1969

Some Problems Connected with the Boltzmann Equation

Francisco Alberto Grunbaum

SOME PROBLEMS CONNECTED WITH THE BOLTZMANN EQUATION

A thesis submitted to the faculty of the Rockefeller University

in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

by

Francisco Alberto Grünbaum

Approved for publication. Henry McKean
Prof. Math. CIMS, NYU.

20 May 1969

The Rockefeller University

New York, New York

ACKNOWLEDGMENT

I am very grateful to Dr. Detlev Bronk and Dr. Frank Brink, Jr. for the hospitality they offered me at Rockefeller University.

Over the lapse of these last three years, I enjoyed and profited from conversations and discussions with Professors Mark Kac, Morris Schreiber and especially Professor Loren Pitt.

To my advisor Professor Henry P. McKean, Jr., I owe the greatest debt of gratitude. For me it has been an invaluable experience to be at his side for these years enjoying his help and his friendship. He guided and encouraged me in every phase of my research and made many personal contributions to make this a more readable work.

I have to thank Mrs. Virginia Early Vilk and Miss Mary Ellen O'Brien for a very fine typing job.

Finally I want to thank my wife, Loli, for an infinite amount of patience and understanding during these last three years.

TABLE OF CONTENTS

INTRODUCTION

CHAPTER 1 THE LINEARIZATION PROBLEM

1. KAC'S MODEL OF THE BOLTZMANN EQUATION
2. EXISTENCE AND UNIQUENESS FOR THE EVEN FLOW
3. THE LINEARIZATION PROBLEM
4. EIGENVALUE INEQUALITIES
5. INTERTWINING OPERATOR
6. REMARKS ON MORE GENERAL CASES

APPENDIX I PRODUCTS OF HERMITE FUNCTIONS

APPENDIX II AN ELEMENTARY INEQUALITY

CHAPTER 2 THE PROPAGATION OF CHAOS

1. THE n-MOLECULE GAS
2. PREPARATIONS
3. McKEAN'S 2-STATE MAXWELLIAN GAS
4. THE GENERAL CASE

APPENDIX I WEAK CONTINUITY OF THE BOLTZMANN FLOW

INTRODUCTION

Consider a dilute gas composed of a very large number of molecules moving in space according to the laws of classical mechanics, and colliding in pairs from time to time. Assume that we can disregard all external effects, such as gravity, so that the motion is completely specified by giving the intermolecular forces.

One is interested in the number of molecules which at time t have position r and velocity v, within $drdv$. This is given by

$$n(t, r, v) = Nf(t, r, v)\, drdv^{*}$$

where f is called the density function. It is clear that this quantity is going to change in time due to the motion of the molecules and to the effect of the collisions.

Boltzmann derived an equation for the rate of change of f with time. It has the form of a non-linear integro-differential equation:

$$(1) \quad \frac{\partial f}{\partial t} + v_1 \frac{\partial f}{\partial r} =$$

$$Bf = \iint [f(v_2^*)f(v_1^*) - f(v_1)f(v_2)]\, |v_1 - v_2|\, I(|v_1 - v_2|, \theta) \sin\theta \, d\theta \, d\phi \, dv_2 .$$

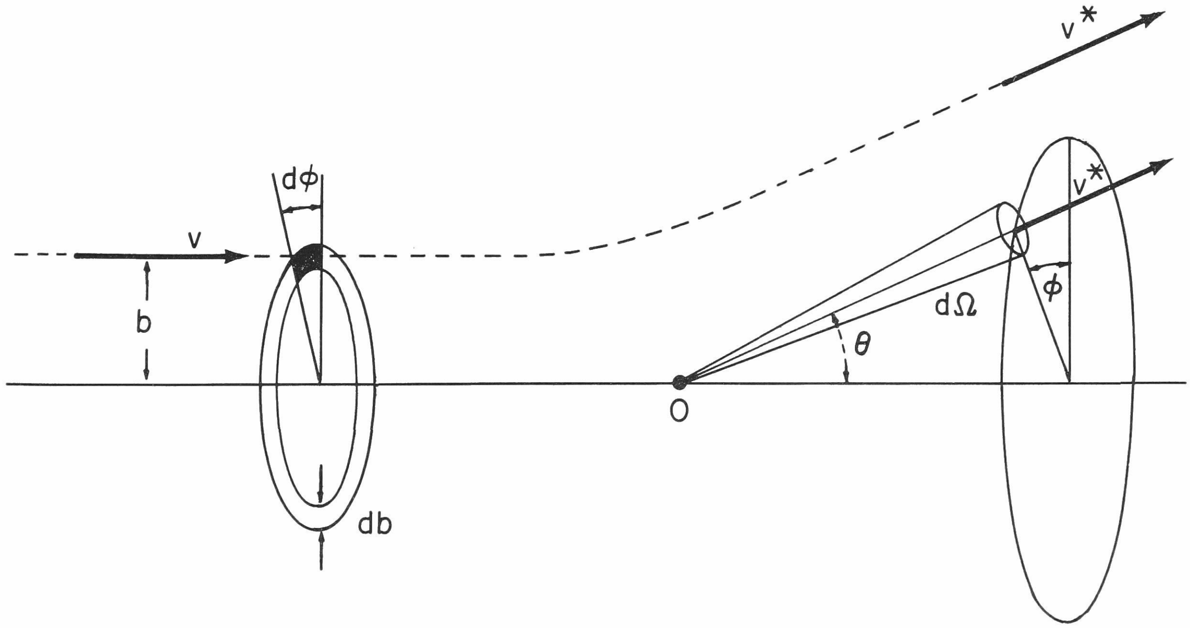
Here f stands for $f(t, r, v_1)$. The integral part, containing the non-linearity reflects the effect of the collisions between molecules;

---

* N is the total number of molecules.

the term $v_1 \frac{\partial f}{\partial r}$ reflects the motion of the molecules between collisions. The integrals on the right will now be explained.

The intermolecular forces are supposed to be given by pair forces, all interactions involving more than two particles being ignored. Therefore, we deal first with the mechanism of an individual collision of two molecules. For the forthcoming description, it is convenient to refer the two colliding molecules, travelling with velocities v_1 and v_2, to the center-of-mass coordinate system. Thus the situation becomes that of a ficticious molecule, travelling with velocity $v = v_2 - v_1$, scattered by a center of force placed at the center-of-mass, denoted by O; see the figure.



Recall that if v_1^*, v_2^* denote the velocities of the molecules after the collision, we have the conservation laws of

momentum: $$v_1 + v_2 = v_1^* + v_2^*$$

and energy: $$|v_1|^2 + |v_2|^2 = |v_1^*|^2 + |v_2^*|^2$$

and their consequence

$$|v_1 - v_2| = |v_1^* - v_2^*|$$

If one also recalls that $v = v_2 - v_1$ and $v^* = v_2^* - v_1^*$ lie in the same plane, then it is clear that the impact parameter b and the angles ϕ and θ in the figure determine the kinematics of a collision completely. The impact parameter b measures the distance from the direction given by v, $0 \leq \phi < 2\pi$ is the azimuth angle and $0 \leq \theta \leq \pi$ is the colatitude which measures the scattering angle formed by v and v^*, see figure.

Think now of a uniform beam of molecules coming from the left with velocity v. Then the dynamical description of the collision process, pictured here as a scattering process, is embodied in the expression

$$|v_1 - v_2|\, b\,db\,d\phi = |v_1 - v_2|\, I(|v_1 - v_2|, \theta) \sin\theta\, d\theta\, d\phi$$

which gives the number of molecules that pass through the shaded area in the figure, per unit time. This formula <u>defines</u> the differential scattering cross section I. Because θ can be computed from b, v, and the intermolecular force, I is uniquely determined by this recipe.

As to the meaning of the integral itself, the number of collisions, in unit time, of a molecule with velocity v_1 with a molecule of (arbitrary) velocity v_2 and parameters b and θ, within $dbd\phi$, is given by

$$(2) \qquad |v_1 - v_2|\, f(v_1, t)\, bdbd\phi .$$

Boltzmann goes one step further and puts for the loss, due to collisions, of molecules with velocities v_1, within dv_1, in unit time,

$$(3) \qquad dv_1 \iiint |v_1 - v_2|\, f(v_1, t) f(v_2, t)\, bdbd\phi\, dv_2 .$$

Now he argues that a direct collision from v_1, v_2 into v_1^*, v_2^* is allied to a restituting collision from v_1^*, v_2^* into v_1, v_2 and that <u>both stand on the same footing.</u> Thus the gain of molecules with velocity v_1, in unit time should be

$$(4) \qquad dv_1^* \iiint |v_1^* - v_2^*|\, f(v_1^*, t) f(v_2^*, t)\, bdbd\phi\, dv_2^* .$$

To get the expression for Bf given by (1), it is now enough to combine (3) and (4) after appropriate use is made of the identity $dv_1 dv_2 = dv_1^* dv_2^*$ and the fact that $|v_1 - v_2|\, I(|v_1 - v_2|, \theta) \sin\theta d\theta d\phi$ remains unchanged if $v_1 v_2$ are interchanged with $v_1^* v_2^*$. This says that <u>a transition from</u> (v_1, v_2) <u>into</u> (v_1^*, v_2^*) <u>has the same chance as one from</u> (v_1^*, v_2^*) <u>into</u> (v_1, v_2), a fact which is used over and over again in this thesis.

We emphasize that we have not presented a derivation from first principles. We computed the Stoss-zahl-ansatz (2) by invoking a dynamical description whose connection with actual molecules in a gas is not completely transparent. Then we accepted (3) on the basis of molecular chaos, and finally from (3) we arrived at (4) on the assumption of microscopic reversibility. The reader will find detailed explanations of these terms in Uhlenbeck-Ford [17] or Grad [3].

To conclude this look at Bf, we remark that the total scattering cross section $\int I \sin\theta d\theta d\phi = \int b db d\phi$ is convergent only if two molecules at a distance larger that some $R < \infty$ cannot feel each other. In that case, we have $\int b db d\phi = \pi R^2$. This kind of restriction is usually referred to as a cutoff interaction.

Although the form of the equation (1) does not depend on the intermolecular force, the quantity $I(|v_1 - v_2|, \theta)$ inside the integral does so. For the moment we mention only the cases of hard spheres and the Maxwellian gas. In the first case, the molecules are hard spheres which do not interact with each other except when they touch. Then they exchange their velocities in a perfectly elastic collision and one gets for $I(|v_1 - v_2|, \theta)$ = a constant multiple of $\sin(\theta/2)$. In the Maxwellian case there is a central potential which is inversely proportional to r^4, and you find that $|v_1 - v_2| I(|v_1 - v_2|, \theta)$ is a function of θ alone with

a pole at $\theta = 0$. The pole is customarily removed by making a cutoff.

There are many features of equation (1) - and (1') below - which do not depend on the intermolecular force, for example:

(a) The Maxwell-Boltzmann densities,

$$g(v) = (2\pi\sigma^2)^{-3/2} \exp(-|v|^2/2\sigma^2)$$

are the only stationary densities for (1').

(b) The first and second moments of the density function $f(t, v)$ are preserved in time, i. e.

$$\int |v| f(t, v)dv \quad \text{and} \quad \int |v|^2 f(t, v)dv$$

are constant.

(c) The H-theorem, which states that the entropy H

$$H(t) = -\int f(t, v) \log f(t, v)dv$$

increases with time.

For all these facts, as well as for related material, the reader can consult Uhlenbeck-Ford [17] or Huang [6].

We will be concerned with a special instance of (1): the so-called spatially homogeneous case in which $f(t, r, v)$ is independent of r, and (1) simplifies to

(1') $$\frac{\partial f}{\partial t} = Bf$$

We also assume the interaction to have a cut-off, so that $\int I \sin\theta d\theta d\phi = \pi R^2 < \infty$.

Under these restrictions the initial value problem for the Boltzmann equation has been much studied. The ideal result here would be that a finite second moment for the initial distribution assures existence and uniqueness of the solution.

The problem turns out to be quite a difficult one. Carleman [1] solved the case of radial solutions for the hard spheres case and Wild [18] the case of a cut-off Maxwellian gas.* Povzner studied a slightly modified equation, which reduces in the spatially homogeneous case to the classical equation with a finite total scattering cross section (cut-off). He proves that a finite second moment for the initial distribution does guarantee existence, but he has to impose a finite fourth moment to get uniqueness. See [13].

We are now in a position to explain the contents of this thesis.

Chapter I is devoted to the problem of approach to equilibrium for (1'). Instead of considering the general Boltzmann equation, we deal

---

* For the space dependent case, the only one of physical interest, consult Grad [4].

only with a one-dimensional caricature of the Maxwellian gas introduced and studied by Kac [7]. This model has all the mathematical features that are present in the actual Maxwellian gas but all manipulations are simpler. From a physical point of view, it has the disadvantage that in a collision, only energy, but not momentum is preserved.

For this simple example you can give a very complete geometrical description of the convergence to the Maxwell-Boltzmann distribution. All consideration are now localized near this Maxwell-Boltzmann distribution, denoted here by g.

First we prove existence and uniqueness of the solution for the non-linear equation in some appropriate L^2 space. Then we see that the rate of approach to g is governed by the first negative eigenvalue of the linearization of B at g, i.e. $\| f_t - g \| \leq \text{constant} \; e^{-vt}$. We also prove that we have a contractive flow, i.e. $\| f_t^{(1)} - f_t^{(2)} \|$ is a decreasing function of t if $f^{(1)}$ and $f^{(2)}$ are close enough to g.

But much more than this can be said. We prove that there is an extremely smooth (actually analytic) non-linear "change of coordinates" around g, so that if the non-linear problem is reexpressed in terms of these new coordinates, then it is exactly the linearized problem. In this picture the result mentioned above about the rate of convergence appears as an innocent manifestation of the fact just mentioned.

Some straight-forward extensions of this to include higher-order collisions are indicated in Appendix II. There one also finds a discussion of the difficulties in trying to extend the results to the general Boltzmann equation (1').

The problem of the rate of approach to equilibrium for the solutions of the (non-linear) Boltzmann equation has been treated by Grad [4] and McKean [9]. Grad shows that for a general spatially homogeneous case, the decay in some appropriate L^2 space is exponential and that the decay exponent can be taken as close as you please to the first negative eigenvalue of the linearized operator, provided the initial distribution is close enough to equilibrium. For Kac's caricature of the Maxwellian gas, McKean gets an exponential decay for the L^1 norm; but his decay constant is much smaller than it should be. On the other hand the L^1 norm, and not the one used by Grad or by us, is the one that makes more sense globally.

To explain the contents of Chapter II, we have to recall some of the history of the subject.

It is well known that after Boltzmann derived his celebrated equation, he spent most of his life involved in various controversies arising from it.

His discovery of the H-theorem, while explaining success-

fully the irreversible approach to an equilibrium distribution, was hard to reconcile with reversible Newtonian mechanics. There was a deep feeling of suspicion of such irreversible behaviour obtained from a time-reversible model. The two main objections were raised by Loschmidt and Zermelo, and one cannot say that Boltzmann was able to settle these points.

Remarkable and far reaching ideas were developed in trying to find a scheme in which Boltzmann's statements would hold free of inconsistency. In this respect the pioneering work of Gibbs and Ehrenfest should be mentioned.

Since then some new insights have been obtained by recognizing that more coherence is gained by converting most of the assertions into probability statements. That this was an essential step was clear in Boltzmann's mind, but not in the minds of most of his contemporaries.

One such development was first proposed by Uhlenbeck in a similar statistical problem in the theory of cosmic ray showers [16]; see also Siegert [14]. This goes by the name of "the master equation", and applies only in the spatially homogeneous case.

Uhlenbeck's idea is to start from scratch talking in probability language and to exploit fully the interpretation of the "Stoss-zahl-ansatz" (2) as a probability affair. He proposes to look at a gas of n molecules

which collide in pairs at random times, the "Stoss-zahl-ansatz" being the law of random collisions. We call this construction "the n-molecule gas." Let $p = p(t, v_1 \dots v_n)$ stand for the density of the distribution of the n velocities. This quantity changes in time due to the collisions according to the master equation

$$\frac{\partial p}{\partial t} = Gp. \tag{5}$$

Here G takes the form

$$Gp(v, \dots, v_n) = \frac{2}{n} \sum_{i<j} \int [p(\dots v_i^* \dots v_j^* \dots) - p(\dots v_i \dots v_j \dots)] \, |v_i - v_j| I(|v_i - v_j|, \theta) \sin\theta \, d\theta \, d\phi$$

where the sum is taken over all possible pairs $i<j$. The map $(v_i^*, v_j^*) \to (v_i, v_j)$ described the result of a collision with scattering angle θ, and I is the differential scattering cross section.

Assume from now on that p is symmetric in $v_1, \dots, v_n$, corresponding to the case of indistinguishable particles. This class of distributions is closed under the flow given by (5) because G commutes with permutations.

Now we want to see how the Boltzmann equation fits into the approach. The master equation describes the joint evolution for n molecules, and Boltzmann's the evolution for one molecule. To connect the two, let us define a family of contracted distributions by means of

$$(6) \qquad p^{(k)}(v_1 \ldots v_k) = \int p(v_1 \ldots v_k, v_{k+1}, \ldots v_n) dv_{k+1} \ldots dv_n \ ,$$

one for each integer $k \leq n$. Here we use the symmetry of p, so that it does not matter which $n-k$ variables are integrated out.

If we now integrate over $v_2, \ldots, v_n$ on both sides of (5) and keep the integral over v_2 on the right hand side explicit, we get

$$\frac{\partial p^{(1)}}{\partial t}(v_1) = \int [p^{(2)}(v_1^*, v_2^*) - p^{(2)}(v_1, v_2)] |v_1 - v_2| I(|v_1 - v_2|, \theta) \sin\theta d\theta d\phi dv_2 \ .$$

Now, this would look like the Boltzmann equation if we had

$$(7) \qquad p^{(2)}(t, v_1, v_2) = p^{(1)}(t, v_1) p^{(1)}(t, v_2).$$

How can this be achieved? We can always start from a "chaotic distribution"

$$p(o, v_1, \ldots, v_n) = p^{(1)}(o, v_1) \ldots p^{(1)}(o, v_n)$$

so that (7) holds at time $t = 0$.

But we cannot assume that this state of affairs persists at later times, because for $t>0$, $p(t, v_1, \ldots, v_n)$ will be determined by the master equation, and (7) <u>is in fact false</u> as soon as we have a bona-fide interaction between the molecules. Kac [7] realized this and suggested a possible mechanism for recovering the validity of (7) by letting the number of molecules go to ∞. The physical reason why

(7) should be true for an infinite gas is that then a molecule that collides with a specific one flies off and is never seen again; hence if no correlations are present at $t = 0$, they cannot develop later either.

Roughly speaking, we prove in Chapter II that _if_ $p(o, v_1 \ldots v_n)$ _is a "chaotic distribution" then this property propagates in time_, i. e. $p(t, v_1, \ldots, v_n)$ _is also chaotic in the limit as_ $n \to \infty$.

Kac substantiated this conjecture for a model of the Maxwellian gas [7].* His proof was later put in a more algebraic form by McKean [9]. Kac's proof requires some estimates which do not hold beyond the Maxwellian case, and that approach cannot be pushed further. In Chapter II we present a method which avoids those estimates. It requires a larger amount of abstraction, and two very believable but still unproven assumptions on the smoothness of the Boltzmann flow, namely:

1) EXISTENCE AND UNIQUENESS hold in the class of all probability measures under the sole condition that the second moment of the initial distribution is finite.

2) SMOOTHNESS: at any fixed $t>0$, the solution of the Boltzmann equation f_t is a smooth (differentiable)** functional of the initial data f_o.

---

* Incidentally, this model is the one we used in Chapter I.

** In a sense to be made precise in Chapter II.

The first assumption appears to be only a technical difficulty. The second one expresses the smoothness of the Boltzmann flow, a fact that should not be doubted.

CHAPTER 1 THE LINEARIZATION PROBLEM

1. KAC'S MODEL OF THE BOLTZMANN EQUATION

The purpose of this section is two-fold. First, we introduce Kac's equation and give a condensed exposition of the $L^1(R)$ theory going with it. This space is the natural place to look at Kac's equation because the solution is a density function and $\int f (=1) < \infty$ is automatic.

Once this is done, we try to prepare the spirit of the reader to take up in sections 2-5 the study of the "change of coordinates" described in the introduction.

For all details omitted in this section and for a very interesting analysis of many questions not touched upon by us, e. g. the central limit theorem for Maxwellian molecules, the reader may consult McKean [9].

We consider the initial value-problem for a probability density function $f(t, a)$

$$\text{(1a)} \qquad \frac{\partial f}{\partial t}(t, a) = \int_R \int_{S^1} [f(t, a^*) f(t, b^*) - f(t, a) f(t, b)] \, db\, dw = Bf^*$$

subject to the conditions

$$\text{(1b)} \qquad f(o, a) = f_o \geq 0, \int f(t, a)\, da \equiv 1, \ \sigma^2(f) = \int a^2 f(t, a)\, da < \infty .$$

---

* S^1 is the circle $0 \leq \theta < 2\pi$, R is the line.

Here $dw = I(\theta)d\theta$ is a probability measure on S^1 <u>with</u> (Lebesgue) <u>density</u> $I(\theta)$, and

$$a^* = a \cos\theta - b \sin\theta$$

$$b^* = a \sin\theta + b \cos\theta$$

Usually the time dependence of f will not be written out explicitly.

McKean adapted a construction first used by Wild [18] for the 3-dimensional Maxwellian gas to express the solution of problem (1) as a weighted sum of "products" of the initial datum f_o with itself.

For that construction, one rewrites the Boltzmann equation in the form

$$\frac{\partial f}{\partial t} = f * f - f \tag{2}$$

where we define

$$(f_1 * f_2)(a) = \iint f_1(a^*) f_2(b^*)\,db\,dw \tag{3}$$

for functions belonging to the class D given by

$$D: \; 0 \leq f, \; \int f = 1, \; \sigma^2(f) = \int f(a)\, a^2\, da < \infty .$$

This product maps $D \times D$ into itself, and for a general $I(\theta)$ is <u>neither commutative nor associative.</u> The first property can be achieved only if one imposes some symmetry on $I(\theta)$. To see this,

one rewrites $f_1 * f_2$ as

$$(f_1 * f_2)(a) = \iint_{a^{*2}+b^{*2}>a^2} \frac{f_1(a^*)f_2(b^*)}{\sqrt{a^{*2} + b^{*2} - a^2}} I(\theta(a^*, b^*))da^*db^* \quad {}^*$$

so that is clear that commutativity is equivalent to $I(\theta(a^*, b^*))$ being a symmetric function of the pair (a^*, b^*). One can see that this amounts to the condition

$$I(\frac{\pi}{4} - \alpha) = I(\frac{\pi}{4} + \alpha) \qquad 0 \leq \alpha < 2\pi.$$

The nonassociativity of the product cannot be remedied by assumptions on $I(\theta)$ and it is an essential ingredient of the problem.**

Without going into the details, see McKean [9], we mention that if the product defined in (3) were associative, one could express the solution of (2) as a Wild's sum

$$(4) \qquad f = \sum_{n=1}^{\infty} e^{-t}(1-e^{-t})^{n-1} f_o * \dots * f_o \quad \text{(n-fold)}$$

In our case, the associativity <u>does not hold</u> and the n-fold products in (4) must be interpreted by putting parentheses between

---

* For a fixed a, $\theta = \theta(a^*, b^*)$ is uniquely determined as the proper rotation needed to go from (a^*, b^*) in R^2 to (a, b) for some b.

** For instance, if $I(\theta) \equiv \frac{1}{2\pi}$, the product given by (3) is commutative but not associative.

the factors in all possible ways and averaging. Each term of Wild's sum is positive and is easy to see that it converges in $L'(R)$ for all $t \geq 0$. One can also see that it gives a solution of (1a). No other solution of (1) exists since Wild's sum can be seen to be the smallest one, and its integral is already 1. One can also check that $\sigma^2(f) = \sigma^2(f_o)$.

This completes our review of the $L'(R)$ situation for (1) and we now turn to the study of the approach to equilibrium.

If the entropy of a probability density function is defined as $H[f] = -\int f \log f$, the H-theorem says that this quantity will increase in time, if f is governed by (1). Gibbs' lemma [9], states that the maximum value of the entropy, computed on the class of all probability density functions with a given second moment, is assumed only by the Gaussian with mean zero.* These two facts have been taken as a proof of the approach of the solutions of (1) to an equilibrium (Gaussian) distribution. See, for instance, Uhlenbeck-Ford [17]. For technical details needed to get a complete proof consult McKean [9] for Kac's model, Carleman [1] for a 3-dimensional gas of hard balls, and Grad [4] for a wide class of cut-off potentials.

A detailed study of this approach to equilibrium for Kac's model is the subject of Chapter I. <u>Our main concern will be to study the non-</u>

---

* For example, $g(a) = \frac{1}{\sqrt{2\pi}} e^{-a^2/2}$ if the second moment is put equal to one. We use g from now on to denote this particular Gaussian density function.

linear evolution given by (1) or (2) "close" to the equilibrium position g and to compare it with the much simpler evolution that one gets by linearizing the equation (1) around this equilibrium position. More explicitly, if we put $f = g + h$ and note that $g * g = g$, we can express (2) as

$$\dot{f} = f * f - f = (g * h + h * g - h) + h * h$$

or equivalently

(5) $$\dot{h} = Ah + h * h$$

A is here the linear operator mapping h into $g * h + h * g - h$, and $h * h$ is a quadratic correction term. The linearized evolution referred to above is given by ignoring $h * h$ in equation (5), thus obtaining

(5') $$h_t = e^{tA} h(0)$$

or

(6) $$f_t = g + h_t$$

This evolution (6) is to be compared, close to $f = g$, with the one given by the actual solution of (2).

Although we have seen that the initial value problem (2) is well posed in $L^1(R)$, we will find it convenient to work in a subspace of $L^1(R)$, namely, $L^2(g^{-1})$.* The fact that this is a Hilbert space will

* $L^2(g^{-1}) = L^2(\sqrt{2\pi}\, e^{a^2/2})$ is the space of all measurable functions $f(a)$ such that $\|f\|^2 = \sqrt{2\pi} \int f^2(a) e^{a^2/2}\, da < \infty$. It is clearly a subspace of $L^1(R)$.

<u>facilitate the comparison between the nonlinear and the linearized flows.</u>

In the next section, we will prove that the problem (1) is also well posed in $L^2(g^{-1})$ and we complete this section with some comments on our choice of this space.

The distance from an arbitrary function in $L^2(g^{-1})$ to the Gaussian $g(a) = \frac{1}{\sqrt{2\pi}} e^{-a^2/2}$ is given by

$$\| f-g \|^2 = \int (f-g)^2(a) g^{-1}(a) da$$

This distance does not make much sense from the point of view of statistical mechanics, see [17]. But we may see that for f close to the Gaussian g it is related in a rather natural way to the entropy of f. Namely, if f satisfies

$$\int f\, da = \int g\, da \quad \text{and} \quad \sigma^2(f) = \sigma^2(g)$$

and we write $f = g + h$, we have

$$\begin{aligned} H[f] &= -\int f \log f = -\int (g+h) \log g(1+\frac{h}{g}) = \\ &= -\int g \log g - \int h \log g - \int (g+h)(-\frac{h}{g} + \frac{1}{2}(\frac{h}{g})^2 - \dots) = \\ &= H[g] + c_1 \int h + c_2 \int h(a) a^2 da + \int h^2 g^{-1} + \text{higher order terms in } h \\ &= H[g] + \frac{1}{2} \int h^2 g^{-1} + \text{"smaller" terms.} \end{aligned}$$

Here we used the fact that both $\int h\, da$ and $\sigma^2(h)$ vanish. Therefore, we

can say that $\| f - g \|^2 = \int h^2 g^{-1}$ gives the first-order correction to the expansion of the entropy of f about the point g.

2. EXISTENCE AND UNIQUENESS FOR THE EVEN FLOW

We consider in this section Kac's model of the Boltzmann equation posed in $L^2(g^{-1})$. In the previous section we noticed that there was a special subset of $L^1(R)$ where the problem was well posed, to wit: positive functions with integral 1 and a fixed variance (1 for instance).

In the same spirit, what we will see now is that Kac's equation is well posed in the subset of $L^2(g^{-1})$ given by those functions with integral 1 and variance 1. It will follow that if we also ask that f be positive, we will still have a well-posed problem.

We state now the theorem to be proved in this section. It constitutes the backbone of all that comes after section 2. For simplicity we assume that $I(\theta) = I(-\theta)$, a natural restriction on physical grounds. In fact, it says that the chance of a collision that takes velocities (a, b) into (a^*, b^*) equals the chance of one with the reversed effect. This feature of the Boltzmann equation was referred to in the introduction as "microscopic reversibility."

THEOREM 1: On the submanifold of $L^2(g^{-1})$ given by

$$\int f(a)da = 1 \qquad \sigma^2(f) = 1$$

the initial value problem

$$\dot{f} = f * f - f$$

is well-posed in a sufficiently small ball C centered about g. This means that there exists only one family of operators Q_t, $t \geq 0$, mapping C into itself and satisfying for each $f_o \in C$

(1) $\quad Q_t f_o \in L^2(g^{-1}),\ \int (Q_t f_o)(a)da \equiv 1,\ \sigma^2(Q_t f_o) \equiv 1,$

(2) $\quad Q_o f_o = f_o,\ Q_{t_1+t_2} f_o = Q_{t_2} Q_{t_1} f_o$ and

$$\text{strong} \lim_{t\to 0} t^{-1}[Q_t f_o - f_o] = f_o * f_o - f_o = Bf_o.$$

Moreover we have

(3) $\quad \| Q_t f_o - g \| \leq \text{constant} \times e^{-vt},$

in which $-v$ is the top (negative) eigenvalue of the linearization of Bf at g,

(4) $\quad \| Q_t f_o^{(1)} - Q_t f_o^{(2)} \| \leq \| f_o^{(1)} - f_o^{(2)} \|$

if $f_o^{(1)}$ and $f_o^{(2)}$ belong to C. The actual proof comes at the end of this section.

To discuss this problem, we make recourse to a formal series expansion of f in terms of the orthogonal basis for the space $L^2(g^{-1})$ given by the Hermite functions

(5) $\quad h_n(a) = e^{-a^2/2} H_n(a) = (-1)^n D^n e^{-a^2/2}, \quad n \geq 0.$

Recall that

$$(6) \qquad \int_{-\infty}^{\infty} h_i(a)h_j(a)\, g^{-1}(a)da = \begin{cases} 0 & \text{for } i \neq j, \\ i!\ \ 2\pi & \text{for } i = j. \end{cases}$$

We express our equation

$$(7) \qquad \dot{f} = f * f - f$$

in terms of the basis h_n $(n \geq 0)$ in a purely formal way at first. Expand $f = f_t$ as

$$(8) \qquad f = \Sigma f_n(t)h_n$$

and assume that it is legitimate to write

$$(9) \qquad \frac{\partial f}{\partial t} = \Sigma \dot{f}_n(t)h_n \ .$$

To get the component-wise version of (7), we use the fundamental relation,* see Kac [7],

$$(10) \qquad (h_i * h_j) = \sqrt{2\pi} \int \cos^i\theta \sin^j\theta\ I(\theta)d\theta\ h_{i+j}$$

satisfied by the Hermite functions in connection with the product defined in (1.3)**.

---

* A proof of this relation and of a generalization of it are found in Appendix I.

** This means equation (3) in section 1, and the same convention is used throughout.

From (9) and (10) we get the desired formula

$$\dot{f}_n = \sqrt{2\pi} \sum_{i=0}^{k} f_i f_{n-i} \int \cos^i \theta \sin^{n-i} \theta \, I(\theta) d\theta - f_n . \tag{11}$$

Observe that because of the assumption on $I(\theta)$ mentioned at the beginning of this section, all integrals of the type

$$\int \cos^{odd} \theta \sin^{odd} \theta \, I(\theta) d\theta \quad \text{and} \quad \int \cos^{even} \theta \sin^{odd} \theta \, I(\theta) d\theta$$

will vanish, and (11) gets simplified into an <u>even system</u>:

$$f_{2n} = \sqrt{2\pi} \sum_{i=0}^{n} f_{2i} f_{2n-2i} \int \cos^{2i} \theta \sin^{2n-2i} \theta \, I(\theta) d\theta - f_{2n} \tag{11a}$$

and an <u>odd system</u>:

$$\dot{f}_{2n+1} = \sqrt{2\pi} \sum_{i=0}^{n} f_{2i+1} f_{2n-2i} \int \cos^{2i+1} \theta \sin^{2n-2i} \theta \, I(\theta) d\theta - f_{2n+1} . \tag{11b}$$

Note that a little more symmetry (esp., $I(\theta) \equiv \frac{1}{2\pi}$) makes all the integrals in the odd system (11a) vanish too. We can therefore summarize the situation for (11a), (11b) as follows:

a) the even components of f evolve by themselves, according to a <u>nonlinear</u> system of the form

$$f_{even} = F(f_{even});$$

b) if we solve that system and insert the resulting even components into the odd system, <u>the odd components evolve linearly according to an equation of form</u>

$$f_{odd} = C(t) f_{odd} .$$

One can prove that all the f_{2n+1} go to zero with t going to $+\infty$; in particular, if we have the extra symmetry mentioned after (11b), we have simply

$$f_{2n+1}(t) = e^{-t} f_{2n+1}(0) . \tag{12}$$

This is the case when $I(\theta) \equiv \frac{1}{2\pi}$ for instance.

The main objective of this chapter is a comparison between the nonlinear flow given by (7) or (11), and its linearization at $f = g = \frac{1}{\sqrt{2\pi}} e^{-a^2/2}$. Therefore, <u>we confine our attention for the rest of this chapter to the even flow given by</u> (11a) <u>and compare it with its linearization at</u> g.

The restrictions, $\int f = 1$ and $\sigma^2(f) \doteq 1$, mentioned at the beginning of this section, are expressed solely in terms of the even components of $f(a)$, namely

$$f_o = \frac{1}{2\pi} \int f(a) e^{-a^2/2} (\sqrt{2\pi}\, e^{a^2/2})\, da = \frac{1}{\sqrt{2\pi}} \;*$$

$$f_2 = \frac{1}{2!\, 2\pi} \int f(a)\, e^{-a^2/2} (a^2 - 1)(\sqrt{2\pi}\; e^{a^2/2})\, da = \frac{1}{4\pi}(\sqrt{2\pi} - \sqrt{2\pi}) = 0$$

---

* One should not confuse this with the initial value of f. This is the only place where confusion may arise.

Because of this we have to concentrate in the even flow given by (11a) on the submanifold described by

(13) $$f_o(t) \equiv \frac{1}{\sqrt{2\pi}}, \quad f_2(t) \equiv 0$$

and we have to prove that this problem is well-posed close enough to $g(a) = \frac{1}{\sqrt{2\pi}} e^{-a^2/2}$.

We first make the linearization explicit by rewriting (11a), subject to (13), as

(14) $$f_{2n} = (\int (\cos^{2n}\theta + \sin^{2n}\theta)I(\theta)d\theta - 1)f_{2n} + \sqrt{2\pi} \sum_{i=2}^{n-2} f_{2i}f_{2n-2i} \int \cos^{2i}\theta \sin^{2n-2i}\theta I(\theta)d\theta \qquad n \geq 2 .$$

Except for the fact that (14) makes only reference to the even components of f with $n \geq 2$, it is clear that it is identical to equation (1.5).* If we use x to denote the vector whose components are the even components of f-g for $n \geq 4$, that is

$$x_o = x_2 = 0; \quad x_{2n+1} \equiv 0; \quad x_{2n} = (f-g)_{2n} \qquad n \geq 2$$

we can reexpress (14) in the succinct form

(14') $$x = Ax + x * x$$

Here A is the diagonal operator acting on the space spanned

by $h_{2n}(n \geq 2)$, by means of the rule

(15) $$Ah_{2n} = (\int(\cos^{2n}\theta + \sin^{2n}\theta)I(\theta)d\theta - 1)h_{2n}$$

$$= \lambda_{2n}h_{2n}, \quad n \geq 2.$$

It is plain that A is selfadjoint with a totally discrete spectrum, ranging from

$$\lambda_4 = \int(\cos^4\theta + \sin^4\theta)I(\theta)d\theta - 1$$

down to -1, which is an accumulation point of the spectrum. For the case $I(\theta) \equiv \frac{1}{2\pi}$ that top negative eigenvalue is $-1/4$. To deal with (11a) subject to (13) rewritten as (14), we need

LEMMA 1: <u>The non-linear part in equation</u> (14) <u>satisfies a Lipschitz condition:</u>

(16) $$\|x * x - y * y\| \leq 2\|x-y\| \max(\|x\|, \|y\|)$$

PROOF: If we could prove

(17) $$\|x * y\| \leq \|x\| \|y\|,$$

then we could get (16) by means of the following string of inequalities

$$\|x * x - y * y\| \leq \|x * (x-y)\| + \|(x-y) * y\| \leq$$

$$\|x-y\| . (\|x\| + \|y\|) \leq 2\|x-y\| \max(\|x\|, \|y\|).$$

The square of the norm of an even vector $x = \sum_{i \geq 2} x_{2i} h_{2i}$ in $L^2(g^{-1})$ is given by

$$\sum_{i \geq 2} x_{2i}^2 (2i)!\ 2\pi$$

and therefore we have from (13)

$$(18) \quad \| x * y \|^2 = \sum_{n=2}^{\infty} (2\pi)(2n)! [\sqrt{2\pi} \sum_{i=2}^{n-2} x_{2i} y_{2n-2i} \int \cos^{2i}\theta \sin^{2n-2i}\theta I(\theta) d\theta]^2$$

$$\leq (2\pi)^2 \sum_{n=2}^{\infty} \left(\sum_{i=2}^{n-2} (2i)! x_{2i}^2 (2n-2i)!\ y_{2n-2i}^2\right) \sum_{j=2}^{n-2} \binom{2n}{2j} [\int \cos^{2j}\theta \sin^{2n-2j}\theta I(\theta) d\theta]^2$$

where we made use of the Schwartz inequality.

Now we notice that

$$\sum_{j=2}^{n-2} \binom{2n}{2j} [\int \cos^{2j}\theta \sin^{2n-2j}\theta\, I(\theta) d\theta]^2 =$$

$$= \sum_{j=2}^{n-2} \binom{2n}{2j} \iint (\cos^{2j}\alpha \sin^{2n-2j}\alpha \cos^{2j}\beta \sin^{2n-2j}\beta) I(\alpha) I(\beta) d\alpha d\beta$$

$$= \iint (\cos\alpha \cos\beta + \sin\alpha \sin\beta)^{2n} I(\alpha) I(\beta) d\alpha d\beta$$

$$= \iint \cos^{2n}(\alpha - \beta) I(\alpha) I(\beta) d\alpha d\beta \leq 1$$

Observe that we made use of the property $I(\theta) = I(-\theta)$ in the third line, to pull the sum inside of the integral. If we apply this to our previous inequality (18), we get

$$\|x*y\| \leq (2\pi)^2 \sum_{n=2}^{\infty} \sum_{i=2}^{n-2} (2i)!\, x_{2i}^2\, (2n-2i)!\, y_{2n-2i}^2 = \|x\|^2 \|y\|^2.$$

This completes the proof of (16).

We are now in a position to prove that (14) is well-posed. Recall that for $I(\theta) \equiv \frac{1}{2\pi}$ the top eigenvalue of A is $-1/4$. We consider this case only, but it should be plain that the proof has to be trivially modified to deal with a different $I(\theta)$.

LEMMA 2: <u>There exists a unique solution to equation</u> (14) <u>if</u> $\|x(o)\|$ <u>is small enough. Moreover</u> $\|x(t)\| \leq Ce^{-t/4}$.

PROOF: Equation (14) is replaced by the integral equation

$$(19) \qquad x(t) = e^{tA}\, x(0) + \int_0^t e^{(t-s)A}\, (x(s) * x(s))\, ds.$$

We are interested in a fixed point of this map because any solution of (19) is such a point, so we must first find a domain mapped into itself by H.

Take E to be the set of those continuous functions $x(t)$ from $[0, \infty)$ into the subspace of $L^2(g^{-1})$ spanned by h_{2n} $(n \geq 2)$, such that $\|x(t)\| \leq \gamma \exp(-t/4)$ for all positive t and $\|x(0)\| \leq \gamma^2$, with a constant γ to be specified later on.

Recalling that the spectrum of A lies to the left of $-1/4$

and using the proof of the previous lemma, we get

$$\|Hx(t)\| \leq e^{-t/4}\|x(0)\| + \int_o^t e^{-(t-s)/4}\|x(s)\|^2\, ds \leq$$

$$\leq e^{-t/4}\|x(0)\| + \gamma^2 \int_o^t e^{-(t-s)/4}\, e^{-s/2}\, ds \leq$$

$$\leq e^{-t/4}\|x(0)\| + 4\gamma^2 e^{-t/4} = e^{-t/4}(\|x(0)\| + 4\gamma^2),$$

so that $\|Hx(t)\| \leq 5\gamma^2 \exp(-t/4)$, and this is smaller than $\gamma \exp(-t/4)$ if $\gamma \leq 1/10$. For any such γ, the corresponding class E is sent into itself by the map H. Note that $Hx(0) = x(0)$.

We now construct a series which formally solves the fixed-point problem $Hx=x$. The task is then to prove its convergence.

Take

$$x^{(1)}(t) = e^{tA}x(0)$$

$$x^{(i)} = Hx^{(i-1)}$$

and define

$$x = \sum_{i=1}^{\infty} [x^{(i+1)} - x^{(i)}] + x^{(1)}. \tag{20}$$

Observe that, for $\gamma \leq 1/10$, each $x^{(i)}$ lies in E whenever $\|x(0)\| \leq \gamma^2$. Define

$$y_i(t) = \max\left(\|x^{(i)}(t)\|, \|x^{(i-1)}(t)\|\right)$$

$$z_i(t) = \max_{t \leq s} \|x^{(i)}(s) - x^{(i-1)}(s)\|.$$

We know that

$$y_i(s) \leq \gamma e^{-s/4}$$

$$z_1(t) = \max_{t \leq s} \|x^{(1)}(t)\| \leq e^{-t/4}\|x(0)\|$$

and

$$z_{i+1}(t) = \max_{t \leq s} \|Hx^i(s) - x^i(s)\| =$$

$$\max_{t \leq s} \Big\| e^{sA}[x^{(i-1)}(0) - x^{(i-2)}(0)] + \int_0^s e^{(s-\alpha)A}(x^{(i-1)}(\alpha)*x^{(i-1)}(\alpha) -$$

$$-x^{(i-2)}(\alpha) * x^{(i-2)}(\alpha))d\alpha \Big\|$$

$$\leq \max_{t \leq s} \Big\|\int_0^s e^{(s-\alpha)A}(x^{(i-1)}(\alpha)*x^{(i-1)}(\alpha) - x^{(i-2)}(\alpha)*x^{(i-2)}(\alpha))d\alpha \Big\|$$

$$\leq 2 \max_{t \leq s} \int_0^s e^{-(s-\alpha)/4} z_i(\alpha)\, y_i(\alpha)\, d\alpha.$$

In the last line we used lemma 1 of this section.

Now we make the inductive hypothesis that there exists a

number q such that

$$z_i(s) \leq \|x(0)\| \, q^{i-1} \, e^{-s/4}.$$

This is true for $i = 1$, and by induction you get

$$z_{i+1}(t) \leq 2\|x(0)\| \, q^{i-1} \gamma \max_{t \leq s} \int_0^s e^{-(s-\alpha)/4} \, e^{-\alpha/2} \, d\alpha \leq$$

$$\leq 8\|x(0)\| \, q^{i-1} \, \gamma \, e^{-t/4},$$

so that if we start with $q = 8\gamma$ (which we know is smaller than one), we would have

$$z_{i+1}(t) \leq \|x(0)\| \, q^i \, e^{-t/4}$$

proving the inductive hypothesis and making our series (20) absolutely convergent:

$$\|x\| \leq \Sigma \|x^{(i+1)} - x^{(i)}\| + \|x^{(1)}\| \leq \|x(0)\| \, e^{-t/4} \sum_{i=0}^{\infty} q^i < \infty.$$

In the same vein, one goes on to prove that the series gives a solution of the original differential equation (14) and that this solution is unique. This ends the proof of lemma 2.

Putting together lemmas 1 and 2 with the expression (12), we have a complete proof of the main theorem of this section for the case

$I(\theta) \equiv 1/2\,\pi$. Lemma 1 is the essential ingredient for part (4) of the theorem.

To prove parts (1) and (2) of that theorem for a non-constant $I(\theta)$, one has to deal also with the <u>linear</u> odd system (11b).

3. THE LINEARIZATION PROBLEM

Suppose we have an equation of the type

$$\dot{x} = Ax + F(x) \tag{1a}$$

where A is linear and F analytic, with neither constant nor linear terms. Very close to $x = 0$, the two vector fields Ax and $Ax + F(x)$ look very much the same. A natural question is how is this reflected in the flows corresponding to (1a) and

$$\dot{x} = Ax. \tag{1b}$$

When x lies in R^n or C^n this question was answered first by Poincaré [12]. He proved that if $\lambda_1, \lambda_2, \ldots, \lambda_n$ are eigenvalues of A and if

a) $\lambda_i \neq \Sigma m_j \lambda_j$ for any positive integers m_j with $\Sigma m_j > 1$,

and

b) all the λ_j lie on one side of a line passing through the origin,

then there exists an analytic map ψ which is locally invertible and linearizes the flow near the origin. This means that if Q_t denotes the map sending the initial data x into the solution $x(t)$ at time t,

for (1a) and if T_t denotes the corresponding map for (1b), then ψ intertwines Q_t and T_t in the sense that

$$(2) \qquad Q_t = \psi^{-1} \circ T_t \circ \psi$$

near the origin.

If we want such a change of coordinates to exist for A irrespective of the F, then conditions a) and b) are probably necessary, as an example by Hartman [5] seems to indicate. Much less, nearly nothing in fact, is required if we are contented with a smooth, but not analytic ψ. See Hartman [5].

In Section 5 we will present such a change of coordinates ψ for the even part of Kac's equation, see (2.11a). Of course the space is now infinite-dimensional. In general, this is bound to produce difficulties. They can be overcome in our case because A has a (negative) pure point spectrum which accumulates at the point $-\int I(\theta)d\theta = -1$. See the remarks after (2.14').

It comes as an unexpected bonus that ψ can be expressed as

$$(3) \qquad \psi = \operatorname*{strong\,lim}_{t \to +\infty} T_{-t} Q_t .$$

A nice and trivial consequence of this fact is that for each small

enough x, there exists a vector $\psi(x)$ so that

$$T_{-t}(T_t\psi(x) - Q_t x)$$

goes to zero as $t \to \infty$. Clearly ψ is characterized by this property. A brief discussion of the relations between (2) and (3) in the case of R^2 will clarify the situation.

For simplicity, let us consider only the case of a symmetric operator A. It is simple to prove that if the limit (3) exists and has an inverse, then the ψ so defined intertwines T_t and Q_t according to (2). On the other hand, simple examples show that not all intertwining maps ψ can be computed by means of (3). Now we will see where the trouble lies and indicate a way to modify (3) to get an intertwining map in general.

First we express Poincare's condition a) in terms of T_t. Take A to be

$$A = \begin{pmatrix} \alpha & o \\ o & \beta \end{pmatrix}$$

and consider the formal power series

$$W(x, y) = (\Sigma a_{ij} x^i y^j, \ \Sigma b_{ij} x^i y^j) = (x, y) + \text{higher order terms.}$$

Then we have

$$T_{-t}WT_t(x,y) = (\Sigma a_{ij}x^iy^je^{(i\alpha+j\beta-\alpha)t},\ \Sigma b_{ij}x^iy^je^{(i\alpha+j\beta-\beta)t})$$

and it is now plain that condition a) is precisely what is needed for an arbitrary W of the above form to be able to split $T_{-t}WT_t$ into I plus a couple of formal power series

$$T_{-t}WT_t = I + W_+(t) + W_-(t)$$

where

(4a) each term of $W_+(t)$ converges to 0 as $t \to +\infty$.

(4b) each term of $W_-(t)$ converges to 0 as $t \to -\infty$.

To understand why (2) will not imply (3) in general, suppose that A satisfies Poincaré's conditions so that there exists an analytic map ψ satisfying (2). Because of the previous comments, we can write

(5) $$T_{-t}Q_t = T_{-t}\psi^{-1}T_t\psi = (I + \psi_+^{-1}(t) + \psi_-^{-1}(t))\,\psi =$$

$$= \psi + \psi_+^{-1}(t)\,\psi + \psi_-^{-1}(t)\,\psi.$$

Clearly (3) will exist only when $\psi_-^{-1}(t)\,\psi \equiv 0$.

Expression (5) also suggests how to get an intertwining map for ψ for the pair T_t, Q_t. One has to split the formal power series

expansion of $T_{-t}Q_t$ into three pieces

(6) $$T_{-t}Q_t = \psi + \psi_1(t) + \psi_2(t)$$

where $\psi_1(t)$ and $\psi_2(t)$ satisfy (4a) and (4b) respectively. <u>To check formally that the</u> ψ <u>so constructed does the trick</u> (2) <u>is straight forward</u>. Thus the real problem is to check convergence of this formal series for ψ.

4. EIGENVALUE INEQUALITIES

In this section we prove an elementary inequality which surprisingly contains the core of the future development. From it we derive important inequalities concerning the eigenvalues of the linear transformation A defined in (2.15).

LEMMA: If $x_1^2 + x_2^2 = 1$, then for every $n, m \geq 2$, we have

$$(-1 + \sum_{i=1}^{2} x_i^{2n+2m}) > (-1 + \sum_{i=1}^{2} x_i^{2n}) + (-1 + \sum_{i=1}^{2} x_i^{2m}). \; * \tag{1}$$

PROOF: For a fixed (x_i) satisfying the assumptions, consider the function

$$h(n) = 1 - \sum_{i=1}^{2} x_i^{2n}.$$

Our aim is to prove that $h(n)$ is subadditive. For that purpose we look at the derivative (with respect to n) of the function $n^{-1}h(n)$:

$$(n^{-1}h(n))' = n^{-2}[\sum_{i=1}^{2} x_i^{2n} - 1 - \sum_{i=1}^{2} x_i^{2n} \log x_i^{2n}] \equiv n^{-2} r(n).$$

One checks that $r(n)$ is a decreasing function of n, so that $n^{-1}h(n)$ will be increasing up to some value of n and decreasing from there on.

---

* We disregard the trivial case when one of the $x_i = 0$. The proof is not the simplest one, and is made so as to extend naturally to the case when there are more than two x_i's; see appendix II.

Assume for a moment that we could prove

$$\frac{h(2)}{2} > \frac{h(4)}{4} . \tag{2}$$

From here and the previous comments, it is clear that $n^{-1}h(n)$ would be decreasing at least from $n = 3$ on. Therefore for $n, m \geq 2$ we would have

$$h(n+m) = \frac{n}{n+m} h(n+m) + \frac{m}{n+m} h(n+m) < n \frac{h(n)}{n} + m \frac{h(m)}{m} = h(n) + h(m)$$

and the lemma would be proved.

Thus we have only to show that (2) holds. This is equivalent to showing that

$$2(1 - x^2 - (1 - x)^2) > 1 - x^4 - (1 - x)^4 \qquad \text{for } 0 < x < 1$$

and this is trivially verified.

Now we can prove the

THEOREM 2: Irrespective of the angular scattering density $I(\theta)$, the eigenvalues of the linear operator A defined in (2.15) satisfy

$$\lambda_{2n} + \lambda_{2m} < \lambda_{2n+2m}. \tag{3}$$

PROOF: Recall that

$$\lambda_{2n} = \int(\cos^{2n}\theta + \sin^{2n}\theta - 1)I(\theta)d\theta,$$

and then notice that except for $\theta = 0, \frac{\pi}{2}, \pi, 3/2\,\pi$ we have strict inequality for the integrands involved in (3). This proves the theorem.

We can say a little more than (3) if we notice that when n grows to $+\infty$, λ_{2n} tends to -1; namely, <u>there exists a</u> $c > 0$ <u>so that</u>

(4) $$\lambda_{2n+2m} - (\lambda_{2n} + \lambda_{2m}) \geq c > 0.$$

<u>independently</u> of $n, m \geq 2$.

5. INTERTWINING OPERATOR

Most of the work up to this point was of a preparatory nature. In section 2, we singled out the even system (11a) describing the nonlinear piece of equation (2.11); it is of the type

(1) $$\dot{x} = Ax + x * x$$

Here x is a vector belonging to the subspace of $L^2(g^{-1})$ spanned by $h_{2n}: n \geq 2$, A is explicitely given in (2.15), and the $*$ product is defined in (1.3).

The point $x = 0$ is critical for (1), and we know from section 2 that solutions exist and are unique nearby. Thus, we can speak of the semigroup Q_t, $t \geq 0$, relating initial data $x = x(0)$ to the solution $x(t)$ at time t, at least if $\|x\|$ is small enough.

A much simpler evolution is obtained by ignoring the nonlinear part $x * x$ in (1):

(2) $$x = Ax.$$

The solution of this problem is given by $T_t = e^{tA}$ acting upon $x = x(0)$. The exponential makes good sense, because A is self-adjoint and negative definite.

The purpose of this section is to prove

THEOREM 3: *If* $\|x\|$ *is small enough, the limit*

$$\psi(x) = \underset{t \to +\infty}{\text{strong lim}}\, T_{-t} Q_t x \tag{3}$$

exists and belongs to the $L^2(g^{-1})$ *span of* h_{2n} $(n \geq 2)$. *Moreover, the map* $x \to \psi(x)$ *is analytic close to* $x = 0$ *with an analytic inverse, and*

$$Q_t = \psi^{-1} T_t \psi. \tag{4}$$

We first introduce some auxiliary material and prove two lemmas.

Given operators K and L acting upon x, define $K * L$ to be the operator

$$K * L : \; x \to (Kx) * (Lx).$$

then (1) and (2) become

$$\dot{Q}_t x = A Q_t x + (Q_t * Q_t) x, \tag{5}$$

$$\dot{T}_t x = A T_t x. \tag{6}$$

Now

$$S_t = T_{-t} Q_t \tag{7}$$

can be easily shown to satisfy the equation

(8) $$\dot{S}_t = e^{-tA}(e^{tA}S_t * e^{tA}S_t), \quad S_o = I.$$

This is equivalent to the integral equation

(8') $$S_t = I + \int_o^t e^{-sA}(e^{sA}S_s * e^{sA}S_s)\,ds,$$

which we use to express S_t as a <u>sum of contributions of different numbers of factors</u>. Namely, we put

(9) $$S_t = \sum_{n=1}^{\infty} R_t^{(n)}$$

where the first three terms are

$$R_t^{(1)} = I,$$

$$R_t^{(2)} = \int_o^t e^{-sA}(e^{sA} * e^{sA})\,ds,$$

$$R_t^{(3)} = \int_o^t e^{-sA}(e^{sA} * e^{sA}\int_o^s e^{-\xi A}(e^{\xi A} * e^{\xi A})d\xi)\,ds +$$

$$+ \int_o^t e^{-sA}(e^{sA}\int_o^s e^{-\xi A}(e^{\xi A} * e^{\xi A})\,d\xi * e^{sA})\,ds\,.$$

Now by induction, if we have defined $R_t^{(i)}$ for all $i < n$, we construct $R_t^{(n)}$ in the following way:

a) split n in all possible ways as $j + (n-j); j = 1, \ldots, n-1$,

b) define $R_t^{(n)}$ as the sum of all possible pieces of the form

$$(10) \qquad \int_0^t e^{-sA}(e^{sA}R_s^{(j)} * e^{sA}R_s^{(n-j)})ds.$$

Having defined $R_t^{(n)}$, $n \geq 1$, in this way we can now check <u>formally</u> that S_t defined by (9) solves (8). In fact

$$e^{-tA}(e^{tA}S_t * e^{tA}S_t) = \sum_{n=1}^{\infty} (\sum_{j=1}^{n-1} e^{-tA}R^j * e^{tA}R^{n-j})) = \dot{S}_t,$$

and clearly

$$S_o = I.$$

To make use (and sense) of the expression

$$S_t = \sum_{n=1}^{\infty} R_t^{(n)},$$

we need to prove that this formal series actually converges in $L^2(g^{-1})$, and therefore gives a bona-fide solution of (7).

For that end we prove

LEMMA 1: <u>There exists a constant</u> $c > 0$ <u>such that for all positive</u> t <u>and all</u> x, y <u>in the</u> $L^2(g^{-1})$ <u>span of</u> $h_{2n} (n \geq 2)$

$$(11) \qquad \| e^{-sA}(e^{sA}x * e^{sA}y\| \leq e^{-cs} \|x\| \, \|y\|.$$

PROOF: By (2.15)

$$e^{sA}h_{2i} * e^{sA}h_{2j} = e^{(\lambda_{2i} + \lambda_{2j})s} h_{2i} * h_{2j} =$$

$$= \sqrt{2\pi} \int \cos^{2i}\theta \sin^{2j}\theta \, I(\theta) d\theta \; e^{(\lambda_{2i} + \lambda_{2j})s} \, h_{2i+2j}.$$

Thus

$$(12) \quad e^{-sA}(e^{sA} h_{2i} * e^{sA} h_{2j}) = 2\pi \int \cos^{2i}\theta \sin^{2j}\theta \, I(\theta) d\theta \, e^{(\lambda_{2i} + \lambda_{2j} - \lambda_{2i+2j})s} h_{2i+2j}.$$

If we now put

$$x = \sum_{i \geq 2} \alpha_{2i} h_{2i}, \qquad y = \sum_{i \geq 2} \beta_{2i} h_{2i},$$

and use (12), we get

$$e^{-sA}(e^{sA} x * e^{sA} y) =$$

$$= \sqrt{2\pi} \sum_{n=2}^{\infty} h_{2n} \sum_{i=2}^{n-2} \alpha_{2i} \beta_{2n-2i} \int \cos^{2i}\theta \sin^{2j}\theta \, I(\theta) d\theta \, e^{(\lambda_{2i} + \lambda_{2n-2i} - \lambda_{2n})s}.$$

Bring in now the inequalities (4.4) to get

$$-\lambda_{2n} + (\lambda_{2i} + \lambda_{2n-2i}) \leq -c.$$

Now this is used to get the bound

$$\| e^{-sA}(e^{sA} x * e^{sA} y \|^2 \leq$$

$$\leq 2\pi \sum_{n=2}^{\infty} (2n)! \left[\sum_{i=2}^{n-2} \sqrt{2\pi} \, \alpha_{2i} \beta_{2n-2i} \int \cos^{2i}\theta \sin^{2n-2i}\theta \, I(\theta) d\theta \right]^2 e^{-2cs} \leq$$

$$\leq e^{-2cs} \|x\|^2 \|y\|^2.$$

In the last line we used the same way of estimating as in (2.18). The lemma is proved.

<u>We are now in a position to prove the crucial</u>

LEMMA 2:

$$\| R_t^{(n)} x \| \le \left(\frac{1-e^{ct}}{c}\right)^{n-1} \|x\|^n \tag{13}$$

PROOF: We get the estimate (13) by induction on n. To begin with, $R_t^{(1)} = I$, so (13) is trivial for $n=1$. Now use (11) and the inductive hypothesis to get from the definition of $R_t^{(n)}$

$$\| R_t^{(n)} x \| = \left\| \sum_{j=1}^{n-1} \int_0^t e^{sA} (e^{-sA} R_s^{(j)} x * e^{-sA} R_s^{(n-j)} x) ds \right\| \le \tag{14}$$

$$\le \sum_{j=1}^{n-1} \int_0^t \| \qquad \| ds \le \sum_{j=1}^{n-1} \int_0^t \left(\frac{1-e^{-cs}}{c}\right)^{n-2} e^{-cs} ds \, \|x\|^n =$$

$$= \left(\frac{1-e^{-ct}}{c}\right)^{n-1} \|x\|^n .$$

The lemma is proved.

Finally, we can use this lemma to estimate the formal sum (9) as follows: if $\|x\| < c$, we have for all $t \ge 0$

$$\| S_t x \| \le \sum_1^{\infty} \| R_t^n x \| \le \sum_1^{\infty} \left(\frac{1-e^{-ct}}{c}\right)^{n-1} \|x\|^n < \infty ,$$

and therefore $S_t x$ <u>is well defined as the sum of an absolutely convergent series.</u> Now we can complete the proof of Theorem 3.

To get the existence of the limit (3), we estimate

$$(15) \qquad \left\| \sum_{n=1}^{\infty} R_{t_2}^{(n)} x - \sum_{n=1}^{\infty} R_{t_1}^{(n)} x \right\|$$

as in (14). The result is that for $t_1 \leq t_2 \leq \infty$, (15) is bounded by

$$\sum_{n=1}^{\infty} (n-1) \int_{t_1}^{t_2} e^{-cs} \left(\frac{1-e^{-cs}}{c}\right)^{n-2} ds \, \|x\|^n =$$

$$= \sum_{n=1}^{\infty} \left(\frac{1-e^{-ct}}{c}\right) \Bigg|_{t=t_1}^{t=t_2} \|x\|^n \leq$$

$$\leq (e^{-ct_1} - e^{-ct_2}) \sum_{n=1}^{\infty} (n-1) \frac{\|x\|^n}{c^{n-1}}$$

in the last step we used the fact that

$$\beta^n - \alpha^n \leq n(\beta - \alpha) \quad \text{if} \quad 0 \leq \alpha \leq \beta \leq 1, \quad n \geq 1.$$

Closely related estimates show that $\psi(x)$ is an analytic function of x, if $\|x\| < c$. This stems from the fact that ψ is expressed as an absolutely convergent power series:

$$(16) \quad \psi(x) = x + \sum_{n \geq 2} \sum_{i=2}^{n-2} x_{2i} x_{2n-2i} \int_0^{\infty} e^{-sA} (e^{sA} h_{2i} * e^{sA} h_{2n-2i}) ds + \ldots\ldots$$

Notice that

$$T_t T_{-(t+s)} Q_{(t+s)} = T_{-s} Q_{t+s} = T_{-s} Q_s Q_t,$$

so that if we keep t fixed, while $s \to \infty$, we get

$$T_t \psi = \psi Q_t. \tag{17}$$

(4) would be immediate from (17) if $\psi(x)$ were invertible. The gradient of $\psi(x)$ at $x = 0$ is the identity map, so the inverse function theorem for analytic functions, see Dieudonné [2], guarantees the existence and analyticity of ψ^{-1} close to $x = 0$. The theorem is proved.

We close this section by observing that the knowledge of ψ comes close to determining the scattering density I. If we perform the indicated integrals in (16), we will get in the denominators of the quadratic part all possible combinations of the type

$$\lambda_{2n} - (\lambda_{2i} + \lambda_{2n-2i}).$$

This is actually enough to determine the spectrum λ_{2i} because we have

$$\lambda_{2n} \to -1 \quad \text{if} \quad n \to \infty.$$

If we recall that

$$\lambda_{2i} = \int (\cos^{2i}\theta + \sin^{2i}\theta - 1) I(\theta) d\theta$$

we see that to find I we have to deal with a classical moment problem; especially, the quadratic part of ψ suffices to determine

$$I(\theta) + I(\pi - \theta) \qquad 0 \leq \theta < \pi. \tag{18}$$

It is clear that the even flow cannot give us any information beyond (18). If we want to recover the whole of I, we have to deal also with the odd eigenvalues λ_{2n+1}:

$$\lambda_{2n+1} = \int (\cos^{2n+1}\theta + \sin^{2n+1}\theta - 1) I(\theta) d\theta.$$

The reader will easily see that the set of all eigenvalues determines the scattering density I unequivocally.

6. REMARKS ON MORE GENERAL CASES

We want to put an end to this chapter by indicating briefly how we could handle two different problems along the same lines.

The first problem might be of physical relevance. The second one is motivated by the hard-sphere model of the Boltzmann equation and poses serious mathematical difficulties not resolved here.

For the first problem, consider a gas undergoing multiple (instead of only binary) collisions. At the end of an exponentially distributed random time, m particles collide together with probability k_m, with $\sum_{m \geq 2}^{n} k_m = 1$.

Consider the vector $(a_1, \ldots, a_m)$ comprising the velocities of the m particles. For simplicity we consider only one-dimensional velocities. The effect of a collision of m particles is a proper m-dimensional rotation

$$0:(a_1, \ldots, a_m) \rightarrow (a_1^*, \ldots, a_m^*).$$

If we now distribute 0 according to the law $I_m(0)d0$, we get a higher-order Boltzmann's equation for the density $f(a_1)$ of the number of particles having velocity a_1

$$\frac{\partial f}{\partial t} = \sum_{m=2}^{n} k_m f^{(m)} - f.$$

Here

$$f^{(m)}(a_1) = \int_{R^{m-1}} \int_{SO(m)} f(a_1^*)\, f(a_2^*) \ldots f(a_m^*)\, I_m(0) d0 da_2 \ldots da_m .$$

The reader should compare this with (1. 1a).

If each of the densities I_m has some symmetry properties (compare with the assumption $I(\theta) = I(-\theta)$ in section 2), we get the same splitting that we had for (2. 11) into an even and odd system. For the case of section 2, we had the numbers

$$\lambda_{2i} = \int (\cos^{2i}\theta + \sin^{2i}\theta - 1)\, I(\theta) d\theta$$

as the eigenvalues for the generator of the linearized flow. The Hermite functions were the eigenfunctions. <u>They are still eigenfunctions for the corresponding linear generator in this case, and the eigenvalues are changed to</u>

(1) $$\lambda_{2i} = \sum_{m=2}^{n} k_m \int_{x_1^2 + \ldots + x_m^2 = 1} (x_1^{2i} + \ldots + x_m^{2i} - 1) I_m(0) d0$$

<u>where</u> $x_i = O_{ii}$ $i = 1, \ldots, m$. This assertion comes from the formula

(2) $$\int \ldots \int h_{i_1}(a_1^*) \ldots h_{i_m}(a_m^*)\, I_m(0) d0\, da_2, \ldots, da_m =$$

$$= \text{constant} \times h_{i_1 + \ldots + i_m}(a_1)$$

which follows from (5) in Appendix I. The restriction

$$x_1^2 + \dots + x_m^2 = 1$$

implies, for all $i, j \geq 2$

(3) $$(-1 + x_1^{2i+2j} + \dots + x_m^{2i+2j}) \geq$$

$$\geq (-1 + x_1^{2i} + \dots + x_m^{2i}) +$$

$$+ (-1 + x_1^{2j} + \dots + x_m^{2j}).$$

For $m = 2$, this is the lemma proved in section 4, the general proof will be found in Appendix II. Using (3) one gets

$$\lambda_{2i} + \lambda_{2j} < \lambda_{2i+2j}$$

for the λ's given by (1), and from here

(4) $$\lambda_{\Sigma 2n_i} > \Sigma \lambda_{2n_i}.$$

Using these inequalities (4), plus the fact that the spectrum of the generator of the linearized flow accumulates at $-\sum_{m=2}^{n} k_m = -1$, one can get a proof of the existence of

$$\psi = \operatorname*{strong\,lim}_{t \to \infty} T_{-t} Q_t$$

in the same fashion as we did in section 5.

This finishes our consideration of the first problem, and we look now at a different situation.

Consider a problem of the form

$$x = Ax + x * x \tag{5}$$

Here x belongs to some Hilbert space and $x * y$ is a bilinear product from $H \times H$ into H. Assume finally that existence and uniqueness can be proved for equation (5), so that it makes sense to speak of the semi-group of operators Q_t satisfying

$$Q_t = AQ_t + Q_t * Q_t .$$

We take up again the problem of comparing $T_t = e^{tA}$ with Q_t. We are mainly interested in finding conditions on A and the $*$ product guaranteeing the existence of

$$\psi = \underset{t \to \infty}{\text{strong lim}}\, T_{-t} Q_t .$$

In the previous sections we have seen how to deal with this problem in the case when A has a purely discrete spectrum (λ_n), with corresponding eigenfunctions f_n: we have to express $f_i * f_j$ in terms of the

set f_n, and if $f_i * f_j$ has a non-zero component in the n-th direction, then we have to be sure that

$$\lambda_n - (\lambda_i + \lambda_j) > c > 0$$

for some c independent of i, j, n.

In the same fashion, <u>if</u> A <u>has the spectral resolution</u> $\int \lambda dE_\lambda$, <u>and if we can find a constant</u> $c > 0$ <u>so that for every</u> λ, μ

$$(6) \qquad (I - E_\lambda) H * (I - E_\mu) H \subset (I - E_{\lambda+\mu+c}) H,$$

<u>then the limit</u>

$$\psi = \text{strong} \lim_{t \to \infty} T_{-t} Q_t$$

<u>exists</u>.

We feel that one is perfectly justified in considering (6) as a rather useless condition. Much more ingenuity is clearly needed to find handier conditions to tackle the problem of the existence of the limit (5) in actual cases such as hard spheres.

APPENDIX I: PRODUCTS OF HERMITE FUNCTIONS

Here we prove a property of the Hermite functions due to Kac [7]. Actually, we generalize the relation he found. Such a generalization was used in section 6. Recall that the Hermite polynomials are defined as

$$(1) \qquad H_n(a) = (-1)^n e^{a^2/2} D^n e^{-a^2/2} \qquad n = 0, 1, 2, \ldots .$$

and that this set is complete and orthogonal in $L^2(g)$ with

$$(2) \qquad \int_{-\infty}^{\infty} e^{-a^2/2} H_i(a) H_j(a)\, da = \begin{cases} 0 & i \neq j \\ i!\sqrt{2\pi} & i = j \end{cases}$$

LEMMA:* <u>If</u> $x_1^2 +, \ldots, +x_n^2 = 1$, <u>then</u>

$$(3) \qquad H_m(x_1 a_1 + \ldots + x_n a_n) =$$

$$= \sum_{\substack{m_1+\ldots+m_n \\ = m}} \binom{m}{m_1 \ldots m_n} x_1^{m_1} \ldots x_n^{m_n} H_{m_1}(a_1) \ldots H_{m_n}(a_n).$$

PROOF: It is convenient to think of the n-tuple $(a_1, \ldots, a_n)$ as a vector in R^n denoted by a. Let O be an orthogonal matrix having the numbers $x_1, \ldots, x_n$ as its first row. We define a map in R^n by means of $a \to Oa$. Setting $b = Oa$, we observe that

---

* This result is well known, but our proof is elementary.

$$\frac{\partial}{\partial b_1} = \Sigma x_i \frac{\partial}{\partial a_i}$$

and that

$$\Sigma a_i^2 = \Sigma b_i^2 .$$

We know that

(4) $$\left(x_1 \frac{\partial}{\partial a_1} + \ldots + x_n \frac{\partial}{\partial a_n}\right)^m =$$

$$= \sum_{\substack{m_1+\ldots+m_n \\ =m}} \binom{m}{m_1 \ldots m_n} x_1^{m_1} \ldots x_n^{m_n} \frac{\partial^{m_1}}{\partial a_1^{m_1}} \cdots \frac{\partial^{m_n}}{\partial a_n^{m_n}} .$$

Therefore, if we apply the right-hand operator in (4) to the function $\exp(-1/2\,|a|^2)$ and then multiply the result by $(-1)^m \exp(1/2\,|a|^2)$, we get the right-hand side of the equality (3).

But now the rest is clear, because the left-hand operator in (4) is $\frac{\partial^m}{\partial b_1^m}$. Thus, if we apply it to

$$\exp(-1/2\,|b|^2) = \exp(-1/2\,|a|^2)$$

and then multiply the result by $(-1)^m \exp(1/2\,|b|^2)$, we get $H_m(b_1)$

on the left-hand side of (3). Finally

$$H_m(b_1) = H_m(x_1 a_1 + \dots + x_n a_n)$$

and the lemma is proved.

Now we are in a position to prove

THEOREM 4: If $O = (x_{ij})$ <u>is an orthogonal matrix and if</u> $a \in R^n$, <u>then</u>

$$(5) \qquad \int \dots \int H_{i_1}((Oa)_1) \dots H_{i_n}((Oa)_n)\, e^{-(a_1^2 + \dots + a_n^2)/2}\, da_2 \dots da_n =$$

$$= (2\pi)^{\frac{n-1}{2}}\, x_{11}^{i_1} \dots x_{n1}^{i_n}\; H_{i_1 + \dots + i_n}(a_1)\, e^{-a_1^2/2}.$$

This result is due to Kac [7] for $n = 2$.

PROOF: Except for the factor $e^{-a_1^2/2}$ the integral in (5) is a polynomial in a_1 alone, and can therefore be written as

$$\Sigma \alpha_k H_k(a_1)\, e^{-a_1^2/2}.$$

We integrate this against $H_m(a_1)da_1$ to pick out α_m. In the left side of (5) we change a into $O^{-1}a = O^*a$ so as to have

$$(6) \qquad \int \dots \int H_{i_1}(a_1) \dots H_{i_n}(a_n)\, H_m((O^*a)_1)\, e^{-(a_1^2 + \dots + a_n^2)/2}\, da_1 \dots da_n = \alpha_m.$$

Now expand $H_m((O^*a)_1)$ using the previous lemma, converting the left side of (6) into a sum of several integrals. Each one of these is an n-fold integral which splits as a product of one-variable integrals.

Using the orthogonality relations (2), we observe that most of these integrals vanish; the only one which gives a contribution is the one coming from the additive decomposition of m as $m = i_1 + \ldots + i_n$. This surviving integral can be easily performed, and the stated formula drops out.

For instance if $n = 2$, which is Kac's case, we get

$$(7) \qquad \int H_i(a^*)\, H_j(b^*)\, e^{-\frac{(a^2+b^2)}{2}}\, db = \sqrt{2\pi}\, \cos^i\theta\, \sin^j\theta\, H_{i+j}(a)\, e^{-a^2/2}.$$

in the notation of sections 1 and 2.

If we integrate against $I(\theta)d\theta$ and use

$$h_m(a) = e^{-a^2/2}\, H_m(a)$$

we find

$$(8) \qquad (h_i * h_j)\,(a) = \sqrt{2\pi}\, \int \cos^i\theta\, \sin^j\theta\, I(\theta)d\theta\, h_{i+j}(a)$$

which coincides with (2.10). Formula (6.2) is derived in the same way.

APPENDIX II: AN ELEMENTARY INEQUALITY

Here we extend the lemma proved in section 4.

LEMMA: If $x_1^2 + \ldots + x_k^2 = 1$ and $n, m \geq 2$, then

$$(1) \qquad \left(-1 + \sum_{i=1}^{k} x_i^{2n+2m}\right) \geq \left(-1 + \sum_{i=1}^{k} x_i^{2n}\right) + \left(-1 + \sum_{i=1}^{k} x_i^{2m}\right).$$

PROOF: A look at the proof presented in section 4 will convince the reader that it is enough to check relation (4.2) i.e. $2h(2) > h(4)$, where $h(n)$ is given in this case by

$$h(n) = 1 - \sum_{i=1}^{k} x_i^{2n}.$$

We will actually prove that, at least for $k \geq 3$,

$$(2) \qquad \frac{h(2)}{2} > \frac{h(3)}{3}$$

which clearly suffices because of the comments of section 4.

Inequality (2) is equivalent to

$$(2') \qquad 3\sum_{i=1}^{k} x_i^4 - 2\sum_{i=1}^{k} x_i^6 < 1$$

under the restraint $\sum_{i=1}^{k} x_i^2 = 1$. Using Lagrange multipliers to maximize the left hand of (2'), we find that at such a maximum, the

quantity

$$\omega = x_i^2(1 - x_i^2) \qquad i = 1, \dots, k$$

has to be independent of i. We want to conclude that all the x_i's have to coincide, and this is seen as follows. As a function of x^2, $x^2(1 - x^2)$ takes the value ω once at a point $y^2 \leq 1/2$ and once at the point $1 - y^2 \geq 1/2$. Because $k \geq 3$ and $\sum_{i=1}^{k} x_i^2 = 1$, the x_i's have to coincide. Thus each one of the x_i^2's takes the value $1/k$, and it is enough to check that

$$3(1/k)^2 k - 2(1/k)^3 k < 1.$$

Thus the lemma is proved.

CHAPTER 2 THE PROPAGATION OF CHAOS

1. THE N-MOLECULE GAS

In this chapter we consider the general 3-dimensional Boltzmann equation. The only restriction imposed on the interaction is that two molecules at a distance larger than some $R < \infty$ cannot feel each other (cut-off).

First we describe the construction of "the n-molecule gas." Then we state what we mean by "propagation of chaos." The proof of the latter will keep us at work for the rest of the chapter, and so the reader may wish to look back at the introduction for moral encouragement, if needed. There one finds the reasons why propagation of chaos is a property worth proving.

Take as state space R^{3n} and let $v = (v_1, \ldots, v_n)$ with $v_i \in R^3$ be a point in R^{3n}. Let T_a indicate a random time distributed according to

$$P(T > t) = e^{-a(n-1)t}.$$

Construct the n-molecule gas as follows: if you are at $v = (v_1, \ldots, v_n)$ at time $t = 0$,

a) pick a pair of indices $i < j$ according to the uniform distribution $\binom{n}{2}^{-1}$;

b) if $v_i = v_j$, you do not jump, but if $v_i \neq v_j$, you wait for time T_a with

$a = |v_i - v_j|$ and then

c) jump from

$$(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) \text{ to } (v_1, \ldots, v_i^*, \ldots, v_j^*, \ldots, v_n)$$

where the scattering parameters (θ, ϕ) are chosen with distribution $I(|v_i - v_j|, \theta) \sin\theta \, d\theta \, d\phi$ properly normalized. *

This recipie is now repeated using a new pair $i<j$, new holding times, and new scattering parameters (θ, ϕ), independently of everything that went on before, etc., etc. Clearly the random process so defined is Markovian; it can be viewed as a sort of "Poisson process."

The joint density $p(t,v) = p(t, v_1, \ldots, v_n)$ of the n-molecules is fully determined as the solution of the equation

$$\frac{\partial p}{\partial t}(v) = G^n p(v) = \tag{1}$$

$$= \frac{n-1}{\binom{n}{2}} \sum_{i<j} |v_i - v_j| \int [p(v_1, \ldots, v_i^*, \ldots, v_j^*, \ldots, v_n) - p(v)] \, I(|v_i - v_j|, \theta) \sin\theta \, d\theta \, d\phi$$

which reduces to the initial density at $t=0$. This is the so-called "master equation."

One notices that G_n commutes with any permutation of the arguments $v_1, \ldots, v_n$, so if $p(0) = p(0, v_1, \ldots, v_n)$ was a symmetric function

---

* Look at the Introduction for the definition of v_i^*, v_j^*.

of $v_1, \ldots, v_n$, the same will hold for $p(t) = p(t, v_1, \ldots, v_n)$ at any later time, and we can restrict our considerations to such densities.

We are now ready for the technical formulation of the propagation of chaos.

Define the exterior product of two measures u and v on R^3 to be the measure on $R^3 \times R^3$ determined by $(v \otimes v)\,(A \times B) = u(A)\, v(B)$ for A, B subsets of R^3. Let

$$\phi_1, \phi_2, \ldots, \phi_m$$

be arbitrary bounded and continuous functions defined on R^3, f_o a measure on R^3, and f_o^n its n-th exterior product $f_o \otimes \cdots \otimes f_o$. Let p be the solution of equation (1) which reduces to (the density of) f_o^n at $t = 0$. Notice that p depends on n. Then the statement of PROPAGATION OF CHAOS is that for each fixed m

$$\int_{R^{3n}} \phi_1(v_1) \ldots \phi_m(v_m)\; p(t, v_1, \ldots, v_m, v_{m+1}, \ldots, v_n)$$

converges as $n \to \infty$ to the m-fold product

$$(\int \phi_1 f) \ldots (\int \phi_m f)$$

of a time-dependent distribution f on R^3. Moreover this distribution is the solution of the Boltzmann equation

$$(2)\quad \frac{\partial f}{\partial t}(v_1) = \iint [f(v_1{*})f(v_2{*}) - f(v_1)f(v_2)]\,|v_1 - v_2|\, I(|v_1 - v_2|, \theta)\sin\theta\, d\theta\, d\phi\, dv_2, \quad *$$

which reduces to f_o at time $t = 0$.

We prove this statement under the assumptions of EXISTENCE, UNIQUENESS, and SMOOTHNESS mentioned in the introduction and in section 4 of this chapter.

---

* Here and in the future, we write the Boltzmann equation making use of a formal density for the distribution f, which is denoted also by f.

2. PREPARATIONS

Consider the set D^n of symmetric probability measures on R^{3n}. The n-molecule gas gives us a flow on D^n. Notice that when we change the n, the place where the motion occurs also changes.

It seems desirable to find a way of visualizing the D^n's as a family of subsets becoming dense in an ambient space; presumably this ambient space ought to be the set of symmetric probability measures on $(R^3)^\infty$.

To do this we start by observing that D^n is a convex set. The extreme points of this set are exactly those measures that charge, say, the point $v = (v_1, \ldots, v_n)$ and all those points in R^{3n} that one gets by permuting $v_1, \ldots, v_n$, giving the same mass to each of these and no mass to any other point. Clearly such a distribution can be visualized as a unit mass placed on a point of R^{3n}/S. Here R^{3n}/S stands for the set of equivalence classes into which R^{3n} splits under the action of the group S of permutations of n letters.

We denote with $\delta(v_i)$ a unit mass located at the point $v_i \in R^3$.

Now we put the set R^{3n}/S is correspondence with a special set of measures in R^3 according to the recipie

(1) $$(v_1, \ldots, v_n) \leftrightarrow \frac{1}{n} \sum_{i=1}^{n} \delta(v_i).$$

We denote with M^n the set of these special measures, and by M^∞ the set of all probability measures on R^3. We have therefore established a one-to-one map between the extreme points of D^n and the set M^n.

Now any element β of D^n is the barycenter of a unique mass distribution Ω_β supported by the set of extreme points α:

$$\beta = \int_{R^{3n}/S} \alpha \, d\Omega_\beta(\alpha). \tag{2}$$

The expression (2) is shorthand for the following state of affairs: if ψ is a bounded continuous function defined on R^{3n}/S, then

$$\int \psi\beta \equiv \psi(\beta) \equiv \beta(\psi) = \int \psi(\alpha) \, d\Omega_\beta(\alpha), \tag{3}$$

$\psi(\alpha)$ being regarded as a numerical function defined on R^{3n}/S or on M^n, as the occasion makes preferable.

The desired visualization of the D^n's is now achieved. D^n has been identified with the set of probability measures on M^n, and for each n this later set is part of the set of probability measures on M^∞.

Now we go back to our propagation of chaos problem and try to get some profit from the identifications made above.

Recall that each n-molecule gas process gives us a motion on D^n. Let $\phi_1, \ldots, \phi_n$ be bounded continuous functions on R^3 and let $p(t)$ be the solution of the master equation that coincides with $p(0) = \beta$ at $t = 0$. Then according to (3),

$$p(t)(\phi_1 \otimes \ldots \otimes \phi_n) = \int_{R^{3n}/S} \alpha(\phi_1 \otimes \ldots \otimes \phi_n) d\Omega_{p(t)}(\alpha). \; ^* \tag{4}$$

We can, in agreement with our identifications, think of this as an integral of the numerical function $\phi(\alpha) = \alpha(\phi_1 \otimes \ldots \otimes \phi_n)$ defined on M^n against a measure that changes in time. But we could equally well try to put this formula into a "dual form"

$$\int_{M^n} (T_t^n \phi)(\alpha) d\Omega_\beta(\alpha) = \int_{M^n} \phi(\alpha) d\Omega_{p(t)}(\alpha), \tag{5}$$

where T_t^n is the mapping dual to the (linear!) map $\Omega_\beta \to \Omega_{p(t)}$, acting on functions defined on M^n.

The strategy of the proof of the propagation of chaos is now as follows:

STEP A: Compute the generator $\tilde{G}^n$ of the semigroup T_t^n introduced in (5);

STEP B: Check that, as $n \to \infty$, this generator converges (in a sense to be specified later) to the generator $\tilde{G}^\infty$ of a semigroup T_t^∞ acting

---

* $(\phi_1 \otimes \ldots \otimes \phi_n)(v_1 \ldots v_n) \equiv \phi_1(v_1) \ldots . \phi_n(v_n)$

on functions defined on M^{∞}. T_t^{∞} will be found to be the translation semigroup associated with the Boltzmann flow on M^{∞}.

STEP C: Check that if $\phi_1, \ldots, \phi_m$ are bounded continuous functions defined on R^3, then the function $\phi^n(\alpha) = \alpha(\phi_1 \otimes \ldots \otimes \phi_m \otimes 1 \otimes \cdots \otimes 1)$ goes to the function $\phi^{\infty}(\alpha) = \alpha(\phi_1) \times \ldots \times \alpha(\phi_m)$, as $n \to \infty$, thinking of the α's now as belonging to $M^n \subset M^{\infty}$.

STEP D: Check that if $\beta = f_o \otimes \ldots \otimes f_o = f_o^n$ with $u \in M^{\infty}$ then $d\Omega_\beta$ converges weakly to the unit mass at $f_o \in M^{\infty}$.

Suppose we were able to show that the convergence of $\tilde{G}^n$ to $\tilde{G}^{\infty}$ implies the convergence of T_t^n to T_t^{∞} on a sufficiently wide class of functions defined on M^{∞}. Then we could try to patch all these steps together to obtain the propagation of chaos as follows: take $\phi_1, \ldots, \phi_m$ and $\beta = f_o^n$ as above, and let f be the solution of Boltzmann's equation which reduces to f_o at $t = 0$. Then we would hope to be able to verify that

$$p(t)(\phi_1 \otimes \ldots \otimes \phi_m \otimes 1 \ldots \otimes 1) = \int T_t^n \phi^n d\Omega_\beta \tag{6}$$

$$\to \int T_t^{\infty} \phi^{\infty} \times \text{the unit mass at } \alpha = f_o$$

$$= f(\phi_1) \times \ldots \times f(\phi_m).$$

Of the program outlined above, the only deep point is to infer the convergence of T_t^n to T_t^{∞} from the convergence of the generators.

This problem has been dealt by Trotter [15], and we reproduce here his result.

Let Ω be a topological space and $X = C(\Omega)$. * For each n, let Ω^n be a subset of Ω and let X^n be the Banach space of restrictions to Ω^n of functions in X. Define

$$\|f\|_n = \sup_{u \in \Omega^n} |f(u)|.$$

and assume that the Ω^n become dense, so that as $n \to \infty$,

$$\|f\|_n \to \|f\| \quad \text{for any} \quad f \in X.$$

We say that a sequence of linear operators A_n acting on X^n converge strongly to A acting on X if $A_n f$ is defined and

$$\lim_{n \to \infty} \|A_n f - Af\|_n = 0$$

for any f in the domain of A.

For each n, let $\exp(tA_n)$ be a strongly continuous contraction semigroup of operators on X_n with infinitesimal generator A_n, and let $\exp(tA)$ be a strongly continuous contraction semigroup of operators on X with generator A.

Trotter's result may now be formulated.

---

* $C(\Omega)$ is the space of continuous functions on Ω, with the supremum norm.

THEOREM:* Suppose C is a dense subset of the domain of A such that

1) the closure of the restriction of A to C coincides with A.

2) for each $f \in C$, $A_n f$ is defined and

$$\lim_{n\to\infty} \|A_n f - Af\|_n = 0.$$

Then $\exp(tA_n)$ converges strongly to $\exp(tA)$.

We close this section by giving a "geometrical picture" of propagation of chaos.

The solution of Boltzmann's equation gives you a deterministic motion on M^∞; this motion is non-linear. However, we can induce from it a linear motion taking place on a class of function F defined on M^∞ by means of

(6) $$(T_t^\infty F)(f_o) = F(f_t).$$

Details are found in section 4. We call such a flow, lifted from the phase space M^∞, a trivial flow.

Now for each $n<\infty$, we construct the n-molecule gas governed by the master equation in M^n. This gives us the linear motion

---

* For a very useful presentation of semigroup theory, the reader can consult Kato [8]. There one finds extensions of this result, due to Trotter and Kato.

$$F \to T_t^n F \tag{7}$$

described in formula (5); this motion is not trivial. Propagation of chaos could be thought as saying that the triviality of the motion is recovered for $n = \infty$, i.e., the sequence of nontrivial motions (7) approximate the trivial one (6).

3. MC KEAN'S 2-STATE MAXWELLIAN GAS

We do now a simple case of the program outlined in steps A, B, C, D in the previous section. For that purpose we study a very much simplified model of Boltzmann's equation, the so-called 2-state Maxwellian gas. McKean introduced it in [10], as an instance of a new class of Markov processes having non-linear generators. In [10] he proved the appropriate propagation of chaos for this model using the kind of estimates employed by Kac [7] for the one-dimensional Maxwellian gas.

There are several reasons to take a look at this model before taking up the actual Boltzmann equation in section 4. The identification of D^n with a set of measures on M^n, presented in section 2 for the general case, becomes extremely clear and suggestive. Then the proof of propagation of chaos can be carried out in detail along the lines indicated in section 2 with a minimum of abstraction. An extra bonus is that we can apply Trotter's result directly without having to make assumptions as to existence, uniqueness and smoothness which are needed for the general case.

The 2-state Maxwellian gas is simpler than the general case in two points:

a) R^3 is replaced by a 2-point space $Z = \{+1, -1\}$,

b) the scattering rule is changed to

$$(e_1^*, e_2^*) = \begin{cases} (e_1, e_1e_2) & \text{with probability } 1/2 \\ (e_1e_2, e_2) & \text{with probability } 1/2 \end{cases}$$

where e denotes an element of Z.

The Boltzmann equation for the 2-state Maxwellian gas takes the form

$$\frac{\partial f}{\partial t}(e_1) = \iint [f(e_1^*)f(e_2^*) - f(e_1)f(e_2)]\, de_2 do$$

$$(1) \qquad \begin{cases} = f(-1)^2 + f(+1)^2 - f(+1) & \text{if} \quad e_1 = +1 \\ = 2f(-1)f(+1) - f(-1) & \text{if} \quad e_1 = -1. \end{cases}$$

Here $\int de_2$ means sum over $e_2 = \pm 1$ and $\int do$ means sum over the two possible outcomes of a collision. The evolution of a measure in Z is described by the evolution of the mass ascribed to $+1$, $f_t(+1) = x(t) = x$:

$$(2) \qquad \dot{x}_t = (1-x_t)^2 + x_t^2 - x_t = (1-x_t)(1-2x_t).$$

Now we go over formulas (2.1), (2.2) and (2.3)*. We observe that an element of Z^n/S is specified by the number of $+$'s, say k, that it contains and may be put in correspondence with the measure

---

* This means formula 3 of section 2. In this chapter there is no reference to chapter I, so no confusion can arise.

$$\frac{k}{n}\delta_{+1} + \frac{n-k}{n}\delta_{-1} = \alpha_k,$$

belonging to $M^n(Z)$ in agreement with the general scheme of (2.1). There are $n+1$ of these extreme symmetric measures in Z^n and the barycentric representation of an arbitrary symmetric measure β reads

(3) $$\beta = \Sigma\, \alpha_k\, \Omega_\beta(\alpha_k)$$

where the summation extends over the grid of points $\{k/n\}$, $k=0,1,\dots,n$ in $[0,1]$. This set is identified with Z^n/S in the obvious way. Here α_k stands for the measure on Z^n charging all those points with k pluses, giving each one of them the mass $\binom{n}{k}^{-1}$. It is useful to note for later use that

(4) $$\Omega_\beta(\alpha_k) = \binom{n}{k}\beta(+1\dots+1, -1\dots-1) \qquad k \text{ pluses, } n-k \text{ minuses.}$$

Let $(i, m-i, Z^{n-m})$ denote the set of points in Z^n having i pluses and $m-i$ minuses among its first m coordinates, and let ψ^n denote the characteristic function of this set. It is clear that we have*

$$\psi^n(\alpha_k) \equiv \alpha_k(i, m-i, Z^{n-m}) = \frac{\binom{n-m}{k-i}}{\binom{n}{k}} \qquad \text{if } i \le k \le n-(m-i)$$

$$= 0 \qquad \text{otherwise.}$$

---

* Using the notation introduced in (2.3).

We have then

$$\beta(i, m-i, Z^{n-m}) = \sum_k \frac{\binom{n-m}{k-i}}{\binom{n}{k}} \Omega_\beta(\alpha_k).$$

Now we start the proof of the propagation of chaos. Recall that the n-molecule gas is a stochastic process in Z^n, and thus the way to achieve the situation described in section 2 is to construct a stochastic process in

$$Z^n/S = \text{the grid } \{k/n\}$$

so that the associated semigroup T_t^n satisfies

(5) $$\Sigma(T_t^n \phi)(\alpha_k)\Omega_\beta(\alpha_k) = \Sigma\phi(\alpha_k)\Omega_{P(t)}(\alpha_k).$$

Next we look closer at the n-molecule gas. Suppose we are at a point of Z^n with k pluses and $n-k$ minuses, indicated by $(k, n-k)$. A collision can send us to any of the following places

$(k+1, n(k+1))$ with probability $\dfrac{\binom{n-k}{2}}{\binom{n}{2}} p(--, +-) \equiv \gamma_k$

$(k, n-k)$ with probability $\dfrac{\binom{k}{2}p(++, ++)+k(n-k)p(+-, +-)}{\binom{n}{2}} \equiv \epsilon_k$

$(k-1, n-(k-1))$ with probability $\dfrac{k(n-k)}{\binom{n}{2}} p(+-, --) \equiv \sigma_k$

We can now write the master equation as

$$\text{(6)} \qquad \frac{\partial P}{\partial t}(k, n-k) = G^n P(k, n-k) =$$

$$= \frac{1}{n}[\sigma_{k+1} P(k+1, n-(k+1)) + \gamma_{k-1} P(k-1, n-(k-1)) + (\epsilon_k - 1)P(k, n-k)],$$

and we are ready to go over steps A, B, C, D of section 2.

STEP A: The simplest way to compute the generator $\tilde{G}^n$ is the following. First we look for a measure m on Z^n/S which is "symmetric" with respect to the process, i.e.,

$$\text{(7)} \qquad m_k \gamma_k = m_{k+1} \sigma_{k+1} \quad \text{for all } k.$$

Then we recall that with respect to such a measure the transition kernel becomes self-adjoint, i.e., if (f, g) denotes $\int fg\,dm$ then $(Kf, g) = (f, Kg)$. The reader can find a neat presentation of this fact in Nelson [11].

One immediately checks that $m_k = \binom{n}{k}$ is a symmetric measure, namely

$$\text{(7')} \qquad \binom{n}{k} \gamma_k = \binom{n}{k+1} \sigma_{k+1}.$$

This fact is a very fortunate one in view of (4). In fact from (4) we get

$$\frac{\partial}{\partial t} \Sigma \phi(\alpha_k) \Omega_{P(t)}(\alpha_k) \Big|_{t=0} = \Sigma \phi(\alpha_k) \binom{n}{k} G^n P(k, n-k)$$

and now using the above mentioned self-adjointness of G^n we get the equivalent expression

$$\frac{\partial}{\partial t}(T_t^n\phi)(\alpha_k)\Omega_{P(t)}(\alpha_k)\Big|_{t=0} = \Sigma(G^n\phi)(\alpha_k)\binom{n}{k}P(k, n-k).$$

Invoking (4) again this is reexpressed as

$$\Sigma(G^n\phi)(\alpha_k)\Omega_\beta(\alpha_k).$$

The moral is that the generator of our process in Z^n/S cannot be distinguished from the G^n of the master equation. At any rate, we denote it by $\tilde{G}^n$ to indicate that it acts on functions and not on measures. See Nelson [11].

Summarizing, we have

$$(8) \qquad (\tilde{G}^n\phi)(\alpha_k) = \frac{1}{n}[\sigma_k\phi(\alpha_{k-1}) + \gamma_k\phi(\alpha_{k+1}) + (\epsilon_k - 1)\phi(\alpha_k)]$$

$$= \frac{n-k}{n}[(n-k-1)(\Delta^2\phi)(\alpha_k) + (n-2k-1)\Delta\phi(\alpha_k)].$$

Here

$$(\Delta\phi)(\alpha_k) = \phi(\alpha_k) - \phi(\alpha_{k-1})$$

and

$$(\Delta^2\phi)(\alpha_k) = \phi(\alpha_{k+1}) - 2\phi(\alpha_k) + \phi(\alpha_{k-1}) .$$

REMARK: Clearly we could reverse the procedure that led us from the master equation

(6') $$\frac{\partial P}{\partial t}(e) = (G^n P)(e) = \frac{n-1}{\binom{n}{2}} \sum_{i<j} \int [P(e^*_{ij}) - P(e)] do \; *$$

to its form (6), and thus go from (8) to the expression

(8') $$(\tilde{G}^n \phi)(\alpha) = \frac{n-1}{\binom{n}{2}} \sum_{i<j} \int [\phi(\alpha^*_{ij}) - \phi(\alpha)] \; do \; *.$$

The only important point in the computation of $\tilde{G}^n$ has been to notice the identities (7'). They are in no way automatic, but rather the consequence of the way in which we built up the n-molecule gas, coupled with the following property of the "scattering rules"

$$Pr(e_1, e_2 \to e_1^*, e_2^*) = Pr(e_1^*, e_2^* \to e_1, e_2).$$

This is of course just the property of "microscopic reversibility" which is present again in the actual Boltzmann equation. This will allow us to write in section 4 an expression similar to (8') for the corresponding $\tilde{G}^n$.

STEP B: Define G^∞ acting on nice functions on $[0,1]$ by means of $G^\infty F(x) = (1-x)(1-2x)\frac{\partial F}{\partial x}$. Clearly for any $F \in C^2[0,1]$ we have

---

* e^*_{ij} and α^*_{ij} are obtained from e and α by changing the i and j components into e^*_i, e^*_j and α^*_i, α^*_j respectively.

$$\sup_k \left|\tilde{G}^n F\left(\frac{k}{n}\right) - \tilde{G}^\infty F\left(\frac{k}{n}\right)\right| \to 0 \quad \text{as} \quad n \to \infty.$$

We will refer to this fact by saying that $\tilde{G}^n$ converges to $\tilde{G}^\infty$ on $C^2[0,1]$.

$\tilde{G}^\infty$ is easily recognized as the infinitesimal generator of the translation semigroup

$$T_t^\infty : F \to F(x_t)$$

with

$$\dot{x}_t = (1 - x_t)(1 - 2x_t)$$

and

$$F \in C[0,1].$$

The flow x_t is precisely the "Boltzmann flow" as the reader will see at once from (2).

STEP C: If ψ^n is the characteristic function of $(i, m-i, Z^{n-m})$, then

$$\psi^n(\alpha_k) = \psi^n\left(\frac{k}{n}\right) = \frac{\binom{n-m}{k-i}}{\binom{n}{k}} =$$

$$= \frac{k}{n} \cdots \frac{(k-i+1)}{(n-i+1)} \times \frac{(n-k)}{(n-i)} \cdots \frac{(n-m-(k-i)+1}{(n-m+1)}$$

converges when $k/n \to x$ to $\psi^{\infty}(x) = x^{i}(1-x)^{m-i}$ uniformly on $[0,1]$, i.e. $\max_{k\leq n} |\psi^{n}(\frac{k}{n}) - \psi^{\infty}(\frac{k}{n})|$ is small for large n.

STEP D: Let $f_o = x_o \delta_+ + (1-x_o)\delta_-$ be a measure on Z and let β denote the measure $f_o^n = f_o \otimes \cdots \otimes f_o$. Then the measure $\Omega_\beta(\alpha_k) = \binom{n}{k} x_o^k (1-x_o)^{n-k}$ converges weakly to the unit mass at x_o. This is immediate from the weak law of large numbers.

Finally we collect all these steps to get a proof of the propagation of chaos, which may be stated in the form

$$e^{tG_n} f_o^n(i, m-i, Z^{n-m}) \to x_t^i (1-x_t)^{m-i} .$$

Rewrite the left hand side as

$$\Sigma \psi^n(\alpha_k)\Omega_{P(t)}(\alpha_k) = \Sigma (T_t^n \psi^n)(\alpha_k)\Omega_\beta(\alpha_k) =$$

$$= \Sigma T_t^n(\psi^n - \psi^\infty)\Omega_\beta(\alpha_k) + \Sigma(T_t^n - T_t^\infty)(\psi^\infty)\Omega_\beta(\alpha_k) + \Sigma T_t^\infty \psi^\infty \Omega_\beta(\alpha_k).$$

On the right side, the first summand goes to zero, as $n \to \infty$, because $\| T_t^n \| \leq 1$ and $\| \psi^n - \psi^\infty \|$ goes to zero by step C. The last summand goes to $x_t^i(1-x_t)^{m-i}$ by step D and the definition of ψ^∞ and T_t^∞.

The middle integral tends to 0 in view of

$$\sup \left| (T_t^n - T_t^{\infty})(\psi^{\infty}) \right| \to 0 \quad \text{as} \quad n \to \infty.$$

This follows from Trotter's result of section 2, and step B. The fact that (with respect to G^{∞}) $C^2[0,1]$ satisfies the conditions of Trotter's theorem is well known.

4. THE GENERAL CASE

To prove propagation of chaos, we perform steps A, B, C, D of section 2. The situation is now technically more involved than for the 2-state Maxwellian gas but the plan is just the same. Previously, the action took place in the Banach space $C[0,1]$. As a preliminary step we introduce a Banach space appropriate for the general case.

THE BANACH SPACE X: Consider the set Ω of all probability measures on R^3 having a finite second moment and provide it with the customary weak topology. Observe that

$$\Omega_a = \left\{ f \in \Omega : \sigma^2(f) = \int |v|^2 f(dv) \leq a \right\}$$

is compact in this topology, and therefore, Ω is exhausted by an increasing family of compact subsets. The compactness of Ω_a is a consequence of the fact that we can control uniformly the tail of any measure in Ω_a, and that if f_n converges weakly to f then

$$\sigma^2(f) \leq \liminf \sigma^2(f_n).$$

For any bounded continuous function h on R^3, we define a linear (continuous) function on Ω by means of the recipe

$$F(f) = \int h(v) f(dv).$$

Taking finite linear combinations of finite products of such functions, we get an algebra of bounded continuous functions (polynomials) on Ω. The Banach space X is the completion of this algebra under the norm

$$\|F\| = \sup_{f \in \Omega} |F(f)| < \infty$$

We retain the symbol F for a generic element of X. We get a sequence of "approximating" Banach spaces X^n by restricting the functions of X to $\Omega^n = M^n \subset \Omega$.* We denote by X_a the set of continuous functions on Ω_a and recall that the Stone-Weierstrass theorem implies that any element of X_a can be uniformly approximated in Ω_a by the above mentioned polynomials. Finally, we introduce the family of subsets $\Omega_a^n = \Omega_a \cap M^n$ of Ω_a and define X_a^n to be the Banach space of functions obtained restricting those in X_a to Ω_a^n.

There is still one more preparatory step before we can go into steps A, B, C, D. To apply Trotter's result we have to prove that all the semigroups involved are strongly continuous and contractive. This is plain for the "approximating" semigroups, and now we deal with the "limit" semigroup T_t^∞ defined by means of translations along the solutions of the Boltzmann equation:

---

* Every measure in M^n has a finite support and therefore belongs to Ω.

(1) $$(T_t^{\infty} F)(f) = F(f_t).$$

To start with, we have to show that if $F \in X_a$ then $f \to F(f_t)$ is in X_a. Once this is done we will know that T_t^{∞} acts on X_a. First of all Ω_a is invariant under the Boltzmann flow, and now from the fact that

$$f \to f_t$$

is weakly continuous*, one infers that $f \to F(f_t)$ is in X_a, first for polynomials and then by the Stone-Weierstrass theorem for any F in X_a.

That T_t^{∞} is a contraction is obvious, because

$$\sup_{f \in \Omega_a} |F(f_t)| \leq \sup_{f \in \Omega_a} |F(f)| \leq \|F\|_{\Omega_a}.$$

To get strong continuity for T_t^{∞} it is now enough to prove

$$\|T_t^{\infty} F - F\| \to 0 \text{ as } t \to 0$$

for each F of our dense algebra of polynomials. T_t^{∞} acts multiplicatively on this algebra:

$$T_t^{\infty}(F_1 F_2) = T_t^{\infty}(F_1)\, T_t^{\infty}(F_2).$$

Thus it is enough to look at linear polynomials. Now the total varia-

---

* See appendix I to this chapter.

tion of Bf is bounded by $2\pi R^2(1 + 2\sigma^2(f))$*. Then it is clear that

$$\sup_{f\in\Omega_a} \left|\int hf_t - \int hf\right| \leq t \sup|h| \sup_{f\in\Omega_a} 2(1+2\sigma^2(f)) \to 0 \quad \text{as} \quad t\to 0,$$

and the strong continuity of T_t^∞ is proved.

Now we are ready to go over steps A, B, C, D of section 2.

STEP A: The n-molecule process, put in dual form in formula (2. 5) gives us a semigroup acting on X^n. The property of "microscopic reversibility" has here the same effect that it had for the simpler 2-state Maxwellian gas, see (3. 8'). Arguing in exactly the same way, one can see that the action of the corresponding generator is given by

$$(2) \qquad \widetilde{G}^n F(f) = \frac{1}{n} \sum_{i<j} |v_i - v_j| \int [F(f_{ij}^*) - F(f)] I(|v_i - v_j|, \theta) \sin\theta \, d\theta \, d\phi$$

for $F \in X^n$, and

$$f = \frac{1}{n} \sum_{k=1}^{n} \delta(v_k)$$

from Ω^n, where

$$f_{ij}^* = \frac{1}{n}(\delta(v_1)+ \ldots + \delta(v_i^*)+ \ldots + \delta(v_j^*)+ \ldots + \delta(v_n)).$$

The semigroup $T_t^n = \exp(t\widetilde{G}^n)$ is contractive and strongly continuous acting on X_a^n because $\sigma^2(f) = \sigma^2(f_{ij}^*)$.

STEP B: We consider now the class of "continuously differentiable"

---

* See appendix I to this chapter.

functions F on Ω_a, such that

$$F(f') - F(f) = \int H_f(v)(f' - f)(dv) + o(\|f' - f\|)^*$$

where H_f is a function on R^3 which is bounded independently of f and $\|f - f'\|$ stands for the total variation of $f - f'$. We refer to this class of functions as C'.

Now we compute $\tilde{G}^n F$ for $F \in C'$, using the notation of (2).

$$\tilde{G}^n F(f) = \frac{1}{n} \sum_{i<j} |v_i - v_j| \int [H_f(v_i^*) - H_f(v_i) + H_f(v_j^*) - H_f(v_j)] I(|v_i - v_j|, \theta) \sin\theta \, d\theta \, d\phi$$

$$+ \frac{1}{n} \sum_{i<j} \int o(\|f_{ij}^* - f\|) \, I(|v_i - v_j|, \theta) \sin\theta \, d\theta \, d\phi \, .$$

The first summand equals the integral of the function H_f against the (signed) measure Bf (from Boltzmann's equation). Therefore, this first summand can be identified with $\tilde{G}^\infty F(f)$, where $\tilde{G}^\infty$ is the infinitesimal generator of the semigroup T_t^∞, acting by translation along the solutions of the Boltzmann equation.

To get an estimate for the second summand observe from the definition of f_{ij}^* that

$$\|f_{ij}^* - f\| \leq \frac{4}{n} \quad , \quad \frac{1}{n} \sum |v_i|^2 = \sigma^2(f) \leq a$$

---

* The "little o" involved in $o(\|f' - f\|)$ is permitted to depend upon f.

and

$$(\pi R^2)^{-1}|v_i - v_j| \int I(|v_i - v_j|, \theta) \sin\theta \, d\theta \, d\phi = |v_i - v_j| \leq 1 + |v_i|^2 + |v_j|^2.$$

Thus, we can bound the second summand in $\widetilde{G}^n F$ by

$$o(1) \chi \frac{1}{n} \sum_{i<j} \frac{4}{n} (1 + |v_i|^2 + |v_j|^2) < o(1) \chi \frac{4}{n^2} [\binom{n}{2} + 2(n+1) \sum_{i=1}^{n} |v_i|^2] = o(1).$$

We therefore get

$$(4) \qquad \sup_{f \in \Omega_a^n} |\widetilde{G}^n F(f) - \widetilde{G} F(f)| \to 0 \quad \text{as} \quad n \to \infty.$$

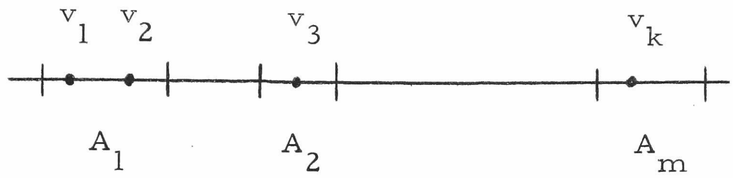
STEP C: Let α be the extreme symmetric probability measure in R^{3n} that charges the class in R^{3n}/S consisting of

n_1 copies of v_1, n_2 copies of $v_2, \ldots, n_k$ copies of v_k

with $n_1 + \ldots + n_k = n$ and all the v_i different among themselves. Recall that according to the scheme of section 2, α is also thought of as the measure on R^3 given by $\frac{1}{n} \Sigma n_i \delta(v_i)$. For fixed bounded continuous functions $\phi_1, \ldots, \phi_m$ we have to study the limiting behaviour of the difference between the functions

$$\phi^n(\alpha) = \alpha(\phi_1 \otimes \ldots \otimes \phi_m \otimes 1 \otimes \ldots \otimes 1) \text{ and } \phi^\infty(\alpha) = \alpha(\phi_1) \ldots \alpha(\phi_m).$$

We assume first that the ϕ_i have pairwise disjoint compact supports A_i. We have to compute the possible ways of arranging our n points that give a contribution to the integral of $\phi_1 \otimes \ldots \otimes \phi_m \otimes 1 \otimes \ldots \otimes 1$ against α. For an illustration, consider the situation described in the picture



In this case the non-zero contributions to the integral come from all the possible rearrangements of the points $v_1, v_3, \ldots, v_k$, plus all the possible rearrangements of the points $v_2, v_3, \ldots, v_k$, and so on.

If we add over all this contributions, we obtain in the general case the value

$$(5) \qquad \alpha(\phi_1 \otimes \ldots \otimes \phi_n) = \alpha(\phi_1) \times \ldots \times \alpha(\phi_n) \frac{n^{m-1}}{(n-1)\ldots(n-m+1)}$$

with $\alpha = \Sigma \frac{n_i}{n} \delta(v_i)$ on the right side. Explicitly, for the case $m = 2$, we get

$$\phi_1(v_1)\phi_1(v_3) \frac{\binom{n-2}{n_1-1,\, n_2,\, n_3-1}}{\binom{n}{n_1,\, n_2,\, n_3}} + \phi_1(v_2)\phi_2(v_3) \frac{\binom{n-2}{n_1,\, n_2-1,\, n_3-1}}{\binom{n}{n_1,\, n_2,\, n_3}} =$$

$$= \frac{n_1\phi_1(v_1)+n_2\phi_1(v_2)}{n} \times \frac{n_3\phi_2(v_3)}{n-1} = \alpha(\phi_1)\alpha(\phi_2) \times \frac{n}{n-1} .$$

With some extra combinatorial complication it is possible to see that if the supports of the ϕ_i are not disjoint, then

$$(6)\quad |\alpha(\phi_1\otimes\dots\otimes\phi_m\otimes 1\otimes\dots\otimes 1)-\alpha(\phi_1)\times\dots\times\alpha(\phi_m)| \leq \epsilon_n|\alpha(\phi_1\otimes\dots\otimes\phi_m\otimes 1\otimes\dots\otimes 1)|$$

Here ϵ_n is arbitrarily small for large n and is independent of α. The same result holds even if $\{\phi_i\}$ do not have compact supports through an appropriate process of approximation. Therefore, for fixed $\phi_1,\dots,\phi_m$,

$$\sup_{\alpha\in M^n} |\phi^n(\alpha) - \phi^\infty(\alpha)| \to 0 \quad \text{as} \quad n\to\infty.$$

STEP D: We have to prove that if F is a bounded continuous function on Ω and $\beta=f\otimes\dots\otimes f$, then

$$\int F(\alpha)\, d\Omega_\beta(\alpha) \to F(f) \quad \text{as} \quad n\to\infty.$$

Observe first that we can restrict the domain of integration to a set Ω_a with an arbitrarily small error because

$$\Big|\int_{\sigma^2(\alpha)\geq a} F(\alpha)d\Omega_\beta(\alpha)\Big| \leq \|F\|\int \frac{\sigma^2(\alpha)}{a}\, d\Omega_\beta(\alpha) = \|F\|\,\frac{\sigma^2(f)}{a}\,.$$

If F is the monomial on M^∞ defined by

$$F_1(\alpha) = \int\phi_1\alpha\times\dots\times\int\phi_m\alpha$$

then it follows from (6) that

$$\left| \int (F_1(\alpha) - \alpha(\phi_1 \otimes \ldots \otimes \phi_m \otimes 1 \otimes \ldots \otimes 1) d\Omega_\beta(\alpha) \right| \leq \epsilon \int d\Omega_\beta(\alpha) = \epsilon .$$

Now $\int \alpha(\phi_1 \otimes \ldots \otimes \phi_m \otimes 1 \ldots \otimes 1) d\Omega_\beta(\alpha)$ is just a complicated way of writing* $\beta(\phi_1 \otimes \ldots \otimes \phi_m \otimes 1 \ldots \otimes 1) = f(\phi_1) \ldots f(\phi_1) = F_1(f)$. Therefore for a polynomial F

$$\int F(\alpha) d\Omega_\beta(\alpha) - F(\alpha)$$

if small for large n.

The proof is now complete because on Ω_a we can approximate any continuous function by means of polynomials.

Steps A, B, C and D being now completed, we put them together to prove our theorem.

This goes essentially in the same way as for the 2-state Maxwellian gas. However, there is one point which cannot be dispensed with so easily as we did then. **

We have proved in step B that $\widetilde{G}_n F \to \widetilde{G}_\infty F$ in an appropriate sense for functions F from the class C'. Trotter's result*** says

---

* See formula (2.3).

** See last paragraph in section 3.

*** See section 2.

that this convergence would imply the convergence of T_t^n to T_t^∞ if we could prove that the closure of $\tilde{G}^\infty$ agrees with its restriction to C'. But the proof of this fact eludes us. Instead, we make a smoothness assumption* on the Boltzmann flow T_t^∞.

<u>SMOOTHNESS ASSUMPTION: If F is a polynomial and if f_t denotes the solution of the Boltzmann equation with initial data f_o, then the map $f_o \to F(f_t)$ is of class C' for each time $t \geq 0$.</u>

A proof of this assumption should come from a very careful look at the proof of existence for the Boltzmann flow.

Using this assumption and step B, we exhibit now a subset of the domain of G^∞ satisfying the requirements of Trotter's theorem. Let A be the class of polynomials on Ω, and let R_1 be the resolvent operator $(1-\tilde{G}^\infty)^{-1}$: $\Phi \to \int_o^\infty e^{-t}\Phi(f_t)dt$. R_1 is bounded and invertible, and therefore maps A 1:1 onto a dense set B.

We prove now that B has the required property. Let Φ be in the domain of $\tilde{G}^\infty$ and approximate $(1-\tilde{G}^\infty)\Phi$ by polynomials $F_n = (1-\tilde{G}^\infty)\Phi_n$ from $A = (1-\tilde{G}^\infty)B$. From the continuity of R_1,

$$\Phi_n = R_1 F_n \to R_1(I-\tilde{G}^\infty)\Phi = \Phi \ .$$

Therefore from

---

* Already mentioned in the introduction.

$$F_n = (I-\widetilde{G}^{\infty})\Phi_n \rightarrow (I-\widetilde{G}^{\infty})\,\Phi \quad ,$$

one concludes that $\widetilde{G}^{\infty}\Phi_n \rightarrow \widetilde{G}^{\infty}\Phi$, completing the proof.

Now we need to know that the kind of convergence described in step B occurs for functions of B, i.e., if F is a polynomial, then

$$\widetilde{G}^{n}\int_0^{\infty} e^{-t}\, F(f_t)dt \rightarrow \widetilde{G}^{\infty}\int_0^{\infty} e^{-t}\, F(f_t)dt,$$

which is evident from step B and the smoothness assumption. At this point, we can invoke Trotter's result to get the strong convergence of T_t^n to T_t^{∞} and complete the proof precisely as in section 3. Clearly

(7) $$\int \phi_1 \otimes \cdots \otimes \phi_m \otimes 1 \otimes \cdots \otimes 1 \; p(t) = \int (T_t^n \phi^n((\alpha) d\Omega_\beta(\alpha) =$$

$$= \int T_t^n(\phi^n - \phi^{\infty})(\alpha)d\Omega_\beta(\alpha) + \int (T_t^n - T_t^{\infty})\phi^{\infty}(\alpha)d\Omega_\beta(\alpha) + \int T_t^{\infty}\phi^{\infty}(\alpha)d\Omega_\beta(\alpha).$$

Each one of these integrals can be restricted to Ω_a with an arbitrary small error because all the semigroups involved are bounded (by 1) and

$$\int_{\sigma^2(f)\geq a} d\Omega_\beta(\alpha) \leq \frac{\sigma^2(f)}{a} \quad .$$

Now we look at each of the integrals in the second line of (7), restricting the domain of integration to Ω_a and making $n \rightarrow \infty$.

The first integral is bounded (in absolute value) by $\|\phi^n - \phi^{\infty}\|$ which goes to 0 according to step C. The second one is bounded by

$\| (T_t^n - T_t^\infty)\phi^\infty \|$ which goes to zero on account of Trotter's result. For the third, observe that

$$\int (T_t^\infty \phi^\infty)(\alpha) d\Omega_\beta(\alpha) = \int \phi^\infty(\alpha_t) d\Omega_\beta(\alpha) =$$

$$= \int \phi_1(\alpha_t) \times \cdots \times \phi_m(\alpha_t) d\Omega_\beta(\alpha)$$

where $\dot{\alpha}_t = B[\alpha_t]$, and $\alpha_o = \alpha$.* The function $\phi^\infty(\alpha_t)$ is, for each t, a bounded continuous function on Ω and therefore we may use step D to conclude that

$$\int (T_t^\infty \phi^\infty)(\alpha) d\Omega_\beta(\alpha) \to \phi^\infty(f_t) = \phi_1(f_t) \times \cdots \times \phi_m(f_t).$$

To sum it up:

THEOREM: Under the present assumptions as to existence, uniqueness and smoothness, if $\phi_1, \ldots, \phi_n$ are bounded continuous functions on R^3, and if f denotes the solution of the Boltzmann equation with initial datum f_o, then

$$\lim_{n\to\infty} \int \phi_1 \otimes \cdots \otimes \phi_m \otimes 1 \otimes \cdots \otimes 1 \; p(t)$$

$$= \int \phi_1 f \times \cdots \times \int \phi_m f.$$

Here $p(t)$ denotes the solution of the master equation with initial data f_o^n.

---

* α_t is the evolution of α_o under the action of the Boltzmann flow.

APPENDIX I: WEAK CONTINUITY OF THE BOLTZMANN FLOW

Here we prove two statements used in section 4, i.e. the useful inequality (1) below and the lemma. From the definition of the Boltzmann operator B, we get

$$(1) \quad \| Bf \| \leq \int (|f(v_1^*)\, f(v_2^*)| + |f(v_1)\, f(v_2)|)\, |v_1 - v_2|\, I(|v_1 - v_2|, \theta) \sin\theta\, d\theta\, d\phi\, dv_1\, dv_2$$

$$\leq 2 \int (1 + |v_1|^2 + |v_2|^2)\, \pi R^2 |f(v_1)\, f(v_2)|\, dv_1\, dv_2$$

$$= 2\pi R^2\, (1 + 2\sigma^2(f)).$$

Now we can prove the

LEMMA: f_t is a continuous function of f_o in the weak topology of Ω_a.*

Take f_o^n in Ω_a converging weakly to f_o^n. Because $\sigma^2(f_t) \leq a$, we can choose for each fixed t a weakly convergent subsequence from $\{f_t^n\}$. By a diagonal process, we may obtain a subsequence (n') so that $f_t^{n'}$ converges weakly to f_t^∞ for all rational t, and part of the task is to show that this holds for all t.

Each $f_t^{n'}$ is a solution of the Boltzmann equation, so that

$$f_{t_2}^{n'} - f_{t_1}^{n'} = \int_{t_1}^{t_2} Bf_s^{n'}\, ds.$$

---

* Ω_a is the set of those measures f with $\sigma^2(f) \leq a$.

Because of the comments preceeding this lemma, the total variation of $f^{n'}_{t_2} - f^{n'}_{t_1}$ is bounded by $|t_2 - t_1| \times c(1 + 2\sigma^2(f_o))$. This permits us to extend the weak convergence of $f^{n'}_t$ to f^{∞}_t to all $t < \infty$.

Now we prove that f^{∞}_t solves the Boltzmann equation with initial datum f_o. By uniqueness this completes the proof of the lemma. For any bounded continuous h

$$\int h(v)\, f^{n'}_t(dv) = \int h(v)\, f^{n'}_o(dv) + \int_o^t dv \int h(v)\,(Bf^{n'}_s)(dv),$$

and to take the limit $n' \to \infty$ we only have to worry about the last integral. This can be split into 4 integrals of which the most complicated is

$$(2) \qquad \int_o^t ds \iint_{R^6} \left[|v_1 - v_2| \int I(|v_1 - v_2|, \theta)\, h(v_1^*) \sin\theta\, d\theta\, d\phi \right] f^{n'}_s(dv_1)\, f^{n'}_s(dv_2).$$

The function

$$\Gamma(v_1, v_2) = |v_1 - v_2| \int I(|v_1 - v_2|, \theta)\, h(v_1^*) \sin\theta\, d\theta\, d\phi$$

is a continuous function of (v_1, v_2) and can be uniformly bounded by a constant multiple of $|v_1 - v_2|$. The second moment of $f^{n'}_s$ is under control, so that we can restrict the R^6 integral to a compact set K with an arbitrary small error. On K, $f^{n'}_s \otimes f^{n'}_s$ converges weakly

to $f_s^{\infty} \otimes f_s^{\infty}$, and thus $\iint_{R^6} \Gamma f_s^{n'} \otimes f_s^{n'}$ converges to what it should.

This last convergence is bounded, and we can pull it inside the time integration in (2) and obtain for every h

$$\int h(v)\, f_t^{\infty}(dv) = \int h(v)\, f_o^{\infty}(dv) + \int_o^t ds \int h(v)(Bf_s^{\infty})(dv), \tag{3}$$

that is to say

$$f_t^{\infty} = f_o^{\infty} + \int_o^t ds\ Bf_s^{\infty} . \tag{3'}$$

This can be written also in the customary form because f_t^{∞} is a continuously differentiable function of time. In fact from (3'), if we use the fact that Bf_s^{∞} has total variation bounded by a constant multiple of $(1+2\sigma^2(f_s^{\infty})) \leq c(1+2a)$, we see that f_t^{∞} depends continuously on t. Therefore, Bf_t^{∞} is continuous and f_t^{∞} has a continuous derivative, allowing us to write

$$\frac{\partial f^{\infty}}{\partial t} = Bf^{\infty}.$$

BIBLIOGRAPHY

[1] CARLEMAN, T.: Problemes Mathematiques dans la theorie cinetique des gaz. Almqvist and Wiksells (1957).

[2] DIEUDONNE, J.: Foundations of Modern Analysis. Academic Press (1960).

[3] GRAD, H.: Principles in the kinetic theory of gases. Handbuch der Physik. Vol. XII, Springer (1958).

[4] GRAD, H.: Asymptotic equivalence of the Navier-Stokes and nonlinear Boltzmann equations. Proceedings of the Symposia in Applied Mathematics. Vol. XVII, American Mathematical Society (1965).

[5] HARTMAN, P.: Ordinary differential equations. John Wiley (1964).

[6] HUANG, K.: Statistical Mechanics. John Wiley (1963).

[7] KAC, M.: Foundations of kinetic theory. Proceedings of the third Berkeley Symposium on Mathematical Statistics and Probability. Vol. III, (1955).

[8] KATO, T.: Perturbation theory for linear operators. Springer-Verlag (1966).

[9] McKEAN, H.: Speed of approach to equilibrium for Kac's caricature of a Maxwellian gas. Arch. Rational Mech. Anal. Vol. 21, 5, p. 343 (1966).

[10] McKEAN, H.: An exponential formula for solving Boltzmann's equation for a Maxwellian gas. J. Combinatorial Theory. Vol. 2, 3, p. 358 (1967).

[11] NELSON, E.: The adjoint Markov process. Duke Math. Journal. Vol. 25, 4, p. 671 (1958).

[12] POINCARÉ, H.: Sur les propiétés des fonctions définies par les équations aux différences partielles. Ouevres, 1. Gauthier-Villars (1929).

[13] POVZNER, A.: On Boltzmann's equation in the kinetic theory of gases. Mat. Sb. Vol. 58, p. 63 (1962).

[14] SIEGERT, A.: On the approach to statistical equilibrium. Phys. Rev. Vol. 76, p. 1708 (1949).

[15] TROTTER, H.: Approximation of semigroups of operators. Pacific J. Math. Vol. 8, 4, p. 887 (1958).

[16] UHLENBECK, G.: On the theory of cosmic-ray showers. Phys. Rev. Vol. 62, 11, p. 467 (1942).

[17] UHLENBECK, G. and FORD, G.: Lectures in statistical mechanics. American Mathematical Society (1963).

[18] WILD, E.: On Boltzmann's equation in the kinetic theory of gases. Proc. Camb. Phil. Soc. Vol. 47, p. 602 (1951).

End