FIFTH LECTURE

It may be useful to say a few words, although it has no direct relevance to the sequel, why all the interest in foundations has arisen. It comes not only from the desire for a more thorough understanding of what is going on. There is also the purely practical desire to see how Boltzmann's theory can be extended. The Boltzmann equation is based on the assumption of a dilute gas. The assumption of dilution comes from assuming only binary collisions. Even in this model I needed the assumption from the very beginning that only pairs of particles can collide and never that three or more of them can come together. Some of the consequences of Boltzmann's and Maxwell's theory were somewhat hard to take. One of them was the independence of viscosity and pressure.

It is probably dear to the heart of some of you, because you have to deal with that case, that viscosity is highly dependent on pressure. But it was an exact conclusion which can be derived for, say, a Maxwell gas that they are independent. This is something which is mathematically proved, a conclusion of a certain assumed model. But it is perfectly well known that there is some dependence of viscosity on density (and hence pressure) in gases. It would be interesting in principle to know how to calculate it. The model would have to be changed, and this correction must clearly come from higher order collisions.

So here's the problem, how to extend the theory to take into account higher order collisions. This nobody really knows yet how to do. If I were to try to do what I did here, I would have to know how to calculate the probability of a triple collision. And, of course, I can't do it without going seriously into the physical situation and examining what happens in a three body problem.

However, serious attempts are made to derive from Liouville's equation by an appropriate method of averaging, the next approximation to the Boltzmann equation. The whole situation is extremely dark. There are schemes which are more or less appealing, more or less convincing. But still only a beginning has been made in that direction. In particular, what seems to be reasonably well established is that you will have to go out of the realm of Markov processes. All we have looked at still have a perfectly good, straight-forward Markovian behavior. But as soon as you begin to introduce the higher order collisions the equations will not be any more of such simple form. That however is a vast and obscure subject.

Now I'd like to take up several related points because although they are perhaps of no great importance they illustrate certain things. They illustrate a technique which might be useful in other connections. The first point I would like to say a few words about is whether there are any chaotic distributions. I mean, it's all good and well to say that chaos propagates and so forth, but you may be actually dealing with an empty situation. There might not be any. So the very first thing I would like to demonstrate to you is that there actually are distributions which have this property. Here is a trivial one: The uniform distribution on the surface of the sphere. I already told you what the first contraction is; it is given by equation (86). But that is not interesting. You would expect the final equilibrium to have molecular chaos because that's really the state where everything is as mixed up as possible. So it's interesting to exhibit other functions which have this behavior. As a matter of fact the easiest and the simplest guess happens to be the right one. Namely, you take an arbitrary function, say $C(x)$ (when I say arbitrary I don't really mean it; you will see what conditions to put on) and define

$$\varphi_n(\vec{R}) = \frac{\prod_{k=1}^{n} C(x_k)}{\int_{S_n} \prod_{k=1}^{n} C(x_k)\, d\sigma} \tag{88}$$

Now I am going to demonstrate that this indeed does have the property of chaos. You can say there is almost nothing to prove, because you really already have a product of the right form. But this is slightly misleading, because don't forget that there is the condition

$$x_1^2 + x_2^2 + \cdots + x_n^2 = n \tag{89}$$

which states that the energy is conserved. This condition ties the x's together and destroys the independence. So in order to prove that chaos is to be propagated, I must calculate the contractions of φ_n and show that the contractions do indeed satisfy condition (80) in the limit as n goes to infinity. The reason why I want to do that is because the method which one uses is commonly in use in other problems of statistical mechanics. It is something which leads to the method of steepest descent, a very neat, simple trick which is useful to know.

First of all, I would like to determine the asymptotic behavior for large n of the denominator. Once I have that it will be very easy to answer the whole thing. Now you define the following function:

$$F_n(r) = \int\limits_{x_1^2+\cdots+x_n^2=r^2} \prod_{k=1}^{n} c(x_k)\, d\sigma \tag{90}$$

and try to calculate essentially the Laplace transform of this function:

$$\int_0^\infty e^{-sr^2} F_n(r)\,dr = \int_{-\infty}^{\infty} dx_1 + \cdots \int_{-\infty}^{\infty} dx_n\, e^{-s(x_1^2+\cdots+x_n^2)}\, c(x_1)\cdots\cdots c(x_n) \tag{91}$$

You see that it's almost a Laplace transform; you could change the variable and actually make it a Laplace transform of a slightly more complicated function. In fact it is just:

$$\frac{1}{2}\int_0^\infty e^{-s\rho}\,\frac{F_n(\sqrt{\rho})}{\sqrt{\rho}}\,d\rho \tag{92}$$

Now I can very easily calculate this -- and it ought to remind you of your days in advanced calculus when you calculated the integral of $e^{-(x^2+y^2)}$ by exactly the same argument. It gives:

$$\frac{1}{2}\int_0^\infty e^{-s\rho}\,\frac{F_n(\sqrt{\rho})}{\sqrt{\rho}}\,d\rho = \left[\int_{-\infty}^{\infty} e^{-sx^2} c(x)\,dx\right]^n \tag{93}$$

Now you have the Laplace transform of some function, and consequently can use the inversion formula. You have all learned it sometime in your career -- I will assume that you know it -- the complex inversion formula. We will make a very nice, neat application of it. The complex inversion formula is one of the most beautiful and most useless things in mathematics, except for just such purposes. The inversion is

$$\frac{F_n(\sqrt{\rho})}{2\sqrt{\rho}} = \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} e^{z\rho}\left[\int_{-\infty}^{\infty} e^{-zx^2} c(x)\,dx\right]^n dz \tag{94}$$

In case the proof escapes you at the moment, you can look it up in any standard text on Laplace transform. Now you look at it and say, well I'm not really interested in this function as such. I am only interested in its value on the sphere $r^2 = n$. That means I want ρ to be n. So now I obtain the formula:

$$F_n(\sqrt{n}) = \frac{\sqrt{n}}{\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\left[e^{z}\int_{-\infty}^{\infty} e^{-zx^2} c(x)\,dx\right]^n dz \tag{95}$$

And it is from this formula -- useless as it usually is -- that I can determine the asymptotic behavior by the method of steepest descent. In fact, the method of steepest descent is always naturally invoked when you have a function raised to a very high power which you integrate. Now how do I determine the saddle point? I must write the integrand in the form of an exponential:

$$e^{n\left\{z+\log\int_{-\infty}^{\infty} e^{-zx^2} c(x)dx\right\}}$$

and differentiate the exponent with respect to z. I will engage here in justifying the method of steepest descent. In this case, it's quite easy although somewhat lengthy. Let me just write down the determining equation for the saddle point. It is just the derivative set equal to zero:

$$1 - \frac{\int_{-\infty}^{\infty} x^2 e^{-z_0 x^2} c(x)\,dx}{\int_{-\infty}^{\infty} e^{-z_0 x^2} c(x)\,dx} = 0 \qquad (96)$$

z_0 is the saddle point. And now you must assume, because I unfortunately cannot prove it, at least in general, that there is a real solution. If there is one, then it's quite easy to prove that it must be unique. I will not worry about that. You can check it for yourself for a lot of functions.

Now what we are going to do is to move the line of integration so as to pass it through the saddle point. That means we put $\gamma = z_0$. Then we change variables and set $z = z_0 + i\frac{\xi}{\sqrt{n}}$. After we do this, we immediately get

$$F_n(\sqrt{n}) = \frac{e^{nz_0}}{\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} e^{-z_0 x^2} e^{i\xi(1-x^2)/\sqrt{n}}\, c(x)dx\right]^n d\xi \qquad (97)$$

Now what you do is to expand in a power series in ξ :

$$e^{i\xi(1-x^2)/\sqrt{n}} = 1 + \frac{i\xi(1-x^2)}{\sqrt{n}} + \frac{(i\xi)^2(1-x^2)^2}{2n} + \cdots\cdots \tag{98}$$

This still has to be multiplied by $C(x)e^{-Z_0x^2}$ and integrated, but we can do it term by term. From the first term you simply get some number

$$A = \int_{-\infty}^{\infty} e^{-Z_0x^2} C(x)\,dx \tag{99}$$

The next integration, using the second term, vanishes because of (96). But the next one doesn't vanish anymore. It is going to be $-\frac{\xi^2}{2n}B$ where

$$B = \int_{-\infty}^{\infty} e^{-Z_0x^2} C(x)x^2\,dx \tag{100}$$

The remaining terms will not contribute anything in the limit as n goes to infinity. We will not bother with them. So the integration over x just gives me $A - \frac{\xi^2}{2n}B$. This thing must now be raised to the nth power:

$$\left(A - \frac{\xi^2}{2n}B\right)^n \sim A^n e^{-B\frac{\xi^2}{n}}$$

This must then be integrated over ξ . But the integral of this is easy so the asymptotic behavior I am after is simply:

$$\frac{F_n(\sqrt{n})}{\sqrt{n}} \sim \frac{(Ae^{z_0})^n}{\pi} \int_{-\infty}^{\infty} e^{-BX^2} dx \tag{101}$$

Looking back at equations (88) and (90) you see that I have only discovered the asymptotic behavior of the denominator. Except for a constant, it behaves like the nth power of something. Now you can get everything else by means of a simple trick. You can make almost exactly the same argument to find the asymptotic behavior of

$$\int_{S_n} g(x_1)\, h(x_2)\, \varphi_n(\vec{R})\, d\sigma$$

where $g(X)$ and $h(X)$ are two arbitrary functions. If I take $h(X_2)$ to be one and $g(X_1)$ to be $\delta(X-X_1)$ then this will give me the one-dimensional contraction, except that I must divide by the denominator of (88) which I now know. I can also choose my arbitrary functions so as to get me the second contracted density $f_2^{(n)}(X, y)$. This is a trick which is very often used. To integrate over all the variables except X_1 and X_2 -- which I have to do to get $f_2(X, y)$ -- may be inconvenient. It's too sharp when you fix $X_1 = X$ and $X_2 = y$, so you simply multiply by arbitrary functions and then integrate. From the result you can gather what the answer is and the calculations are much easier.

Now you perform exactly the same calculation; the only modification is going to be that these functions $h(X)$ and $g(X)$ come in. You can repeat the argument without me, because it is just the same. And believe it or not, what you are going to come out with is the following:

$$\frac{1}{\sqrt{n}}\int_{S_n} g(x_1)h(x_2)\varphi_n(\vec{R})\,d\sigma \sim \frac{(Ae^{z_0})^n}{\pi}\cdot\int_{-\infty}^{\infty} e^{-Bx^2}dx$$

$$\cdot\left[\int_{-\infty}^{\infty} g(x)e^{-z_0x^2}c(x)\,dx\right]\cdot\left[\int_{-\infty}^{\infty} h(y)e^{-z_0y^2}c(y)\,dy\right]\Bigg/\left[\int_{-\infty}^{\infty} e^{-z_0x^2}c(x)\,dx\right]^2$$

So this gives me the asymptotic behavior. Now all you have to notice is that you are through. Because putting h equal to one and making g the delta function, and dividing by the denominator (101) you get

$$c^*(x) = \frac{e^{-z_0x^2}c(x)}{\int_{-\infty}^{\infty} e^{-z_0x^2}c(x)\,dx} \tag{102}$$

That's the one-dimensional contraction. Now you can make h and g delta functions -- that means you integrate over all variables but two -- and you find that the two-dimensional contraction is simply the product:

$$c^*(x)\,c^*(y) \tag{103}$$

The three-dimensional contraction, and the other higher order contractions, can be found in the same way. They are just given by (80), so my distribution (88) does indeed have the Boltzmann property.

Now it's an interesting thing that, as you see, the contraction of such a distribution is not simply $C(X)/\int_{-\infty}^{\infty} C(x)\,dX$. But clearly it couldn't be this because I could take $C(X)$ to be a function which goes to infinity fast enough that the integral doesn't converge. But, after all, the contraction must be a density function; and that means a function whose integral happens to be one. So something must save it. And it is very interesting that the thing which saves it is $e^{-Z_0 X^2}$. Z_0 is derived by solving the equation (96). This equation has a very vivid physical significance, as you can now see. In fact it means that the contracted distribution must be such that the variance is equal to the average energy per particle.

The second point which I wish to discuss is much more interesting. We already saw that if I start with chaos I can replace, for the purpose of my study, the Master equation by the non-linear Boltzmann equation. We claimed, and demonstrated to our satisfaction, that this was due to starting with a very special initial distribution, namely the chaotic one. But there are many other starting distributions. There is one very interesting one which I will now briefly discuss with you. I will start at time zero with a distribution

$$\varphi_n(\vec{R}, 0) = \frac{C(x_1)}{\int_{S_n} C(x_1)\,d\sigma} \tag{104}$$

Let us see what it means. This distribution depends only on one variable -- as far as all the other variables are concerned it's a constant. It's a perfectly good function. It is positive or non-negative if I have chosen it to be so; and certainly its integral is one. Physically it means, roughly speaking, that all the particles but the first one are already in thermal equilibrium. You simply take one particle out of equilibrium. Let's say you shoot an extremely fast particle in, while the rest of the ensemble, the rest of the gas, is in equilibrium.

It's now a very easy calculation to find the contractions of this distribution. You simply need the surface area of the $(n-1)$ -dimensional sphere of radius $\sqrt{n-x^2}$. Let me just tell you what the contractions are, because you can see it very easily. The contractions are:

$$f_1^{(n)}(x_1,0) = \frac{c(x_1)\left(1-\frac{x_1}{n}\right)^{\frac{n-3}{2}}}{\int_{-\infty}^{\infty} c(x_1)\left(1-\frac{x_1}{n}\right)^{\frac{n-3}{2}} dx} \sim \frac{c(x_1)\, e^{-x_1^2/2}}{\int_{-\infty}^{\infty} c(x_1)\, e^{-x_1^2/2}\, dx} \tag{105}$$

$$f_1^{(n)}(x_2,0) = \frac{\left(1-\frac{x_2}{n}\right)^{\frac{n-3}{2}}}{\int_{-\infty}^{\infty} \left(1-\frac{x_2}{n}\right)^{\frac{n-3}{2}} dx} \sim \frac{1}{\sqrt{2\pi}}\, e^{-x_2^2/2} \tag{106}$$

For the other particles, the contractions are the same as this one for x_2 . It's already the Maxwell distribution you see. When I say that one particle is out of equilibrium, that's not quite true. Remember that the function $\varphi(\vec{R})$ refers to a whole swarm of possible systems. On the energy surface $\sum x_k^2 = n$

each point corresponds to a system. And $\varphi(\vec{R})$ actually measures the probability of picking, or choosing, a particular one of these systems. The way to think about it is that I have a lot of boxes, each filled with gas, and all in equilibrium. In each of these boxes you put in one particle, a stranger. This is always the difficulty in statistical mechanics. After all, you are always interested in what happens to one particular gas, but you agree to consider ensembles instead. So the density $\varphi(\vec{R})$ really refers to many systems -- it is a distribution over the ensemble. In loose language, what we have in this extreme case is simply a gas in equilibrium into which a particle is shot which has not quite the proper average velocity. But I cannot speak of one particle not having the right average velocity and so you simply think in terms of an ensemble.

Now we want to see how equilibrium gets itself established. You can still say that the Master equation is valid. The only difference between this situation and the other one is that they have different initial conditions. On the other hand, the whole interest will clearly center about the contraction of the first particle. Because essentially nothing happens to all the others. Now what is then the equation which governs the evolution of the first contracted distribution? That's also a Boltzmann' equation -- at least it is known as a Boltzmann equation -- but this time it is linear. You can easily derive it by again integrating the Master equation (72). There is no difficulty at all. I will drop the subscript one and take the limit $n \to \infty$. It then reads:

$$\frac{\partial f(x,t)}{\partial t} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dy \int_{-\pi}^{\pi} g(\theta) \left\{ f(x\cos\theta + y\sin\theta, t) e^{-\frac{(-x\sin\theta + y\cos\theta)^2}{2}} - f(x,t) e^{-y^2/2} \right\} d\theta \qquad (107)$$

That's the equation, the linear Boltzmann equation. You see that it looks very much like the other one, in a way. But now you see you have an entirely different situation. Because now you watch particle number one. And while its own distribution of velocity changes it has no effect on the medium. In fact the distribution of velocities of all the particles in the medium remains Maxwellian regardless of the collisions. That means there is no effect on the medium. After all, there is only one particle which is out of step and $n-1$ already nicely in equilibrium. In this whole mess, having one out of step doesn't make any difference. Whereas in the case where you started with a chaotic distribution, each time you perform a collision you also change the distribution of the medium.

Now this linear Boltzmann equation belongs to a very revered class of equations. For instance you encounter such equations in diffusion problems. Also in Brownian motion, except there the operator is much simpler due to the fact that the Brownian particle is usually very heavy. Hence when it suffers a collision with the light particles of a gas the velocity does not change very radically. But here the velocity in a collision changes from x to $x\cos\theta + y\sin\theta$ when it collides with a particle of velocity y. Look at the tremendous change in velocity for some angles of collision! This is what you might call a violent change. But if our particle were very heavy and the particle with which it collides very light, then of course the laws of conservation of momentum and energy in a collision will produce an operator in which the changes in velocity will be small. It's perfectly clear that if a heavy sluggish thing is occasionally being pricked by a little one, it will not change its velocity very much. In the limit, as the ratio of the masses goes to infinity, such an operator becomes the second derivative plus certain first derivative terms which are related to friction. This of course then leads to a diffusion equation. A diffusion equation is always a limiting case of a Boltzmann equation for a linear situation, when the ratio of the masses becomes infinite. It is not really a new, separate equation.

So the linear Boltzmann equation is another one included in this pattern, the pattern of development from the Master equation. Finally, in the literature on the Boltzmann equation, a third Boltzmann equation enters. It's known as the linearized Boltzmann equation -- not linear but linearized. It comes from the non-linear equation I spoke about yesterday by means of a device which I will show you. As a matter of fact, I am going to show you also a very interesting error that is being made by everybody. This linearized equation can be introduced entirely outside the context of the Master equation by simply looking at the Boltzmann equation:

$$\frac{\partial f}{\partial t} = \int_{-\infty}^{\infty} dy \int_{-\pi}^{\pi} d\theta\, g(\theta) \left\{ f(x\cos\theta + y\sin\theta, t)\, f(-x\sin\theta + y\cos\theta, t) - f(x,t)\, f(y,t) \right\} \quad (108)$$

Now we make the usual argument that physicists are so fond of. We know from the H-theorem that $f(x,t)$ approaches, as time goes to infinity, the Maxwell-Boltzmann distribution. Now if you are near equilibrium, if you are very near it, you can write $f(x,t)$ as follows:

$$f(x,t) = f_0(x)\left[1 + p(x,t)\right]$$

Where $f_0(x)$ is the Maxwell-Boltzmann distribution and $p(x,t)$ is a correction which ought to be small compared to one. This is then put in the non-linear Boltzmann equation (108) and second order terms are neglected. So a linear equation results for $p(x,t)$. This linearized equation looks as follows:

$$\frac{\partial p}{\partial t} = \int_{-\infty}^{\infty} f_0(y)\,dy \int_{-\pi}^{\pi} d\theta\, g(\theta) \Big\{ p(x\cos\theta + y\sin\theta, t) + p(-x\sin\theta + y\cos\theta, t) - p(x,t) - p(y,t) \Big\} \tag{109}$$

We have with abandon and pleasure thrown away all the second order terms and now have a linear equation which we can handle.

At last, we are in the realm of linear operators and can speak about eigenfunctions and eigenvalues! In fact, my example was so nicely designed that the eigenfunctions come out to be Hermite functions. While one can calculate them, one can easily see it before hand, by "pure thought," that it will be so. Looking at the Master equation, you see that the eigenfunctions there are spherical harmonics on the n-dimensional sphere. We are dealing with the contraction, and the contraction of a spherical harmonic is a Gegenbauer polynomial. In the limit $n \longrightarrow \infty$ the Gegenbauer functions when properly normalized are known to go into Hermite functions. This application of "pure thought" is quite impressive, but really very simple.

Now the interesting thing about the linearized Boltzmann equation:

$$\frac{\partial p}{\partial t} = \Lambda_1 p \tag{110}$$

is this operator I call Λ. It determines the decay in time. You might say the last stages of decay, because you are near equilibrium and you are a tired old man. That's the behavior being described here. Now one can calculate the spectrum for this operator and it turns out to be a very simple one. It's a discrete spectrum, but one which approaches a finite limit. That's an interesting point. The finite limit simply happens to be:

$$-\int_{-\pi}^{\pi} g(\theta)\, d\theta \tag{111}$$

The eigenvalues congregate or accumulate at this value, the total cross-section, if you like. Zero belongs to the spectrum, which you can immediately guess from physical grounds. There are actually two eigenfunctions in this case. Zero is a degenerate eigenvalue. It's a lovely observation, which I think is due to Uhlenbeck, that the degree of degeneracy of the zero eigenvalue is always equal to the number of conservation laws. It is very trivial to prove once one notices it. In this case, two things are conserved: The number of particles and the energy. And consequently the eigenvalue zero for this equation will be of double degeneracy. For the real case, for the Maxwell gas, there is five-fold degeneracy. Because there we have exactly five conservation laws: particles, energy, and the three components of momentum. For the linear Boltzmann equation (107) zero is also an eigenvalue. It has to be, because that corresponds to the final equilibrium distribution. But it is a simple eigenvalue because only the number of particles is conserved. The energy is not conserved, at least in single collisions. The total energy is conserved, all right, at least on the average. But that doesn't matter because you're watching only this one particle.

Anyway, in this linearized case, zero is doubly degenerate. And this will give you two decay modes. Now the following question arises. After all, we know that somehow in the general non-linear case you get a decay toward equilibrium. So what is the relation between the decay in the final stages and the decay of the system far from equilibrium? You might think, because linearization is an approximate procedure, that the relaxation times will only be approximations. The interesting thing is that this is not so, at least for the first few relaxation times. This is quite easy to see on this example, although I don't quite know how to prove it in general. Suppose the eigenvalues of the linearized operator Λ_1 are $\mu_1, \mu_2, \ldots$ (zero I exclude). Then from the solution of equation (110) we see that $f(x,t)$ is a linear combination of exponentials with the eigenvalues in the exponents:

$$f(x,t) = \varphi_0(x) + \sum_k \varphi_k(x) e^{\mu_k t} \tag{112}$$

The first term corresponds to equilibrium. Then there is the slowest decaying term, that's $\varphi_1(x) e^{\mu_1 t}$. And so it goes on, with the faster decaying terms coming later in the series.

The solution of the non-linear equation (108) is also a combination of exponentials. But alas, more exponents will enter; in fact, all the linear combinations of the μ_k with non-negative integral coefficients. That's easy to see formally, as I will show you in the next lecture. In our particular case it can be rigorously justified because you can calculate everything explicitely.

Now notice what this means. The lowest decay mode, namely the one described by μ_1 , is the same. But the next slowest may not be the one with μ_2 . It may, instead, be the one with a time constant $2\mu_1$. Now this has an interesting implication. It makes no sense to use the full solution of the linearized equation as an approximation to the solution of the non-linear one. Only the first two terms -- the equilibrium one and the slowest decaying mode -- are for sure the same. Beyond that point there may be linear combinations which contribute more than the terms you keep. The non-linearity begins to set in after $2\mu_1$. It is inconsistent to keep any μ_k beyond this point. This is a mistake made very often, even by very good people.