FOURTH LECTURE

Let me go back for a moment to the old derivation of Boltzmann's equation, which Max Dresden probably gave you. I would like to go over it because I will need a certain critique of it. Boltzmann started with the spatially homogeneous monatomic gas. So in a large volume V we have gas particles , and let us assume for the sake of definiteness that the particles are rigid spheres. Their diameters are all equal to, say, δ . They can only suffer binary collisions and that's the only way to exchange energy. That is the model. Boltzmann then derived his famous integro-differential equation. The probability (although he always called it the number of particles) of finding a particle in $d\vec{r}\, d\vec{v}$ at time t will be denoted by $f(\vec{r},\vec{v},t)\, d\vec{r}\, d\vec{v}$. Let's say that there is no external field of force except at the boundary of V . Then the Boltzmann equation reads

$$\frac{\partial f}{\partial t} + \vec{v}\cdot\nabla_r f = \frac{\delta^2}{2}\int d\vec{\omega}\int d\vec{\ell}\,\left[\tilde{f}\tilde{f}_1 - f f_1\right]\left|(\vec{\omega}-\vec{v})\cdot\vec{\ell}\,\right| \tag{61}$$

And now I will explain what it means. You assume that a particle with velocity $\vec{v}$ (within $d\vec{v}$) and a particle with velocity $\vec{\omega}$ (within $d\vec{\omega}$) collide at the point $\vec{r}$. At collision, the center line of the particles is in the direction $\vec{\ell}$ (within $d\vec{\ell}$) which is a unit vector. $f(\vec{r}, \vec{v}, t)$ is the probability density that the first particle is at $\vec{r}$ with velocity $\vec{v}$ at time t. $f_1(\vec{r}, \vec{\omega}, t)$ refers in the same way to the second particle. The wiggles mean that you substitute for $\vec{v}$ and $\vec{\omega}$ the velocities after the collision. They are of course completely determined by the momentum and energy conservation laws. The integral over $d\vec{\ell}$ is a surface integral over the unit sphere. The factor $\left|(\vec{\omega}-\vec{v})\cdot\vec{\ell}\,\right|$ comes simply because we use elastic spheres, as Max Dresden has probably told you. If you have particles repelling each other according to some force law, then this thing becomes some function of the $|(\vec{w}-\vec{v})|$ and $|(\vec{w}-\vec{v})\cdot\vec{\ell}|$. The term $\vec{V}\cdot\nabla_{\vec{r}} f$ is called the streaming term. The term on the right hand side is called the collision term.

Now I would like to make some preliminary remarks. The equation we have written is the full Boltzmann equation -- the Boltzmann equation in phase space. Boltzmann actually derived this equation in two steps. He first derived what you might call the equation in velocity space alone. That is, assuming that the distribution in space is uniform. If this is so, then the gradient in (61) with respect to $\vec{r}$ is going to be zero. So the term $\vec{V}\cdot\nabla_{\vec{r}} f$ will not be there, and $\vec{r}$ becomes simply a parameter. You can just cross it out, and get then an equation involving only $\vec{V}$ and t:

$$\frac{\partial f}{\partial t} = \frac{\delta^2}{2}\int d\vec{w}\int d\vec{\ell}\left[\tilde{f}\tilde{f}_1 - f f_1\right]\left|(\vec{w}-\vec{v})\cdot\vec{\ell}\right| \qquad (62)$$

This is the Boltzmann equation in velocity space, which is valid only for spatially homogeneous systems. That means that the probability of finding a particle anywhere in the volume V is the same. This, of course, is a very uninteresting case from the point of view of hydrodynamics. In the case of hydrodynamics the primary purpose is to really show how the mass of the gas moves. But it was from this equation that Boltzmann derived the H-theorem. That covers the approach to thermal equilibrium, if the gas is already in spatial equilibrium.

It's a very simple but very interesting derivation which I will repeat in order to show you the analogy with what I have done for the simple model. Now what does $\frac{\partial f}{\partial t}$ represent? According to Boltzmann it is the total rate of change of the number of particles in a little volume of phase space. Now this change is due to two causes. One cause is streaming, and one cause is collision. Consequently, the total is the sum of the two. Now as you know very well, there are a lot of operations by which to combine them. You could multiply, divide, take logarithms or something else; why is it then that one takes the sum of the streaming and the collision terms? It's entirely unclear to me, and upon closer questioning of my physicist friends, it is also unclear to them. One simply assumes it. Actually you can easily see, if you think for a moment, that it cannot be true. Because streaming and collision cannot be really separated. After all, what are collisions in a mechanistic model? You have certain short range forces, and when two particles come close together a violent event takes place. That's a collision. But streaming is also a motion under the influence of the same forces, only in the range where the forces are somewhat less sharp, less powerful. Why you should separate them into these two phenomena which are clearly related to each other and make a sum is really not clear at all. In fact, there are other indications, mostly through work which to me is completely dark. The only man who understands it is Uhlenbeck, and in fact he even wrote out in detail the theory by Bogoliuboff, who derived a Boltzmann equation which has certain coupling terms between streaming and collision. This was done in some very formal way which as I said, I do not pretend to understand. Hence I do not intend to impose my ignorance on you. But it means that you should not get exactly a sum, but an extra term as well. If you follow the derivation that Max Dresden gave, it will somewhat change the equations of hydrodynamics. Probably, for very small flows and very small gradients this coupling term will be small and hence nothing will be changed. The really interesting problem, at least to me, is that I am unable to find a probabilistic model which will lead to the full Boltzmann equation. I will show you how one can very easily be led to the equation in velocity space, however.

And now I will betray a secret of why one can do this. Once we have spatial homogeneity, then we have a lack of specification and position. And consequently we have wide freedom to average over all possible positions. If you don't have spatial homogeneity, then the problem becomes over-determined. There's absolutely no room, or at least I can't find any room, to introduce a stochastic element. I don't know what's random anymore, and so I cannot find a stochastic model which will lead to the full Boltzmann equation. That's actually one of the problems that I think is very interesting, but which nobody takes very seriously because people want to draw conclusions from equations before they understand them. In my opinion it is an important problem to really understand in what sense (61) is a probabilistic equation.

In the case of Brownian motion, which I will speak about briefly sometime tomorrow, the collision operator is much simplier. It becomes the diffusion operator. It's perfectly understandable how the streaming is introduced; and that, of course, you might say can be used as an analogy. Since it's OK in Brownian motion, it must be all right for the similar equation here with a more complicated collision operator. But really the problem is now to cleanly derive equation (61) from a well defined stochastic model. I think it would be of some importance to do so, but since I don't know how to do it, I'm going to devote myself to the spatially homogeneous case.

I will show you how from a very simple stochastic model, which was already treated by Boltzmann, equation (62) can be derived. The most interesting thing is that the fundamental equation we are going to write -- the Master equation -- will be linear. Yet we can get the non-linear equation (62) from it

In fact, as you will see, the non-linearity is a fake in a certain sense. The equation only looks non-linear and one often wonders whether there are not many other non-linear equations which in the same sense are fakes.

Now I would like to set up for you the purely statistical approach to this problem of a spatially homogeneous gas not in thermal equilibrium. I have n particles, and let their velocities be $\vec{V}_1, \vec{V}_2, \ldots, \vec{V}_n$. Because the energy is exchanged only through elastic collisions, the kinetic energy stays fixed all the time -- since energy is conserved in collisions. And so we can say that

$$E = \sum_{j=1}^{n} \vec{V}_j^{\,2} \tag{63}$$

where I am going to assume that $E = n\sigma^2$. Now actually, this is already a bit of an assumption. It means that we assume, roughly speaking, that the energy per particle is fixed. It is important to notice that if you look upon each velocity as having three components, then (63) is the equation of a $3n$-dimensional sphere. Not being clairvoyant, I will draw it as a circle of radius $\sqrt{E}$:

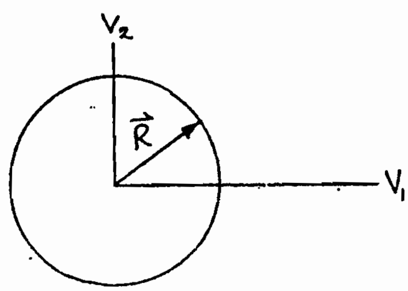

Figure 2

A point on this sphere defines completely the state of my system, because the state now is defined by giving only the velocities.

Now we must analyze what can happen to this point on the sphere. Most of the time nothing happens to it; but every once in a while a collision between two particles occurs which will change the state. Let me combine all the velocities into a great big vector, actually a $3n-$ dimensional vector,

$$\vec{R} = \begin{bmatrix} \vec{V}_1 \\ \vdots \\ \vec{V}_n \end{bmatrix}.$$

Now most of the time, $\vec{R}$ simply goes into $\vec{R}$, which means that nothing happens. But sometimes $\vec{R}$ will go into $\vec{R'}$, let's say, as a result of a collision between two particles. And of course they can collide in many different ways, so actually this $\vec{R'}$ covers a multitude of sins. The transformation $\vec{R} \rightarrow \vec{R'}$ can be quite a complicated thing.

Let us now try to calculate the probability of the transition $\vec{R} \rightarrow \vec{R'}$ I take a differential time element dt, a very short time, and will calculate the probability that the ith and jth particles $(i < j)$ will collide during this time. And moreover, that they will collide in such a way that their line of centers will be in a direction $\vec{\ell}$ lying in the solid angle $d\vec{\ell}$. This is a standard calculation of Boltzmann, and although I have not seen the notes of Max Dresden's lectures, he undoubtedly drew you a collision cylinder. It's almost impossible to give lectures on Boltzmann's equation without drawing a collision cylinder. I will not bore you with the details again, since all we need to do is to find the volume of the collision cylinder divided by the total volume V. It is

$$\psi_{ij}\, d\vec{\ell}\, dt = \frac{\delta^2}{V} \frac{\left|(\vec{V}_j - \vec{V}_i)\cdot\vec{\ell}\right| - (\vec{V}_j - \vec{V}_i)\cdot\vec{\ell}}{2}\, d\vec{\ell}\, dt \tag{64}$$

You might notice that $\left|(\vec{V}_j - \vec{V}_i)\cdot\vec{\ell}\right| - (\vec{V}_j - \vec{V}_i)\cdot\vec{\ell}$ is either zero or simply $2\left|(\vec{V}_j - \vec{V}_i)\cdot\vec{\ell}\right|$. It corresponds to the fact that if the velocities happen to point in the wrong way no collision will take place. Max Dresden may have written this as $\cos\theta$. It might be worthwhile to remind you that Boltzmann did not interpret this as a probability. He simply claimed that after multiplying by N_i and N_j this was the actual number of collisions taking place. But in reality it is a probability, because what do you do? You say that for the collision to occur the particle has to be in its collision cylinder. And since I assume that the spatial distribution is uniform then the ratio of the volume of the collision cylinder to the total volume of the gas is just the collision probability.

Notice that I have done exactly what I did on the other model. I assumed that a transition can take place, and I calculated the probability of this elementary event. Notice that I have already performed an averaging -- in calculating the probability (64). Now I want to find the probability that no collision occurs during the time dt. First of all let me integrate over $d\vec{l}$ to get the probability that a collision will occur between the ith and the jth particles regardless of where the line of center points is. Then if I sum this over all the pairs of particles I get the probability that a collision will occur between some pair of particles:

$$a\,dt = \sum_{i \leq j \leq n} \int \psi_{ij}\, d\vec{l}\, dt \tag{65}$$

Where I have called all this great big sum of integrals a. And hence $1 - a\,dt$ is the probability that no collision will occur.

Moreover, the probability that something happens is of the order of magnitude dt. That means that a collision is indeed a rare event. Stochastic processes in which you have the situation that something happens with probability proportional to dt are referred to as Poisson - like. The simplest such process is the Poisson stochastic process which you meet for instance in radioactive disintegration. Here the probability is $a\,dt$ that a particle will be emitted and is $1 - a\,dt$ that nothing happens.

Now let me write out a little bit more precisely what are my transitions, and then you will see the complete analogy with the previous model. I can say that I have the following situation: In time dt either $\vec{R}$ goes into $\vec{R}$ (into itself, no transition) with probability $1 - a\,dt$, or $\vec{R}$ goes to some other state. We must now write out what happens in such a transition. If the i and j particles collide then the transition is

$$\vec{R} \longrightarrow \begin{bmatrix} \vec{v}_1 \\ \vdots \\ \vec{v}_i + (\vec{v}_j - \vec{v}_i)\cdot\vec{\ell}\,\vec{\ell} \\ \vec{v}_j + (\vec{v}_j - \vec{v}_i)\cdot\vec{\ell}\,\vec{\ell} \\ \vec{v}_n \end{bmatrix} = \vec{R}' \tag{66}$$

What I have written in the ith and the jth places are simply the velocities after the collision. None of the other particles are affected. How do we find the velocities after collision? We simply solve the equations of conservation of momentum and conservation of energy, remembering that they are colliding in the direction $\vec{\ell}$. The relation (66) I will write

$$\vec{R}' = A_{ij}(\vec{\ell})\,\vec{R} \tag{67}$$

Where $A_{ij}(\vec{\ell})$ is simply the transition operator. It expresses what I am to do to $\vec{R}$ in order to get the transition to $\vec{R}'$.

An interesting and very simple observation is that this $A_{ij}(\vec{\ell})$ is a rotation. It is a rotation of a great, big sphere -- that is all it is. Actually, that's very easy to prove because the sum of squares (63) is the same before and after collision. I should have said that I have another conservation law here, namely that the total momentum is conserved. So really it is not a $3n$ - dimensional sphere but rather a $(3n-3)$-dimensional sphere. This is a little bit irksome, because you have walls of the container and collisions with the walls do not conserve momentum, as you very well know. Otherwise there would be no pressure on the walls. But that is a very minor point, and to avoid trouble all you have to do is a little work. Whenever you come to the wall you simply artifically re-introduce the particle back in the interior with the same velocity. If you wanted to put in the wall effect it would be simply too much writing. And besides, I am going to very soon consider with you a simplified model in which I'm going to violate conservation of momentum. Not because I don't like it, but because mathematically it's a complication and contributes comparatively little to the understanding of the general picture. So $A_{ij}(\vec{\ell})$ is a rotation of the $(3n-3)$ -dimensional sphere.

And now you can see what an individual gas does. It starts from some point on the sphere and occasionally a violent rotation will take place and it jumps to another point. But most of the time nothing will happen. Then after a long time it is again going to jump. Then again for a long time nothing will happen. You can look upon this particular scheme of evolution of a perfect gas as a random walk on a $(3n-3)$ -dimensional sphere. The random walk is described precisely by the probabilities I have calculated which tell me what the probability is of an elementary step.

Now what is the problem? It is the following: I am given $\varphi(\vec{R}, 0)$ the initial distribution of points (systems). For instance, if I know rather precisely the velocity then I can have a very sharp probability density around some point of the sphere. The question is, of course, how to find the distribution $\varphi(\vec{R}, t)$ at time t. You can see that the analogy with the previous model is almost complete. All I have to do is to simply write the analog of what I called the Master equation, which is nothing but the equation of propagation of probability. You remember, though, that my time variable was discreet, so that I was led to a difference equation. But now I have a continuous time variable so I am going to have a differential equation of the first order in time. This is namely the following

$$\frac{\partial \varphi(\vec{R},t)}{\partial t} = \sum_{1 \leqslant i < j \leqslant n} \int d\vec{\ell} \left\{ \varphi\left[A_{ij}(\vec{\ell})\vec{R}, t\right] - \varphi(\vec{R},t) \right\} \psi_{ij}(\vec{\ell}) \tag{68}$$

And this is the Master equation.

It's a very interesting equation. I especially appeal to the mathematicians in the crowd because it is an equation in which the operator acts on the independent variable inside the function. It's a rather strange looking equation and there are many very interesting properties, as you will see. But already you can see one thing intuitively. If I start with an arbitrary distribution subject to some smoothness conditions (that is necessary) then due to this random jiggling of the sphere I am going to see it spread out. In the limit, as time goes to infinity, the limiting distribution ought to be uniform over the sphere. I will anticipate myself a little bit and tell you that this is, in fact, the ergodic theorem specialized to this particular model. To show that it will eventually become uniformly smeared out, you have to prove something about these very special rotations $A_{ij}(\vec{\ell})$. They are actually six dimensional rotations because each collision only involves two particles. What you have to show is that they generate essentially the whole rotation group of the $(3n-3)$ -dimensional sphere. To be technically quite correct, they generate a transitive sub-group of the rotation group: Which means that you're able to get essentially from every point on the sphere to every other point on the sphere, or arbitrarily close to every other point on the sphere, by a combination of such rotations. That's a very well known condition in the theory of Markov chains. A Markov chain is called ergodic if you can go from every state to every other state. Here we must be able to go from one point on the sphere to any other point on the sphere performing, however, only these strange little rotations. That can be demonstrated, and I will speak about it a little bit later.

At the moment I want to call your attention to the following remarkable facts, just to show you what wonders can happen. Our Master equation (68) is a perfectly linear equation. There is nothing non-linear about it. Moreover, this linear equation embodies all the assumptions that Boltzmann ever used. These assumptions sit, of course, only in the formula for $\psi_{i\gamma}(\vec{\ell})$. On the other hand, Boltzmann came up with this equation (62). A non-linear equation! And the problem which then arises is how are they related? What is the relation between $\varphi(\vec{R}, t)$ and $f(\vec{v}, t)$? And what is the relation between the linear Master equation in many variables and the non-linear Boltzmann equation in very few variables?

In order to answer this question, I will have to simplify the model somewhat to produce a similar situation where I can prove everything rigorously. Unfortunately, not everything I am going to say for the simplified model I can prove for the real one. There are mathematical difficulties. But nobody doubts that with greater ingenuity than I have been able to show up to today it could probably be carried out. I will try to maintain most of the essential features, if not all the essential features, of this problem. But at the same time I will reduce the problem to one I can really analyze. It's a time honored procedure. If you can't solve the problem that you set out to solve, then try to simplify it -- but without throwing away the baby with the basket. That is the only condition: you must not over-simplify it.

My simplified problem is the following. The first simplification is that rather than three-dimensional velocities I will have one-dimensional velocities. That's a rather unimportant simplification. I will call the velocity of the ith particle x_i. And I will take for conservation of energy the condition:

$$x_1^2 + x_2^2 + \cdot \cdot \cdot \cdot \cdot + x_n^2 = n \tag{69}$$

So now the state of my system is described by a point $\vec{R} = \{x_1, \cdots, x_n\}$ lying on the n-dimensional sphere (69). The second simplification is that my transitions are of the form

$$\vec{R} \longrightarrow \vec{R}' = \begin{bmatrix} x_1 \\ \vdots \\ x_i \cos\theta + x_j \sin\theta \\ -x_i \sin\theta + x_j \cos\theta \\ \vdots \\ x_n \end{bmatrix} \tag{70}$$

which I will also write in the form $\vec{R}' = A_{ij}(\theta)\vec{R}$.

Now you see that I have changed the physical collision. Instead of a complicated six dimensional rotation my collision now produces a very simple two dimensional rotation. What I have written is that if particles i and j collide then the resulting velocities will be what you get by rotating through an angle theta. The angle theta plays the role of $\vec{\ell}$. And it is here that I violate the conservation of momentum. The energy is still conserved, but the momentum is not, except on the average. That is simply because this is essentially a one dimensional model; and one dimension is too poverty stricken for two conservation laws to hold at the same time.

And now I will make a real simplification, which makes this gas almost a Maxwell gas. I will say that the probability of the transition $\vec{R} \longrightarrow \vec{R}'$ is only a function of the angle theta, namely:

$$\text{Prob}\left\{\vec{R} \rightarrow \vec{R}'\right\} = \frac{1}{n} g(\theta)\, d\theta\, dt \tag{71}$$

where $g(\theta) \geq 0$. On occasion we may also assume that $g(\theta) = g(-\theta)$.
That's the usual assumption -- which is known in physics under the fantastic name of the principle of microscopic reversibility -- but for most purposes it is not needed, at least not for the mathematical development.

Now I am going to write down the Master equation, which assumes an extraordinarily simple form:

$$\frac{\partial \varphi}{\partial t} = \frac{1}{n} \sum_{1 \leq i < j \leq n} \int_{-\pi}^{\pi} g(\theta) \left\{ \varphi\left[A_{ij}(\theta)\vec{R}, t\right] - \varphi(\vec{R}, t) \right\} d\theta \tag{72}$$

The n in the denominator comes from the assumed probability (71). Of course, that's a parameter, you might say; but it is very important to include it here. Because I have to maintain the analogy with equation (64) which has the volume V in the denominator. Now the volume, of course, is proportional to the number of particles -- it is simply the number of particles times what's called the specific volume. So I always have in the denominator something proportional to the number of particles.

I can't give you anything that collides this way. I am simply imitating by mathematics the more complicated situation we described earlier. This artificial gas which I have constructed Professor Uhlenbeck once referred to in a lecture as a "caricature of a gas." This it is; but a caricature implies resemblance or else it would not be a good caricature. I am going to discuss it with you because one can really understand much better what is going on when some of the mathematical difficulties are dispensed with. Now before

I proceed, remember that the whole thing takes place on the sphere $\sum x_i^2 = n$. My first goal is to show how out of this Master equation I can get a non-linear equation; and then also to indicate what are the mathematical difficulties in treating the similar problem for the real case.

First of all I must define what $f(x,t)$ is. Recall that in the Boltzmannian language $f(x,t)\,dx$ is the probability that a particle has velocity x within the differential volume dx at time t. But when I say a particle that means I'm not allowed to distinguish particles. I'm not saying particle number 17, because it must be the same for all particles. Consequently, in order to place myself in an advantageous and perfectly realistic position, I will have to assume that at least at time $t = 0$ the particles are indistinguishable. And that means that $\varphi(\vec{R}, 0)$ is symmetric in the x's. It is easy to show, and you will certainly believe me, that if it is symmetric at time zero, it will remain so for all time. So it will follow also that $\varphi(\vec{R},t)$ is symmetric in the x's for all t. Now I am going to define the contracted distributions or contracted densities. The first contraction is

$$f_1^{(n)}(x,t) = \int_{x_2^2+\cdots+x_n^2 = n-x^2} \varphi(\vec{R},t)\,d\sigma_1 \qquad (73)$$

Let me explain what this means. I fix the velocity x of the first particle, and I integrate $\varphi(\vec{R},t)$ over what's left, the remaining sphere. $d\sigma_1$ is the surface element on this $(n-1)$-dimensional sphere; so that $\varphi(\vec{R},t)\,d\sigma_1$ is just the probability of finding the other velocities in $d\sigma_1$ when the velocity of the first particle is x. So if I integrate over all these other velocities I simply get the probability density that $X_1 = X$. Now you can define the second contraction

$$f_2^{(n)}(x,y,t) = \int\limits_{x_3^2+\cdots+x_n^2 = n-x^2-y^2} \varphi(\vec{R},t)\, d\sigma_2 \tag{74}$$

which is simply a joint probability density. Roughly speaking, it is the probability that one particle has velocity x and another particle has velocity y at time t.

All Boltzmann was interested in was $f_1(x,t)$, so he only tried to get an equation for it. How am I going to get such an equation? All I have to do is to integrate the Master equation (72) over all the variables but one. I will fix X_1 to be x and integrate over the complimentary sphere. The integration is entirely elementary and is really hardly worth bothering with. You obtain the following equation:

$$\frac{\partial f_1^{(n)}(x,t)}{\partial t} = \frac{n-1}{n}\int_{-\sqrt{n-x^2}}^{\sqrt{n-x^2}} dy \int_{-\pi}^{\pi} d\theta\, g(\theta)\left\{ f_2^{(n)}(x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta, t) - f_2^{(n)}(x,y,t)\right\} \tag{75}$$

I could go on, of course, and derive an equation for the joint density $f_2^{(n)}(x,y,t)$. I merely have to integrate the Master equation over all the variables but two. If I do this, I find that the equation for $f_2^{(n)}(x,y,t)$ involves the contracted density $f^{(3)}(x,y,z,t)$. Now I would like to call your attention to the following interesting feature which is the plague of the statistical mechanics of non-equilibrium phenomena. That is: The recursion goes the wrong way. Usually the recursion is from something complicated to something simpler. But here, to calculate f_1 you need f_2 ; if you want f_2 you need f_3 ; if you want f_3 you need f_4 ; and so forth. Instead of biting it's own tail, so to speak, and closing, it moves in the wrong direction. In turbulence, when you calculate simple correlations, double correlations, etc. you find the same phenomenon -- the double correlation involves the triple correlation; the triple involves the quadruple. And then people simply out of sheer desperation say let the quadruple be zero. Because they finally get somewhat impatient with the whole thing -- it's a never ending affair. But mathematics is one science where you are not allowed to become impatient; and you at least have to find out what's going on.

First of all let me notice the following. Suppose I let n go to infinity. Then I can erase the factor $\frac{n-1}{n}$, which simply becomes one. The integration over y is perhaps a little bit ticklish because the limits can be anything. But let us be optimistic and suppose they become $-\infty$ and $+\infty$. Then equation (75) will become:

$$\frac{\partial f_1(x,t)}{\partial t} = \int_{-\infty}^{+\infty} dy \int_{-\pi}^{\pi} d\theta\, g(\theta) \left\{ f_2(x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta, t) - f_2(x,y,t) \right\} \qquad (76)$$

And that looks extraordinarily like the Boltzmann equation. Except to get the Boltzmann equation out of this, you have to replace f_2 by a product of f_1's. If you now make this assumption that $f_2(x,y,t)$ is for some inexplicable reason given by

$$f_2(x,y,t) = f_1(x,t) \cdot f_1(y,t) \qquad (77)$$

then you can substitute this in the integral and there is your non-linear equation. And that's exactly the Boltzmann equation for this model. If you want the real Boltzmann equation, the honest-to-goodness one, then you can obtain it in the same way by integrating my old Master equation (68). Instead if these simple two-dimensional rotations you have the real rotations and that is all.

Now the question is, are we justified in making this assumption (77). Here I would like to call your attention to one very important fact. My Master equation from which everything has to be derived is not only a linear equation but it is also first order in time. It can be symbolically written in the following form:

$$\frac{\partial \varphi}{\partial t} = \Omega \varphi \tag{78}$$

It is perfectly well known how to write a formal solution for such an equation. It is simply written as

$$\varphi(\vec{R},t) = e^{t\Omega}\varphi(\vec{R},0) \tag{79}$$

with the usual understanding that you simply expand the exponential in a power series and interpret the powers of the operators in the usual way. Now this has the following immediate consequence: That once you have decided on $\varphi(\vec{R},0)$ then everything is completely and uniquely determined. Consequently you are not allowed to assume anything at time t. So you cannot make (77) an assumption. But you are allowed to assume it at time $t = 0$ because presumably the initial situation is up to you. So suppose I happen to be so clever that I have started with a distribution $\varphi(\vec{R},0)$ which has the property (77) at time zero. Then the question is, will this property maintain itself? And that is a crucial question. Unless the operator Ω is such that it will maintain the "factorizability" of the distribution, there is no possibility of getting the Boltzmann equation.

This question is answered in a theorem which is known by a very high sounding name, namely the "theorem of the propagation of chaos." And I will now state this theorem but will not prove it. First, I will call a distribution chaotic -- actually I prefer to say it has the Boltzmann property -- if the following holds for the contracted densities:

$$\lim_{n\to\infty} f_k^{(n)}(x_1,\ldots,x_k;0)=\prod_{\ell=1}^{k}\lim_{n\to\infty} f_1^{(n)}(x_\ell;0) \tag{80}$$

For those with a mathematical conscience, I will have to define what I mean by a limit of functions. I will not go into that in any detail; the easiest way to deal with it is to say that convergence is understood in the weak sense.

Now I will also say that the sequence of density functions has the Boltzmann property, or the property of chaos, if (80) holds for every k. Of course, the first question which arises is, are there such distributions? The answer is yes, there are. In fact I will tell you how to construct a big class of them. Then the theorem is: chaos persists forever. This means simply that you can replace 0 by t in relation (80) and still have it right. If you take care to establish the property at time $t = 0$, then it will maintain itself forever. And that, remarkably enough, is difficult to prove in the actual physical case. In our case, for the caricature of a gas, it isn't difficult to prove but is tedious. The theorem is undoubtedly correct in general and nobody doubts it. However, I would like to warn you about one thing. When someone says that chaos propagates you might say, well, certainly. After all, if you start with something which is chaotic and all that happens is that some collisions take place which, if anything, shake the whole thing up some more, then why shouldn't it propagate? But that is simply a verbal argument; and one is verbally misled. Chaos does not mean lack of order; it is a very specific property of the initial distribution and really means asymptotic independence -- because the content of (80) is simply that for very large n the velocities of the particles are essentially independent. And then the fact that this particular operator Ω is such that it perserves this property is, of course, an extraordinarily fortunate thing. One ought to be grateful for it, but one still ought to be surprised that it is so.

If you believe this theorem that chaos persists for all time, then of course it becomes reasonably easy to simply go with n to infinity and end up with the Boltzmann equation for all times. (We work again with weak convergence.) Now this is interesting. At first I actually thought that it had a greater significance than it has unfortunately proven to have. You see, the non-linearity of this equation is due simply to a very special choice of the initial condition. It is not something inherent in the problem. The fundamental problem is linear, but with a tremendous number of variables. The reason why we get a non-linear equation here is not because there is something non-linear in the mechanism. Rather, it is because we insist on starting from the initial distribution which has a very special structure (and only because we want to reconstruct Boltzmann's theory).

You might say that this immediately gives me a way of solving the non-linear Boltzmann equation. Because I certainly know what the solution of the Master equation is, at least formally. Then all I have to do is to be sure I prepare myself a proper initial distribution and then integrate the formal solution (79). That is correct; that is one way to get a solution of this thing. However, you don't gain very much by it because it is almost as difficult to xecute it as to solve the Boltzmann equation directly. Except, of course, that or this caricature of a gas the equation was so constructed that you can solve t explicitly if you want to. I am not going to go into the details because he calculations are somewhat laborious. But I want to underline the fact, hich really is pedagogically and mathematically most interesting, that we have ere a new origin of non-linearity. A man-made non-linearity! You didn't have o have it at all: It's the price you pay for having a contracted equation.

Incidentally, it goes to show how careful one must be if one wants o close such chains of equations (75). Because what one usually tries to do, hen one gets tired, is to say, all right, I will assume that the fourth one is xpressible in terms of the second one. But then you had better prove that if ou assume this in time $t = 0$ then it is always so; that is, it propagates. his is a point which is often overlooked. People make assumptions for the sake f getting an answer. But the really hard point is to prove that the equations ake sense.

On the other hand, all the general conclusions one wants to draw bout the approach to equilibrium can be gotten from the Master equation. You on't have to go over to the Boltzmann equation. In particular I will show you a proof of the H- theorem using the Master equation. It's a very special H-theorem, because it holds only for distributions which initially had the chaotic property. As a matter of fact, for mathematicians, the proof is immediate. ou must simply notice that the operator Ω is self-adjoint and negative efinite. That is all. Now let me prove it for you. (This proof goes in general; it goes also for the other Master equation (68), the realistic one.)

First, I will assume that all my functions are, to use the mathematical language, square integrable:

$$\int_{S_n} \left[\varphi(\vec{R})\right]^2 d\sigma < \infty \tag{81}$$

This already excludes initial distributions which are too detailed. For example, the delta function is not square integrable. I need some spread, some slight amount of fuzziness, so I put my initial distribution in L_2. The inner product is then simply defined as the integral over the surface of the n-dimensional sphere:

$$(\varphi, \psi) = \int_{S_n} \varphi(\vec{R})\psi(\vec{R})\,d\sigma \tag{82}$$

Now let me calculate $(\Omega\varphi, \psi)$ and prove that Ω is self-adjoint, that is $(\Omega\varphi, \psi) = (\varphi, \Omega\psi)$. For this purpose I will need to assume that $g(\theta) = g(-\theta)$ which, you remember, is called microscopic reversibility. Then we get

$$(\Omega\varphi, \psi) = \frac{1}{n}\int_{S_n} d\sigma \sum_{1\leq i\leq j\leq n} \int_{-\pi}^{\pi} g(\theta)\left\{\varphi(A_{ij}(\theta)\vec{R}) - \varphi(\vec{R})\right\} d\theta\, \psi(\vec{R})$$

$$= \frac{1}{n}\sum_{1\leq i\leq j\leq n} \int_{-\pi}^{\pi} d\theta\, g(\theta)\left\{\int_{S_n} d\sigma\, \psi(\vec{R})\varphi(A_{ij}(\theta)\vec{R}) - \int_{S_n} d\sigma\, \psi(\vec{R})\varphi(\vec{R})\right\}$$

Now I'm going to change variables. $A_{ij}(\theta)\vec{R}$ I'm going to call $\vec{R}'$. Then $\vec{R} = A_{ij}^{-1}(\theta)\vec{R}' = A_{ij}(-\theta)\vec{R}'$. No change in $d\sigma$ is needed because I'm simply making a Euclidian change of variable --- $A_{ij}(\theta)$ is a rigid rotation which preserves the element of integration. So, making these changes, we simply come out with:

$$(\Omega\varphi, \psi) = \frac{1}{n}\sum_{i\leq j} \int_{-\pi}^{\pi} d\theta\, g(\theta)\left\{\int_{S_n} d\sigma\, \psi(A_{ij}(-\theta)\vec{R}')\varphi(\vec{R}') - \int_{S_n} \psi(\vec{R})\varphi(\vec{R})\,d\sigma\right\}$$

$$(\Omega\varphi,\psi)=\frac{1}{n}\sum_{i\leq j}\int_{-\pi}^{\pi}d\theta\, g(\theta)\left\{\int_{S_n}d\sigma\left[\psi(A_{ij}(\theta)\vec{R}')\varphi(\vec{R}')-\psi(\vec{R}')\varphi(\vec{R}')\right]\right\}=(\varphi,\Omega\psi) \tag{83}$$

So the operator Ω is self-adjoint. In the more realistic case (see equation (68)) the proof goes through in the same way. In this case the $A_{ij}(\vec{\ell})$ are involutions (they are their own inverses) and that makes the proof even simpler.

Now why is Ω negative definite? You will see why in a moment. First of all, to prove this property, all I have to do is show that

$$(\Omega\varphi,\varphi)=\frac{1}{n}\int_{S_n}d\sigma\sum_{i<j}\int_{-\pi}^{\pi}g(\theta)\,d\theta\left\{\varphi(A_{ij}(\theta)\vec{R})\varphi(\vec{R})-\varphi^2(\vec{R})\right\}$$

is never positive. This requires a slight trick, but very slight. I am going to replace $\varphi^2(\vec{R})$ by $\frac{1}{2}\left[\varphi^2(\vec{R})+\varphi^2(A_{ij}(\theta)\vec{R})\right]$. Now I claim I haven't changed anything, because the integral of $\varphi^2(A_{ij}(\theta)\vec{R})$ is just the same as the integral of $\varphi^2(\vec{R})$. It is simply the same change of variables that we made before, just a rigid rotation. But now notice that in having done this, I have a negative square, $-\frac{1}{2}\left[\varphi(A_{ij}(\theta)\vec{R})-\varphi(\vec{R})\right]^2$ for my integrand. Of course, I have to multiply by $g(\theta)$, but this is non-negative. Consequently, I can conclude that

$$(\Omega\varphi,\varphi)\leqslant 0 \tag{84}$$

If you really watch it carefully, you will notice that I didn't need here the principle of microscopic reversibility. I only needed it to prove that the operator was self-adjoint.

Now this immediately implies the H-theorem. Take the Master equation (78), multiply by $\varphi(\vec{R})$, and integrate over the sphere. Then you get

$$\frac{1}{2}\frac{d}{dt}\int_{S_n}\varphi^2(\vec{R})\,d\sigma = (\Omega\varphi,\varphi) \leqslant 0 \tag{85}$$

And that proves it. It tells you that there is a quantity, H if you wish, whose time evolution is one-directional.

By a somewhat more complicated argument one can show that not only (85) holds but more generally

$$\frac{d}{dt}\int_{S_n}\varphi^{\alpha}(\vec{R};t)\,d\sigma \leqslant 0 \ , \quad \alpha > 1$$

and hence also

$$\frac{d}{dt}\int_{S_n}\varphi(\vec{R};t)\log\varphi(\vec{R};t)\,d\sigma \leqslant 0$$

If φ has the property of chaos the latter relation is intuitively equivalent to the usual H-theorem (2). Unfortunately, I am unable to establish the equivalence rigorously.

We now return to the operator Ω which we have shown is self-adjoint and also negative definite. As everybody knows the thing of interest is its spectrum. Now the spectrum is real, since the operator is self-adjoint; and, because the operator is negative definite, it must lie to the left of zero. Now zero is surely a member of the spectrum, because you can very easily see that $\varphi = C$, a constant, is an eigenfunction belonging to the eigenvalue zero. One can make some other statements about this spectrum, but nobody knows it completely. In particular it's not known -- although one can almost bet one's last dollar on it -- that as n goes to infinity the spectrum does not close up on zero. If it doesn't, if it is really cut off, then by simply looking at the operator solution you can say much more immediately. Because then it immediately follows that this solution (79) will decay exponentially to a constant. The constant comes from the fact that it is the eigenfunction with eigenvalue zero. That's the equilibrium state. The next eigenvalue, the one closest to zero, gives you the speed of the approach to equilibrium.

Unfortunately, neither I nor anybody else has been able to prove that zero is not an accumulation point of the spectrum. We are able to prove it very indirectly if you start with a factorized distribution, i.e., your initial distribution is chaotic. But that may be due to the fact that a factorized distribution is automatically orthogonal to all the modes with eigenvalues lying in the vicinity of zero. It is simply not known whether there are such modes or not. In any case, for this special choice of the initial distribution, we can show rigorously that the decay is exponential. To what? To a constant, which of course means the uniform distribution on the sphere.

Now the question is, what is the one-dimensional contraction of the uniform distribution, the equilibrium distribution. If you calculate it -- and it's an old calculation already made by Maxwell -- you get

$$f_1^{(n)}(x) = \frac{\left(1-\frac{x^2}{n}\right)^{\frac{n-3}{2}}}{\int_{-\sqrt{n}}^{\sqrt{n}}\left(1-\frac{x^2}{n}\right)^{\frac{n-3}{2}}dx} \tag{86}$$

The calculation is entirely trivial; you only have to know a little bit about the geometry of the sphere. And now, as n goes to infinity, look what happens to it:

$$\frac{\left(1-\frac{x^2}{n}\right)^{\frac{n-3}{2}}}{\int^{\sqrt{n}}\left(1-\frac{x^2}{n}\right)^{\frac{n-3}{2}}dx} \longrightarrow \frac{1}{\sqrt{2\pi}}e^{-x^2/2} \tag{87}$$

which is just the Maxwell distribution. This is probably by far the most satisfying derivation of the Maxwell distribution. You simply ask for the one-dimensional contraction of the uniform distribution on the sphere. The two-dimensional contraction, needless to say, will simply involve the product of $e^{-x^2/2}$ and $e^{-y^2/2}$. All that will remain to do in the next lecture will be to make a few general remarks about what kind of distributions are chaotic, and how to construct them.

Now I would like to make a few final comments, and answer some of your questions. I have tried to build up, in a perfectly consistent way, the early stages of kinetic theory on a simple stochastic model. This proved to be interesting because it betrayed the nature of the non-linearity in the Boltzmann equation. It also made this equation philosophically rather peculiar. Because if you believe in it you must ask yourself why nature prepares for you at time zero such a strange factorized distribution. Because otherwise you can't get Boltzmann's equation. You must somehow reconcile yourself that for some reason the systems with which you deal are already so prepared as to have this property. There is a current theory, which nobody can prove, because nobody can even properly state it, that most distributions are already factorized -- at least most symmetric distributions. This would mean that it's really very difficult to have a distribution on the n-dimensional sphere which is symmetric in all the variables and which is not at least approximately factorized. That of course would answer it, if it were so. But you must say what you mean by "most" of them, and I don't know the answer. There are statements made in the literature that somehow any distribution decays very rapidly into a factorized one. From there on, of course, it would go by the

Boltzmann equation. But that is wrong. It was proved by Dr. R. Brout, one of my young colleagues at Cornell in physics, that if you start with a non-factorized distribution you'll never decay into the factorized kind. At least the time it takes you is of the same order of magnitude as the time to reach equilibrium. You will not produce chaos out of order unless you wait long enough. You have to wait very long. But if you have chaos in the beginning, it will stay with you.