SIXTH LECTURE

I start again with the non-linear Boltzmann equation (108). To solve this equation we will use a scheme which Hilbert used on the full equation. The idea is simply to write formally

$$f(x,t)=f_0(x)\left\{1+\epsilon p_1(x,t)+\epsilon^2 p_2(x,t)+\cdots\right\} \tag{113}$$

f_0 is simply the equilibrium distribution $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$. ϵ is an artificial parameter. Now you feed this expansion into both sides of the equation and compare coefficients of like powers. When you do that you are going to obtain a hierarchy of equations. For the first one you get

$$\frac{\partial p_1}{\partial t}=\Lambda p_1 \tag{114}$$

I will also write out the next one:

$$\frac{\partial p_2}{\partial t}=\Lambda p_2+\left[p_1,p_1\right] \tag{114'}$$

where I have used what I call the Boltzmann bracket. It is defined by

$$\left[\varphi,\psi\right]=\int_{-\infty}^{\infty}f_0(y)dy\int_{-\pi}^{\pi}g(\theta)d\theta\left[\varphi(x\cos\theta+y\sin\theta)\psi(-x\sin\theta+y\cos\theta)-\varphi(x)\psi(y)\right] \tag{115}$$

From here on, all the equations are of a very similar nature. For example,

$$\frac{\partial P_3}{\partial t} = \Lambda P_3 + \text{linear combination of } [P_1, P_1], [P_1, P_2], [P_2, P_2] \tag{116}$$

Now you must put some initial conditions. The simplest is that the initial distribution $f(x, 0)$ is given. We will write this in the form

$$f(x, 0) = f_0(x)\{1 + h(x)\} \tag{117}$$

which is simply a definition of $h(x)$. Following Hilbert, we are going to put ϵ equal to one. So we can state our initial conditions as:

$$\begin{aligned} P_1(x, 0) &= h(x), \\ P_k(x, 0) &= 0, \quad k \geq 2 \end{aligned} \tag{118}$$

Now what I said about these time constants will emerge perfectly obviously. I can solve equation (114) operationally speaking:

$$P_1(x,t) = e^{t\Lambda} h(x) = \sum_k e^{\mu_k t} \varphi_k(x) \tag{119}$$

The $\varphi_k(x)$ are eigenfunctions of my operator Λ ; and the μ_k are the eigenvalues. In my case, of course, the spectrum is discrete. Now what happens if you solve the equation for $P_2(x,t)$? The first equation was homogeneous, but this one is not. You see also that the inhomogeneous term is $[P_1, P_1]$

And this involves essentially the square of $p_1(x,t)$. When you square it out what time constant will occur? Clearly all the ones of the form $\mu_k + \mu_\ell$ will be there. Then, when you solve equation (115) for $p_3(x,t)$, what are you going to find? That $p_3(x,t)$ involves $\mu_k + \mu_\ell + \mu_m$.

To solve for f you must calculate $p_1(x,t)$, $p_2(x,t)$, $p_3(x,t)$.... That will now involve μ_k, $\mu_k+\mu_\ell$, $\mu_k+\mu_\ell+\mu_m$, ... And so it goes on. Every p_j will have, for time constants, the μ's, then sums of two μ's, sums of three, etc.

Now the method says that after you have found all the p_j to put them in and set $\epsilon = 1$. You put ϵ in to begin with only to know which coefficients to compare. Eventually, ϵ is not there. The method is supposed to yield a solution, which formally it does. If the series (112) makes any sense and you feed it back in, then formally it is a solution. That is clear. But if the method makes any sense it must lead to a series which is in some sense convergent. Then, of course, the conclusion I mentioned is certainly so -- you will involve all the linear combinations of the μ's. It can be verified in the case of my simple model that the series does indeed make sense.

A point I would like to make is that this method, which seems so strange, is really a perturbation method. It's very easy to decipher what the background of it is, from the Master equation. The Master equation really betrays what the meaning of the method is. Since we are dealing with the Boltzmann equation we already, implicitly, have assumed chaos at time zero. That means that for the Master equation my initial distribution must be factorizable, or approximately factorizable and each factor must contract to $f(x,0)$. Very roughly, without going into detail, I can write

$$\varphi^{(n)}(\vec{R},0) \sim \prod_{k=1}^{n} f(x_k,0) = \prod_{k=1}^{n} f_0(x_k)\left\{1 + h(x_k)\right\} \tag{120}$$

Now the $f_0(x_k)$, when multiplied out, become very simple. Because you get simply

$$\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{\frac{-(x_1^2+x_2^2+\cdots+x_n^2)}{2}} = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{n}{2}} \tag{121}$$

which is a constant. So it can be incorporated back in the proportionality factor. So we then get

$$\varphi^{(n)}(\vec{R},0) \sim 1 + \sum_k h(x_k) + \sum_{k,l} h(x_k)h(x_l) + \cdots \tag{122}$$

Now what must I do? I must act with the operator $e^{t\Omega}$ where Ω is the master operator. Then I must contract. When I do this the first term, the constant, simply remains. The next term, $\sum_k h(x_k)$, gives exactly $P_1(x,t)$. From the next one you get exactly $P_2(x,t)$, etc. It's simply a matter of direct verification. So really, the method of Hilbert is a contraction of this very simple expansion. The merit here is that you don't have to talk about an artificial ϵ -- you simply have to make the

expansion (120) and operate upon it. And, as you see, the Master equation together with the chaos property gives you a certain understanding of the Hilbert procedure. This in turn tells you all about the time constants of the solution.

This is about all I'm going to say about the Boltzmann equation and related problems. You have seen that there was much more analysis than probability theory. The probability theory ended long ago when we wrote out the Master equation. The rest was purely a verification that one gets the results one would expect either from intuition or from prior knowledge of elementary kinetic theory. As a matter of fact, this is true of most probability theories on this level. The probability disappears after the first stage, and from there on some kind of analysis begins.

(This concludes the discussion of kinetic theory. In the following material Professor Kac discusses stochastic models.)

We will consider a very simple stochastic model, a random walk. Unfortunately, this model is little known. It has very interesting features and leads not to a diffusion equation but to a hyperbolic one. The model first appeared in the literature in a paper by Sidney Goldstein, known to you mostly because of his work in fluid dynamics. The model had first been proposed by G. I. Taylor -- I think in an abortive, or at least not very successful, attempt to treat turbulent diffusion. But the model itself proved to be very interesting.

The problem is the following: Suppose you have a lattice of points. I mean discrete, equidistant points as in Figure 1.

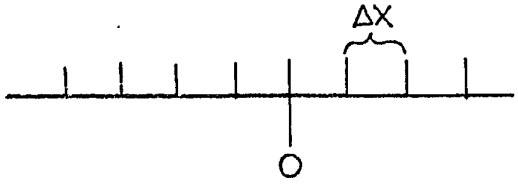

Figure 1.

Now I start a particle from the origin $X=0$ and the particle always moves with speed V . It can move either in the positive direction or in the negative direction. I flip a coin, let's say, to determine which. Each step is of duration Δt and covers a distance ΔX . So we have $\Delta X = V\Delta t$. Each time you arrive at a lattice point there is a probability of reversal of direction. I assume that $a\Delta t$ is to be this probability. Then, of course, $1-a\Delta t$ is the probability that the direction of motion will be maintained.

So actually what happens is that for a time you move in the direction you have chosen. And then, all of a sudden, you flip over. For a time you move in the new direction, until again disaster overtakes you. And so you will oscillate. As is usual in such problems, what is wanted is the probability that after a certain time t the particle is at a certain interval.

My notation will be a little strange for a discrete model, but it will be convenient for me later. Let X now stand only for abscissas of discrete points, the lattice points. And let me call the displacement after n steps S_n . This is the displacement if I start from the origin. It is the displacement after time $n\Delta t$. Now I will take a function $\varphi(X)$, an "arbitrary" function. And I will ask for the average $\langle \varphi(X+S_n) \rangle$. This will really give me all I want -- for example, $\varphi(X)$ could be the characteristic function of an interval. In that case this average will simply be the probability of finding the particle in that interval after n steps if it started at the point X. But instead of taking such a special function I will take a more general one. It's really no harder.

Now let me analyze the problem a little bit. I introduce the following random variable:

$$\mathcal{E} = \begin{cases} 1 \text{ with probability } 1 - a\Delta t \\ -1 \text{ with probability } a\Delta t \end{cases} \tag{1}$$

and I consider a sequence of such independent random variables $\mathcal{E}_1, \mathcal{E}_2, \cdots, \mathcal{E}_{n-1}$. Each of them has this strange distribution (1) and they are all independent. In other words, I have a coin, an extremely biased coin and the $\mathcal{E}'s$ are now the result of n independent tosses. Now I can very easily write out the displacement. If I start in the positive direction from the origin then it will be

$$S_n = v\Delta t \left(1 + \mathcal{E}_1 + \mathcal{E}_1\mathcal{E}_2 + \cdots + \mathcal{E}_1\mathcal{E}_2 \cdots \mathcal{E}_{n-1}\right) \tag{2}$$

Indeed, the first step will certainly take me a distance $v\Delta t$ in the positive direction. Now I must toss my coin and find what will happen to the velocity. It will change from v into $\mathcal{E}_1 v$ i.e., it will be maintained or else it will reverse according to the outcome of the toss. So in the next step I will move an additional distance $\mathcal{E}_1 v\Delta t$. And so it goes on, and you see how (2) comes about. If I had started in the negative direction then the displacement would have been

$$S_n' = -v\Delta t \left(1 + \mathcal{E}_1 + \mathcal{E}_1\mathcal{E}_2 + \cdots + \mathcal{E}_1\mathcal{E}_2 \cdots \mathcal{E}_{n-1}\right) = -S_n \tag{3}$$

I could combine the formulas together by saying that initially I have chosen my direction at random. But let me not even do that. Instead let me consider the two functions

$$F_n^+(x) = \langle \varphi(x+S_n) \rangle \tag{4}$$

$$F_n^-(x) = \langle \varphi(x-S_n) \rangle \tag{5}$$

And now, as is usually done, I am going to write a recursion formula for these things. First of all, let me write

$$F_n^+(x) = \langle \varphi\left[x + v\Delta t + v\Delta t\,\varepsilon_1(1+\varepsilon_2+\varepsilon_2\varepsilon_3+\cdots+\varepsilon_2\varepsilon_3\cdots\varepsilon_{n-1})\right] \rangle \tag{6}$$

You notice I have factored out ε_1. Now the averaging is really just a weighted sum over all possible sequences of ε's. The weights are dictated by the probability distribution. But I can perform the averaging in two different steps. I can first perform the average on ε_1, and then on all the remaining ε's. So let me first of all average on ε_1. This variable can assume the value -1 with probability $a\Delta t$; and it can assume the value $+1$ with probability $1-a\Delta t$. So I can simply write

$$\begin{aligned} F_n^+(x) = {} & a\Delta t \langle \varphi\left[x + v\Delta t - v\Delta t(1+\varepsilon_2+\varepsilon_2\varepsilon_3+\cdots)\right] \rangle \\ & + (1-a\Delta t) \langle \varphi\left[x + v\Delta t + v\Delta t(1+\varepsilon_2+\varepsilon_2\varepsilon_3+\cdots)\right] \rangle \end{aligned} \tag{7}$$

But now look at this. The averages have exactly the same form as before -- except that x is replaced by $x+v\Delta t$ and n is replaced by $n-1$. This gives me the formula

$$F_n^+(x) = a\Delta t\, F_{n-1}^-(x+v\Delta t) + (1-a\Delta t) F_{n-1}^+(x+v\Delta t) \tag{8}$$

In exactly the same way I can obtain another relation using F_n^- . It is

$$F_n^-(x) = a\Delta t\, F_{n-1}^+(x - v\Delta t) + (1 - a\Delta t)\, F_{n-1}^-(x - v\Delta t) \tag{9}$$

So now I have a system of recursion relations.

Now the standard time-honored way is to pass from these difference equations to a differential equation in the limit $\Delta t \to 0$. I will assume that all the mathematical difficulties in passing to the limit can be overcome. They're usually quite a nuisance. This is only an introduction so I will assume that all the formal steps are justified. In order to pass from the discrete to the continuous, notice first of all that n measures time. Actually, n is the number of steps and $n\Delta t$ is the time. The limit I have to perform is $\Delta t \to 0$, but $n\Delta t$ must be kept equal to my time t. Now let me re-write relation (8):

$$\frac{F_n^+(x) - F_{n-1}^+(x)}{\Delta t} = \frac{F_{n-1}^+(x + v\Delta t) - F_{n-1}^+(x)}{\Delta t} - aF_{n-1}^+(x + v\Delta t) + a\,F_{n-1}^-(x + v\Delta t) \tag{10}$$

And now I can pass to the limit to get

$$\frac{\partial F^+}{\partial t} = v\frac{\partial F^+}{\partial x} - aF^+ + aF^- \tag{11}$$

There is no n anymore, because I went to the limit. From the other relation, (9), I get in a similar way

$$\frac{\partial F^-}{\partial t} = -v\frac{\partial F^-}{\partial x} + aF^+ - aF^- \tag{12}$$

There is an analogy between these two equations and the linear Boltzmann equation. I will not develop this analogy, but you can yourself pick out streaming terms, and terms representing collision with the medium. Really, these equations and the Boltzmann's equations express conservation laws. You simply write, in a clever way, that the particles don't get lost.

Now the amazing thing is that these two linear equations of first order can be combined into a hyperbolic equation. For this purpose I will introduce two new functions:

$$F = \frac{1}{2}\left(F^{+} + F^{-}\right) \quad \text{and} \quad G = \frac{1}{2}\left(F^{+} - F^{-}\right) \tag{13}$$

Now, add up equations (11) and (12). Then you are going to get, in this new notation,

$$\frac{\partial F}{\partial t} = V \frac{\partial G}{\partial x} \tag{14}$$

Now subtract (12) from (11) to get

$$\frac{\partial G}{\partial t} = V \frac{\partial F}{\partial x} - 2aG \tag{15}$$

Now the problem is to eliminate G. To do this, differentiate (14) with respect to t and (15) with respect to X. Everything then becomes obvious, and I obtain

$$\frac{1}{V}\frac{\partial^2 F}{\partial t^2} = V \frac{\partial^2 F}{\partial x^2} - \frac{2a}{V}\frac{\partial F}{\partial t} \tag{16}$$

This is a very well-known equation, namely the telegrapher's equation. We now need to show what the initial conditions are. Remember that F^+ came from F_n^+. This, in turn, comes from (4). Now S_n is the displacement after a time $n\Delta t$; and we want this time to be zero. So that, in the limit, $F_n^+(X)$ simply becomes $\varphi(X)$. The same is true of $F_n^-(X)$, so we get

$$F(X,0) = \varphi(X) \tag{17}$$

Now what about the derivative with respect to time? This can be deciphered from the first order equations. But it's a little bit cumbersome to see it, so I will simply state what it is. We will later get this result from a somewhat different point of view. For the moment, I will ask you to believe me that

$$\left(\frac{\partial F}{\partial t}\right)_{t=0} = 0 \tag{18}$$

So now we have our initial conditions.

It is actually an accident, in a way, that we came out with a differential equation. Nevertheless, we will take advantage of this accident and discuss a few points. First of all, there is one limiting case which is extremely easy. That's when $a = 0$. Then, of course, the probability of reversing direction is zero. If you start moving in one direction, you never stop. What would $F(X,t)$ be? There are no reversals of direction and no random variables. So from (4) you see that $F_n^+(x) = \varphi(X + nv\Delta t)$ and from (5) that $F_n^-(X) = \varphi(X - nv\Delta t)$. So it follows that

$$F(X,t) = \frac{\varphi(X+vt) + \varphi(X-vt)}{2} \tag{19}$$

And that, of course, is a well-known classical case of the vibrating string.

That's all very fine, but not very interesting. You can get something better if you let $a \to \infty$ and $V \to \infty$ in such a way that $2a/V^2$ remains constant, say $1/D$. This can always be done, and I am allowed to choose D anyway I want to. This limiting case of equation (16) then becomes the diffusion equation:

$$\frac{1}{D}\frac{\partial F}{\partial t} = \frac{\partial^2 F}{\partial x^2} \tag{20}$$

Why I must let a and V go to infinity is easy to see. Because everybody knows that diffusion, or Brownian motion, can be looked on as a random walk. But in the standard model the probability of a move to the right or to the left is one-half. Now you see the probabilities in our model are either extremely small or extremely large. The only way they can be brought to where they will be one-half and one-half is to let a approach infinity as Δt goes to zero. If a does not go to infinity, there will always be a drift. You know, also, from the random walk model that the velocity of a particle is infinite in the limit. So we have to let V also go to infinity.

Before I proceed, let me tell you that this method of deriving the telegrapher's equation gives you what is now popularly known as a Monte Carlo way of solving it. What you are going to do is to go to your computer. You are going to store random numbers in the computer, or else you generate them as you need them. You start a lot of particles walking from X, let's say half of them in one direction and half in the other. At every time step you "flip a coin," using the random numbers. The "coin" is weighted so the probability is $a\Delta t$ that a particle will reverse direction and $1-a\Delta t$ that its direction will stay the same. At each time t you look to see where all the

particles are. And you calculate the value of the function φ for each particle, add them up, and divide by the number of particles. Then that's the approximate solution at time t of the telegrapher's equation.

This is a completely ridiculous scheme. Because in order to have good accuracy Δt had better be small. I don't know how small, but certainly you ought to discretize reasonably well. But that means the probability of reversing directions is very close to zero. It is a very unlikely event. This is the type of situation which is very difficult to handle by any kind of Monte Carlo technique. Because with such a small probability you would have to have an enormous number of particles. You would need a really ridiculous number. Otherwise, the fluctuation will be enormous.

Consequently, we'll have to be clever. What I want to do is to completely cut across the first phase of deriving the difference equation. I will treat the whole thing as a process with continuous time. The discretization will be avoided. In so doing, I will first of all discover a very neat way of writing a solution of this telegrapher's equation. Also, it will be extremely suggestive as to how to pierce out in different directions, mathematical and physical. I will assume then, that I have a continuous motion of my particle. During each time interval dt there is a spontaneous probability of changing direction $a\,dt$. The probability of not changing direction is then $1-a\,dt$. Now this is reminiscent of Mr. Poisson, whom I mentioned earlier. I will now define the Poisson process for you. I cannot, of course, go into some of the more delicate, purely mathematical difficulties. These involve some measure-theoretical points. I will have to stick throughout these lectures to a more intuitive presentation.

We suppose that $N(t)$ is a random variable (technically, a measurable function) for each given time t. It isn't a well-defined number -- it is something which has a distribution. Moreover, $N(t)$ can assume only integral values: 0, 1, 2, etc. The probability that $N(t)$ is equal to k at time t is given by the famous Poisson formula:

$$\text{Prob}\{N(t) = k\} = e^{-at}\frac{(at)^k}{k!} \tag{21}$$

That's one condition. A second, extremely important condition is that if you take a finite number of time points arranged in increasing order:

$$t_1 < t_2 < t_3 \cdots < t_n \tag{22}$$

then the increments:

$$N(t_2)-N(t_1),\ N(t_3)-N(t_2),\ \ldots\ldots,\ N(t_n)-N(t_{n-1}) \tag{23}$$

are independent. Then $N(t)$ is a Poisson process.

Think of $N(t)$ as representing the number of radioactive particles emitted up to time t. Then $N(t_2)-N(t_1)$ represents the number of radioactive particles emitted in the time interval (t_1, t_2). Now look at another time interval, say (t_3, t_4) which does not overlap the first one. The number of particles emitted during this interval is clearly independent of the number emitted in the other one. That is extremely intuitive.

Now if you assume -- and this is one of the standard derivations in all elementary textbooks -- that you have an event with probability $a\,dt$ of happening in $t+dt$ and $1-a\,dt$ of not happening, then the number of events which occur up to time t is the Poisson process. This, in fact, can

be made the basic definition. In particular, the number of collisions my particle undergoes up to time t is just a Poisson process. You will recall that every time I suffer a collision I reverse the velocity. So what is the plot of my velocity? The velocity can only be either V or $-V$. If I start in the positive direction, then for some time my velocity will be just V. Then I suffer a collision, and it changes to $-V$. It remains the same until I suffer another collision, and so forth. So my graph will look something like the one in Figure 2.

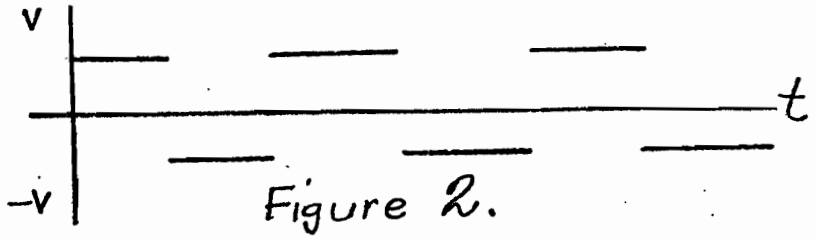

Figure 2.

Now the question is how to relate the velocity to the Poisson process. That's perfectly obvious. Because the number of collisions up to time t is just $N(t)$, and the velocity changes sign at each collision. So one way of writing it is:

$$V(t) = V \cdot (-1)^{N(t)} \tag{24}$$

This simply says that after an even number of collisions I have my old velocity. After an odd number I have just the negative of it. The displacement $X(t)$ is simply:

$$X(t) = \int_0^t V(\tau)\, d\tau = V \int_0^t (-1)^{N(\tau)}\, d\tau \tag{25}$$

This is the continuous analogue of my S_n.

The analogy with the discrete case is suggestive. I would expect that the solution of the telegrapher's equation (16) with the conditions (17) and (18) is simply

$$F(x,t) = \frac{1}{2}\left\langle \varphi\left(x+v\int_0^t (-1)^{N(\tau)}\,d\tau\right)\right\rangle + \frac{1}{2}\left\langle \varphi\left(x-v\int_0^t (-1)^{N(\tau)}\,d\tau\right)\right\rangle \qquad (26)$$

This is certainly what the whole thing suggests. Because what have I done? I have merely replaced a discrete random walk by a continuous one. And (26) is just what I found for the discrete case written out for this one. It is easy to prove that it is so. It can be done directly. I will sketch the proof later on, but it is not the proof which is so interesting. What is interesting is this very elegant way of writing the solution of telegrapher's equation in terms of the Poisson process.

First, I would like to call to your attention that in this form it is entirely feasible to use the Monte Carlo method. You have, no more, any difficulty with small probabilities. In the other, discrete, version I was plagued with them from the very beginning. All you need here is a machine or source of radioactive material which will produce a Poisson process. Then you simply take a hundred samples, say, of the Poisson process. For each one you calculate the integral (25) and then simply perform the averaging. Thus you can have the same problem formulated in two different ways, one of which is useful and the other not.

The second observation is really extremely amusing and shows that if one hits upon the right formulation one always gets more than one has bargained for. Our solution, in the form (26), is extremely reminiscent of the solution

of the equation of the vibrating string. Remember that for the vibrating string I had simply $\frac{1}{2}\left[\varphi(x+vt)+\varphi(x-vt)\right]$ and there was no average. Now these two differ in only one respect. Time t is replaced by this "randomized time" $\int_0^t (-1)^{N(\tau)} d\tau$. And then, because you don't know what it is going to be exactly, you must average.

This amusing observation persists for all equations of this form in any number of dimensions. Take for instance the case of propagation of radio waves:

$$\frac{1}{V^2}\frac{\partial^2 F}{\partial t^2}+\frac{2a}{V^2}\frac{\partial F}{\partial t}=\Delta F \tag{27}$$

And again suppose you want to solve the problem with the initial conditions:

$$F(x,y,0)=\varphi(x,y)$$
$$\left(\frac{\partial F}{\partial t}\right)_{t=0}=0 \tag{28}$$

Now the rule, for all dimensions, is the following. Forget about the bothersome terms -- take just the wave equation. Write down any solution you know. Solutions are very well known in all dimensions. Then, wherever you see time, replace it by this randomized version and average. This gives you the desired solution.

This seems surprising, but it really should not be. Some of the things which seem strange purely from the point of view of differential equations become obvious from another interpretation. For example, consider the theory of a cable where the equation (16) is applicable. If you put a charge in one place and let go then after a time you will have delta functions at two points with a certain continuous distribution in between. The delta functions,

which attenuate exponentially, simply correspond to particles which have not yet suffered a collision. And, of course, the continuous part simply corresponds to the particles which have suffered a lot of collisions and became completely mixed up. Our solution (26) appears very natural in this setting.