m,n}$ where

$$\phi_{l,m,n}(v) := \mathscr{Y}_{l,m}(v)\, L^{l+1/2}_n(|v|^2) \tag{2.23}$$

for $l \geqslant 0$, $-l \leqslant M \leqslant l$, and $n \geqslant 0$.

These eigenfunctions are not normalized, but it is easy to see that the norm of $\phi_{l,m,n}$ is independent of m. Using the identity

$$\int_0^\infty e^{-y} y^\gamma L^\gamma_m(y)\, dy = \delta_{m,n} \left(\frac{\Gamma(\gamma+1)\, \Gamma(n+3/2)}{\Gamma(n+1)\, \Gamma(\gamma+1)} \right)$$

one easily obtains

$$c_{l,n} := \|\phi_{l,m,n}\|^2_{L^2(M_{1/2}(v)\, d^3v)} = \frac{1}{2l+1} \frac{\sqrt{\pi}}{2} \left(\frac{\Gamma(\gamma+1)\, \Gamma(n+3/2)}{\Gamma(n+1)\, \Gamma(\gamma+1)} \right) \tag{2.24}$$

Next, we observe that since C and D are rotationally invariant, their eigenvalues do not depend on m. Thus, we may define $\lambda_{l,n}$ and $\mu_{l,n}$ by

$$C\phi_{l,m,n} = \lambda_{l,n} \phi_{l,m,n} \qquad \text{and} \qquad D\phi_{l,m,n} = \mu_{l,n} \phi_{l,m,n} \tag{2.25}$$

What we are required to show is that

$$\lambda_{l,n} = \mu_{l,n} \quad \text{for } l+n=1 \qquad \text{and} \qquad \lambda_{l,n} \leqslant (13/15)\,\mu_{l,n} \quad \text{for } l+n \geqslant 2 \tag{2.26}$$

Actually, it is much easier, and suffices for our purposes, to work with only a part of C, and thus with only a part of $\lambda_{l,n}$. Define

$$\langle \phi, C^{\|}\phi \rangle_{L^2(M(v)\,d^3v)} = \mathscr{E}^{\|}(\phi, \phi)$$

and in the same manner define $C^{\perp}$. It is easy to see that C, $C^{\|}$ and $C^{\perp}$ are all diagonalized by the $\phi_{l,m,n}$ basis, and evidently, if we denote the eigenvalues corresponding to $C^{\|}$ and $C^{\perp}$, respectively, by $\lambda^{\|}_{l,n}$ and $\lambda^{\perp}_{l,n}$, then

$$\lambda_{l,n} = \lambda^{\|}_{l,n} + \lambda^{\perp}_{l,n} \tag{2.27}$$

The analysis that led to $\lambda_{l,n} \leqslant \mu_{l,n}$ clearly shows that

$$\lambda^{\|}_{l,n} \leqslant (1/3)\,\mu_{l,n} \qquad \text{and} \qquad \lambda^{\perp}_{l,n} \leqslant (2/3)\,\mu_{l,n}$$

Thus, to establish the inequality in (2.26), it suffices to show that

$$\lambda^{\|}_{l,n} \leqslant (1/5)\,\mu_{l,n} \qquad \text{for} \quad l+n \geqslant 2 \tag{2.28}$$

We now show that the inequality in (2.28) holds whenever $l \geqslant 2$; this reduces the demonstration of (2.26) to the computation of the eigenvalues $\lambda^{\|}_{l,n}$ with $l=0$ and $l=1$.

Simple group-theoretic considerations expedite the computations now before us. Fix any $\omega \in S^2$, and some rotation R_ω taking the vector $e_3 := (1, 0, 0)$ to ω. Introduce new coordinates w_1, w_2, and w_3 by $w = R_\omega^{-1} v$. In these coordinates, $\omega \cdot \nabla = \partial/\partial w_3$.

Note that

$$\begin{aligned}\phi_{l,0,n}(v) &= \phi_{l,0,n}(R_\omega w) = \mathscr{Y}_{l,m}(Rw)\, L_n^{l+1/2}(|w|^2) \\ &= \sum_{m=-l}^{l} R^{(l)}_{0,m}(\omega)\, \mathscr{Y}_{l,m}(w)\, L_n^{l+1/2}(|w|^2)\end{aligned} \tag{2.29}$$

where $R^{(l)}_{j,k}(\omega)$ is the $(2l+1) \times (2l+1)$ matrix representing R_ω. The first step in computing $\mathscr{E}^{\|}(\phi_{l,0,n}, \phi_{l,0,n})$ is to perform the average

$$\frac{1}{\pi} \int_{\mathbb{R}^2} \phi_{l,0,n}(v)\, e^{-(w_2^2 + w_3^2)}\, dw_1\, dw_2 \tag{2.30}$$

of $\phi_{l,0,n}(v)$ in the plane perpendicular to ω. Writing (2.30) out using (2.29), it is clear that the only term in the sum that survives is the $m=0$ term. Thus, the quantity in (2.30) is

$$\frac{1}{\pi} R_{0,0}^{(l)}(\omega) \int_{\mathbb{R}^2} \phi_{l,0,n}(w)\, e^{-(w_2^2+w_3^2)}\, dw_1\, dw_2 := R_{0,0}^{(l)}(\omega)\, F_{l,n}(w_3) \quad (2.31)$$

The second step is to differentiate this with respect to w_3, square the result, and integrate against $(1/\sqrt{\pi})\exp(-w_3^2)$. The third and final step is to average with respect to ω. Since all of the ω dependence is in the $R_{0,0}^{(l)}(\omega)$ factor, and since $\int_{S^2} |R_{0,0}^{(l)}(\omega)|^2\, d\omega = 1/(2l+1)$, we obtain

$$\mathscr{E}^{||}(\phi_{l,0,n}, \phi_{l,0,n}) = \frac{1}{2l+1} \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} \left| \frac{d}{dw_3} F_{l,n}(w_3) \right|^2 e^{-w_3^2}\, dw_3 \quad (2.32)$$

The by now familiar argument based on Jensen's inequality shows that

$$\frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} \left| \frac{d}{dw_3} F_{l,n}(w_3) \right|^2 e^{-w_3^2}\, dw_3$$

$$\leqslant \int_{\mathbb{R}^3} \left| \frac{\partial}{\partial w_3} \phi_{l,0,n}(w) \right|^2 M(v)\, d^3w \leqslant \int_{\mathbb{R}^3} |\nabla \phi_{l,0,n}(w)|^2\, M(v)\, d^3w \quad (2.33)$$

Combining (2.32) and (2.33), we have that $\lambda_{l,n}^{||} \leqslant [1/(2l+1)]\, \mu_{l,n}^{||}$, and thus (2.28) is established in all cases in which $l \geqslant 2$.

In order to compute the $l=0$ and the $l=1$ series of eignevalues of $C^{||}$, it is useful to introduce a generating function for the corresponding eigenfunctions. Fix any l and define the generating function $g_l(v, t)$ by

$$g_l(v, t) = \sum_{n=0}^{\infty} \phi_{l,0,n}(v)\, t^n \quad (2.34)$$

By (2.22) we have

$$g_l(v, t) = \mathscr{Y}_{l,0}(v)(1-t)^{-(l+3/2)}\, e^{-|v|^2 t/(1-t)}$$

Clearly

$$\langle g_l(\cdot, t), C^{||} g_l(\cdot, t) \rangle_{L^2(M(v)\, d^3v)} = \sum_{n=0}^{\infty} c_{l,n} \lambda_{l,n}^{||} t^{2n} \quad (2.35)$$

Since for each fixed t, $g_l(\cdot, t)$ is the product of an explicit polynomial and an explicit Gaussian function, the inner product on the left in (2.35) is readily computed and expanded as a power series. Equating coefficients then determines the $\lambda_{l,n}''$.

To carry out the computation, we introduce coordinates w_1, w_2, and w_3 adapted to ω as before. Then, as before,

$$g_l(v, t) = g_l(R_\omega w, t)$$
$$= \sum_{m=-l}^{l} R_{j,k}^{(l)}(\omega)\, \mathcal{Y}_{l,m}(w)(1-t)^{-(l+3/2)}\, e^{-|v|^2 t/(1-t)} \tag{2.36}$$

The first step in computing the left side of (2.35) is to perform the average

$$\frac{1}{\pi}\int_{\mathbb{R}^2} g_l(v, t)\, e^{-(w_2^2+w_3^2)}\, dw_1\, dw_2 \tag{2.37}$$

of $g_l(\cdot, t)$ in the plane perpendicular to ω, and, as before, the only term in the sum that survives is the $m=0$ term. Thus, the quantity in (2.37) is

$$\frac{1}{\pi} R_{0,0}^{(l)}(\omega) \int_{\mathbb{R}^2} g_l(w, t)\, e^{-(w_2^2+w_3^2)}\, dw_1\, dw_2 := R_{0,0}^{(l)}(\omega)\, G_l(w_3) \tag{2.38}$$

The second step is to differentiate this with respect to w_3, square the result, and integrate against $(1/\sqrt{\pi})\exp(-w_3^2)$. The third and final step is to average with respect to ω. As before, we obtain

$$\mathscr{E}^{\|}(g_l(\cdot, t), g_l(\cdot, t)) = \frac{1}{2l+1}\frac{1}{\sqrt{\pi}}\int_{\mathbb{R}} \left|\frac{d}{dw_3} G_l(w_3)\right|^2 e^{-w_3^2}\, dw_3 \tag{2.39}$$

By direct computation, one finds that

$$\mathscr{E}^{\|}(g_0(\cdot, t), g_0(\cdot, t)) = 2t^2(1-t^2)^{-3/2} = \sum_{n=0}^{\infty} \frac{(2n+1)!!}{2^n n!}\, t^{2n} \tag{2.40}$$

Comparison with (2.35) reveals that

$$\lambda_{0,n}^{\|} = 4n/(2n+1) \tag{2.41}$$

In the same way, one finds that $\mathscr{E}^{\|}(g_1(\cdot, t), g_1(\cdot, t)) = (1+2t^2)(1-t^2)^{-5/2}$, and thus, that

$$\lambda_{1,n}^{\|} = (4n+2)/(2n+3) \tag{2.42}$$

Of course, through the relation of D to the quantum mechanical harmonic oscillator mentioned above, it is well known that

$$\mu_{l,n} = 2l + 4n \tag{2.43}$$

This may also be directly established by the generating function method that we have just applied to the computation of $\lambda_{l,n}$.

Together, (2.41)–(2.43) establish the validity of (2.26) in the remaining cases. ∎

Proof of Lemma 2.3. We begin with certain square-integrability bounds that pave the way for the application of Lemma 2.4. First, integrating (2.4), we obtain

$$I(f) = \int_{\mathbb{R}^3} |\nabla \ln f(v)|^2 f(v)\, d^3v \leqslant 10\lambda^{-1} \tag{2.44}$$

For the rest of the proof we fix $\alpha := (1 - e^{-2\lambda})/2$. By (2.3), $f(v)/M_\alpha(v) \geqslant [10^{-1} e^{-4/(5\lambda)}]$ pointwise, and thus

$$\int_{\mathbb{R}^3} |\nabla \ln f(v)|^2 M_\alpha(v)\, d^3v \leqslant [10^{-1} e^{-4/(5\lambda)}]^{-1} I(f)$$

We now define $\phi(v) := \ln f(v)$, and rewrite the last estimate using (2.3) as

$$\int_{\mathbb{R}^3} |\nabla \phi(v)|^2 M_\alpha(v)\, d^3v \leqslant [10^{-1} e^{-4/(5\lambda)}]^{-1}\, 10\lambda^{-1} \tag{2.45}$$

Some further notation must be introduced. This notation corresponds to that used in the proof of Lemma 3.2 of ref. 12, which is a qualitative version of the present Lemma 2.3. For each fixed $\omega \in S^2$, define

$$G_\omega(v, v') = f^{1/2}(\tilde{v}) f^{1/2}(\tilde{v}') \quad \text{and} \quad g_\omega(v) = \left[\int_{\mathbb{R}^3} G_\omega(v, v')^2\, d^3v'\right]^{1/2} \tag{2.46}$$

We also define

$$f \circ_\omega f = \int_{\mathbb{R}^3} f(\tilde{v}) f(\tilde{v}')\, d^3v'$$

Note that $f \circ_\omega f = g_\omega(v)^2$, and hence,

$$f \circ f(v) = \int_{S^2} f \circ_\omega f\, d\omega = \int_{S^2} g_\omega(v)^2\, d\omega \tag{2.47}$$

As in the proof of Lemma 3.2 of ref. 12, we define

$$\|\nabla g_\omega\|^2 = \int_{\mathbb{R}^3} |\nabla g_\omega(v)|\, d^3v$$

$$\|\nabla_v G_\omega\|^2 = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} |\nabla_v G_\omega(v, v')|^2\, d^3v\, d^3v'$$

and note that

$$4\,\|\nabla g_\omega\|^2 = I(f \circ_\omega f), \qquad 4\,\|\nabla_v G_\omega\|^2 = I(f) \tag{2.48}$$

Then since

$$3 = I(M^f) \leqslant I(f \circ_\omega f) \leqslant I(f) \leqslant 10\lambda^{-1} \tag{2.49}$$

we have

$$\frac{3}{I(f)} \leqslant \frac{\|\nabla g_\omega\|^2}{\|\nabla_v G_\omega\|^2} \leqslant 1 \tag{2.50}$$

We have shown in Lemma 3.3 of ref. 12 that

$$\begin{aligned}&\|\nabla_v G_\omega\|^2 - \|\nabla g_\omega\|^2\\ &\quad = 2\,\frac{\|\nabla g_\omega\|^2}{\|\nabla_v G_\omega\|^2}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\left|\frac{\|\nabla_v G_\omega\|^2}{\|\nabla g_\omega\|^2}\nabla \ln g_\omega(v) - \nabla_v \ln G_\omega(v, v')\right|^2\\ &\quad\quad \times G_\omega(v, v')^2\, d^3v\, d^3v'\end{aligned} \tag{2.51}$$

The left side of (2.51) is $1/4[I(f) - I(f \circ_\omega f)]$. To express more simply the right side, we define

$$\psi_\omega(v) = \frac{\|\nabla_v G_\omega\|^2}{\|\nabla g_\omega\|^2} \ln g_\omega(v)$$

and note that on account of (2.48) and (2.49), $\nabla\psi_\omega$ is square integrable. Also, note that

$$\nabla_v \ln G_\omega(v, v') = \tfrac{1}{2}\nabla_v \ln f(\tilde{v}) + \tfrac{1}{2}\nabla_v \ln f(\tilde{v}') = (P_\omega^\perp \nabla\phi)(\tilde{v}) + (P_\omega^\parallel \nabla\phi)(\tilde{v}')$$

where the projections P_ω and $P_\omega^\perp$ are those defined in (2.15).

We can now rewrite (2.51) as

$$\begin{aligned}&I(f) - I(f \circ_\omega f)\\ &\quad \geqslant \frac{12}{I(f)}\,[10^{-1} e^{-4/(5\lambda)}]^2 \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} |\nabla\psi_\omega(v) - [(P_\omega^\perp \nabla\phi)(\tilde{v}) + (P_\omega^\parallel \nabla\phi)(\tilde{v}')]|^2\\ &\quad\quad \times M_\alpha(v)\, M_\alpha(v')\, d^3v\, d^3v'\end{aligned} \tag{2.52}$$

where we have used (2.3) to replace $G_\omega(v, v')^2$ with $[10^{-1} e^{-4/(5\lambda)}]^2$ $M_\alpha(v)\, M_\alpha(v')$.

We use Lemma 2.4 to estimate the cross term in this integral: First,

$$2\left|\int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\{\nabla\psi_\omega(v)\cdot[(P_\omega^\perp\nabla\phi)(\tilde{v})+(P_\omega^{\|}\nabla\phi)(\tilde{v}')]\}^2 M_\alpha(v)\,M_\alpha(v')\,d^3v\,d^3v'\right|$$

$$\leqslant\int_{\mathbb{R}^3}|\nabla\psi_\omega(v)|^2 M_\alpha(v)\,d^3v$$

$$+\int_{\mathbb{R}^3}\left|\int_{\mathbb{R}^3}[(P_\omega^\perp\nabla\phi)(\tilde{v})+(P_\omega^{\|}\nabla\phi)(\tilde{v}')]\,M_\alpha(v')\,d^3v'\right|^2 M_\alpha(v)\,d^3v \quad (2.53)$$

(We have first carried out the v' integration, and *then* used the arithmetic-geometric mean inequality to estimate the effect of the v integration.) Thus the integral in (2.53) is no less than

$$\int_{\mathbb{R}^3}|\nabla\phi(v)|^2 M_\alpha(v)\,d^3v$$

$$-\int_{\mathbb{R}^3}\left|\int_{\mathbb{R}^3}[(P_\omega^\perp\nabla\phi)(\tilde{v})+(P_\omega^{\|}\nabla\phi)(\tilde{v}')]\,M_\alpha(v')\,d^3v'\right|^2 M_\alpha(v)\,d^3v \quad (2.54)$$

Next since $f\mapsto I(f)$ is convex,

$$I(f)-I(f\circ f)\geqslant\int_{S^2}(I(f)-I(f\circ_\omega f))\,d\omega$$

Thus, from (2.52) and (2.54),

$$I(f)-I(f\circ f)$$

$$\geqslant 10^{-1}\lambda[10^{-1}e^{-4/(5\lambda)}]^2\left[\int_{\mathbb{R}^3}|\nabla\phi(v)|^2 M_\alpha(v)\,d^3v-\langle\phi,C_\alpha\phi\rangle_{L^2(M_\alpha(v)\,d^3v)}\right]$$

$$\geqslant I(f)^{-1}\,\lambda[10^{-1}e^{-4/(5\lambda)}]^2\,12\left[\frac{2}{15}\int_{\mathbb{R}^3}|\nabla(\phi(v)-\Pi_{\mathcal{K}_\alpha}\phi(v))|^2 M_\alpha(v)\,d^3v\right]$$

where $\Pi_{\mathcal{K}_\alpha}$ is the orthogonal onto $\mathcal{K}_\alpha$. Clearly, $\nabla(\Pi_{\mathcal{K}_\alpha}\phi(v))=cv+d$ for some c in $\mathbb{R}$ and some d in $\mathbb{R}^3$. Since

$$|d|+3c^2\alpha=\int_{\mathbb{R}^3}|\nabla(\Pi_{\mathcal{K}_\alpha}(v))|^2 M_\alpha(v)\,d^3v$$

$$\leqslant\int_{\mathbb{R}^3}|\nabla\phi(v)|^2 M_\alpha(v)\,d^3v\leqslant[10^{-1}e^{-4/(5\lambda)}]\,I(f)$$

by (2.44), we have the inequality (2.12) and the bounds (2.13) and (2.14). ∎

Lemma 2.5 (Information dissipation bounded below in terms of relative information for regularized densities). Let f be a velocity density with zero bulk velocity and unit temperature. Suppose further that $f = \mathscr{P}_\lambda g$ for some other such density g and some λ with $0 < \lambda \leqslant 1/10$. Then,

$$I(f) - I(f \circ f) \geqslant \Gamma_\psi(\lambda, J(f)) \tag{2.55}$$

where

$$\Gamma_\psi(\lambda, \varepsilon) = [10^{-1} e^{-4/(5\lambda)}]^2 \, e^{-R_\psi(\varepsilon, \lambda)^2/\lambda} \left(\frac{\varepsilon}{6 + 2\varepsilon}\right) \tag{2.56}$$

and

$$\begin{aligned} R_\psi(\varepsilon, \lambda) = \inf\{ R \mid (2\psi_{g^{(\lambda)}}(R/2) + \psi_{M_{3\lambda}}(R/2) + \psi_{M_{(1-e^{-2\lambda})}}(R/2)) \\ \leqslant (\varepsilon/(3+\varepsilon)) \, 10^{-3} \lambda [10^{-1} e^{-4/(5\lambda)}] \} \end{aligned} \tag{2.57}$$

Proof. By direct computation we have

$$\int_{\mathbb{R}^3} |av + b + \nabla \ln f(v)|^2 f(v) \, d^3v = 3(a-1)^2 + |b|^2 + J(f) \geqslant J(f) \tag{2.58}$$

for all a and b.

Next, note that for any $R > 1$,

$$\begin{aligned} &\int_{|v| \geqslant R} |av + b + \nabla \ln f(v)|^2 f(v) \, d^3v \\ &\leqslant 3a^2 \int_{|v| \geqslant R} |v|^2 f(v) \, d^3v + 3 \, |b|^2 \int_{|v| \geqslant R} f(v) \, d^3v \\ &\quad + 3 \int_{|v| \geqslant R} |\nabla \ln f(v)|^2 f(v) \, d^3v \\ &\leqslant 3(a^2 + |b|^2) \, \psi_f(R) + 30\lambda^{-1} \int_{|v| \geqslant R} M_{3\lambda} * g^{(\lambda)}(v) \, d^3v \\ &\leqslant 3\lambda^{-1}([10^{-1} e^{-4/(5\lambda)}] \, I(f) \, \psi_f(R) \, \psi_f(R) + 10\psi_{M_{3\lambda} * g^{(\lambda)}}(R)) \\ &\leqslant 150\lambda^{-1} [10^{-1} e^{-4/(5\lambda)}] \, I(f) [2\psi_{g^{(\lambda)}}(R/2) + \psi_{M_{3\lambda}}(R/2) + \psi_{M_{(1-e^{-2\lambda})}}(R/2)] \end{aligned} \tag{2.59}$$

where we have used (2.4) of Lemma 2.2 in the second inequality, and the bound on a and $|b|$ provided by (2.13). Then clearly

$$\int_{|v| \leqslant R_\psi(J(f), \lambda)} |av + b + \nabla \ln f(v)|^2 f(v) \, d^3v \geqslant \frac{J(f)}{2}$$

where $R_\psi(\cdot, \lambda)$ is defined in (2.57).

Finally,

$$\int_{\mathbb{R}^3} |av+b+\nabla \ln f(v)|^2 M_{\lambda^{1/2}/2}(v)\, d^3v$$

$$\geqslant \int_{|v| \leqslant R_\psi(J(f),\lambda)} |av+b+\nabla \ln f(v)|^2 M_{\lambda^{1/2}/2}(v)\, d^3v$$

$$\geqslant e^{-R_\psi(J(f),\lambda)^2/\lambda} \int_{|v| \leqslant R_\psi(J(f),\lambda)} |av+b+\nabla \ln f(v)|^2 f(v)\, d^3v$$

$$\geqslant e^{-R_\psi(J(f),\lambda)^2/\lambda} \frac{J(f)}{2}$$

since

$$M_{\lambda^{1/2}/2}/f \geqslant M_{\lambda^{1/2}/2}(R_\psi(J(f),\lambda))/\|f\|_\infty$$

on the region of integration, and we have the uniform bound (2.3) on f.

The result now follows from Lemma 2.4, the fact that $I(f) = 3 + J(f)$, and (2.58). ∎

Proof of Theorem 1.1. Suppose that $D(f) \geqslant \varepsilon$. Then by Lemma 2.1, $J(\mathscr{P}_\lambda f) \geqslant \varepsilon$ for all $\lambda \leqslant \Lambda(\chi, \varepsilon)$, and we have from the discussion in the introduction, the preceding lemma, and the obvious monotonicity properties of the function $\Gamma_{\lambda,\psi}(\varepsilon)$ defined there that

$$H(f \circ f) - H(f) = \int_0^\infty [I(\mathscr{P}_\lambda f) - I(\mathscr{P}_\lambda f \circ \mathscr{P}_\lambda f)]\, d\lambda$$

$$\geqslant \int_{\Lambda(\chi,\varepsilon)/2}^{\Lambda(\chi,\varepsilon)} [I(\mathscr{P}_\lambda f) - I(\mathscr{P}_\lambda f \circ \mathscr{P}_\lambda f)]\, d\lambda$$

$$\geqslant [\Lambda(\chi,\varepsilon)\varepsilon/4]\, \Gamma_\psi(\Lambda(\chi,\varepsilon)/2, \varepsilon) \quad ∎$$

Theorem 2.6 (Bound on the strong L^1 rate of approach to equilibrium). Let f_0 be a velocity density with zero bulk velocity and unit temperature and finite entropy. Suppose further that for some function χ increasing from zero, we have $\chi_{f_0} \leqslant \chi$. Suppose also that for some $p > 2$

$$\int_{\mathbb{R}^3} |v|^p f_0(v)\, d^3v < \infty \tag{2.60}$$

Then the solution of

$$\frac{\partial}{\partial t} f_t(v) = \nu[f_t \circ f_t(v) - f_t(v)] \tag{2.61}$$

starting from f_0 satisfies

$$\|f_t - M^{f_0}\|^2_{L^1(\mathbb{R}^3)} \leqslant D(f_t) \leqslant \Psi(t) \tag{2.62}$$

where $\Psi(t)$ is an explicitly computable function specified below which satisfies $\lim_{t\to\infty} \Psi(t) = 0$, and which depends on f_0 only through $D(f_0)$, χ, and $\int_{\mathbb{R}^3} |v|^p f_0(v)\, d^3v$.

Proof. We have shown in ref. 12 that

$$\chi_{f_t} \leqslant \chi_{f_0} \leqslant \chi \tag{2.63}$$

Furthermore, by a result of Elmroth,[19]

$$\int_{\mathbb{R}^3} |v|^p f_t(v)\, d^3v < C\left(\int_{\mathbb{R}^3} |v|^p f_0(v)\, d^3v\right)$$

uniformly in t, where C is a constant depending only on $\int_{\mathbb{R}^3} |v|^p f_0(v)\, d^3v$ as indicated. Therefore if we define ψ by

$$\psi(R) = C\left(\int_{\mathbb{R}^3} |v|^p f_0(v)\, d^3v\right)\Big/ R^{p-2}$$

it follows that

$$\psi_{f_t} \leqslant \psi \tag{2.64}$$

uniformly in t. Now define $\Psi(t)$ to be the solution of the ordinary differential equation

$$\frac{d}{dt}\Psi(t) = -\nu\Phi_{\psi,\chi}(\Psi(t)), \qquad \text{with} \quad \Psi(0) = D(f_0)$$

By (2.62), (2.63), and Theorem 1.1,

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

The inequality (2.61) now follows from a standard comparison argument, and it is evident that $\Psi(t)$ decreases to zero monotonically. ∎

3. ENTROPY PRODUCTION ESTIMATES FOR THE HARD-SPHERE COLLISION KERNEL

As indicated in the introduction, we will use the fact that the entropy production is monotone in the rate function to apply the results of the previous section to physical collision kernels. This entails the consideration of several rate functions at the same time. We shall write $\mathscr{Q}(f, f; \hat{b})$ to denote the collision kernel with rate function $\hat{b}$ when it is necessary to make this dependence explicit.

In the rest of this section, b denotes the hard-sphere rate function defined in (1.4). Fix any $\nu > 0$, and define two more rate functions

$$b^{(\nu)}(v, v', \omega) = \begin{cases} b(v, v', \omega) & \text{for } \; b(v, v', \omega) \geqslant \nu \\ \nu & \text{for } \; b(v, v', \omega) \leqslant \nu \end{cases} \tag{3.1}$$

and

$$b_{(\nu)}(v, v', \omega) = \begin{cases} 0 & \text{for } \; b(v, v', \omega) \geqslant \nu \\ \nu - b(v, v', \omega) & \text{for } \; b(v, b', \omega) \leqslant \nu \end{cases} \tag{3.2}$$

Notice that $b^{(\nu)} = b + b_{(\nu)}$, and hence,

$$-\int_{\mathbb{R}^3} \ln f \, \mathscr{Q}(f, f; b) \, d^3v = -\int_{\mathbb{R}^3} \ln f \, \mathscr{Q}(f, f; b^{(\nu)}) \, d^3v - \left[-\int_{\mathbb{R}^3} \ln f \, \mathscr{Q}(f, f; b_{(\nu)}) \, d^3v \right]$$

Since $b^{(\nu)} \geqslant \nu$, we have from Theorem 1.1 that

$$-\int_{\mathbb{R}^3} \ln f \, \mathscr{Q}(f, f; b^{(\nu)}) \, d^3v \geqslant -\int_{\mathbb{R}^3} \ln f \, \mathscr{Q}(f, f; \nu) \, d^3v \geqslant \nu \Phi_{\psi_f, \chi_f}(D(f)) \tag{3.3}$$

The reason that this bound is useful to us is that we shall be able to show that with $\phi_f(\nu)$ defined by

$$\phi_f(\nu) := -\int_{\mathbb{R}^3} |\ln f \, \mathscr{Q}(f, f; b_{(\nu)})| \, d^3v \tag{3.4}$$

we have $\phi_f(\nu) = o(\nu)$. Thus, for ν small enough, $\phi_f(\nu) \leqslant (1/2) \, \nu \Phi_{\psi_f, \chi_f}(D(f))$ and

$$-\int_{\mathbb{R}^3} \ln f \, \mathscr{Q}(f, f; b) \, d^3v \geqslant (1/2) \, \nu \Phi_{\psi_f, \chi_f}(D(f)) \tag{3.5}$$

Of course, to apply (3.5) to the study of solutions of the hard-sphere Boltzmann equation f_t, we shall need bounds which allow us control the choice of ν for $f = f_t$ explicitly in terms of the initial data f_0. It is easy to see why such bounds will hold for a wide class of initial data: Note that

$\|b_{(\nu)}\|_\infty = \nu$ and that the support of $b_{(\nu)}$ decreases to the empty set as ν decreases to zero. Thus, any sort of estimate which asserts the uniform integrability of the functions

$$\ln f_t(v)[f_t(\tilde{v}) f_t(\tilde{v}') - f_t(v) f_t(v')]$$

with respect to Lebesgue measure on $\mathbb{R}^3 \times \mathbb{R}^3$ gives us control, uniform in t, on the rate at which $\phi_{f_t}(\nu)$ tends to zero with decreasing ν. The first bounds of this type were obtained by Carleman, and his pioneering work has since been extended and complemented by many authors, as we shall shortly explain.

Other bounds are needed to control χ_{f_t} and ψ_{f_t} for solutions of the hard-sphere Boltzmann equation uniformly in t. While moment bounds may be used to control ψ_{f_t} exactly as in the last section, it is no longer the case that either χ_{f_t} or $I(f_t)$ is monotone decreasing in t, as they were in the constant-rate-function case. However, estimates on the modulus of continuity of $f_t(\cdot)$ that are uniform in t, again a type of estimate considered by Carleman, can be used to deduce the control we require on χ_{f_t}.

We now collect these estimates on solutions of the hard-sphere Boltzmann equation in the following theorem. We call the inequalities stated there "Carleman estimates" because they are, as we have indicated, all of a type first considered by Carleman.[9,10] After stating the theorem, we make a long series of remarks explaining how the theorem, as we have stated it, may be extracted from published literature. The result we present is only a convenient case of what is known. But our goal is to illustrate in the clearest manner possible how the results of the last section may be applied to the hard-sphere collision kernel. We do this by focusing on a special case—albeit one that is sufficiently broad to be of real physical interest. After treating this special case, we explain how recent refinements of Carleman's pioneering work may be used to broaden greatly the applicability of our analysis.

Nowhere in the published literature are explicit constants to be found, and neither shall we provide them here. While all of the constants are explicitly computable (without even too much book-keeping effort if one uses methods close to those originally employed by Carleman), *our aim is to show how the size of which constants affect the rate of approach to equilibrium in which quantitative way.*

Lemma 3.1 (Carleman estimates). Let f_0 be a density with zero bulk velocity and unit temperature satisfying the bounds

$$f_0(v) \leqslant \tilde{C}(1 + |v|^2)^{-5} \tag{3.6}$$

and

$$|f_0(v) - f_0(w)| \leqslant \tilde{B}\, |v - w| \qquad \text{for all} \quad v, w \in \mathbb{R}^3 \tag{3.7}$$

Then there exists a unique solution of the hard-sphere Boltzmann equation with initial data f_0. Furthermore, there exist computable constants C and B, depending only on $\tilde{B}$ and $\tilde{C}$ so that

$$f_t(v) \leqslant C(1 + |v|^2)^{-5} \tag{3.8}$$

and

$$|f_t(v) - f_t(w)| \leqslant B\, |v - w| \qquad \text{for all} \quad v, w \in \mathbb{R}^3 \tag{3.9}$$

uniformly in $t > 0$.

Finally, there exists a constant A depending only on $\tilde{B}$ and $\tilde{C}$ such that

$$|\ln f_t(v)| \leqslant A(1 + |v|^2)^{3/2} \tag{3.10}$$

uniformly in $t > 1$.

Remarks and References. The bounds (3.8) and (3.9) under these conditions are guaranteed by a special case of the theorem on p. 58 of Carleman's book.[10] To see that the constants are computable and how to compute them, one may consult the proof that is given only in the original paper, and there only in the special case that the density is a radial function. However, as stated there, the methods apply without this assumption. A more general result is proved *in full detail* by Arkeryd.[3]

The bound (3.10) calls for more extensive comment. Carleman proves a bound of this type in Theorem III of his paper (ref. 9, p. 119), *but only on the time interval* $0 < t_0 \leqslant t \leqslant t_1 < \infty$, and of course, only for radial densities. Again, the assumption that the density be radial is inconsequential (see ref. 26 for an explicit consideration of this point), but the finiteness of the time interval would be a problem.

However, inspection of the proof reveals that apart from depending on t_0 and t_1, the constant A in Carleman's bound depends only on a lower bound on $f_0(v)$ in some neighborhood $\{v \mid |v - v_0| < d\}$ of some point v_0. Since

$$\int_{|v| \leqslant \sqrt{2}} f_t(v)\, d^3v \geqslant 1/2$$

for each t, there must be some v_0 with $|v_0| \leqslant \sqrt{2}$ such that $f_t(v_0) \geqslant 3/(4\pi)$. Then, since (3.9) holds, we have

$$f_t(v) \geqslant 10^{-1} \quad \text{uniformly on} \quad \{v \mid |v - v_0| < (10B)^{-1}\}$$

Replacing f_0 by f_{t-1} and taking $t_0 = 1$ and $t_1 = 2$, we find that Carleman's argument now gives us

$$|\ln f_s(v)| \leqslant A(1+|v|^2)^{3/2} \qquad \text{for} \quad t < s < t+1$$

where A is independent of $t > 1$. In this way, the bound (3.10) may be extracted from published proofs.

Lemma 3.2 (Upper bound on entropy production for rate function $b_{(\nu)}$). Let f be a finite-entropy velocity distribution with zero bulk velocity and unit temperature. Suppose further that f satisfies the Carleman estimates (3.8)–(3.10). Then

$$\int_{S^2} \left[\int_{\mathbb{R}^3} \int_{\mathbb{R}^3} |\ln f(v)| \cdot |f(\tilde{v}) f(\tilde{v}') - f(v) f(v')| \, b_{(\nu)}(v, v', \omega) \, d^3v \, d^3v' \right] d\omega \leqslant 10^2 AC\nu^{28/25} \tag{3.11}$$

where A and C are the constants of (3.8) and (3.10).

Proof. For each fixed ω, define $E(R, \nu)$ by

$$E(R, \nu) := \{(v, v') \mid |(v - v') \cdot \omega| \geqslant \nu \text{ and } |v|^2 + |v'|^2 \leqslant R^2\}$$

Then

$$\left| \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \ln f(v) [f(\tilde{v}) f(\tilde{v}') - f(v) f(v')] \, b_{(\nu)}(v, v', \omega) \, d^3v \, d^3v' \right|$$
$$\leqslant 2AC \iint_{E(R,\nu)} (1 + |v|^2 + |v'|^2)^{3/2} (1 + |v|^2 + |v'|^2)^{-5} \, d^3v \, d^3v'$$
$$+ 2AC \iint_{|v|^2 + |v'|^2 \geqslant R^2} (1 + |v|^2 + |v'|^2)^{3/2} (1 + |v|^2 + |v'|^2)^{-5} \, d^3v \, d^3v' \tag{3.12}$$

The first integral in (3.12) is no greater than $2AC$ times the Lebesgue measure of $E(R, \nu)$. The second integral in (3.12) is no greater than $(8\pi^2/3) R^{-1}$.

We now estimate the Lebesgue measure of $E(R, \nu)$. We shall denote the Lebesgue measure of a measurable set E by $|E|$. First, fix ω and v. If $|(v - v') \cdot \omega| \leqslant \nu$, then for any α, either

$$|\cos \vartheta| \leqslant \nu^{\alpha} \qquad \text{or} \qquad |v - v'| \leqslant \nu^{(1-\alpha)}$$

where ϑ is the angle between ω and $v - v'$.

Clearly

$$|\{v \mid |v|^2 \leqslant R^2 \text{ and } |\cos \vartheta| \leqslant \nu^\alpha\}| \leqslant \pi R^3 \nu^\alpha$$

and thus

$$|\{(v, v') \mid |v^2| + |v'|^2 \leqslant R^2 \text{ and } |\cos \vartheta| \leqslant \nu^\alpha\}| \leqslant (4\pi^2/3)\, R^3 \nu^\alpha$$

Evidently,

$$|\{(v, v') \mid |v|^2 + |v'|^2 \leqslant R^2 \text{ and } |v - v'| \leqslant \nu^{(1-\alpha)}\}| \leqslant (4\pi/3)^2\, R^6 \nu^{(1-\alpha)}$$

Altogether,

$$\left| \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \ln f(v) [f(\tilde{v}) f(\tilde{v}') - f(v) f(v')]\, b_{(\nu)}(v, v'\omega)\, d^3v\, d^3v' \right|$$
$$\leqslant \nu A C((8\pi^2/3)\, R^{-1} + (4\pi^2/3)\, R^3 \nu\alpha + (4\pi/3)^2\, R^6 \nu^{(1-\alpha)})$$

Now choosing $R = \nu^{-\alpha/7}$ and $\alpha = 21/25$, so that each of the three terms above contains the same power of ν, and averaging over ω, we obtain the stated result. ∎

We now use the Carleman estimates to control χ_{f_t} uniformly in t. We shall make use of the fact that when f satisfies the bound (3.8), then f^p is integrable for all $p > 3/10$. (In fact, an easy argument using Hölder's inequality shows that f^p is integrable for all $p > 3/5$ whenever f is a finite-temperature density. Thus, what we now do can be extended to a much more general setting, as we later explain.)

We shall denote $[\int_{\mathbb{R}^3} f^p(v)\, d^3v]^{1/p}$ by $\|f\|_p$ for all $p > 0$; the fact that $\|\cdot\|_p$ is only a norm for $p \geqslant 1$ shall be without consequence below.

Lemma 3.3 (Continuity estimate for the entropy). Let f and g be two velocity densities with zero bulk velocity and unit temperature which satisfy the bound (3.8). Then

$$|H(f) - H(g)| \leqslant [10(C + e) \ln(C + e) + 10^3 C^{1/2}]\, \|f - g\|_1^{2/9} \quad (3.13)$$

whenever $\|f - g\|_1 \leqslant 1$, and where C is the constant in (3.8).

Proof. We have

$$|f \ln f - g \ln g| \leqslant |f^{1/2} - g^{1/2}| \cdot |f^{1/2} \ln f|$$
$$+ |g^{1/2}| \cdot |f^{1/2} \ln f - g^{1/2} \ln g| \quad (3.14)$$

We shall separately estimate the integrals of each of these terms over the following sets, defined for any fixed $c > 0$:

$$\Omega_1 = \{v \mid f(v), g(v) \geqslant c\}$$

$$\Omega_2 = \{v \mid f(v) \geqslant 2c, c \geqslant g(v)\} \cup \{v \mid g(v) \geqslant 2c, c \geqslant f(v)\}$$

$$\Omega_3 = \{v \mid 2c \geqslant f(v), g(v)\}$$

We shall bound the integral over Ω_1 using the fact that $|f-g|$ is small on Ω_1 when $\|f-g\|_1$ is small. We shall bound the integral over Ω_2 using the fact that Ω_2 is a small set when $\|f-g\|_1$ is small. We shall bound the integral over Ω_3 using the fact that f and g are small on Ω_1 when $\|f-g\|_1$ is small, together with the fact that $\|f\|_{1/2}$ and $\|f\|_{1/2}$ are controlled by (3.8).

First, note that $\|(f^{1/2} \ln f)\|_\infty \leqslant 2 + \ln C$. Also, on Ω_1, $|f^{1/2} - g^{1/2}| \leqslant \frac{1}{2} c^{-1/2} |f-g|$, and

$$|f^{1/2} \ln f - g^{1/2} \ln g| \leqslant (1 + c^{-1/2}) |f-g|$$

Thus,

$$\int_{\Omega_1} [|f^{1/2}(v) - g^{1/2}(v)| \cdot |f^{1/2} \ln f| + |g^{1/2}| \cdot |f^{1/2} \ln f(v) - g^{1/2} \ln g(v)|] \, d^3v$$
$$\leqslant [(1 + \ln C/2) c^{-1/2} + C^{1/2} \ln C(1 + c^{-1/2})] \|f-g\|_1 \tag{3.15}$$

Next, note that

$$\|f-g\|_1 \geqslant \int_{\Omega_2} |f(v) - g(v)| \, d^3v \geqslant \int_{\Omega_2} c \, d^3v = c\, |\Omega_2|$$

Therefore,

$$\int_{\Omega_2} [|f^{1/2}(v) - g^{1/2}(v)| \cdot |f^{1/2} \ln f| + |g^{1/2}| \cdot |f^{1/2} \ln f(v) - g^{1/2} \ln g(v)|] \, d^3v$$
$$\leqslant 4(C \ln C) c^{-1} \|f-g\|_1 \tag{3.16}$$

Finally, note that on Ω_3, for all $c \leqslant 1$, $g^{1/2} \ln g$, $f^{1/2} \ln f \leqslant (1/\gamma) c^{(1-\gamma)}$ for any $\gamma > 0$. Therefore,

$$\int_{\Omega_3} [|f^{1/2}(v) - g^{1/2}(v)| \cdot |f^{1/2} \ln f| + |g^{1/2}| \cdot |f^{1/2} \ln f(v) - g^{1/2} \ln g(v)|] \, d^3v$$
$$\leqslant (1/\gamma) c^{(1/2-\gamma)} (\|f\|_{1/2}^{1/2} + 3 \|g\|_{1/2}^{1/2}) \tag{3.17}$$

By summing (3.15)–(3.17) and then choosing $c = \|f-g\|_1^{2/3}$ and $\gamma = 1/6$, we obtain the result. ▌

Lemma 3.4 (Bounds on χ_f for densities satisfying Carleman estimates). Let f be a velocity density with zero bulk velocity and unit temperature. Suppose further that f satisfies the Carleman estimates (3.8) and (3.9). Then for all $\lambda \leqslant 10^{-2}$,

$$\mathscr{P}_\lambda f(v) \leqslant K_1(C)(1+|v|^2)^{-5} \tag{3.18}$$

where $K_1(C)$ depends computably on the constant C in (3.8). Moreover,

$$\chi_f(\lambda) \leqslant K_2(B, C)\, \lambda^{1/20} \tag{3.19}$$

where $K_2(B, C)$ depends computably on the constants B and C of (3.8) and (3.9).

Proof. We recall the notation $g_{(\lambda)}(v) = e^{3\lambda} g(e^\lambda v)$. Then, using $*$ to denote convolution, we may rewrite (1.14) as

$$\mathscr{P}_\lambda f = (M_{(1-e^{-2\lambda})e^{2\lambda}} * f)_{(\lambda)} \tag{3.20}$$

Fix any $\alpha > 0$, any v, and any $R < |v|$. Then

$$\begin{aligned}
&\int_{\mathbb{R}^3} M_\alpha(w)\, f(v-w)\, d^3w \\
&\quad = \int_{|w| \leqslant R} M_\alpha(w)\, f(v-w)\, d^3w + \int_{|w| \geqslant R} M_\alpha(w)\, f(v-w)\, d^3w \\
&\quad \leqslant C \left\{ [1 + (|v|-R)^2]^{-5} + \int_{|w| \geqslant R} M_\alpha(w)\, d^3w \right\} \\
&\quad \leqslant C\{ [1 + (|v|-R)^2]^{-5} + \pi^{-1/2}[x + (2/x)]\, e^{-x^2} \}
\end{aligned}$$

where $x^2 = R^2/2\alpha$. We choose R so that $x^2 = 6 \ln(1+|v|^2)$. Then

$$\pi^{-1/2}[x + (2/x)]\, e^{-x^2} \leqslant 18\pi^{-1/2}[1 + (|v|-R)^2]^{-5}$$

Next, we have $R \leqslant (12\alpha)^{1/2}\, |v|$. In our application, $\alpha = e^{2\lambda} - 1 \leqslant 2\lambda e^{2\lambda} \leqslant (1/50)\, e^{1/50}$. Thus, $R \leqslant (1/4)\, |v|$, and hence $[1 + (|v|-R)^2]^{-5} \leqslant 20(1+|v|^2)^{-5}$. This establishes (3.18) for $|v| \geqslant 2^{-1/2}$. Since we clearly have $\mathscr{P}_\lambda f(v) \leqslant C$, for all v, (3.18) holds for $|v| \leqslant 2^{-1/2}$. As the first step in proving (3.19), note that by an elementary computation, $H(\mathscr{P}_\lambda f) = H(M_{e^{2\lambda}-1} * f) - 3\lambda$, we have that

$$\chi_f(\lambda) \leqslant H(M_{e^{2\lambda}-1} * f) - H(f) \tag{3.21}$$

By (3.19), we may apply Lemma 3.3 provided we can control $\|M_{e^{2\lambda}-1} * f - f\|_{L^1(d^3v)}$. To do this, we introduce the $L^1(d^3v)$ modulus of continuity $\delta_f(\cdot)$:

$$\delta_f(r) = \sup\left\{\int_{\mathbb{R}^3} |f(v+w) - f(v)|\, d^3v \mid |w| \leqslant r\right\} \tag{3.22}$$

Note that for $|w| \leqslant 1$ and $R \geqslant 2$,

$$\int_{\mathbb{R}^3} |f(v+w) - f(v)|\, d^3v \leqslant (4\pi/3)\, BR^3\, |w| + C(8\pi/7)\, 2^{10} R^{-7}$$

Choose $R = 2\, |w|^{-1/10}$ and obtain

$$\delta_f(r) \leqslant (10^2 B + 10C)\, r^{7/10} \tag{3.23}$$

Also, since f is a density, $\delta_f(r) \leqslant 2$ for all r.

Then since

$$\begin{aligned} &\|M_{e^{2\lambda}-1} * f - f\|_{L^1(d^3v)} \\ &\quad \leqslant \int_{\mathbb{R}^3} M_{e^{2\lambda}-1}(w) \left(\int_{\mathbb{R}^3} |f(v+w) - f(v)|\, d^3v\right) d^3w \\ &\quad \leqslant (4\pi/3)\, R^3 \delta_f(R) + 2 \int_{|w| \geqslant R} M_{e^{2\lambda}-1}(w)\, d^3w \; le \\ &\quad\quad + (4\pi/3)\, R^3 (10^2 B + 10C)\, R^{7/10} + 6(e^{2\lambda} - 1)\, R^{-2} \end{aligned}$$

Choose $R = 2(e^{2\lambda} - 1)^{1/3}$, and finally obtain

$$\|M_{e^{2\lambda}-1} * f - f\|_{L^1(d^3v)} \leqslant K_3(B, C)\, \lambda^{7/30}$$

The result (3.20) is now a consquence of Lemma 3.3. ∎

As explained at the beginning of this section, the following theorem is a direct consequence of Theorem 1.1 and the lemmas proved here.

Theorem 3.4 (Entropy production bounds for the hard-sphere collision kernel). Let f be a density with zero bulk velocity, unit temperature, and finite entropy satisfying the bounds (3.8)–(3.10). Then there is a strictly increasing function $\Phi_{A,B,C}\colon \mathbb{R}_+ \to \mathbb{R}_+$ depending only on the constants A, B, and C in (3.8)–(3.10) such that

$$-\int_{\mathbb{R}^3} \ln f(v)\, \mathcal{Q}(f, f)(v)\, d^3v \geqslant \Phi_{A,B,C}(D(f)) \tag{3.24}$$

where $\mathcal{Q}$ is the Boltzmann collision kernel for hard spheres.

Because of Lemma 3.1, we now obtain an analog of Theorem 2.7 for the hard-sphere Boltzmann equation for initial data satisfying bounds of the type (3.5) and (3.6).

ACKNOWLEDGMENTS

The work of E.A.C. was partially supported by U.S. NSF grant DMS 92-07703; that of M.C.C. by FLA.

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, L^∞ estimates for the space-homogeneous Boltzmann equation, *J. Stat. Phys.* **31**:347 (1983).
4. L. Arkeryd, Stability in L^1 for the spatially homogeneous Boltzmann equation, *Arch. Rat. Mech. Anal.* **103**:151–168 (1988).
5. A. R. Barron, Entropy and the central limit theorem, *Ann. Prob.* **14**:336–342 (1986).
6. N. M. Blachman, The convolution inequality for entropy powers, *IEEE Trans. Inform. Theory* **2**:267–271 (1965).
7. A. N. Bobylëv, The theory of the non-linear spatially uniform Boltzmann equation for Maxwellian molecules, *Math. Phys. Rev.* **7**:111–223.
8. L. D. Brown, A proof of the central limit theorem motivated by 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.
9. T. Carleman, Sur la solution de l'équation intégrodifférential de Boltzmann, *Acta Math.* **60**:91–146 (1933).
10. T. Carleman, *Problèmes Mathématique dans la théorie cinétique des gaz* (Almqvist Wiksells, Uppsala, 1957).
11. E. A. Carlen, Superadditivity of Fisher's information and logarithmic Sobolev inequalities, *J. Funct. Anal.* **101**:194–211 (1991).
12. E. A. Carlen and M. C. Carvalho, Strict entropy production bounds and stability of the rate of convergence to equilibrium for the Boltzmann equation, *J. Stat. Phys.* **67**:575–608 (1992).
13. E. A. Carlen and A. Soffer, Entropy production by block variable summation central limit theorems, *Commun. Math. Phys.* **140**:339–371 (1991).
14. C. Cercignani, *H*-theorem and trend to equilibrium in the kinetic theory of gases, *Arch. Mech.* **34**:231–241 (1982).
15. C. Cercignani, *The Boltzmann Equation and Its Applications* (Springer-Verlag, New York, 1988).
16. I. Csiszar, Informationstheoretische Konvergenenzbegriffe im Raum der Wahrscheinlichkeitsverteilungen, *Pub. Math. Inst. Hung. Acad. Sci.* **VII**(Series A):137–157 (1962).
17. L. Desvillettes, Entropy dissipation rate and convergence in kinetic equations, *Commun. Math. Phys.* **123**:687–702 (1989).

18. M. Dresden, *Kinetic Theory Applied to Hydrodynamics* (Colloquium Lectures in Pure and Applied Science Series 2, Socony Mobil Oil Co. Field Research Laboratory, Dallas, Texas, 1956).
19. 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).
20. T. Elmroth, On the H-function and convergence towards equilibrium for a space-homogeneous molecular density, *SIAM J. Appl. Math.* **44**:150–159 (1984).
21. H. Grad, Asymptotic theory of the Boltzmann equation, *Phys. Fluids* **6**:147–181 (1963).
22. L. Gross, Logarithmic Sobolev inequalities, *Am. J. Math.* **97**:1061–1083 (1975).
23. T. Gustafsson, Global L^p properties for the spatially homogeneous Boltzmann equation, *Arch. Rat. Mech. Anal.* **103**:1–38 (1988).
24. J. Lebowitz and E. Montroll, *Nonequilibrium Phenomena I: The Boltzmann Equation* (North-Holland, Amsterdam, 1984).
25. Ju. V. Linnik, An information theoretic proof of the central limit theorem with Lindeberg conditions, *Theory Prob. Appl.* **4**:288–299 (1959).
26. N. Maslova and Yu. Romanovskii, Mathematical methods of studying the Boltzmann equation, *St. Petersburg Math. J.* **1**:1–41 (1992).
27. J. C. Maxwell, *The Scientific Papers by J. C. Maxwell* (Dover, New York, 1965).
28. H. P. McKean, Speed of approach to equilibrium for Kac's caricature a Maxwellian gas, *Arch. Rat. Mech. Anal.* **21**:343–367 (1966).
29. 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).
30. D. Morgenstern, Analytical studies related to the Maxwell–Boltzmann equation, *J. Rat. Mech. Anal.* **4**:154–183 (1955).
31. C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, Illinois, 1949).
32. A. Stam, Some inequalities satisfied by the quantities of information of Fisher and Shannon, *Inform. Control* **2**:101–112 (1959).
33. G. Toscani, New *a priori* estimates for the spatially homogeneous Boltzmann equation, *Cont. Mech. Thermodyn.*, in press.
34. G. Toscani, Strong convergence towards equilibrium for a gas of Maxwellian pseudomolecules, *Cont. Mech. Thermodyn.*, in press.
35. G. Toscani, Lyapunov Functionals for a Maxwell Gas, Ferrara preprint (1992).
36. 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).
37. G. Uhlenbeck, The statistical mechanics of non-equilibrium phenomena, 1954 Higgins Lectures, Mimeographed notes in Fine Hall Library, Princeton University.
38. B. Wennberg, Stability and exponential convergence in L^p for the space homogeneous Boltzmann equation, *Nonlinear Analysis T.M.A.*, to appear.
39. E. Wild, On Boltzmann's equation in the kinetic theory of gases, *Proc. Camb. Phil. Soc.* **47**:602–609 (1951).