FIELD RESEARCH LABORATORY

SOCONY MOBIL OIL COMPANY, INC.

COLLOQUIUM LECTURES IN THE PURE AND APPLIED SCIENCES

NO. 2

OCTOBER, 1956

"SOME STOCHASTIC PROBLEMS IN PHYSICS AND MATHEMATICS"

BY PROFESSOR MARK KAC
THE ROCKEFELLER INSTITUTE

These lecture notes were transcribed from the tape recordings by Dr. M. R. Foster of the Field Research Laboratory, Socony Mobil Oil Co., Inc., Dallas, Texas. The transcript was reviewed and edited by Professor Kac. Dr. S. M. Foulks and Dr. S. R. Faris prepared the final manuscript for printing.

INTRODUCTION

These notes were prepared from tape recordings with comparatively little editing. They furnish once again a conclusive proof that the arts of speaking and writing are vastly different and, perhaps, even incompatible. In writing one doesn't have the recourse to the most powerful instrument of a speaker--the use of hands. In speaking one cannot conveniently refer to the learned and amusing treatise of Mr. Fowler on the proper usage of English.

If the only price for preservation of the immediacy of a live lecture is the number of split infinitives and garbled or unfinished sentences, then perhaps it is worth paying.

FIRST LECTURE

This series of lectures will be devoted to a subject which is becoming more and more popular and useful -- probability theory. I will not take the textbook line of development -- I will rather illustrate how some of the notions and methods have originated in important, simple problems in physical science. I will start with things which look very simple and progress towards more and more complicated things towards the end.

I would like to introduce one possible innovation in colloquia. To speak for an hour and a half is reasonably easy, especially if you are a college professor and are used to it. But to listen for an hour and a half is difficult. In fact, I think the span of attention of a human being is somewhat around fifty minutes. After that you may speak in the most flowing terms, tell the most beautiful jokes, and it is simply impossible for the listener to absorb it. Consequently, after about fifty minutes we will have a short break, so we can stretch ourselves for a few minutes. I would like also to devote some of the time, perhaps twenty minutes, to questions which might arise in the audience. The material will, perhaps, be somewhat strange to you, and it would be much better to settle some of the difficulties right off the bat. Do not be afraid to ask a question. Everybody has the fear that he may appear terribly ignorant. We all are ignorant -- the only difference is the degree of ignorance. Some of your questions I assure you I will not be able to answer; sometimes I may be able to give you a reference; sometimes I will go to the library and look it up and give you the answer tomorrow; and sometimes I will give you the answer right off the bat. The difference between you and me is that I have spent fifteen years on this subject, and you are going to spend probably only one week. Of course, I hope to convert you to the subject, so that you will be spending more and more time from now on.

I would like to start, after these preliminaries, with a little history. As you probably know, the theory originated in the very lowly problems of gambling. But that's not really how it entered science. It entered serious scientific thought in the latter part of the nineteenth century -- although some indications were already present early in the nineteenth century -- in connection with the kinetic theory of matter. Since the kinetic theory of matter is now so well known, it is interesting to tell you that it was by no means easy to convince people that atoms are something which exist. In fact, Boltzmann was severely attacked for empty speculation by the reviewers of his book. There are actually many reasons for that. One of the serious reasons was that the subject was plagued by certain paradoxes. It was in trying to resolve these paradoxes that the probability theory -- we might even say the statistical way of thinking -- came into being, at least in physics.

Now where did these paradoxes originate? That is a very interesting story, very educational, which I will tell you by way of an introduction to the subject. You probably know that Boltzmann, following people like Clausius and Meyer, tried to explain the behavior of gases on the basis of a mechanistic model. A gas was regarded as a system of a large number of particles, and the laws of mechanics were used to derive the equation of state and other things. The crowning achievement of the work of Boltzmann and of Maxwell is the derivation of the so-called H theorem. This is one of the things Professor Dresden spoke about when he discussed the Boltzmann equation and how it is related to hydrodynamics. We will devote a certain amount of time to Boltzmann' equation from a somewhat different point of view. We assume to begin with that

we have a spatially homogeneous gas. This means that the spatial distribution of particles is uniform. This is not the case Professor Dresden treated, because for hydrodynamics the whole problem is how the spatial distribution changes.

Now let $f(\vec{v},t)d\vec{v}$ be the velocity distribution in velocity space. Boltzmann actually interpreted this as the number of particles having velocity $\vec{v}$ at time t within the volume element $d\vec{v}$. He wrote out a complicated integro-differential equation from which he derived the H theorem:

$$\frac{d}{dt}\int f(\vec{v},t)\log f(\vec{v},t)\,d\vec{v} \leqslant 0 \qquad (1)$$

The integral (it is actually a triple integral) is denoted by H. Consequently,

$$H = \int f(\vec{v},t)\log f(\vec{v},t)\,d\vec{v} \qquad (2)$$

decreases, or at least does not increase in time. This was a remarkable achievement which particularly pleased Boltzmann because H clearly was some kind of an analog of the negative entropy. It was well known from classical thermodynamics that this function of state has the remarkable property that it never decreases. Here Boltzmann had managed to construct a mechanistic quantity which had a similar behavior.

All was well until it was pointed out that this was clearly untenable, since it was in contradiction with mechanics. The objections were crystallized in two paradoxes. One was the reversibility paradox of Loschmidt (around 1876) and the second was the recurrence paradox due to Zermelo and Poincare'. That

came a little later, after 1900, I think. The reversibility paradox is the simpler of the two and in a way more basic. It simply says the following: All equations of mechanics are time-reversible. This means that if you perform the transformation $t \to -t$ (that is, replace time by minus time) then the equations do not change. This is simply due to the fact that in mechanics all derivatives with respect to time are of second order. There is no way to distinguish the equations written with time going forward and time going backward. Using a more philosophically appealing terminology, there is no mechanical experiment that will tell you which way time is flowing. Consequently, said Loschmidt, something must be wrong, because this H presumably is a quantity which can be derived from a mechanistic description of the system. But now, if I change t to minus t, the quantity instead of decreasing, will increase. Thus from a purely reversible model, we draw an irreversible conclusion and something clearly is wrong.

It's interesting that in mathematics one contradiction is enough; it is even enough to think that something is wrong, to have doubts. But in physics, you must have several paradoxes before people will believe that something is wrong. So, to provide another one, Zermelo (who at the turn of the century was also interested in the logical foundations of this subject) recalled a theorem of Poincare'. It's a very famous and very beautiful theorem. It says that any conservative, closed, dynamical system is such that (I will first state it very loosely) if you start from somewhere, then unless you are extremely unlucky in choosing your starting point, you are bound to come back arbitrarily close to the starting point. To put it another way: conservative dynamical systems with finite energy are quasi-periodic. That is, states tend to recur. Consequently,

if this quantity H were a mechanical quantity then it would have to oscillate. Starting from a certain value, it would eventually have to come back arbitrarily near to that value. This would certainly contradict the result that H changes in only one direction.

Before I proceed to elucidate this paradox, I would like to prove the Poincare' theorem for you. It is very simple, and it illustrates a rather important pedagogical and fundamental point about general mathematical thought. I will emphasize this particular point many times during the conference. Suppose you have a system of mechanical particles with some strange arbitrary forces between them. Then everybody knows that the motion of the system can be described in terms of Hamilton's equations of motion. I don't even need to know the exact form of the Hamilton equations of motion. All I need to know is that there is a certain function H of the positions (generalized positions, if you wish) and momenta, and that in terms of this function certain differential equations can be written which describe the motion of the particles. Since the Hamiltonian depends on the positions and momenta, this suggests that it would be nice to look at a 6n-dimensional space, the coordinates being the q's and the p's. As most of you know, this space is known to us as the phase space.

The Hamiltonian equations of motion can be written in an extremely simple, concise way, under the name of the Liouville theorem. It simply says the following: suppose you take a little region or set A. (Refer to Figure 1 where the phase space is visualized as a plane).

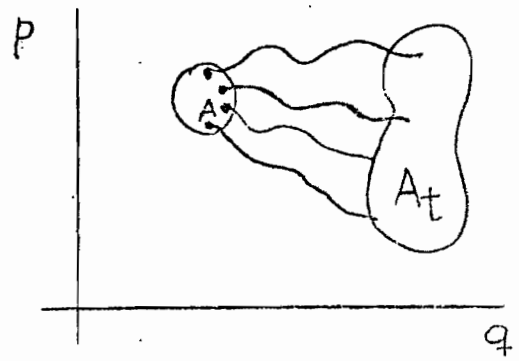


Figure 1

Now every point in this set A can be looked upon as a starting point for my system. And the trajectory will then describe the motion of the system according to the equations of motion. I now consider all the points in this little set A as the initial values for my system and look to see what happens after time t. After time t I collect them again into a set which I am going to call A_t. This is the set we get from A after time t due to the motion of the points. You might think of this set A_t as a very simply shaped one, something like a circle or a sphere. Actually, because of the complexity of the motion in the 6n-dimensional space, the shape of the thing would be tremendously complex. But the important thing is that its volume remains the same. And that's the Liouville theorem. It states that the volume of A is an invariant measure. This both implies and is implied by Hamilton's equations of motion. So, if you want to remember once and for all the simplest way to describe the motion of a system of particles, it is simply that the volume in phase space is conserved by the motion.

Since I mentioned that the system is conservative, we must take one further step. Not all phase space is being used, since the energy is being conserved. The Hamiltonian plays the role of the energy, so you are really not just anywhere in phase space, you are confined to the surface $H(\vec{q}, \vec{p}) = E$. This surface has to be assumed to be bounded. It is part of the assumptions of Poincare's theorem that this so-called energy surface is bounded. However, it is extremely easy to use Liouville's theorem (I will skip the details because that's unimportant for our purposes) to see what happens on the surface itself. All you have to do is to take a nearby surface and look at a little cylindrical

volume bounded by these two surfaces. Liouville's theorem tells us that this volume is invariant under the motion. We need only transform this knowledge to what happens on the surface by letting the two surfaces approach each other.

When you do that, you discover the following: If you take a surface element $d\sigma$, that is, a small element on the energy surface, and watch it move, then after time t the surface element is unfortunately not the same. What does remain the same is $d\sigma$ divided by the square root of the absolute value of the gradient of H. If you now take a set A which lies on the energy surface, then the set A_t also is on the energy surface -- this is because the energy is preserved. What we can say now is that

$$\int_A \frac{d\sigma}{|\text{grad}H|} = \int_{A_t} \frac{d\sigma}{|\text{grad}\,H|} \tag{3}$$

It is not a new fact. It is a simple, concise statement of the dynamics for closed, conservative systems.

Now the mathematician, if he is worth his name, extracts from this what is essential. What is essential here, for the whole picture, is not that we are dealing with a mechanical system pursuing the Hamilton equations of motion, but rather that we have the following very simple situation. Consider an abstract set, call it Ω, which will be the analog of the energy surface. I will call the points of this set ω. Suppose I also have a one-parameter family of transformations T_t of Ω -- this is an arbitrary family and depends only on time. The only condition these transformations satisfy is that $T_t \cdot T_s = T_{t+s}$. You can see this immediately: Where you are after time t + s is determined by where you were at the time t and, starting from there, the motion for

an additional time s. This property is called the semi-group property. That is just a name and it doesn't matter. The important thing is that we have such a one-parameter family of transformations which correspond to the motion. If ω is my starting point then $T_t(\omega)$ is where the starting point will be t seconds later.

Now finally comes the heart of the matter: I have a measure on this space. Now when I say measure I don't want to go into the matter in extreme detail; all I need is a class of sub-sets to which I know how to attach a number like volume. So for a large class of sub-sets there is a certain measure which I am going to call $\mu(A)$, or $|A|$. If you so wish, this corresponds simply to the integral:

$$\mu(A) = \int_A \frac{d\sigma}{|\text{grad } H|} \qquad (4)$$

Finally, what is very, very important, is that the transformations are measure-preserving. Now, what does this mean? This means that if we take a set A and the transformed set A_t then the measure of A is the same as the measure of A_t.

Now I have abstracted all I know from mechanics: I abstracted the energy surface, I abstracted the motion, I abstracted the statement of the Liouville theorem, which simply is that the measure is preserved. With this I am going to prove to you in three lines the Poincare' theorem. Rather than to consider continuous time, I am simply going to discretize it and to select T_1, T_2, T_3, ..., which are the T_t's after one, two, three seconds, etc. We can write that T_1 is equal to T, T_2 equals T^2, T_4 equals T^4, etc. An important condition, which we have mentioned, is that the measure of the whole space is

finite. This is an abstraction of the assumption that we are dealing with a bounded energy surface.

Now let me take a set A, and let me take a point ω from it and see what happens to this point. I start with ω ; this is transformed to $T(\omega)$, then to $T^2(\omega)$, and to $T^3(\omega)$, etc. This set of points is called an orbit. Now what I want to prove is that unless I was very unlucky in choosing my initial point ω , then eventually I will come back into A. That is, I want to prove that some $T^k(\omega)$ is again in the set A. This will help us decide also what it means to be unlucky.

Suppose that this is not so, that some points are such that they never return, or rather that their images never return. We are going to consider a set B, which is the set of points in A which never return. I must somehow show that set B must be very small: It must be an accident to choose a starting point from B. If this is so, then most of the starting points will be such that eventually I will return.

I now operate very simply. I take B, the image of B, the image of the image of B, etc. These are the sets B, $T(B)$, $T^2(B)$, etc. Now I claim that these sets do not overlap. Why not? Well, suppose for the sake of argument, that $T(B)$ and $T^3(B)$ overlapped; this means that there is a point which belongs to both of them. (I should have said that the transformation T is invertible, that it is uniquely invertible. This is certainly so in mechanics, since if you know where you are now, you can certainly determine where you were ten minutes ago.)

Now suppose that P is a point common to $T(B)$ and $T^3(B)$. Consider $T^{-1}(P)$: We are simply taking the point back one step. The result must be a

point of B since we started with a point in T(B). It must also be a point in $T^2(B)$ since the point P was in $T^3(B)$. This argument produces a point lying in both $T^2(B)$ and in B. But this can't be: B was a set, none of whose points returned to B. Thus, our assumption that T(B) and $T^3(B)$ had a point in common must be rejected. The same argument can be applied to show that none of our sets overlap.

Now we are through -- because now I have infinitely many sets which do not overlap. They all have the same measure, because the transformation preserves measure. Now all these sets lie on the energy surface, which is a set of finite measure. If the measure of each of these sets were positive, then we would have infinitely many non-overlapping chunks, each having positive measure, to be included in a set of finite measure. This is clearly impossible. That means that the measure of the set B must be zero. Otherwise these infinitely many sets could not be made to fit. When I say that I have to be terribly unlucky to choose a starting point which behaved badly, what I mean is that the measure of the set of such points is zero. And that's exactly the meaning of Poincaré's recurrence theorem.

I want to show you how very often mathematical abstraction--that is, throwing away what is non-essential, and keeping only what is pertinent to the problem--simplifies. Really, there is absolutely no problem at all in proof of the Poincare' theorem, once you know this. You don't have to think of a swarm of points moving in a complicated way in three-dimensional space. The whole thing is reduced to a very simple, very intuitive notion. That of the motion of a single point in phase space, the motion being described by means of a measure-preserving transformation.

Coming back to the paradoxes, we now have two. One was the reversibility paradox, and the other the recurrence paradox based on this theorem. Looking at these paradoxes, we are faced with a disturbing situation. On the one hand, it was extremely appealing to have the H theorem, which, so to speak, links thermodynamics with mechanics. On the other hand, it was extremely disquieting that the link was not entirely correct. Consequently, the problem was how to reconcile the proof of the H theorem -- which everybody believes contains some truth--with the difficulties exemplified by the paradoxes I have described. Boltzmann himself, as well as other people, has proposed that the solution of these difficulties would be found in a probabilistic treatment. We should say, well, it is not always so, but with overwhelming probability, it is so. To quote Gilbert and Sullivan, "What, never? No, never. What, never? Well, hardly ever." The intent was to express the immutable statement "in every case" into some kind of a probabilistic setting. Boltzmann himself and his followers of that time were not quite clear in their explanations.

As a matter of fact, even now we do not know the answers to all the questions. But at least we can give a consistent description without any contradictions. I will do so, at least with the case of the ideal gas, to show you how this can be done. By means of an extremely simple model I will now explain to you again both the illness and the cure. Actually, the model which I am going to describe to you is a simplification of another model which was proposed for similar purposes many years ago by Ehrenfest. I think I will skip the Ehrenfest model, because the one I am going to describe will have all the features. Moreover, it will be very easy to calculate. Now again, this is a

time-honored device in science: that you are trying to extract from an extremely complicated situation, a simple situation which has the difficulty of the old one, but not the morass of details. We want an example which will have similar difficulties, but from which we can compute, calculate, discuss, or do anything we want. I am going to give you such a simple model with which I can follow what Boltzmann did, and also what Gibbs did. I will show you all the difficulties which come up -- exactly the same two paradoxes -- and how finally one can resolve them, at least to some satisfaction. It will also give us the opportunity to go into the actual Boltzmann equation and to connect it with other things.

The model I am going to consider has nothing whatsoever to do with a gas or mechanical system. But don't let that worry you, because I will show you all the analogies. The model is the following: Take a circle, and on the circle take n equidistant points (vertices of a regular n-gon). Of these a certain number is marked, and I am going to say that m of them are marked. These m points form a set which I will call S, just for abbreviation. Now on each of the $\overline{n}$ points there sits a ball. This can be either black or white. In order to be completely definite, let us say that at time $t=0$, all the balls are black. Now comes the dynamics of this model. During each time unit (the model will be discrete in time, and the length of the elementary time interval will be unity) every ball moves one step counterclockwise. But with the following proviso: A ball moving from a point in the set S changes color. If the point isn't in S, the ball doesn't blush; it maintains its own color. At the next step some new balls will now be in S and again move counterclockwise. All those which are in S change color and those which aren't in S don't change color. So you go on.

Now the question is, what is going to happen in the long run? You can first of all argue quite intuitively that if the set S, where the blushing, the changing of color occurs, is sufficiently irregular, then eventually we are going to have roughly half white and half black ones. I can assure you that if you take a poll of public opinion, this is the answer you will always get. Let us now analyze this answer. Our analysis will be a direct imitation of what Boltzmann did, and you will see where one gets the paradox. Let me denote by $N_w(t)$ the number of white balls at time t, and by $N_b(t)$ the number of black balls at time t. At the beginning I have the condition that $N_b(0)$ is equal to n. That is my initial condition. Let $N_w(S,t)$ and $N_b(S,t)$ be the number of white and black balls in the set S at time t. We have the conservation law:

$$N_w(S,t) + N_b(S,t) = m \tag{5}$$

Now I am going to write the other conservation law for this model. I will follow Boltzmann, religiously copying what he did for the case of the gas. What can I say about $N_b(t+1)$, the number of black balls at time t+1? It is certainly the number of black balls at time t plus the ones I gain minus the ones I lose. Now what do I gain? I gain all the white ones which are in the set S. I lose all the black ones in the set S. Therefore:

$$N_b(t+1) = N_b(t) + N_w(S,t) - N_b(S,t) \tag{6}$$

$$N_w(t+1) = N_w(t) + N_b(S,t) - N_w(S,t) \tag{7}$$

To go further, we must make a special assumption. After some time has gone by, and if set S is pretty irregular, the balls should be fairly distributed. That is, the proportion of white balls in the set S ought to be in the same proportion as in the set of all the balls. That is, a reasonable approximation ought to be:

$$N_b(S,t) \cong \frac{m}{n} N_b(t) \tag{8}$$

$$N_w(S,t) \cong \frac{m}{n} N_w(t) \tag{9}$$

No sane person would argue with this. I assure you that if a situation like this came up in your daily work you would unflinchingly make this assumption. As a matter of fact, I wouldn't blame you, because if we didn't make such assumptions we would be working against the progress of science -- because it is clearly, obviously so.

We can now begin to solve the whole thing. For subtracting equation (7) from equation (6) and making use of our assumptions, we get:

$$N_b(t+1) - N_w(t+1) \cong \left(1 - 2\frac{m}{n}\right)\left[N_b(t) - N_w(t)\right] \tag{8'}$$

This is the simplest possible difference equation; it simply relates what happens at time t+1 to the situation at time t. It will take you exactly a second to see from this that the excess of black balls over white balls at time t is given by:

$$\frac{N_b(t) - N_w(t)}{n} \cong \frac{N_b(0) - N_w(0)}{n}\left(1 - 2\frac{m}{n}\right)^t \tag{9'}$$

I will now make the assumption that $2m$ is less than n ; in fact, I am going to put μ equal to $\frac{m}{n}$, and assume that μ is less than 1/2. Now if we believe all this, then the excess of black over white balls will decay exponentially, and hence, after a long enough time, they will nearly equalize. And moreover, we have the feature of the H theorem: not only will the numbers equalize, but the equalization will be irreversible.

This corresponds so well to our intuition that we are apt to overlook the difficulties. So now we take a second look at the model. First of all, the model is completely reversible. Again, I have exactly the same thing as in dynamics, because suppose that it were an exact conclusion that the numbers equalize. I start with n black balls and take, say, a thousand steps. If I rotate the set S one unit counterclockwise and then calling every black ball white and every white ball black, I make the balls go around clockwise instead of counterclockwise. The model doesn't know the difference, so the excess would still keep decreasing exponentially. But after 999 steps I will have to come back to where I started. So clearly, the conclusion is untenable. That's the reversibility paradox applied to the model.

The second paradox is very simple here, because you see that after $2n$ steps you come exactly to the initial situation. The model is so designed that it is strictly periodic. Let us represent the color by either +1 or -1; let's say that black is +1 and white is -1. Then in n steps each ball, no matter where it started, will have passed through all the points in the set S. The color will, therefore, have changed m times, so that the final color will be $(-1)^m$. If m were odd, then the black ball would have come out white.

However, if you run over the course again you again change color m times so that you certainly come back to the initial color. So, the model is completely periodic with period $2m$, and you have exactly -- or certainly a very strong analog with -- the situation of an actual gas, with the exponential decay in the role of the H theorem.

Now the question is, where is the offender? This model is very nice, because the offender sticks out. The offender clearly sits in the assumptions (8) and (9), because there is no doubt that the conservation laws are valid, and that's all we've done. From there on, it was the simplest possible elementary algebra. This assumption, or the analog thereof, was already made by Boltzmann -- only he didn't know it was an assumption. It is now so recognized and is referred to usually by the German word "stosszahlansatz". There seems to be no good English translation, but we may call it the "collision number postulate". In the original derivation of Boltzmann, it seemed that this was not a new assumption, but that it was somehow inherent in the model, so to speak. However, we now see clearly that it is an assumption and must try to justify it in some way.

If we want to maintain our conclusion -- which we think is more or less reasonable -- then we must somehow re-formulate or re-interpret the model. Perhaps by calling things by different names we can save our result. This is also a time-honored scientific practice: if you can't beat them, join them. If there is something wrong with the derivation, but the result seems to be more or less reasonable, the problem is to re-label and reorganize things to make sense. And that is what I will show you, that we can get out of all these difficulties by simply thinking statistically.

Some of the objections can be answered almost immediately. After all, with the recurrence paradox, we only run into difficulties at extremely long times, of the order of n. If n is the Avogadro number, $\sim 10^{23}$, and if it takes a tenth, or hundredth, or a millionth of a second to perform an experiment, then it will be after all quite a long time before this happens. Maybe we can save our results by saying that they hold only for time small compared to n. If that is correct, then we will discover how this proviso has to be made. But the other objection, namely of going for awhile in one direction, then violently reversing the model and coming back, still holds. That's a short time business. It holds for one step or for two steps. There is nothing to prevent us from moving just a few steps counterclockwise, shifting the set S one unit counterclockwise, changing the colors of the balls, and going back. So clearly it cannot be tenable to say that the excess of black balls is strictly decreasing as our formula indicates. Now the question is, how to extricate ourselves in a natural way from this situation and still end up with a result which is at least reminiscent of this one. And that is where the probabilistic analysis comes in.

Now let us be slightly more sophisticated -- the sophistication will increase as we go along. I'm going to get a little formalism going, which I will start now and finish in the next lecture. Let me introduce the following notation: I'm going to number the points from 1 to n; p will be a number between 1 and n;

$$\varepsilon_p = \begin{cases} -1 & \text{if } p \text{ is in } S \\ +1 & \text{if } p \text{ is not in } S \end{cases} \qquad (10)$$

$$\eta_p(t) = \begin{cases} +1 & \text{if the ball at } p \text{ at time } t \text{ is black} \\ -1 & \text{if the ball at } p \text{ at time } t \text{ is white} \end{cases} \tag{11}$$

This last quantity, of course, changes from time to time -- at various times the balls will be black or white. Now let me write out what I know about the model.

$\eta_p(t+1)$ is the color of the ball at the point p and time t+1. The ball came there from p-1, and it changed color or it did not change color depending on whether p-1 was in the set S or wasn't in the set S. So it immediately follows that:

$$\eta_p(t+1) = \varepsilon_{p-1}\eta_{p-1}(t) \tag{12}$$

This is merely the symbolic way of writing the dynamical description of my model. It says that the color of the ball at p at time t+1, is the color of the ball at the previous point at the previous time multiplied by -1 or +1 depending on whether the color changed there or did not change there. If we continue this recursion we get, very easily, that

$$\eta_p(t) = \eta_{p-t}(0)\varepsilon_{p-1}\varepsilon_{p-2}\cdots\varepsilon_{p-t} \tag{13}$$

That's simply repeating this thing, p times. Since I assumed that at time zero all the balls are black then $\eta_{p-t}(0)$ is simply 1, because at time zero all the η's are equal to one. That was my simplifying assumption. Now take the sum of all the η_p :

$$\sum_p \eta_p(t) = \sum_p \varepsilon_{p-1}\varepsilon_{p-2}\cdots\varepsilon_{p-t} \tag{14}$$

What is the meaning of this sum? η_p is +1 for a black ball and -1 for a white ball, so this sum is simply the number of black balls minus the number of white balls. Now I will divide the whole thing by n and speak then of the proportional excess:

$$\frac{1}{n}\left[N_b(t)-N_w(t)\right]=\frac{1}{n}\sum_p \varepsilon_{p-1}\varepsilon_{p-2}\cdots\cdot\varepsilon_{p-t} \qquad (15)$$

You see again that this is as far as I can go. But now I'm going to say to myself the following: Suppose that all I know about this set S is that it has exactly m elements But I have no idea where they are located. Then for each possible location of the set S I can follow through and find this number which I call the proportional excess. At the end, I can look at all these numbers. In other words, I will not study one individual set S, but I will rather study all possible sets S. I will look at the numbers

$$\frac{1}{n}\left[N_b(t)-N_w(t)\right]_S \qquad (16)$$

one for each possible set S. And I hope that I will notice the remarkable phenomena that most of them cluster around $(1-2\mu)^t$. True, we are going to have a few of them away from this value, but most of them, in fact an overwhelming proportion of them, I hope to be very close to $(1-2\mu)^t$.

But first I will aim for the more modest result that the average of all these numbers -- that is, the average over all possible positions of the set S -- is precisely $(1-2\mu)^t$. So the problem now is to perform this average, and I will demonstrate with complete precision in the next lecture that the result is exactly the one we want. The operation is not completely specified,

because I must tell you how to average. And that's always an element of arbitrariness which, in my opinion, can never be entirely disposed of in any physical situation. Somewhere you are going to get to a point where you will have to decide how to perform the averaging. For the immediate purposes I am going to assume that all positions of the set S are equally probable. That means, there is no way to distinguish, to prefer, one position over another. Our results must then be qualified by saying that they are obtained under the assumption that the positions of the set S are equally probable.

SECOND LECTURE

Now I want to find the average:

$$\left\langle \frac{1}{n}\left[N_b(t)-N_w(t)\right]\right\rangle_{\{S\}} = \left\langle \frac{1}{n}\sum_p \varepsilon_{p-1}\varepsilon_{p-2}\cdots\varepsilon_{p-t}\right\rangle_{\{S\}} \qquad (17)$$

Whatever else about averages can be said, an average of a sum is always the sum of the averages. So now I put the average inside the sum (it will be understood that we are averaging over all positions of S, so I can suppress the symbol S.):

$$\left\langle \frac{1}{n}\left[N_b(t)-N_w(t)\right]\right\rangle = \frac{1}{n}\sum_p \left\langle \varepsilon_{p-1}\varepsilon_{p-2}\cdots\varepsilon_{p-t}\right\rangle \qquad (18)$$

All you have to notice now is that all these averages are the same. Because it doesn't really matter what p is. You can simply take any point -- and moreover you can run the summation in the opposite order. Therefore,

$$\left\langle \frac{1}{n}\left[N_b(t)-N_w(t)\right]\right\rangle = \left\langle \varepsilon_1\varepsilon_2\cdots\varepsilon_t\right\rangle \qquad (19)$$

How to calculate this average? So far, I haven't done anything -- I simply used the most rudimentary properties of the averaging operation. But now we must define a little bit of what we mean by this. We mean by this the following: That all the ε_p are either +1 or -1 subject to a condition. They are subject to a condition because I know that I have exactly m elements in this set S. So what is the condition? It is that if I sum all the ε_p on p then I get m minuses and n-m pluses, or a total of n-2m. Thus the condition which expresses the fact that I have exactly m elements in the set S is

$$\sum_p \varepsilon_p = n-2m \qquad (20)$$

We can now say what the average means. It simply means to sum over all possible

sequences $\varepsilon_1, \varepsilon_2, \cdots \varepsilon_t$ subject to condition (20) and then divide by the number of all possible sequences. That's the definition of the average.

The number of possible sequences is very easy to get because I have m elements to choose out of n objects. That's the binomial coefficient $\binom{n}{m}$, so we have the denominator out of the way. The numerator is a little bit more complicated. It is the quantity $\varepsilon_1 \varepsilon_2 \cdots \varepsilon_t$ summed over all possible sequences subject to the condition (20). I will now show you a trick for evaluating this sum. Notice that it is the condition on the ε_p which is the nuisance. Because if we were free to sum over all possible sequences, then it would be very, very easy. In fact, it would be extraordinarily easy because you would simply have ε_1 summed over all possible values, times ε_2 summed over all possible values, times ε_3 etc. Since the ε_p can only be either plus or minus one, these sums are trivial.

Our trouble then is with this condition (20). This type of thing occurs over and over again in all probability problems -- this is a very elementary example of it. We use the following trick: Noticing that all numbers involved here are integers, consider the following formula:

$$\frac{1}{2\pi i} \oint \frac{dz}{z^{l+1}} = \begin{cases} 0 & \text{if } l \neq 0 \\ 1 & \text{if } l = 0 \end{cases} \tag{21}$$

This is simply a standard formula in complex variable theory written in a somewhat un-standard form. The contour of integration is a circle around the origin. Now in this formula, put $l = \sum_p \varepsilon_p - n + 2m$. Let us then write down this formula:

$$\frac{1}{2\pi i}\oint \frac{dz}{z^{\varepsilon_1+\varepsilon_2+\cdots+\varepsilon_n-n+2m+1}} = \begin{cases} 0 \text{ if } \sum_p \varepsilon_p \neq n-2m \\ 1 \text{ if } \sum_p \varepsilon_p = n-2m \end{cases} \tag{22}$$

But now the sum I am seeking is just

$$\sum_{\text{over all } \varepsilon's} \varepsilon_1\varepsilon_2\cdots\varepsilon_t \frac{1}{2\pi i}\oint \frac{dz}{z^{\varepsilon_1+\cdots+\varepsilon_n-n+2m+1}} \tag{23}$$

Where now the sum is over all epsilons, since our nuisance condition is now taken care of by the integral.

And now you interchange the integral with the summation:

$$\frac{1}{2\pi i}\oint \frac{dz}{z^{2m-n+1}} \sum_{\text{over all } \varepsilon's} \frac{\varepsilon_1\varepsilon_2\cdots\varepsilon_t}{z^{\varepsilon_1+\cdots+\varepsilon_n}} \tag{24}$$

And the nice thing is that this sum I can calculate very easily. Because you see that the general term in this summation can be written as follows:

$$\left(\prod_{k=1}^{t} \varepsilon_k / z^{\varepsilon_k}\right)\left(\prod_{k=t+1}^{n} 1/z^{\varepsilon_k}\right) \tag{25}$$

Where I assume that t is less than n. The general term splits up into a product of factors -- and the factors all depend on different subscripts. So you can simply sum each factor separately, and then multiply them out. That is, you can first perform the summation over ε_1 , then over ε_2 , etc. Now what happens if you sum the first factor over ε_1 ? ε_1 is either plus one or minus one, so the result is simply $(\frac{1}{z} - z)$. How many such things are there? All together, t of them. The last factor will behave differently,

because there is no ε in the numerator. It gives $(\frac{1}{z} + z)$, and there are n-t such terms. So all the sum simply becomes $(\frac{1}{z}-z)^t (\frac{1}{z}+z)^{n-t}$. And now I can express my sum as a very neat integral:

$$\frac{1}{2\pi i}\oint \frac{dz}{z^{2m-n+1}}\left(\frac{1}{z}-z\right)^t\left(\frac{1}{z}+z\right)^{n-t} \tag{26}$$

This is a very simple, but a very common trick, and it is important to understand all the steps. Now when you clean this up a little bit algebraically you get

$$\frac{1}{2\pi i}\oint \frac{1}{z}\left(\frac{1-z^2}{1+z^2}\right)^t \frac{(1+z^2)^n}{z^{2m}}\,dz \tag{27}$$

Remember that I must take this expression and divide the whole thing by $\binom{n}{m}$. Now as a matter of fact, we can also write a very complicated formula for $\binom{n}{m}$ which will be of the same form. Indeed,

$$\binom{n}{m} = \sum_{\text{over all } \varepsilon's} \frac{1}{2\pi i}\oint \frac{dz}{z^{\varepsilon_1+\cdots+\varepsilon_n-n+2m+1}} \tag{28}$$

Because I'm simply counting each time $\sum_p \varepsilon_p$ is n-2m, and otherwise I don't count at all. The expression on the right is the same as formula (23) when t = 0. So it ought to be that

$$\binom{n}{m} = \frac{1}{2\pi i}\oint \frac{1}{z}\frac{(1+z^2)^n}{z^{2m}}\,dz \tag{29}$$

which is simply what you get from formula (27) by putting t = 0. And it is

very easy to verify that this is indeed so. I will now write out, in its full glory, the average which we want:

$$\left\langle \frac{N_b(t)-N_w(t)}{n} \right\rangle = \frac{\oint \frac{1}{z}\left(\frac{1-z^2}{1+z^2}\right)^t \frac{(1+z^2)^n}{z^{2m}}\,dz}{\oint \frac{1}{z}\,\frac{(1+z^2)^n}{z^{2m}}\,dz} \tag{30}$$

That is a nice compact way of writing it, if you like it. But it also gives me a way to calculate. It will be a nice simple exercise in the method of steepest descent. I am interested in the case where n is very large, and I will assume that $\frac{m}{n}$ approaches μ as n approaches infinity. Moreover, I assume that 2μ is less than one. Now I hope you are familiar with the principle of the method of steepest descent. We will apply it first to the integral in the numerator. The first step is to write the integrand in exponential form. It will be enough to consider only the factor that involves n and m. The rest is a perfectly well defined function which doesn't change with n and m. So we have that

$$\frac{(1+z^2)^n}{z^{2m}} = e^{n\log(1+z^2)-2m\log z} \tag{31}$$

There is no point in worrying about the branch of the logarithm -- it doesn't affect the result. Now we must find the place where the derivative of the exponent is zero -- that is the saddle point. Now what is the derivative of this? It is $\frac{2nz}{1+z^2} - \frac{2m}{z}$. For this to be zero, simply means that $\frac{z_0^2}{1-z_0^2} = \frac{m}{n}$, which I can say is μ. (I may as well assume it is exactly equal to μ). Now if I subtract from one, I get $1-\mu = \frac{1}{1+z_0^2}$ -- and I think that is all I

need to find Z_s. There are two values and we will take the positive one, although it doesn't matter. Now, having found the saddle point, you substitute for Z wherever you see it, the value of the saddle point.

What I need to calculate now is the saddle point for the integral in the denominator. But the equation determining the saddle point is just the one we have already written. The part of the integrand involving n and m has the same asymptotic behavior as before. The $\frac{1}{Z}$ is also exactly the same. So all that is left is simply the value of $\left[(1-Z^2)/(1+Z^2)\right]^t$ at the saddle point. And what is it? It is just $(1-2\mu)^t$. Hence, by the saddle point method the asymptotic behavior is

$$\left\langle \frac{N_p(t)-N_w(t)}{n} \right\rangle \sim (1-2\mu)^t \qquad (32)$$

This simply means that the average over all possible positions of the set S -- if these positions are assumed equally probable -- agrees exactly with what we had before. However, notice that in performing this limiting operation, I make certain assumptions. I kept t fixed while n and m went to infinity. The important thing is that t was fixed while n went to infinity, which in terms of physics means that t must be small compared to n. Otherwise you could not use the method of steepest descent. If t were of the order of magnitude of n, then of course, the whole thing would not be justified. I would have to include the factor involving t in writing out the exponent, and the position of the saddle point would be entirely different -- the asymptotic behavior would be entirely different.

So it's only if t is small compared to n -- or better yet, if t is fixed while the number of points tends to ∞ our statement can be maintained.

Now this is completely in agreement with what is usually thought in statistical mechanics and kinetic theory, that you can only believe in the conclusions of kinetic theory or statistical mechanics if the times of observation are short compared to the underlying Poincare' cycle. (Remember I spoke earlier of Poincaré's theorem; the Poincare' theorem tells me that I eventually have to come back close to my starting point in phase space. The mean time I have to wait before coming back -- this can be very long -- is called Poincare' cycle). In our model, the Poincare' cycle is 2n because the whole thing is periodic with this period.

Let us think a little bit about what all this means. Suppose I'm going to plot $\frac{1}{n}\left[N_b(t)-N_w(t)\right]$ against t. How many curves am I going to get? I am going to get $\binom{n}{m}$ curves because I get a different one for each choice of the set S. Each of these curves starts always with one, and is periodic with period 2n. Now I am going to fix myself a t -- I'm going to look at these curves only at one point. Think of n as being very large compared to t -- think of n being 10^{23}, and of t being 10^6. Now at the time t each of the curves has some value, and all these values concentrate very strongly near $(1-2\mu)^t$. If I were to draw the exponential curve $(1-2\mu)^t$, then I will observe that at any fixed t most of the curves lie very close to it. I haven't demonstrated this yet. All I have demonstrated is that the average is exactly $(1-2\mu)^t$. Now I need something slightly more refined, namely the variance. I will skip this calculation -- it's elementary, but it's much more lengthy. What you do is to calculate the average of this thing:

$$\left[\frac{N_b(t)-N_w(t)}{n}-(1-2\mu)^t\right]^2 \qquad (33).$$

over all possible positions of the set S. This is known as the variance; and if you take the square root you get the standard deviation. If you calculate the standard deviation you will discover (and this is very interesting) that with t fixed as n goes to infinity, the standard deviation is of order $1/\sqrt{n}$. Roughly speaking, using the statistical 3σ business, this means that only a negligible portion of my values lie more than three standard deviations away from the mean $(1-2\mu)^t$. Now remember that n went to infinity, so the larger n is the more these values concentrate at the mean. So really, it takes extremely bad luck to observe a sizable deviation from this $(1-2\mu)^t$.

However, in spite of the convincingness of this argument, one should remember that an arbitrary assumption is floating around. Namely, in performing my average I count all the sets S as being equal -- as being peers of each other. This is something which one cannot justify. It seems natural, because why should one set S be better than another one? But still, you can certainly choose a set S for which the whole thing will deviate very much -- all you have to do is to pick a set S very regularly. But the point that is made obvious is that such a set must be very special. Because they are really proportionately extremely small in number. The predominant number of sets will lead to this nice decreasing exponential curve. And that's how thermodynamics or thermodynamical conclusions ought to be understood. An appropriate way of stating it is that for most of the configurations you're going to observe what you think you ought to observe. Why nature is so kind that it agrees with that, is a philosophical question that I am not prepared to discuss with you here, because I really don't know.

Since we have already learned something from this model, I would like to push it somewhat further and finally to connect directly with the Boltzmann equation and some of the more refined ways of discussiong stochastic phenomena in physics. The first step is to find some way to avoid the integrals. These integrals plague me because I was such an honest fellow and stuck laboriously and religiously to sets with exactly m elements. Now everybody knows that I could allow certain leeway in the number of elements, so I'm going to simplify my life in the following way. I will assume that only on the average does the set S have m elements. The set S may have more or it may have less, but on the average it ought to have m. I did not make this assumption before, because one ought to go through the calculation once, at least, and see that nothing bad happens. Besides you have seen a little trick which may prove useful to you in the future.

But now I'm going to change the model slightly, and assume the following. I am not going to determine my set S by placing m points. Instead I am going to have a coin which has the probability μ of falling heads and the probability $1-\mu$ of falling tails. And at each point of my circle I am going to toss this coin. If it is heads, I'm going to put the point in the set S. If it is tails, I'm not going to put it in the set S. So now the ε's themselves are not any more the well defined quantities of one or minus one whether you're in the set S or not. They now depend on the outcome of the flip of the coin -- or in technical jargon, they become random variables. Each ε_j is -1 with probability μ , and +1 with probability $1-\mu$. Moreover, they are completely independent, because I assume that each time I move from one point to another, I'm going to flip this coin again, independently of the previous toss.

Now, how many elements are there in the set S? Well, that is not a well defined quantity anymore -- it is clearly also a random variable. The number of elements in the set S can be found from the summation of all the ε_p. That won't quite give it to me because the sum is $n-2m$ just as before. So you see that

$$m = \frac{n - \sum_p \varepsilon_p}{2} \tag{34}$$

Now what is the average number of elements in the set S? This is just

$$\langle m \rangle = \frac{1}{2}\left(n - \sum_p \langle \varepsilon_p \rangle\right) \tag{35}$$

and the average $\langle \varepsilon_p \rangle$ is simply $1-2\mu$. So by the time you figure this out, you get

$$\langle m \rangle = n\mu \tag{36}$$

and μ was $\frac{m}{n}$ you remember, so on the average I have the right number of elements. Also you can show that if you take the actual number of elements in S minus the average number, if you square this, average it, and then take the square root to get the standard deviation, then this is of the order of the square root of n. You might say that the number of elements in S is roughly $n\mu$ with an error of the order of $\sqrt{n}$. That means very close to $n\mu$, so we'll expect that a result obtained from this model ought to be exactly the same as the result obtained from the other model. And this one is going to be much easier.

Those of you who have studied a little bit of the standard statistical mechanics may have heard of the grand canonical ensemble. The grand canonical

ensemble is the one in which you allow the number of particles in your system to vary. Yet the results of calculations using the grand canonical ensemble are exactly the same -- because although you allow the number of particles to vary, the mean is the prescribed one and the variations are very small. What I am doing right now, really, is replacing a canonical by a grand canonical ensemble. If you look at the text books on statistical mechanics, you will see that with the canonical ensemble you always have the steepest descent calculation, always these complicated integrals. With the grand canonical ensemble you avoid them. Of course, you have to prove the equivalence, and that is usually done only roughly. Sometimes they aren't equivalent, but I am not going to worry you with that. In simple cases, it stands to reason that if you allow the number of elements to vary, very little, the results should not change very much.

Now, had I chosen this model from the beginning, I'd have no difficulties at all. You remember that I had to calculate

$$\langle \varepsilon_1 \varepsilon_2 \cdots\cdots \varepsilon_t \rangle \quad . \tag{37}$$

But now the averaging is different. We not only have many possible positions of the set S, but now we have also the variability in the number of elements. In fact, the averaging is now simply the kind of averaging you would use in playing this game of tossing a coin. And everything is extremely simple because the tosses are independent; I assumed so. The probability theory that everybody knows tells us that the average of a product of independent things is the product of the averages. Therefore,

$$\langle \varepsilon_1 \varepsilon_2 \cdots\cdots \varepsilon_t \rangle = \langle \varepsilon_1 \rangle \langle \varepsilon_2 \rangle \cdots\cdots \langle \varepsilon_t \rangle \,. \tag{38}$$

Since all the averages are equal, this is simply the average $\langle \varepsilon \rangle^t$. But we have seen that $\langle \varepsilon \rangle = 1-2\mu$, so again you find the result

$$\langle \varepsilon_1 \varepsilon_2 \cdots \varepsilon_t \rangle = (1-2\mu)^t \qquad (39)$$

We are not surprised, but at the same time we ought to be mildly pleased. Since we also got this result from the exact calculation, there clearly is no question of the agreement within the canonical and grand canonical treatments in this particular case. And we have here the wonderful thing that this calculation was immediate. Now with this change in model, I am going to rediscuss the whole problem.

(Remark: I always have to assume that μ is less than 1/2. If it is equal to 1/2, then nothing will ever change from the average because $1-2\mu$ is zero. If μ is bigger than 1/2 it is very interesting, because the whole thing oscillates. We have to have μ less than 1/2 so that the change of color will not be too common an event. That's all the assumption I need, and I will assume from now on that this is so. Of course, if μ is close to 1/2, then you will change colors so actively, and so often, that the equalization will take place very quickly.)

I would like now to tune up the whole problem anew and bring out further analogies with thermodynamics. In this treatment we'll try to parallel that of Gibbs. The idea of Gibbs was the following: Initially, at time t=0, you are given a certain distribution of systems in phase space. Then as time progresses, this distribution evolves in a certain way. You try to prove that in some sense this distribution becomes more and more uniform on the energy surface. To put

it another way, if you have a localized distribution at time t=0 -- the probability is concentrated mainly over a small region of space -- then to begin with you know where your system is, more or less. As time goes on, you know less and less of where the system is. The system, so to speak, wanders around, dissolves. And the knowledge as to where it is gets less and less because of the motion. This is not quite correct, because you really have to introduce all sorts of auxiliary notions such as the coarse-grained density, etc. But one can imitate still, to some extent, what Gibbs tried to do and then see where one runs into difficulties again. Fortunately, for this particular model, one can solve everything. The behavior is simple enough that one can see easily what happens.

The model is the one with the modification that the $\mathcal{E}_p$ are now independent (chosen independently by tossing this μ-coin, you might say). Let us produce a dictionary of terms. First, what is phase space for our model? Well, it's a very simple space in this case. In mechanics, it's the space of all coordinates and momenta -- that means all the attributes required to define the system uniquely. What do I have to do to define my system uniquely? I have to know which ball is black and which is white on each point. That is, I have to know a sequence of plus and minus ones. So the phase space, here, is simply a finite set of points, each point being a sequence of n plus or minus ones. There are 2^n such sequences -- that is, 2^n points in my phase space.

Now suppose that initially at time zero I have a certain distribution $\rho(\vec{\eta}, 0)$ ($\vec{\eta}$ is a point in my phase space, that is a sequence of $\pm$ 1's, which I abbreviate by a vector). $\vec{\eta}$ can assume 2^n different values so that $\rho(\vec{\eta}, 0)$ is really a set of numbers. These numbers give me a distribution if

none are negative, and if they all add up to one. That is, if

$$\rho(\vec{\eta},0) \geq 0 \quad \text{and} \quad \sum_{\vec{\eta}} \rho(\vec{\eta},0) = 1 \tag{40}$$

I have a distribution. For example, you can say that $\rho(\vec{\eta},0)$ is 1 for the vector $\vec{\eta} = (1,1,1,\cdots,1)$, and is 0 for all other ones. That would mean that I would know initially with complete certainty that all the balls are black. But I don't have to take this particular distribution. I can say that I have only partial knowledge at the beginning, then take any distribution I like for $\rho(\vec{\eta},0)$.

The question is, how does the distribution evolve in time? That depends, of course, on the position of the set S. The distribution evolves in time according to the equation

$$\rho(\eta_1,\cdots,\eta_n;t+1) = \rho(\varepsilon_1\eta_2,\varepsilon_2\eta_3,\cdots,\varepsilon_n\eta_1;t) \tag{41}$$

This is the equation of evolution of $\rho(\vec{\eta},t)$. For those of you who are familiar with the terminology, this is the Liouville equation for the model. It tells you how the initial density of an ensemble of systems evolves in time, if the individual systems follow the equations of motion. In statistical mechanics, the Liouville equation unfortunately cannot be solved explicitly. This model, of course, was so designed as to be able to solve it. Let me now decide that since I already have probability in the whole business, I will treat the problem probabilistically.

Let me completely forget that there is a deterministic way in which the whole thing evolves. I can set up what will soon emerge as a very simple

Markov chain and argue as follows. The state of a system is defined entirely by a vector such as $\vec{\eta}$. At each time step I make a transition from this vector to another vector depending on what my set S is. What I know is that if at time t my model is in the state $\vec{\delta}$ then at time t+1 it will be in the state $\vec{\eta}$ where $\vec{\eta}$ is simply related to $\vec{\delta}$. Namely,

$$\begin{aligned} \delta_1 &= \varepsilon_1 \eta_2 \\ \delta_2 &= \varepsilon_2 \eta_3 \\ &\vdots \\ \delta_n &= \varepsilon_n \eta_1 \end{aligned} \tag{42}$$

If I know what my set S is, then I know the ε's -- so I know the exact transition. That's perfectly clear. But in this model, the ε's are random variables. So I must ask, what is the probability of this transition? That is, what is the probability that I make a specified transition in one step. Let me call this transition probability $P(\vec{\delta}|\vec{\eta})$. It is given by

$$P(\vec{\delta}|\vec{\eta}) = \text{Prob}\left\{\varepsilon_1 = \delta_1\eta_2\,,\ \varepsilon_2 = \delta_2\eta_3\,,\ \cdots,\ \varepsilon_n = \delta_n\eta_1\right\} \tag{43}$$

(remember that all the ε's , η's , and δ's are either plus one or minus one, so whether you write ε or $\frac{1}{\varepsilon}$, η or $\frac{1}{\eta}$, δ or $\frac{1}{\delta}$, it doesn't matter. It is a very nice algebra). Now this can be easily calculated because I know that the ε's are independent. To say that the epsilons are chosen independently means exactly that

$$\text{Prob}\left\{\varepsilon_1 = \delta_1\eta_2\,,\ \varepsilon_2 = \delta_2\eta_3\,,\ \cdots\cdots,\ \varepsilon_n = \delta_n\eta_1\right\} = \tag{44}$$

$$\prod_{k=1}^{n} \text{Prob}\left\{\varepsilon_k = \delta_k\eta_{k+1}\right\}$$

Now there comes a little bit of a problem --- a very minor one. I don't know whether $\delta_k \eta_{k+1}$ is plus one or minus one. It can be either. So the best way to write it is

$$\text{Prob}\left\{\varepsilon_k = \delta_k \eta_{k+1}\right\} = \frac{1}{2} + \frac{1-2\mu}{2}\,\delta_k \eta_{k+1} \tag{44'}$$

Because now if $\delta_k \eta_{k+1}$ is equal to plus one, then I'm going to get $1-\mu$ for the probability that ε_k is one. If $\delta_k \eta_{k+1}$ is minus one, then I am going to get μ exactly, so this works very nicely. So now the transition probability is just

$$P\left(\vec{\delta}\,|\,\vec{\eta}\right) = \prod_{k=1}^{n} \left\{\frac{1}{2} + \frac{1-2\mu}{2}\,\delta_k \eta_{k+1}\right\} \tag{43'}$$

Now that I have the probability of a transition from a state $\vec{\delta}$ to a state $\vec{\eta}$ I'm really in a position to get somewhere. Because now I can make use of a very well developed theory -- the theory of Markov chains.

Now remember what the initial distribution is. It's 2^n positive numbers which add up to one. I can think of the initial distribution as a vector with 2^n components. In a similar way, I can think of the transition probability $P(\vec{\delta}\,|\,\vec{\eta})$ as a matrix of order $2^n \times 2^n$. To see what the distribution after one step is we just multiply the vector $\vec{\rho}(0)$ (that is, the initial distribution) by the matrix. To see what happens after two steps we multiply again by the same matrix, which I will call P. Consequently, to see what happens after t steps, we simply need to raise the matrix to the power t:

$$\vec{\rho}(t) = P^t \vec{\rho}(0) \tag{45}$$

So, if I accept this theory, then all I have to do is find P^t and apply it to the initial distribution. This gives me the distribution at time t. In particular, what I want to show is that no matter where I start at the beginning -- that is, no matter what $\vec{\rho}(0)$ I choose -- if t becomes large enough, then the components of the vector $\vec{\rho}(t)$ approach the same values. So if I am allowed to replace my model by a Markov chain, then everything will depend on the properties of the matrix P. I hope that with a certain amount of cleverness I will be able to decipher these properties.

This method of approach is now known in the trade as the Master equation approach. I will tell you why it is called that in a moment. If you simply write what happens in the transition from time t to time t plus 1:

$$\vec{\rho}(t+1) = P\vec{\rho}(t) \tag{46}$$

then this is known as the "Master equation." This term was proposed by Uhlenbeck many years ago and it was thought initially that the "Master" referred to him. But the "Master" in this terminology refers to this equation which gives you all the information you need. For instance, in the equation we had before, where I considered the number of black balls minus the number of white balls, the information was only partial. Even though I knew how the excess of black balls was going to behave, I didn't really know what the probability was that a black ball sits in place 17 after a time t. But if I am so incredibly insistent on having such information, I can decipher it from this Master equation because it gives me precisely, at each time, the probability of any given situation. That's why it's called the Master equation.

Now the ordinary procedure in physics is to simply assume the Master equation. You say that between every pair of states I have a possible transition.

Somehow I discover the transition probability, and then from there on I just study this equation. It is quite interesting pedagogically -- and I will show it to you, although I will have to skip many details -- that I can derive such an equation for a gas. And then, from this equation, which is linear, derive the non-linear equation of Boltzmann. So it will be an interesting example where a non-linear equation is a consequence of a linear one -- and that is an interesting phenomenon which ought to be looked at a little bit.

But the one remaining problem now facing many people -- and a great deal of investigation is being done on it -- is whether or not this approach can be justified. On our example we can test whether it is justifiable or not. Notice what I have done. I get only the transition probability for one step in time. I then use the resulting matrix to propagate the distribution forever. But in reality, what I should do is solve the Liouville equation (equation 41) and then do the averaging at the end. In both cases, I am averaging over the positions of the set S. Only in one case I do the averaging and then the propagation, and in the other the propagation and then the averaging. In fact, for those of you who are familiar with the terminology, the first procedure which goes with the Master equation is called repeated averaging. Whereas, what really makes sense, is to solve the Liouville equation for time t and then perform the averaging. Now it's not at all clear that these things are interchangeable, that these procedures are equivalent. Because notice what one really does. One can put it picturesquely, although not quite correctly, as follows. The one relation between $\rho(t)$ and $\rho(0)$ we can symbolize by

$$\rho(t) = \langle L\, \rho(t-1) \rangle = \langle L \rangle \rho(t-1) = \langle L \rangle^t \rho(0) \qquad (47)$$

You get the propagation by repeated averaging. The other corresponding to

$$\rho(t+1) = \langle L^t \rho(0) \rangle = \langle L^t \rangle \rho(0) \qquad (47')$$

would be the modified Gibbs approach, to solve the Liouville equation if you can (you almost never can) and then average.

People usually manage to give at least a justification, but not quite a proof, that the results are essentially the same for the dynamical system. It's a very profound fact, because it makes life so much simpler. Because you don't have to solve Liouville's equation, which is tantamount to solving the equations of motion. So if you could once and for all prove that it doesn't matter if you'd average the t^{th} power of the operator of if you average the operator and raise it to the t^{th} power, then you'd be in very good shape. As I say, in most of the present day applications of probability, statistical mechanics, and kinetic theory everyone always makes this assumption anyway. He takes what happens during a very short time, averages the transition operator over that short time, and uses this average operator as the one which propagates. That's done not only in classical statistical mechanics, but in quantum statistical mechanics as well. For instance, in the supposed derivation of the transport equation, first done by Pauli and Fierz. You always say, well there are some random phases; and after a short time you average over them. Then you go for a little time and you again average over them. You constantly keep averaging, whereas in reality you ought to wait until the end of the calculation.

Now in the next lecture I will show you first of all how one can write exactly the solution for both of these approaches. Then we will see how the two things differ, and try to talk ourselves out of this difference. Then I will show you, with some detail, how a corresponding Master equation can be written for an ideal gas with binary collisions. The Master equation will be a linear equation from which all general conclusions can be derived: the H-theorem and various

other things. I will show you how it is related to the non-linear Boltzmann equation which was discussed with you by Professor Dresden. And then with this background we can go on and look at some other stochastic models which, although different, are not unrelated. For instance, in the theory of Brownian motion a very similar treatment can be done. The operators are more complicated and the Master equation becomes a diffusion equation. But, still, fundamentally the same ideas persist.

THIRD LECTURE

To remind you where we were, I tried to discuss this problem from what one might say is a purely probabilistic point of view. For I simply said to myself that my system can be in any state described by a vector $\vec{\eta}$ whose components are either plus ones or minus ones. And in each elementary step I can perform a transition from a state $\vec{\delta}$ to a state $\vec{\eta}$, and I would like to calculate the probability of such a transition. In fact, I have already calculated it. Now I might as well write out the equation completely and simply hope for the best:

$$\rho(\vec{\eta}, t+1) = \sum_{\vec{\delta}} \rho(\vec{\delta}, t)\, P(\vec{\delta}|\vec{\eta}) \tag{48}$$

This, as a matter of fact, is exactly the matrix equation (46) written out in its full glory. In this connection it is called the Master equation. In the mathematical literature it is referred to as the Chapman-Kolmogoroff equation. But of course it's just the perfectly obvious consequence of our assumptions.

I am first going to provide myself with a convenient basis in my phase space. Remember that my phase space is a very simple set whose elements are vectors with components that are plus one or minus one. Consider the following functions:

$$1;\ \eta_k;\ \eta_k\eta_\ell;\ \eta_k\eta_\ell\eta_m;\ \ldots\ldots;\ \eta_1\eta_2\cdots\eta_n \qquad k \leqslant \ell \leqslant m \leqslant \cdots$$

How many such quantities do I have? It's perfectly obvious that I have n single η's. I have $\binom{n}{2}$ $\eta_k\eta_\ell$, $k \leqslant \ell$ etc. So there are exactly

$$1 + n + \binom{n}{2} + \binom{n}{3} + \cdots\cdots + \binom{n}{n} = 2^n$$

of these things. I claim that they form a complete set. First of all, what does η_1 mean? η_1 ought to be looked at as a function defined on my space

of all possible sequences of ones and minus ones. It means to always take the first component. That's the function. And $\eta_k \eta_\ell$ always means to take the product of the kth component and the ℓth component. So we have here 2^n functions defined on the set of all possible sequences of plus ones and minus ones of length n.

Now I claim that every function of the vector $\vec{\eta}$ -- and in particular $\rho(\vec{\eta},0)$ -- can be written in terms of these functions. In fact,

$$\rho(\vec{\eta},0) = \frac{1}{2^n} + \sum C_k \eta_k + \sum C_{k\ell} \eta_k \eta_\ell + \cdots + C_{12\ldots n} \eta_1 \eta_2 \cdots \eta_n \quad (49)$$

This is similar to expanding a function into a series of harmonics. In fact, you can easily find the coefficients by Fourier's formula, if you wish. For instance,

$$C_{k\ell} = \frac{1}{2^n} \sum_{\vec{\eta}} \rho(\vec{\eta}) \eta_k \eta_\ell \quad (50)$$

This formula, and in fact the validity of the expansion, follows from the fact that my functions are orthogonal. The reason I have $\frac{1}{2^n}$ in formula (49) simply goes back to the condition that the sum of all the ρ values must be one. Except for this constant all the terms sum to zero. So the constant must be such that when it is multiplied by 2^n it will give you one; and then, of course, it's $1/2^n$.

And now the question is, what happens if I operate with the matrix on each individual term of the expansion? Let's take a typical term, a very simple one -- let's take $\eta_1 \eta_2$. It's a vector of 2^n components. Now let's apply the matrix P to it and see what happens. The result is

$$P\{\eta_1 \eta_2\} = \sum_{\vec{\delta}} \delta_1 \delta_2 \prod_{k=1}^{n} \left[\frac{1}{2} + \frac{1-2\mu}{2} \delta_k \eta_{k+1} \right] \quad (51)$$

where we have made use of equation (43'). What does it mean to sum over $\vec{\delta}$? It means summing over all δ_1 , δ_2 , . . ., δ_n . But now look. What you are summing is a product of independent functions. Hence, in summing, you can simply sum things separately and then multiply to get the answer. It's again like integrating a function of n variables which happens to be a product of functions of individual variables. You know very well that such a thing is just a product of integrals. Now the summation over δ_1 and δ_2 is evidently special. But what is the summation of one of the other factors? When δ_k is + 1 you get

$$\frac{1}{2} + \frac{1-2\mu}{2}\eta_{k+1}$$

and when δ_k is - 1 you get

$$\frac{1}{2} - \frac{1-2\mu}{2}\eta_{k+1}$$

So, if you add them together, you get one. What, however, about the factor that has the δ_1 ? There are only two values for δ_1 , plus one and minus one, so it's going to be

$$(1-2\mu)\eta_2$$

So finally when you perform the whole sum, believe it or not, you're going to get

$$P\{\eta_1\eta_2\} = (1-2\mu)^2\eta_2\eta_3 \qquad (52)$$

Now all this is actually something which one ought to go through by oneself. It's extremely simple, and in fact the symbolism makes it look much deeper than it really is. Yet, if you did not have the symbolism, it would take me three blackboards to write it all out. It is something you can, with the greatest of ease, check for yourself. When the operator P is applied to any of

the terms in the expansion (49) the following two things happen: It multiplies them by $(1-2\mu)$, that means it decreases their length in that proportion. And then it shifts the indices by one. This is in fact a description of the operator P. You saw this happen in the example we went through, and it is easy to show that it is generally so. Now it is very interesting to see what happens if you apply the operator twice. It means multiplying by $(1-2\mu)^4$ and shifting the indices by two. If I apply it t times, then I am simply going to have $(1-2\mu)^{2t}$ and the indices will be shifted by t. There is the usual understanding, that we mentioned before, that if by any chance some number becomes bigger than n, then you simply reduce it modulo n to get back within the range. Now we can immediately write the solution of the Master equation. Everything was constructed so that no trouble could possibly arise and the solution is

$$\rho(\vec{\eta},t)=\frac{1}{2^n}+(1-2\mu)^t\sum_k C_k\eta_{k+t}+(1-2\mu)^{2t}\sum_{k<l}C_{kl}\eta_{k+t}\eta_{l+t}+\cdots$$

$$\cdots+(1-2\mu)^{nt}C_{1,\ldots,n}\eta_{1+t}\eta_{2+t}\cdots\eta_{n+t} \qquad (53)$$

(refer to equation 49).

Now let's take a look at it. This is a very interesting result, because it agrees completely with our intuition. Notice that all the terms except the first one have an exponential tacked to them. So all of them vanish exponentially, as t goes to infinity. And in the limit, you get the distribution $\frac{1}{2^n}$ which is the uniform distribution. This is actually a general feature of all such equations where the matrix P is a stochastic matrix (there are exceptions, but they are minor exceptions) -- If a matrix has all its elements non-negative, and if the sum of the elements in each row is one, such a matrix is called a stochastic matrix. Barring certain exceptions which I am not going to

go into, a high power of such an operator destroys everything except one fundamental eigenvector. All other eigenvectors are annihilated as you easily can see from (53). You have here a very good example of the disappearance of information. My initial distribution $\rho(\vec{\eta},0)$ could have been concentrated on one point -- I could have said it was exactly equal to one at one specific vector, and zero otherwise. So I would know precisely the state of the system at time zero. But at the end, all the states become equally probable. Then I know as little as possible about the system. I started from complete knowledge and ended up with complete ignorance. (This is not unlike the behavior of students, except that their initial state is somewhat different.)

However, all this was based on definitely treating the whole thing by probability methods. Because I simply assumed that I could replace my model by a stochastic model. I said to myself, I don't care what the detailed mechanism is. I will replace my model by a model in which I have at each time interval a transition from a state $\vec{\delta}$ to a state $\vec{\eta}$, with the probability given by a complicated formula. And the only place I used the dynamics of the model was to guess, or to derive if you wish, the formula for the transition probability. I emphasize this, because I am going to follow the same procedure in deriving the Boltzmann equation. I am going to do exactly the same thing.

On this model we can go farther, however. Because we can ask ourselves, can we really justify this probability approach? To answer this question, I must go back to the fundamental problem, whether $\langle L^t \rangle$ is the same as $\langle L \rangle^t$. In general, this is an extremely complicated question. There is now some hope, because there are some partial results. Both in this country and in Russia there are many people who have tried to approach this particular

problem for actual dynamical systems. But the difficulty is that you can never solve the analog of equation (41) -- that is, the Liouville equation which gives you the evolution of the distribution in phase space. It involves all the complications of the equations of motion of a complicated dynamical system. Since it is impossible to solve it rigorously, it seems very difficult at first to see how one could possibly justify the Boltzmann procedure. Here, however, everything is extremely easy. Because the Liouville equation can be solved exactly and I can perform the averaging at the end.

And now let us see what happens. I will still start with exactly the same initial distribution, but now I will use equation (41) and write out the exact formula for $\rho(\vec{\eta},t)$. Here you will discover

$$\rho(\vec{\eta},t) = \frac{1}{2^n} + \sum_p C_p \eta_{p+t} \varepsilon_p \varepsilon_{p+1} \cdots \varepsilon_{p+t-1} + \sum_{p<q} C_{pq} \eta_{p+t} \eta_{q+t} (\varepsilon_p \cdots \varepsilon_{p+t-1})(\varepsilon_q \cdots \varepsilon_{q+t-1}) + \cdots \qquad (54)$$

Now of course, the coefficients which are attached to these terms do not decrease exponentially. They are either plus one or minus one, because remember that each ε_p is either plus one or minus one. But that's not surprising after all, because I have not yet performed the average over the positions of the set S. What I have written so far simply represents L^t, the operator raised to the power t. So I must now perform the average over the positions of the set S:

$$\langle \rho(\vec{\eta},t) \rangle = \frac{1}{2^n} + \sum_p C_p \eta_{p+t} \langle \varepsilon_p \varepsilon_{p+1} \cdots \varepsilon_{p+t-1} \rangle + \sum_{p<q} C_{pq} \eta_{p+t} \eta_{q+t} \langle (\varepsilon_p \cdots \varepsilon_{p+t-1})(\varepsilon_q \cdots \varepsilon_{q+t-1}) \rangle + \cdots \qquad (55)$$

And now you discover a very interesting thing -- that the first two terms are just the same as before. Because $\langle \varepsilon_p \cdots \varepsilon_{p+t-1} \rangle$ is just an average of a

product. Since I have assumed that the $\mathcal{E}_p$ were chosen independently (remember I flipped a coin to determine at each point p). Consequently,

$$\langle \mathcal{E}_p \cdots \mathcal{E}_{p+t-1} \rangle = \langle \mathcal{E}_p \rangle \cdots \langle \mathcal{E}_{p+t-1} \rangle = (1-2\mu)^t \tag{56}$$

So the first two terms are exactly what I had before. The trouble begins with the next term. In the next term, notice that you have $(\mathcal{E}_p \cdots \mathcal{E}_{p+t-1})$ and $(\mathcal{E}_q \cdots \mathcal{E}_{q+t-1})$. If these two groups were entirely separate, if they did not have any $\mathcal{E}$ in common, we could again say that the average of the product is the product of the averages and you would get what you should: $(1-2\mu)^{2t}$. Unfortunately, however, these groups will not be non-overlapping for all of the terms. Let me define

$$\Delta(p,q,t) = \begin{cases} 2t \text{ if } q-p > t \\ 2(p-q) \text{ if } q-p \leq t \end{cases} \tag{57}$$

Now if you calculate this average, you get the following:

$$\langle (\mathcal{E}_p \cdots \mathcal{E}_{p+t-1})(\mathcal{E}_q \cdots \mathcal{E}_{q+t-1}) \rangle = (1-2\mu)^{\Delta(p,q,t)} \tag{58}$$

What this means is that if p and q are close together the decay factor is $(1-2\mu)^{2(p-q)}$ which does not decay. It stays the same as p goes to infinity. To add insult to injury, the very last term, $C_{12\cdots n}\eta_1 \cdots \eta_n$ won't decay at all. Nothing will happen to it.

Now what is the trouble, and how do we deal with it? One should try to solve Liouville's equation, and one should average. This is definitely correct it is certainly the way one should do the problem. But, lo and behold, you do not get the same result as with the more convenient Boltzmann approach. So this is the real difficulty -- and it occurs also in the actual physical case: What you really want is to use the Boltzmann approach which you know how to handle and apply. But all you really have is Liouville's equation with the

possibility of some averaging. And here you are certainly not led to quite the same result. Fortunately, it is not very serious in this case.

You can talk your way out of it as follows (there are several ways of talking oneself out of it, but we will only look at one of them). Suppose that we start from a symmetric distribution; this means that the function $\rho(\vec{\eta}, 0)$ is a symmetric function in the arguments $\eta_1, \cdots, \eta_n$. In other words, there is no distinction between the points. You don't know which is which. If $\rho(\vec{\eta}, 0)$ is to be a symmetric distribution, then all the C_k's are the same, all the C_{kl}'s are the same, etc. Now even if you start with a symmetric distribution, as you should, then as the thing evolves even that is destroyed. This is because some of the coefficients will be multiplied by $(1-2\mu)^{2t}$ and some will be multiplied by $(1-2\mu)^{p-q}$. Hence, you might say that, in time, the original indistinguishability of sites gets destroyed. In the language of gases, if you wish: At time zero you may not be able to tell which particle is which but by the time they collide a few times it is possible to re-establish their identity. That's not a very pleasant perspective. Consequently, what I should really do is not only to follow the motion in time and then perform an average over the positions of the set S. But in addition to that, I should perform a symmetrization on all the variables η_k.

This is not a logical proof which is independent of how persuasive I am. What I'm saying right now is really a matter of belief, because I have no logical compulsion for it. But you can see that since it is really impossible to follow each particle separately, and know where it is all the time, and what it's color is, that this indistinguishability should be maintained all the time. So now if you also symmetrize -- that is, if you average over all possible

arrangements of the $\eta_1, \ldots, \eta_n$ -- then you will discover the following interesting things. First of all, the two terms we like are not affected. They are just

$$\frac{1}{2^n} + (1-2\mu)^t C_1 \sum_p \eta_p \tag{59}$$

(since all the C_k's are the same, I can take $C_1 = C_k$ outside the sum). Now I want to show you what happens to the terms with which we had difficulty. They become

$$\frac{C_{12} \sum_{1 \leq p \leq q \leq t} (1-2\mu)^{\Delta(p,q,t)}}{\binom{n}{2}} \sum_{1 \leq p \leq q \leq t} \eta_{p+t} \eta_{q+t} \tag{60}$$

Now we like the terms for which $\Delta(p,q,t)$ is equal to $2t$. The terms which we don't like are the ones for which $\Delta(p,q,t)$ is less than $2t$. The number of such terms is only of order n. But the denominator is of order n^2. Consequently, the proportion of objectionable terms is very small -- (remember that n has to go to infinity).

Now the same thing happens for every term. In the next one, we're going to have $\binom{n}{3}$ in the denominator, while the number of undesirables is only going to be of order n^2. Consequently you can say to yourself, if you are a physicist: all right, I am simply going to neglect the ones I don't like because there are so few of them in comparison to the others. However, nothing can be done to the very last term in (55), because it simply does not change at all. No matter what you do with this model, the last term simply insists on staying. No symmetrization will help you -- nothing will help you except to simply forbid terms like this. You can think of it as follows: If $C_{12\ldots n}$ were sizable then my initial distribution would have a lot of this extraordinarily high harmonic component. It would be an extremely wiggly thing. If

you think of an ordinary Fourier series, where for instance a coefficient of a very high harmonic, say 10,000, is sizable then that means that you have components in the curve of frequency 10,000. It is terribly wiggly, so that you have to know the curve with great precision over very small time intervals. That's the Fourier's series case. In this case, I must simply assume that $\rho(\vec{\eta}, 0)$ is a relatively smooth function; that is, I cannot allow my initial distribution to be too fine-grained. In particular, I would not be allowed to start with all the balls black. That's too sharp, and the $C_{12\cdots n}$ coefficient is simply one. In fact, all of the coefficients are one.

So, I must first of all, assume that $\rho(\vec{\eta}, 0)$ is smooth so that I can neglect the last coefficient and others like it. Secondly, I must symmetrize. Even then I still have a problem. Because there are a lot of terms in equation (55), and even though in each individual term I am throwing away only a small proportion of the members, it could be that the cumulative error I am going to make is going to hurt me. Actually, I'm not even sure on this model whether I can show this or not. In the physical literature, where one tries to imitate this for the real Liouville equation for a gas, all that is done, and not even that quite precisely, is to show that in the few early terms the error is of small order. Whether or not they will accumulate is difficult to say.

On the other hand, you can say to yourself that I'm not going to be so ambitious as to justify the full Master equation. Rather, I'm going to stick to physical statements about single particles, pairs of particles, or triples of particles. Don't forget that the Master equation gives you a description of the whole statistics of all the sites of particles. If I am only going to ask questions which will never involve correlations or joint distributions of

more than three particles, then, of course, I will never go beyond the triple product of the η's. The neglect of the terms which don't decay will then be of course perfectly justified.

Consequently, this already sheds some light on how far we should trust the Master equation. We certainly ought not to trust it when it comes to conclusions involving numbers of particles comparable to the total number. But presumably you may as well accept everything else. So we may replace our deterministic model -- you might call it a deterministic model with final averaging -- by a stochastic one.

There are still various inconveniences involved. We may have to deal with infinite matrices, for instance. Because really, if you want to be a mathematician, you would have to perform the limit $n \to \infty$. But then you have an infinite matrix and the purely mathematical nuisances become very cumbersome to deal with. So close your eyes to it a little bit, and simply keep n fixed. But consider t small compared to n -- for instance, you allow t to go as far as the square root of n. I will postpone further discussion of these points until I come to the Boltzmann equation where the same difficulty arises.

This model has been useful to us. On it you have seen in excruciating detail essentially all the difficulties, all the problems and all the approaches that one uses in the kinetic theory of non-equilibrium phenomena. It perhaps looked a little more formidable than it really was, because the formulas were so long. I will now see what can be done by honestly assuming from the very beginning what you might call a stochastic model. Let me first say a few introductory words. In classical physics -- (that's discounting quantum mechanics where there are various added features) -- the use of probability is twofold.

There are some theories which, from the very beginning, cry out for a probability model. Without ever saying why or without bothering to justify it, one simply makes an assumption that the process under consideration is what we will call a random process. One then proceeds with an analysis of that. Examples of such problems would be Brownian motion, noise in electrical circuits, and similar problems. In all of these you are willing to accept a stochastic model from the very beginning.

On the other hand, there is a much larger and older body of physical theory -- revolving, roughly speaking, around kinetic theory and statistical mechanics -- where the use of probability theory has to be justified. In fact the old masters of the game, primarily Boltzmann, felt very uncomfortable about using probability. Even now, the majority of physicists, when faced with matters involving gases or liquids, are very reluctant to use it. They will tell you that there must be some way in which they can treat it purely dynamically.

In the example which I discussed at such great length I have tried to show you at least that it's possible to justify a probability model. Except that even when this is done, there still remains this necessity of averaging. In every probabilistic model in physics and in all other sciences there must be some lack of specification over which you can average. In the example I considered it was the set S. That's the whole problem as to how probability can be introduced in kinetic theories of mechanics. It's not completely solved yet; it's in a state of flux. But what I would like to show you is that at least as far as the theory of the ideal gas is concerned, the most elementary portion of Boltzmann's work can be put into a probabilistic scheme quite consistently. In just the same way we wrote out the Master equation for our model, we can also

write a Master equation for an ideal gas. We can then study it as a mathematical entity and see whether conclusions we derive from it do or do not agree with the standard conclusions of other theories. This point of view leads to a variety of problems both mathematical and physical. Some of the mathematical problems, remarkably enough, have not been properly settled. Some of them are of a very curious nature which I will discuss with you, where progress really would be of considerable and fundamental interest, I think.

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

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

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 $\vec{V}_1, \vec{V}_2, \cdots, \vec{V}_n$. 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 $\sqrt{E}$:

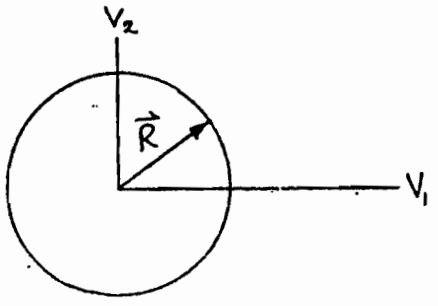


Figure 2

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 $\vec{l}$ lying in the solid angle $d\vec{l}$. 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

$$\psi_{ij}\, d\vec{l}\, dt = \frac{\delta^2}{V} \frac{|(\vec{v}_j - \vec{v}_i)\cdot\vec{l}| - (\vec{v}_j - \vec{v}_i)\cdot\vec{l}}{2} d\vec{l}\, dt \qquad (64)$$

You might notice that $|(\vec{v}_j - \vec{v}_i)\cdot\vec{l}| - (\vec{v}_j - \vec{v}_i)\cdot\vec{l}$ is either zero or simply $2|(\vec{v}_j - \vec{v}_i)\cdot\vec{l}|$. 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:

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

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

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

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

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

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 $\psi_{ij}(\vec{l})$. 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:

$$X_1^2 + X_2^2 + \cdots + X_n^2 = n \tag{69}$$

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

$$\vec{R} \longrightarrow \vec{R}' = \begin{bmatrix} X_1 \\ \vdots \\ X_i \cos\theta + X_j \sin\theta \\ \vdots \\ -X_i \sin\theta + X_j \cos\theta \\ \vdots \\ X_n \end{bmatrix} \tag{70}$$

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 $\vec{l}$. 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:

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

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

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

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

$$f_2^{(n)}(x,y,t) = \int_{X_3^2+\cdots+X_n^2 = n-x^2-y^2} \varphi(\vec{R},t)\, d\sigma_2 \tag{74}$$

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:

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

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

$$f_2(x,y,t) = f_1(x,t)\cdot f_1(y,t) \tag{77}$$

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

of 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:

$$\lim_{n\to\infty} f_k^{(n)}(X_1, \ldots, X_k; 0) = \prod_{l=1}^{k} \lim_{n\to\infty} f_1^{(n)}(X_l; 0) \tag{80}$$

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 ıere 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 rou 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 nake sense.

On the other hand, all the general conclusions one wants to draw ıbout the approach to equilibrium can be gotten from the Master equation. You ion't have to go over to the Boltzmann equation. In particular I will show you ı 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)roperty. As a matter of fact, for mathematicians, the proof is immediate. fou must simply notice that the operator Ω is self-adjoint and negative lefinite. 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

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

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

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:

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

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

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

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

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

$$f_1^{(n)}(X) = \frac{\left(1-\frac{X^2}{n}\right)^{\frac{n-3}{2}}}{\int_{-\sqrt{n}}^{\sqrt{n}} \left(1-\frac{X^2}{n}\right)^{\frac{n-3}{2}} dX} \qquad (86)$$

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:

$$\frac{\left(1-\frac{X^2}{n}\right)^{\frac{n-3}{2}}}{\int_{-\sqrt{n}}^{\sqrt{n}} \left(1-\frac{X^2}{n}\right)^{\frac{n-3}{2}} dX} \longrightarrow \frac{1}{\sqrt{2\pi}} e^{-X^2/2} \qquad (87)$$

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.

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:

$$F_n(r) = \int_{X_1^2+\cdots+X_n^2=r^2} \prod_{k=1}^{n} C(X_k)\, d\sigma \tag{90}$$

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

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

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:

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

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:

$$1 - \frac{\int_{-\infty}^{\infty} x^2 e^{-z_0x^2}C(x)\,dx}{\int_{-\infty}^{\infty} e^{-z_0x^2}C(x)\,dx} = 0 \qquad (96)$$

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

$$F_n(\sqrt{n}) = \frac{e^{nz_0}}{\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} e^{-z_0x^2}e^{i\xi(1-x^2)/\sqrt{n}}C(x)dx\right]^n d\xi \qquad (97)$$

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

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

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:

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

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:

$$\frac{1}{\sqrt{n}} \int_{S_n} g(x_1) h(x_2) \varphi_n(\vec{R}) d\sigma \sim \frac{(Ae^{z_0})^n}{\pi} \cdot \int_{-\infty}^{\infty} e^{-Bx^2} dx \cdot \left[\int_{-\infty}^{\infty} g(x) e^{-z_0 x^2} C(x) dx \right] \cdot \left[\int_{-\infty}^{\infty} h(y) e^{-z_0 y^2} C(y) dy \right] \Big/ \left[\int_{-\infty}^{\infty} e^{-z_0 x^2} C(x) dx \right]^2$$

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

$$C^*(x) = \frac{e^{-z_0 x^2} C(x)}{\int_{-\infty}^{\infty} e^{-z_0 x^2} C(x)\, dx} \tag{102}$$

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 $C(X) \Big/ \int_{-\infty}^{\infty} C(X)\, dX$. 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

$$\varphi_n(\vec{R}, 0) = \frac{C(X_1)}{\int_{S_n} C(X_1)\, d\sigma} \qquad (104)$$

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

erfectly 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:

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

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

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:

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

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 $h-1$ 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 $\eta \longrightarrow \infty$ 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 $\mu_1, \mu_2, \ldots\ldots$ (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.

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,

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

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

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

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:

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

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

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

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_γ 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_γ 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.

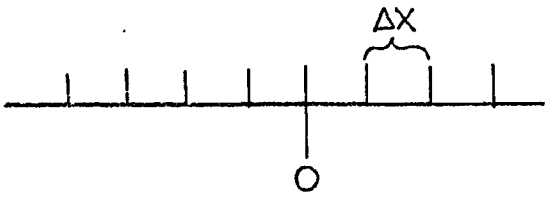


Figure 1.

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):

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

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.

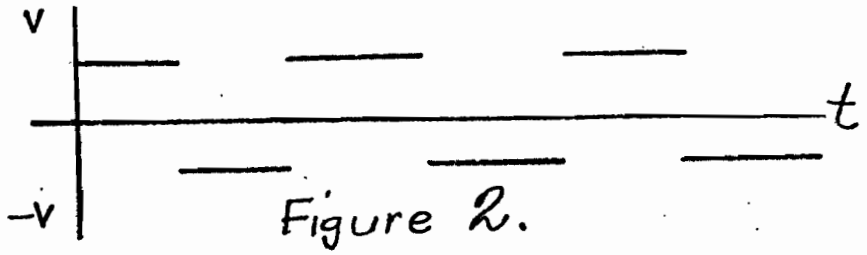


Figure 2.

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.

SEVENTH LECTURE

I will now sketch a proof -- one which does not really satisfy me entirely. It's a rather ugly thing, because one has to compute too much. One feels that one shouldn't have to compute anything at all. Such a statement ought to be provable by "pure thought." I would also like to call your attention to the fact that we have not accomplished very much, except to find an interesting way of writing the solution. We will later see a similar treatment for certain parabolic equations -- with the difference that in that case one can actually use the new probabilistic form to draw significant analytical conclusions.

The calculations which we are going to perform will not be wholly wasted, because we will need to do similar things later on. I will only prove the statement (26) for functions φ which are reasonably "decent." In particular, I will assume that my function can be written as a Fourier integral:

$$\varphi(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\xi x}\Phi(\xi)\,d\xi \tag{29}$$

Now if you substitute this in the statement we are trying to prove you get:

$$F(x,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi(\xi)e^{-i\xi x}\left\langle \cos\left(\xi v\int_0^t (-1)^{N(\tau)}\,d\tau\right)\right\rangle d\xi \tag{30}$$

Now you look at this cosine and expand it in a power series:

$$1-\frac{(\xi V)^2}{2!}\left\langle\left(\int_0^t(-1)^{N(\tau)}d\tau\right)^2\right\rangle+\frac{(\xi V)^4}{4!}\left\langle\left(\int_0^t(-1)^{N(\tau)}d\tau\right)^4\right\rangle+\cdots \qquad (31)$$

These averages, as you very well know, are called moments -- they are the moments of this strange "randomized" time. In this case, only the even moments enter. However, let me just show you first how one calculates the first moment. This can actually be done immediately because you can interchange the averaging and integration:

$$\left\langle\int_0^t(-1)^{N(\tau)}d\tau\right\rangle=\int_0^t\left\langle(-1)^{N(\tau)}\right\rangle d\tau \qquad (32)$$

(This can be easily justified. The averaging is also an integration, an integration over the space of all functions $N(t)$. So it is merely a question of interchanging the order of integration.) Now the path is clear, because I know what the distribution of $N(\tau)$ is. It is given by (21). Then, from the definition of an average -- or mathematical expectation, if you wish -- we find that:

$$\left\langle(-1)^{N(\tau)}\right\rangle=\sum_{k=0}^{\infty}(-1)^k\cdot e^{-a\tau}\frac{(a\tau)^k}{k!} \qquad (33)$$

This is an easy series, isn't it? If I pull $e^{-a\tau}$ out then what I have left is just the series for $e^{-a\tau}$. So the whole business is just $e^{-2a\tau}$. And so we can find the first moment:

$$\mu_1(t) = \left\langle \int_0^t (-1)^{N(\tau)} d\tau \right\rangle = \int_0^t e^{-2a\tau} d\tau \qquad (34)$$

It is only a little bit more complicated to calculate the second moment. It's a very common trick, which is used over and over again, to write the square of an integral as a double integral by introducing two variables. Let me do that:

$$\left\langle \left(\int_0^t (-1)^{N(\tau)} d\tau \right)^2 \right\rangle = \left\langle \int_0^t \int_0^t (-1)^{N(\tau_1)} (-1)^{N(\tau_2)} d\tau_1 d\tau_2 \right\rangle \qquad (35)$$

The expression on the right is an integral over a square. Since it is completely symmetrical in the two variables τ_1 and τ_2, I can integrate over only half the square and multiply by two.

$$2! \left\langle \iint_{0 \le \tau_1 \le \tau_2 \le t} (-1)^{N(\tau_1)} (-1)^{N(\tau_2)} d\tau_1 d\tau_2 \right\rangle = 2! \iint_{0 \le \tau_1 \le \tau_2 \le t} \left\langle (-1)^{N(\tau_1)+N(\tau_2)} \right\rangle d\tau_1 d\tau_2$$

There is a reason why I put $2!$ rather than simply 2: I am anticipating the results for the higher moments. And now you perform a very simple trick. You merely write $N(\tau_2) = N(\tau_1) + \left[N(\tau_2) - N(\tau_1) \right]$. There's a point in writing it this way, because I have separated $N(\tau_2)$ into a sum of two things which are

independent. In the case of the higher moments you do the same thing, only there will be more terms in the decomposition. Making now this substitution we have:

$$2! \int\int_{0 \leq \tau_1 \leq \tau_2 \leq t} \left\langle (-1)^{2N(\tau_1)+N(\tau_2)-N(\tau_1)} \right\rangle d\tau_1 \, d\tau_2 \tag{36}$$

Now this becomes greatly simplified, because $2N(\tau_1)$, whatever else it is, is an even number. Therefore,

$$\left\langle (-1)^{2N(\tau_1)+N(\tau_2)-N(\tau_1)} \right\rangle = \left\langle (-1)^{N(\tau_2)-N(\tau_1)} \right\rangle \tag{37}$$

which, from the meaning of the average, is nothing more than:

$$\sum_{k=0}^{\infty} (-1)^k \cdot e^{-a(\tau_2-\tau_1)} \frac{[a(\tau_2-\tau_1)]^k}{k!} = e^{-2a(\tau_2-\tau_1)} \tag{38}$$

Consequently, the second moment finally becomes:

$$\mu_2(t) = \left\langle \left(\int_0^t (-1)^{N(\tau)} d\tau \right)^2 \right\rangle = 2! \int_0^t d\tau_2 \int_0^{\tau_2} d\tau_1 \, e^{-2a(\tau_2-\tau_1)} \tag{39}$$

Now this last is just the integral of a convolution. Why does the convolution come in? Precisely because of the decomposition we made of $N(\tau_2)$ into two independent parts. It shouldn't take much imagination to see that this will also happen when you go to the higher moments. Indeed, this happens in all stochastic processes with independent increments. You always have such

convolutions coming in. The natural thing to do, then, is to take the Laplace transform. Things will be much simpler, because the Laplace transform of a convolution is the product of transforms. Now I can write (39) in the form of a double convolution by introducing the Heaviside function:

$$\Delta(t) = \begin{cases} 0 & \text{if } t < 0 \\ 1 & \text{if } t > 0 \end{cases} \tag{40}$$

It becomes just:

$$\mu_2(t) = 2! \int_0^\infty \Delta(t-\tau_2) \int_0^\infty \Delta(\tau_2-\tau_1) e^{-2a\tau_1} d\tau_1 = \Delta^* \Delta^* e^{-2at} \tag{41}$$

Then we get the Laplace transform immediately:

$$\int_0^\infty e^{-st} \frac{\mu_2(t)}{2!} dt = \frac{1}{s} \cdot \frac{1}{s} \cdot \frac{1}{s+2a} \tag{42}$$

The general formula turns out to be different for n even and odd. It's a nice exercise for you to verify that:

$$\int_0^\infty e^{-st} \frac{\mu_n(t)}{n!} dt = \begin{cases} \dfrac{1}{s^{\frac{n+1}{2}}} \cdot \dfrac{1}{(s+2a)^{\frac{n+1}{2}}}, & \text{for } n \text{ odd} \\ \dfrac{1}{s^{\frac{n}{2}+1}} \cdot \dfrac{1}{(s+2a)^{n/2}}, & \text{for } n \text{ even} \end{cases} \tag{43}$$

Now the moments are rather messy, because they are inverse Laplace transforms of these things. That suggests that rather than to work with $F(X,t)$ itself, we may better work with its Laplace transform. So let us calculate this Laplace transform.

$$\int_0^\infty e^{-st} F(X,t)dt = \frac{1}{2\pi s}\int_{-\infty}^{\infty} \Phi(\xi) e^{-i\xi X}\left(\sum_{n=0}^{\infty}\left[\frac{(\xi V)^2}{s(s+2a)}\right]^n\right)d\xi \qquad (44)$$

We have here made use of formula (30) and the moments we have calculated. The series can be summed. It's a very well-known one. You then get an extremely simple formula:

$$\int_0^\infty e^{-st} F(X,t)dt = \frac{1}{2\pi s}\int_{-\infty}^{\infty} \Phi(\xi) e^{-i\xi X}\left[1+\frac{(\xi V)^2}{s(s+2a)}\right]^{-1}d\xi \qquad (45)$$

And now it is easy to see, at least formally, what transpires. The Laplace transform of the telegrapher's equation is

$$V\frac{d^2 f}{dx^2} - \frac{s(s+2a)}{V} f - \frac{s+2a}{V} F(X,0) = 0 \qquad (46)$$

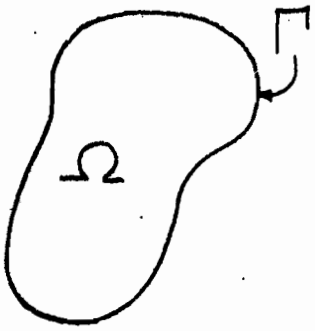
This form of the equation, as a matter of fact, is often used in solving the telegrapher's equation. But now you can directly verify that $f(X,s)$, the Laplace transform of my function $F(X,t)$, satisfies this one. You simply substitute it in. It follows then that my function $F(X,t)$ satisfies the telegrapher's equation.

This is not really a very nice proof. It is very inelegant -- although perhaps one should never speak of elegance among people who are engaged at least part of their time in applied work. The lack of elegance here is that in simply verifying this formula you are essentially solving the equation. It's sort of cheap to simply verify something by brute force. It would be much nicer if one could see it directly. But I don't want you to take me seriously. I don't want to stop doing things because they don't adhere to certain principles of elegance. Boltzmann used to say, when he was criticized that his work was inelegant, that elegance should be left to shoemakers and tailors. Perhaps this is really true. But this proof is a little bit aesthetically dissatisfying, I would say.

This same proof goes also for a higher number of dimensions. Again it's simply a matter of writing the Laplace transform and verifying the same formula. There are certain disappointments in connection with this. Because really one learns comparatively little. At least I haven't learned anything really startling by doing it this way. The situation changes radically if you go to other differential equations, those of parabolic and elliptic type. When you study them from the point of view of stochastic problems from which they arose, then a considerable amount of new knowledge and a new approach results.

Let me first give you one more example of the great advantage which can accrue if one looks at the same thing from a different point of view. I will consider a very classical problem: The asymptotic behavior of eigenvalues of the Laplacian. You have some region Ω with a boundary Γ. These are in the plane, let us say. Everything is assumed to be suitably smooth

so that classical analysis is applicable. The problem we are considering can then be stated:



$$\frac{1}{2}\Delta u + \lambda u = 0 \quad \text{in } \Omega \tag{47}$$

$$u = 0 \quad \text{on } \Gamma \tag{48}$$

You will recognize this as the vibrating membrane problem. I now consider the eigenvalues in their increasing order: $\lambda_1, \lambda_2, \lambda_3, \ldots$. And I denote by $A(\lambda)$ the number of eigenvalues less than λ. In the plane as λ goes to infinity,

$$A(\lambda) \sim \frac{|\Omega|}{2\pi}\lambda \tag{49}$$

The symbol $|\Omega|$ means the area of the region. In three dimensions it becomes the volume and the lambda gets raised to the power 2/3. The curious thing is that the constant depends only on the area, not on the shape at all.

There is a very heroic story connected with it. Let me tell you the story. This theorem, in three-dimensional space, was conjectured for the first time by the great H. A. Lorentz in 1908. This was during a meething in Goettingen devoted to the new quantum theory -- it wasn't quantum mechanics yet. There was, by that time, a new theory of specific heats which Debye proposed. It extended the older theory of Einstein and was one of the first great triumphs of the new quantum theory. By playing around with Debye's theory

ou can deduce some properties of the heat content. Now it's perfectly natural hat the heat content should be proportional to the volume. And certainly it annot depend on the shape. So there was the conjecture. Three years later, n 1911, the late Herman Weyl proved this. In so doing he made a tremendously nteresting contribution to this branch of mathematics. It is in this connec- ion that he first introduced the famous variational characterization of the igenvalues, the so-called minimax characterization. From there on the method as applied very successfully to different problems in many different ways. Weyl's original proof was not difficult, but not entirely "understandable." lthough you understand the steps you still don't quite know what makes it tick.

I will now give you a proof which has the advantage that one understands very well how it works. In fact, it makes the theorem appear relatively superficial. The way it is stated now it has a certain appearance of depth. But we will look at it from a somewhat different point of view -- the point of view connected with probability and diffusion. What I'm going to give you will not be a complete proof, because there are several delicate points which have to be justified. This can be done and it has been done.

We will look at this theorem in a different way, in what I think is the proper way. We regard the equation (47) as arising not from the vibrating membrane or anything of that sort, but from the diffusion equation:

$$\frac{\partial P}{\partial t} = \frac{1}{2}\Delta P \tag{50}$$

You know perfectly well how this goes. I will, of course, require that

$$P(x,y,t) = 0 \quad \text{on } \Gamma \tag{51}$$

and, in addition, I want to make the following assumption:

$$P(x,y,t) \longrightarrow \delta(x-x_0)\delta(y-y_0) \text{ as } t \longrightarrow 0 \qquad (52)$$

So, if you look on $P(x,y,t)$ as being the concentration of some diffusing stuff, it's initially all concentrated at (x_0,y_0).

If you take the equation (50) subject to the initial condition (52) and the boundary condition (51), then it is well known from entirely classical stuff how to write at least a formal solution. You simply make a separation of variables and find that you can write

$$P(x_0,y_0\,|\,x,y,t) = \sum_{j=1}^{\infty} e^{-\lambda_j t}\varphi_j(x,y)\varphi_j(x_0,y_0) \qquad (53)$$

The $\varphi_j(x,y)$ are the normalized eigenfunctions of the operator $\frac{1}{2}\Delta$; the λ_j are the eigenvalues, the same ones as before. This is one of the most standard results in classical mathematical physics.

Now let us try to interpret this $P(x_0,y_0|x,y,t)$. The condition (51) means that Γ is an absorbing barrier -- any of this diffusing stuff which reaches it is eaten up. So it's perfectly normal that $P(x_0,y_0|x,y,t)$, the concentration, is going to approach zero as you approach the boundary. Initially all the mass is at the point (x_0,y_0), but as time goes on it diffuses out and gets absorbed.

Now suppose that t is very small. Place yourself in the position of the diffusing stuff -- you are there together with all the other particles that are going to diffuse. Now you are going to move for a very short time.

In such a short time you don't know what will happen on the boundary, because you haven't had a chance to discover what horrible disasters are going to befall you. The knowledge of your fate, that you are going to be eaten at the boundary, has not yet reached you. The smaller t is, the less knowledge of this you possess. Consequently, for small t, a good approximation ought to be given by the solution of the same diffusion equation without regard to the boundary. Of course, it's an intuitive principle, this principle of not feeling the boundary. But it leads me to think that early in the game the solution will be that of the unrestricted diffusion problem:

$$\frac{1}{2\pi t} e^{-\frac{[(x-x_0)^2+(y-y_0)^2]}{2t}} \tag{54}$$

which is perfectly well known.

If we are so courageous to think that this is a good approximation for small t, then perhaps it also holds when $X=X_0$ and $y=y_0$. It then becomes $1/2\pi t$. But then that gets us, using (53),

$$\sum_j e^{-\lambda_j t} \varphi_j^2(x_0,y_0) \sim \frac{1}{2\pi t} \quad \text{as } t \to 0 \tag{55}$$

Now integrate both sides of this asymptotic equality over the region. Using the fact that the eigenfunctions were normalized -- the integral of $\varphi_j^2(x,y)$ is one -- you get:

$$\sum_j e^{-\lambda_j t} \sim \frac{|\Omega|}{2\pi t} \tag{56}$$

And now we are almost through. All we have to do is know a little bit of mathematics. Because I can write the sum as a Stieltjes integral, a very convenient way of writing it:

$$\sum_j e^{-\lambda_j t} = \int_0^\infty e^{-\lambda t} dA(\lambda) \tag{57}$$

$\frac{1}{t}$ can also be written as a similar integral, so you get from (56):

$$\int_0^\infty e^{-\lambda t} dA(\lambda) \sim \frac{|\Omega|}{2\pi} \int_0^\infty e^{-\lambda t} d\lambda \tag{58}$$

which is true as t goes to zero. Now you know about this function $A(\lambda)$ that it is non-decreasing. After all, the bigger λ is, the more λ_j's you include. Now there is a theorem that says that from such an asymptotic equality, for t going to zero, there results another asymptotic equality for λ going to infinity:

$$A(\lambda) \sim \frac{|\Omega|}{2\pi}\lambda \quad \text{as } \lambda \longrightarrow \infty \tag{59}$$

This theorem is known as the Hardy-Littlewood-Karamata Tauberian theorem. It produces the result (59) we were after.

Now this proof has one tremendous advantage. (Actually, it isn't a proof yet because we have done a certain amount of skullduggery.) It is intuitively completely appealing. Moreover, it is unforgettable. The basic principle, this not feeling the boundary for a short time, is only visible

if you look upon the problem from the point of view of the diffusion equation. From the wave equation viewpoint there is no such simple interpretation. This principle of not feeling the boundary has by now been exploited many more times in similar connections. Even here it can tell you something more. It tells you that it doesn't matter what boundary condition you put on. The same asymptotic behavior ought to be found for any homogeneous boundary condition. That, of course, is part of Weyl's theorem.

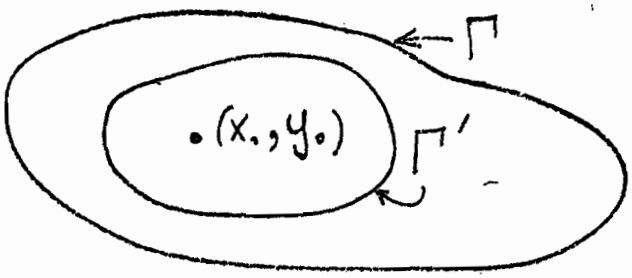
This is an example of the advantage accrued from looking at a mathematical equation in different physical contexts. If you look at it from one point of view it may be much more revealing than if you look at it from another. This is why one should try to formulate even familiar things in as many different ways as possible. You undoubtedly will learn something in the process. Just to finish up this story, let me show you that our principle not only illuminates the result, it also suggests the detailed proof. I claim that immediately I can have an inequality,

$$P(x_0, y_0 | x, y, t) \leq \frac{1}{2\pi t} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2t}} \tag{60}$$

From the physics of the problem this is a yelling triviality. Because what does it mean? It means that the concentration when there is an absorbing barrier is less than when there is not.

To find an inequality which goes the other way is only a little bit more difficult. Let us draw around this point (x_0, y_0) another curve,

call it Γ'. We make Γ' lie entirely within Γ:



Consider now the same problem with Γ replaced by Γ'. You put the absorbing barrier at Γ'. I will denote the solution by $P'(x_0, y_0 | x, y, t)$. Now I claim that P must be greater than or equal to P':

$$P'(x_0, y_0 | x, y, t) \leqslant P(x_0, y_0 | x, y, t) \tag{61}$$

Because, after all, the stuff is going to be eaten earlier. In the first problem I still have the possibility of diffusing outside Γ' and then coming back in without having been eaten. These possibilities are denied to me if I have the boundary at Γ'. So P is bigger than P'. What is the obvious thing to do? To select the boundary Γ' so that you can solve for P'. Either a circle or a rectangle will do. So now you have inequalities going both ways, inequalities which certainly hold for all x and y. So you can put $x = x_0$ and $y = y_0$ to get.

$$P'(x, y | x, y, t) \leqslant \sum_j' e^{-\lambda_j t} \varphi_j^2(x, y) \leqslant \frac{1}{2\pi t} \tag{62}$$

This inequality is true for every point (x, y) provided I surround it by a certain square. Now think of the whole region as being covered by little squares. At the center of each such square you have the inequality. Then you

just integrate out. Of course, to finish this up, you have to know how the left hand side behaves as t goes to zero. But I know the solution explicitly for a square, so it can be simply verified. So this is a very brief sketch of essentially the full proof.

I wish to make one observation. Since one tries to sell the method, one may as well be honest. The reason the proof works so nicely and neatly is because you have all these inequalities which are properties of parabolic equations. One should distinguish between being able to see things clearly and being able to prove them. If you wanted to do the same problem for the biharmonic equation then you could still enunciate the principle of not feeling the boundary. You could still get some answer, but you would not be quite sure how to prove it. There are no corresponding inequalities for the biharmonic heat equation, for instance. In our diffusion problem $P(x_0, y_0 | x, y, t)$ remains always positive. The inequalities are quite obvious, and their proof is intuitively quite clear.

Now this is actually a very interesting question. What equations do and what equations don't have the properties that we need? If you have an elliptic operator Λ of second order then for the corresponding equation $\frac{\partial P}{\partial t} = \Lambda P$ will have all the necessary inequalities. For higher order equations I don't think anybody really knows what to do in this particular way.

But anyway, here is an example of the very classical theorem, done in all the textbooks, which can be treated by a probability approach. Our problem was one in differential equations, a purely mathematical problem. But that it came from a physical situation is not something to be sneezed at.

Because the knowledge of where it comes from gives you a way of approaching the problem. You also see that taking it as coming from one part of physics may be more useful, more illuminating, than taking it as coming from another part of physics. It's a real advantage, both to look at an equation from different points of view and to be aware of the physical interpretation. That is the reason why I believe that if any real breakthrough ever comes in non-linear differential equations, it will come only when a certain amount of physical knowledge will be amassed, so that some intuition will be developed.

The rest of the time I will devote to discussing problems arising from the simplest cases of Brownian motion. I was hoping, when I planned the lectures, to be able to do much more with Brownian motion. But it turns out that one always is too optimistic as to how much time one has. Consequently, I will have to confine myself to the so-called Brownian motion of a free particle. This is already familiar to you in the classical theory of Einstein and Smoluchowski. But I will look at it from a somewhat different point of view, a point of view you might call integration in function spaces.

Now how does the original theory of Einstein or Smoluchowski approach the simplest case of Brownian motion? Let's take a straight line and say to ourselves that we have a Brownian particle starting at $X = X_0$. I assume from the very beginning that, at best, I will be able to predict the probability density $P(X_0|X,t)$. If I multiply by dX then this is simply the probability of finding a particle between X and $X+dX$ at time t if I start from X_0 . And now one makes an assumption of the past being independent of the future. Mathematically, it can be formulated as follows:

$$P(X_0|X,t) = \int_{-\infty}^{\infty} dy\, P(X_0|y,t')\,P(y|X,t-t') \tag{63}$$

To see this, you notice that $P(X_0|X,t)$ is the probability of coming from X_0 to X in time t. But at any intermediate time, say t', I have to be somewhere. Let's say that at time t' I have the position y. So I have made a transition from X_0 to y in the time t'. But then I must make a transition from y to X in the time $t-t'$. What is the probability of this? Well, since the past is independent of the future, the probability is the product of two probabilities. In fact it is just the integrand in equation (63). But since y could be anywhere, since I don't know what it is, I have to integrate over all the possibilities.

This is a very famous equation. It is known under a different name, depending on whether you are a mathematician or a physicist. In the mathematical literature it became known as the Chapman-Kolmogoroff equation, although I do not quite know why Chapman was attached to it. Among the physicists, it is known as the Smoluchowski equation, because he considered it in great detail. It is interesting to remark that a similar equation can also be written in quantum mechanics. Except that it is not anymore the probability density that is involved. Instead it's a complex-valued function known as the probability amplitude. One of the reasons this is so is because in quantum mechanics you cannot make the same argument. You cannot say that if you are at one place at time t and another at time t' then you must have been somewhere in the meantime. This is wrong for a very interesting reason. In quantum mechanics one must always look upon things operationally. I must be able to perform an

experiment to find the intermediate position. But if I do that it will change where I end up. Any experiment you can think of will disturb the particle in some way.

Now if you don't make any further assumptions, there are slews of solutions of this equation (63). This is even true in the spatially homogeneous case. This is the case where $P(X_0 | X, t)$ depends only on the difference $X - X_0$, not on X and X_0 separately. I can go further even than that. I can consider the symmetric case where $P(X_0|X,t)$ depends only on the absolute value $|X - X_0|$:

$$P(|X-X_0|,t) = \int_{-\infty}^{\infty} P(|x-y|,t')\,P(|y-x_0|,t-t')\,dy \qquad (64)$$

The probabilities of going in one direction or the other are completely equal. Even then there are a tremendous number of solutions of this equation. For instance, a somewhat unusual solution, not perhaps known to you all, would be

$$\frac{t}{\pi}\,\frac{1}{t^2+(X-X_0)^2} \qquad (65)$$

If you substitute this into the equation you will verify that it is a solution. And there are many, many others.

But if you assume, in addition to the symmetry, that the second moment is finite, then there is only one solution. Or, rather, one form of a solution. It is namely the Gaussian distribution:

$$\frac{1}{\sqrt{2\pi t}\,\sigma}\cdot e^{-\frac{(X-X_0)^2}{2\sigma^2 t}} \qquad (66)$$

σ^2 is the second moment, the mean square displacement:

$$\int_{-\infty}^{\infty} (X-X_0)^2 P(X-X_0, t)\, dX \quad < \infty \tag{67}$$

It is really the finiteness of this moment that forces the solution (65). You recognize it as the solution of the diffusion equation. I will show you formally how equation (64) can be reduced back to the diffusion equation when the second moment (66) is finite.

EIGHTH LECTURE

In the symmetric, spatially homogeneous case with a finite second moment we have found that the only solution of equation (63) was:

$$\frac{1}{\sqrt{2\pi t}} e^{-\frac{(X-X_0)^2}{2t}} \tag{68}$$

At least you believed me when I told you so. In rewriting it I have put σ equal to one. This function is, of course, the well-known fundamental solution of

$$\frac{\partial P}{\partial t} = \frac{1}{2} \frac{\partial^2 P}{\partial X^2} \tag{69}$$

In writing the fundamental equation (63) I used the assumption about the past being independent of the future. This can be extended to answer a more complicated question. You can ask, for instance, for the probability of finding a particle between the limits (α_1, β_1) at time t_1, between (α_2, β_2) at time t_2, . . . , and between (α_n, β_n) at time t_n. This problem is immediately soluble because the answer can simply be expressed in terms of the elementary transition probability $P(X_0|X,t)$. This is again because of the assumption of the independence of the past and the future. In fact, the probability is simply

$$\int_{\alpha_1}^{\beta_1} \cdots \int_{\alpha_n}^{\beta_n} P(X_0|X_1,t_1) P(X_1|X_2,t_2-t_1) \cdots P(X_{n-1}|X_n, t_n-t_{n-1}) \tag{70}$$

(We assume, of course that $0<t_1<t_2<\cdots<t_n$.)

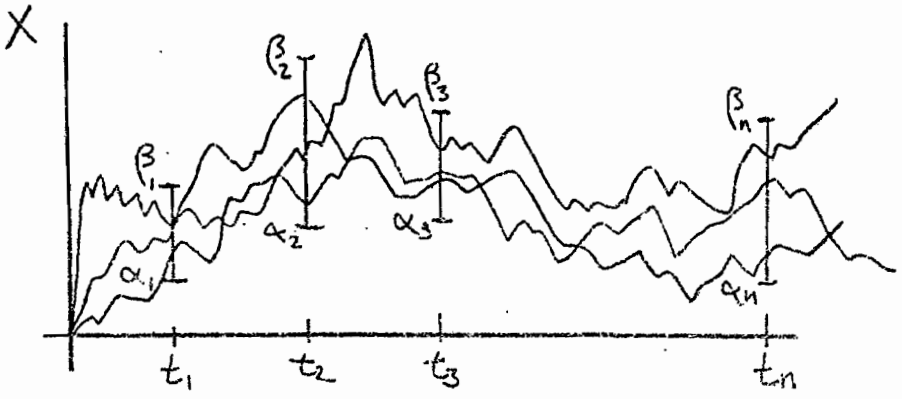
In the integrand I have a product of probabilities. That I can always separate the probability into such a product is referred to as the Markovian property. In our case, it's actually simpler than the Markovian property. It's simply the independence of the increments during non-overlapping time intervals.

You might say that nothing new has been added. This probability is simply expressible in terms of the basic transition probability. However, I can now make this a beginning of a measure theory. Because I can take the following point of view. Let me think of the path of the particle in space-time:

Indeed, let me think of all possible paths the particle can take. Some may even be discontinuous. A priori, I don't know. So I will consider the set of all possible functions $X(t)$ such that $X(0)=0$. Let's call this set of functions, or paths, S.

Now in this space S I want to introduce a notion of a measure. This simply means that I want some analogue of volumes or areas. I have made a beginning because I know already measures of certain sets. Let me explain this more carefully using Figure 3.

Figure 3.



I have marked the points $t_1 < t_2 < \cdots < t_n$ on the time axis. Above these points I have marked the intervals $(\alpha_1, \beta_1), \ldots\ldots, (\alpha_n, \beta_n)$. These are known as windows. Now you consider the set of all paths which pass through these windows. This is of course a sub-set of the original set. To this sub-set I will assign the measure given by formula (70).

Let me remind you of what one does in ordinary measure theory. In the plane, for example, one first takes certain elementary figures, say rectangles, with sides parallel to the coordinate axes. One then assigns a measure to these elementary sets, ordinary areas if you wish. Then from these sets you can build up more complicated sets. For instance, you can build a circle from an infinite number of such rectangles. Of non-overlapping rectangles, if I do it carefully. Then I simply add the axiom that the measure of any set which consists of such disjoint sets becomes simply the sum of the measures. In this way I define the measure of sets which are more interesting than mere rectangles.

In Figure 3 I can choose my windows in many, many different ways. Each choice defines for me a set of paths, namely the paths that pass through the windows. These sets play the role of the rectangles. They are the elementary sets. Their measure is given to me by the elementary theory of Brownian motion -- in fact, by formula (70). Having this, why not imitate the measure theory in the plane and ask ourselves for the measures of more interesting sets. For instance, the set of all those paths for which

$$\int_0^t x^2(\tau)\,d\tau < \alpha \tag{71}$$

Before we answer such specific questions we must ask ourselves some other ones. Although this measure enjoys most of the properties of ordinary measure, it also has some very interesting peculiarities. Suppose I take the set of all continuous paths. Now I want to find the measure of this thing. I am not going to go into all the details. It is a rather complicated affair. The bare facts are that this set C, the set of all continuous paths, is not measurable. To be somewhat technical, the outer measure of C is one and the inner measure is zero. To find the outer measure, you simply cover the set of continuous functions by combinations of sets that are measurable and consider the measure of this combination. Of course, you can do this in many ways so you take the least upper bound of all the possible measures. You just get 1. To find the inner measure you do the same thing with the complement of C and subtract from one. It turns out to be zero.

This is very unpleasant. It's certainly a pathological feature. One knows intuitively that Brownian paths ought to be continuous. So what one does is simply to add one more set, namely the set C of continuous functions and to say from the beginning that its measure is one. Then I have to re-do the whole thing and check that the measure I obtain is a reasonable measure. That can be done. And in this way you obtain what is now known as the Wiener measure*. It was first introduced by Norbert Wiener in a somewhat different

---

* The Wiener measure was introduced around 1922. It appeared in a whole series of papers in the MIT Journal of Mathematics and Physics. These papers were not very understandable, so for a long time it was overlooked. These were remarkable papers, because at the time he wrote them very few people in America even knew the ordinary measure theory. So no one foresaw the interest and importance of this. There is a rather good exposition in the book by Paley and Wiener. You will find it in the chapter on random functions. Now, these things are done much more abstractly. If you can swallow some of the abstractions then you will find it done in great generality in the book by Doob on stochastic processes. All methods of introducing this measure are somewhat pedagogically unsatisfying. To my mind the best is still the one which Paley and Wiener used. But it has also many severe disadvantages. Doob's is the easiest to formulate, but involves a lot of tedious details.
It is regretable that in more modern literature (notably in Doob's book) the great pioneering work of Wiener is not quite given the recognition it deserves.

way. It is a measure in the space of continuous functions built upon these elementary sets, quasi-intervals as they are often called.

Once you have this measure many other very interesting questions can be answered. You can now ask for the measure of the set of differentiable functions. This set turns out to be measurable and the measure is zero. That means, in a sense, that almost every Brownian path is non-differentiable. This is a pleasing result because anybody who observes the Brownian motion under a microscope sees a very erratic motion. In the mathematical idealization this particular measure actually gives it to you. In fact, it proves much more. It proves, for instance, that almost every path is of unbounded variation. That is to say, almost every path has infinite length. This is perhaps somewhat less pleasing, but it is a rigorous conclusion from your measure theory and nothing can be done about it. As a matter of fact, it turns out that the set of functions satisfying (71) is measurable. Its measure can actually be calculated. You get a complicated formula involving elliptic functions which I will not trouble you with at the moment.

Now you might think that such a pathological creature as this Wiener measure ought to be of comparatively little use. Hence it will probably surprise you that it has rather remarkable uses in the theory of differential equations. It's a rather interesting route, this starting from a simple stochastic process in physics, passing through a process of abstraction to a measure theory, and then coming back to look at familiar classical differential equations from a strange point of view. To illustrate it I am going to consider with you a very basic problem. I will treat it more or less heuristically, because rigorous proofs can be found in the literature. We will concentrate on the story behind the proof.

I am going to take a function $V(x)$, a reasonable one. And I am going to consider the following expression:

$$\int_0^t V(X(\tau))\,d\tau \tag{72}$$

If $V(x)$ happens to be x^2, then that's exactly the same thing as (71). Such an expression is called a functional because the independent variable is a function, the whole path $X(t)$. It is not a function of one variable or even of a finite number of variables. To find its value you must know the whole continuous path $X(t)$. Such things are familiar in the calculus of variations. There you try to minimize or maximize not a function, but a functional.

This functional (72) can assume different values depending on the choice of the Brownian path. It depends on how your particle decides to move. As in every statistical theory, one is clearly interested in the distribution of these values. This distribution now is perfectly well defined. It's simply the measure of the set of those paths for which (72) is less than α. I will call it $\sigma(\alpha,t)$:

$$\sigma(\alpha,t)=\mu\left\{\int_0^t V(X(\tau))\,d\tau < \alpha\right\} \tag{73}$$

The Wiener measure I have called μ. It is the same, of course, as the probability. The two words are synonymous.

Now I can go a little farther. I can use tools familiar to everyone who works even in elementary statistics. Everybody knows that in order to calculate a distribution function all you need are its moments. But rather than to use the moments I will use what's called the moment generating function.

For this purpose I will assume that $V(x)$ is positive. It's not necessary, but we will avoid getting into any unnecessary details. Then the moment generating function is given by:

$$\int e^{-u\int_0^t V(x(\tau))d\tau}\, d\mu \tag{74}$$

The integral is with respect to the Wiener measure. Once you have a measure defined then integration becomes defined also. In probability theory we abbreviate this:

$$E\left\{e^{-u\int_0^t V(x(\tau))d\tau}\right\} \tag{75}$$

The symbol E means the mathematical expectation. I can also write it in the form:

$$E\left\{e^{-u\int_0^t V(x(\tau))d\tau}\right\} = \int_0^\infty e^{-u\alpha}\, d_\alpha\sigma(\alpha,t) \tag{76}$$

So you see that the moment generating function is simply the Laplace transform of the probability distribution.

Now the interesting thing is that a calculation of this functional is related very closely -- I will tell you in a moment how it is related -- to solving the differential equation

$$\frac{\partial Q}{\partial t} = \frac{1}{2}\frac{\partial^2 Q}{\partial X^2} - uV(X)Q \qquad (77)$$

In fact, if I know how to solve this equation with appropriate conditions, then I will know how to find (76). I will show this to you by a heuristic argument. A rigorous proof is not entirely easy. My original proof was very lengthy with lots of nuisances. But at that time I didn't know any better. By now there are a lot of rigorous proofs. Perhaps this is good. As a friend of mine said: "by the time there are two independent proofs I believe the theorem."

Why are these things related? To bring this out, I will consider not a simple Brownian motion, but a Brownian motion with the following modification. As before, a particle starts from zero and begins to move. But then, if the particle happens to be at X at time t, then it will have a probability $uV(x)dt$ of disappearing. A friend of mine once called it "random walk with manholes." There is the probability $uV(x)dt$, if you are at X at time t, that you will absolutely fall into the manhole and never be seen, never be met with again.

What is then the interpretation of (79)? I claim that this is the probability of survival if you have chosen the path $X(\tau)$. That is perfectly clear, because the probability of surviving for a time dt is just

$$1 - uV(x)dt \approx e^{-uV(x)dt} \qquad (78)$$

What happens to you in non-overlapping time intervals still is going to be independent. So the probability that you will survive for the full time t will be a product of such things. It will be a product of exponentials. This will produce me a sum in the exponent, in other words an integral. So I will get simply

$$e^{-u\int_0^t V(X(\tau))\,d\tau} \tag{79}$$

for the probability of survival if you travel the path $X(\tau)$. This is the probability that someone will shake your hand at time t, that you will be there. Now the average of this or the expectation gives the over-all probability of survival, regardless of what path you are taking.

Now it is intuitively obvious, at least without "manholes," that Brownian motion and diffusion, ordinary classical diffusion, are mathematically equivalent. Remember that the function (68) was also the fundamental solution of the diffusion equation. Now how do you make this equivalence physically obvious? You assume that at time zero you start a lot of particles out at $X = X_0$. Each of them will take a different path. Each will undergo Brownian motion. Then, at the end of time t, you look at the interval $(X, X+dX)$ on the X axis. The number of particles there is roughly given by:

$$n \cdot \frac{1}{\sqrt{2\pi t}} e^{-\frac{(X-X_0)^2}{2t}} dX \tag{80}$$

where n is the number that originally started. So you see that diffusion is the macroscopic manifestation of Brownian motion. If you look at a

diffusing substance on a molecular scale then each particle performs a Brownian motion. If you then collect them up at time t, their number in $(x, x+dx)$ is very well approximated by the probability times n. In fact, the error will be on the order of $\sqrt{n}$. So if you have essentially infinitely many particles to begin with there is almost no error at all. So diffusion is what you see if you watch the Brownian motion not of one particle but of a lot of them.

Now how about this Brownian motion with manholes? What is the macroscopic counterpart? It is evidently diffusion with a distribution of sinks. The density of sinks, so to speak, is given by the function $V(x)$. The Brownian motion with manholes, with this possibility of disappearance, is macroscopically equivalent to diffusion with sinks. And this is exactly what the differential equation (77) describes.

Of course, we must have conditions on the solution of the differential equation. Everything must start from $x=0$ when $t=0$. This means that

$$Q(x,t) \longrightarrow \delta(x) \quad \text{as } t \to 0 \tag{81}$$

This is the only condition I have. Then $Q(x,t)$ is the concentration at time t at x. But, since my initial concentration integrates to one, $Q(x,t)dx$ is the probability of not having disappeared and being found at x which is within dx . So if you want to know the probability that you are somewhere at time t, that you have not disappeared, you simply integrate on x. This gives you the survival probability. So there is no particular surprise in the relation

$$E\left\{e^{-u\int_0^t V(x(\tau))d\tau}\right\} = \int_{-\infty}^{\infty} Q(x,t)\,dx \qquad (82)$$

I do not want to create the impression that this is a proof. It is merely an intuitive argument. I must show rigorously that my description of Brownian motion through this measure really does produce the result. But, as I said, there are proofs available in the literature.

One can extend this slightly. I will need a somewhat more refined formula than (82). First of all, I don't have to start at $X=0$. I would like to be able to start from anywhere. To do that I will simply add X_0 to $X(\tau)$. (I will always assume that $X(0)$ is zero.) $Q(x,t)$, which has a singularity at zero, is merely replaced by $Q(X_0|X,t)$ that has a singularity at X_0. Because now you want the condition:

$$Q(X_0|X,t) \longrightarrow \delta(x-x_0) \qquad (83)$$
$$\text{as} \quad t \longrightarrow 0$$

So, starting the particle at X_0 rather than at zero gives

$$E\left\{e^{-u\int_0^\infty V(X_0+X(\tau))d\tau}\right\} = \int_{-\infty}^{\infty} Q(X_0|X,t)\,dx \qquad (84)$$

Now what about the limits of the integral? Why did I have $-\infty$ to $+\infty$? Because I did not ask where I end up. I only asked for the probability of survival. But suppose I were to make these limits from a to b. How would I

have to modify the left hand side? You see, I am asking not only for the probability of survival, but for the probability of survival and of ending up between a and b. It can be written

$$E\left\{e^{-u\int_0^\infty V(X_0+X(\tau))d\tau};\, a<X_0+X(t)<b\right\}=\int_a^b Q(X_0|X,t)\,dX \qquad (85)$$

The meaning of this notation is that my integration will not be over all the paths, but only over paths ending up between a and b.

Let me now take $a=X-\epsilon$ and $b=X+\epsilon$. And let me divide both sides of the equation by the integral

$$\int_a^b \frac{e^{\frac{-(u-X_0)^2}{2t}}}{\sqrt{2\pi t}}\,du \qquad (86)$$

In a moment you will see why I divide by this integral. The result then is

$$\frac{E\left\{e^{-u\int_0^\infty V(X_0+X(\tau))d\tau};\, X-\epsilon<X_0+X(t)<X+\epsilon\right\}}{\frac{1}{\sqrt{2\pi t}}\int_{X-\epsilon}^{X+\epsilon} e^{-\frac{(X_0-u)^2}{2t}}\,du}=\frac{\int_{X-\epsilon}^{X+\epsilon} Q(X_0|X,t)\,dX}{\frac{1}{\sqrt{2\pi t}}\int_{X-\epsilon}^{X+\epsilon} e^{\frac{-(X_0-u)^2}{2t}}\,du} \qquad (87)$$

On the left hand side is a very understandable expression. The denominator is simply the probability that $X_0 + X(\tau)$ lies between $X-\epsilon$ and $X+\epsilon$ That comes from my original, simple theory. The numerator is an integral over a certain set of paths. The set of paths that end up between the same limits $X-\epsilon$, $X+\epsilon$. So the left hand side is just the conditional expectation. There is a symbol for this. It is written:

$$E\left\{ e^{-u\int_0^t V(X_0+X(\tau))\,d\tau} \;\middle|\; X-\epsilon < X_0+X(t) < X+\epsilon \right\} \tag{88}$$

Now let epsilon go to zero:

$$E\left\{ e^{-u\int_0^t V(X_0+X(\tau))\,d\tau} \;\middle|\; X(t)+X_0=X \right\} = \frac{Q(X_0|X,t)}{\frac{1}{\sqrt{2\pi t}} e^{-\frac{(X-X_0)^2}{2t}}} \tag{89}$$

It requires a certain amount of talking to make it entirely rigorous, but you can see it intuitively. Finally, multiplying out, I get the formula

$$Q(X_0|X,t) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(X-X_0)^2}{2t}} E\left\{ e^{-u\int_0^t V(X_0+X(\tau))\,d\tau} \;\middle|\; X(t)=X-X_0 \right\} \tag{90}$$

Now you can say to me, what have you gained? You have taken a classical expression, the fundamental solution of a certain differential equation, and laboriously written it in the form of an average over a space of functions. Where does this get you? To show you what has been gained I will prove a theorem which is a very important theorem in quantum mechanics. This will be reminiscent of the proof I gave you in the last lecture. You will see how this formulation almost carries you to the answer.

The theorem concerns the distribution of eigenvalues of the Schroedinger equation. To prove it I will need to make further assumptions about the function $V(x)$. So far it's completely general. Let me assume that $V(x)$ is symmetric, that is, that $V(x) = V(-x)$. That's not very important. Also, that $V(x)$ goes to ∞ as x goes to ∞. That is important. Let me also put u equal to one, since it doesn't really matter. Our equation then reads:

$$\frac{\partial Q}{\partial t} = \frac{1}{2}\frac{\partial^2 Q}{\partial x^2} - V(x)Q \tag{91}$$

Now you associate with this equation the corresponding eigenvalue problem. This you get by separating variables in the usual way. It is

$$\frac{1}{2}\frac{d^2\psi}{dx^2} - V(x)\psi = -\lambda\psi \tag{92}$$

This is exactly Schroedinger's equation with the potential $V(x)$.

It has been proved that the spectrum forms a discrete set. That is why we put restrictions on $V(x)$. If $V(x)$ does not approach infinity on both sides then you may have a continuous spectrum. All sorts of horrible things can then happen. So we will stick to the discrete case.

Now, if you have the discrete eigenvalues $\lambda_1, \lambda_2, \ldots\ldots$ and the corresponding normalized eigenfunctions $\psi_1, \psi_2, \ldots\ldots$ then you can write a very simple expression. It is:

$$Q(X_0|X,t) = \sum_j e^{-\lambda_j t} \psi_j(X_0)\psi_j(X) \tag{93}$$

This is, again, an entirely classical result. It is a solution of equation (91). This equation is not the same as the time-dependent Schroedinger equation. The time-dependent Schroedinger equation has an i floating around. But if I do a separation of variables I get just equation (92). So the time-independent Schroedinger equation (92) can be looked upon as arising from the purely parabolic equation (91). And it is this time-independent equation which gives you the eigenfunctions and eigenvalues.

Now I have an expression for $Q(X_0|X,t)$, namely (90). If I put $X = X_0$, then I get:

$$\sum_j e^{-\lambda_j t}\psi_j^2(X) = \frac{1}{\sqrt{2\pi t}} E\left\{ e^{-\int_0^t V(X+x(\tau))d\tau} \,\middle|\, x(t)=0 \right\} \tag{94}$$

Now, to eliminate $\psi_j(X)$ -- for the time being they are of no interest to me -- I will integrate out the whole thing on X. Then, because the integral of $\psi_j^2(X)$ is one (because they are normalized) I finally obtain:

$$\sum_{j=1}^{\infty} e^{-\lambda_j t} = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} dx\, E\left\{ e^{-\int_0^t V(X+x(\tau))d\tau} \,\middle|\, x(t)=0 \right\} \tag{95}$$

And that's the basic formula.

Now I'm almost through. At least the intuitive argument will get me there almost immediately. To put in all the steps is not entirely trivial. What you do is to again introduce the function

$$A(\lambda) = \text{the number of } \lambda_j\text{'s} < \lambda .$$

Then the sum can be written as an integral:

$$\sum_{j=1}^{\infty} e^{-\lambda_j t} = \int_0^{\infty} e^{-\lambda t}\, dA(\lambda) \tag{96}$$

which is a very convenient way to write it. Now I already showed you in the last lecture that to find the behavior of $A(\lambda)$ for large λ I need the behavior of the integral for small t. And that's where the great advantage of this point of view comes in. Because now the behavior for small t is almost immediately visible. Just think about it for a moment and you will see what happens. This $X(\tau)$ is a Brownian path. You recall that at time $t = 0$ it starts from $X = 0$. That's my assumption. Also, I am only considering those Brownian paths which at time t come back to zero. That's the meaning of the conditional expectation. So at time zero it is zero and almost immediately after, it again is zero. You will agree with me that it is very unlikely that $X(\tau)$ will become very large in the meantime. In other words, as complicated as the Brownian path is, a good first approximation would be to simply replace it by the chord connecting its end points. Because the function cannot deviate very much from it. Granted that the curve is terribly wiggly, but I'm not approximating a derivative -- I am only approximating the function! What I have, of

course, to prove is that the probability of a sizeable deviation in a short time is very small. That is something which one has to do. But intuitively you certainly believe it.

Then all I do, I simply cross this $X(\tau)$ out. That's my approximation. The right hand side of (95) then reads

$$\frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} dx E\left\{ e^{-tV(x)} \,\middle|\, x(t) = o \right\} \tag{97}$$

Now what about this conditional expectation? Now, you see, I am simply taking the conditional expectation of a constant. I am integrating a constant over a set of measure one. So it doesn't do anything, and so I have for small t

$$\frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} dx e^{-tV(x)} \tag{98}$$

I have still to assume, in addition to all the other assumptions, that this integral makes sense. $V(x)$ must grow sufficiently fast for the integral to exist. For instance, $V(x) = \log|x|$ won't work.

Now let me perform a simple trick. Namely, I will notice that

$$\frac{1}{\sqrt{2\pi t}} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-tp^2/2} dp \tag{99}$$

That is merely a complicated way of writing $\frac{1}{\sqrt{2\pi t}}$. Then (98) becomes

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-t\left(\frac{p^2}{2}+V(x)\right)}\,dp\,dx \tag{100}$$

Now if you are quantum mechanics conscious -- or classical mechanics conscious -- you will recognize $p^2/2 + V(x)$. I even used p for p. It is simply the Hamiltonian of a particle moving in the potential $V(x)$. Of course, the mass has been normalized to one.

Now let me write this integral in a very illuminating form. Take the region

$$\frac{p^2}{2} + V(x) < \lambda \tag{101}$$

That's some region in phase space. For instance, if I were to deal with a harmonic oscillator where $V(x) = x^2/2$, then this would be a circle with radius $\sqrt{2\lambda}$. I will call the area of this region $B(\lambda)$. This can be calculated.

Now I can write the double integral (100) in a very nice way:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-t\left(\frac{p^2}{2}+V(x)\right)}\,dp\,dx = \frac{1}{2\pi}\int_{0}^{\infty} e^{-t\lambda}\,dB(\lambda) \tag{102}$$

And now you compare what I am interested in and what I end up with. They are asymptotically equal as t goes to zero:

$$\int_0^\infty e^{-\lambda t}\, dA(\lambda) \sim \frac{1}{2\pi} \int_0^\infty e^{-\lambda t}\, dB(\lambda) \tag{103}$$

Now I have to apply a Tauberian theorem just as before. To be completely rigorous one must make some further assumptions on $V(x)$. But, anyway, the argument makes it intuitively almost clear that

$$A(\lambda) \sim \frac{1}{2\pi} B(\lambda) \quad \text{as} \quad \lambda \to \infty \tag{104}$$

In words, the number of eigenvalues less than λ is asymptotically given by the area in phase space where the Hamiltonian of a system is less than λ. Except that I must divide it by 2π.

This result is well known in quantum mechanics. It is always used in justifying that, in the limit of high quantum number, classical mechanics and quantum mechanics coincide. In fact, this goes back much further than that. It goes back to the old quantum theory, long before Schroedinger's equation. You see, in the usual classical mechanics all energies are allowed. But Planck decided only to allow those energies for which the area $B(\lambda)$ is a multiple of a certain fundamental unit. That's how he quantized.

Some of the formal steps which I gave here are not at all trivial to justify. They have been done in complete and excruciating detail by Dr. D. Ray in his doctoral thesis. He went much farther and proved much more.

But the strength of the intuitive argument is amazing. It becomes immediately quite clear what the approximation ought to be. It is otherwise so difficult, because you don't know anything at all from relation (93). But when I write $Q(X_0|X,t)$ in the form (90) the road is straight and you can see for miles Geometrically what it amounts to is that in a short time the path of the Brownian motion can be approximated by the straight line. And then, all you need is that the asymptotic behavior for high eigenvalues is related to the short time behavior of your path.

To finish up this particular story let me show you where the difficulty comes in with the low eigenvalues. If you go back to formula (82) and put in for $Q(X,t)$ the expansion (93) in eigenfunctions you get:

$$E\left\{e^{-\int_0^t V(x(\tau))d\tau}\right\} = \sum_j e^{-\lambda_j t}\psi_j(0)\int_{-\infty}^{\infty}\psi_j(x)\,dx \qquad (105)$$

Now look at the series of exponentials on the right hand side. As t goes to infinity the term that predominates is the first one. So you immediately get that:

$$\lambda_1 = -\lim_{t\to 0}\frac{\log E\left\{e^{-\int_0^t V(x(\tau))d\tau}\right\}}{t} \qquad (106)$$

This formula has some interesting features. The first feature is that in order to find the lowest eigenvalue you have to let t approach infinity.

That means you have to make a long time observation of Brownian paths. They have time to become fully developed and do all the horrible things they can do. But, at least theoretically, the expression (106) gives rise to a numerical method for finding the lowest eigenvalue. In fact, a sampling method that would now be known by the name of Monte Carlo. In principle, what one ought to do is to observe a large number of Brownian paths, calculate $\exp -\int_0^t V(x(\tau))\,d\tau$ for each of these paths, take the average over all these paths, take the logarithm, and then divide by t. That would be the approximation. Of course, you cannot observe a continuous Brownian path. But you know perfectly well that Brownian motion can be approximated by random walk. So you discretize the Brownian motion and state it in terms of a random walk. That actually was done, in 1949, for two cases -- for the potential $V(x) = x^2$ and for the potential $V(x) = |x|$. Both are extremely classical cases. The results initially were very encouraging. If I remember correctly, we ran a hundred walks, each walk of duration about 50 or 60 steps. In any case it was not a very major operation and the results for this were amazingly good. In fact, it agreed on the nose for some strange reason. The results were within 3 or 4 per cent which is about the best accuracy one could expect without going to expensive further labor. We tried the next summer and the summer after that to do a similar thing for more dimensions and for certain singular potentials. For the hydrogen atom, where again the answer is known, we again hit it almost on the nose. But it was purely a matter of luck. We used 300 walks and the values we got from most of them were very small. There were three enormous ones, three out of 300! And the enormous ones gave the right answer. That's clearly an accident. So, in this case, the method is already not very good. For helium it was completely off.

You might compare the formula (106) with the more usual one:

$$\lambda_1 = \inf \frac{\int_{-\infty}^{\infty} [\psi'^2 - V(x)\psi^2]\, dx}{\int_{-\infty}^{\infty} \psi^2(x)\, dx} \qquad (107)$$

This formula comes from the variational principle in quantum mechanics. It serves as a basis for the Rayleigh-Ritz method. To get an approximation to λ_1 you simply take a trial function $\psi(x)$ and put it in. This variational method, applied to helium, gave the first great triumph of quantum mechanics. For hydrogen, everybody used Bohr's theory.

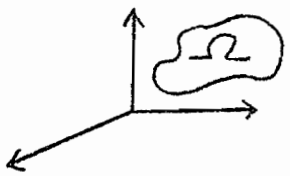
Another observation in this connection is that by a most elementary trial function you can already compute a good estimate. It takes no work at all and you already get it within five per cent. Whereas with the Monte Carlo we couldn't even get it within a thousand per cent. And this was with the help of computing machines! The whole thing blew up completely. That's due to the unfortunate nature of the potentials. After all, it's a three-body problem and the potentials become singular. The interesting thing is that, qualitatively, you do get some kind of an over-all picture. If you look at the random walk of the two electrons around the nucleus of a helium atom, you see that most of the time they are far apart and a certain part of the time they are close together. General features of the quantum mechanical picture can be read off, and this the Monte Carlo confirmed. Mind you that it's a three-dimensional walk! We could take only a very few walks and even that was a major undertaking. We knew from a statistical analysis of errors that

we should have many more samples. But the random walk technique was still good enough to give you the general features, but not good enough to give you such a fine point as the lowest eigenvalue.

This departs a little from the theme of the first half of the lectures, because I'm not using stochastic phenomena in physics per se. Instead, I am using them as a way of looking at certain things which are not stochastic to begin with. Our equation (91) is a perfectly good equation. You don't need any probability theory at all to look upon it. Likewise, the telegrapher's equation (16) can stand by itself. But the point is, that by this stochastic approach, you are led not only to a new kind of mathematics which could be reward enough in itself; you also gain a way of looking at things which is quite powerful, which gives you all the results cheaper and also gives new results. I will wind up this series of lectures by showing you how this can be applied to potential theory.

NINTH LECTURE

This lecture I will devote to giving you a little bit of an idea how the method of integration in function spaces can be applied to the classical potential theory. We will start just as before. Only now I will move myself in three-dimensional space. The reason why we go to three-dimensional space is because in two dimensions there are some very special features. Now let me have a region Ω and consider a three-dimensional Brownian motion.



I imagine that the particle starts from the origin $\vec{r} = 0$. Now remember what I did with $X(t)$, the Wiener process. I constructed certain basic sets of paths and assigned probabilities to them. Here it's a very similar thing that I do, in fact a pure analogy. $X(t)$ now becomes the vector $\vec{r}(t)$. Again you take n time points $t_1 < t_2 < \cdots < t_n$. And now, instead of using intervals, I will take n arbitrary open sets, $\Omega_1, \Omega_2, \ldots, \Omega_n$. What is the probability, the measure, of those paths for which $\vec{r}(t_1) \in \Omega_1$, $\vec{r}(t_2) \in \Omega_2$, . . . , and $\vec{r}(t_n) \in \Omega_n$? This again is going to be a multiple integral:

$$\int_{\Omega_1} d\vec{r}_1 \int_{\Omega_2} d\vec{r}_2 \cdots \int_{\Omega_n} d\vec{r}_n \, P(0|\vec{r}_1, t_1) P(\vec{r}_1|\vec{r}_2, t_2 - t_1) \cdots P(\vec{r}_{n-1}|\vec{r}_n, t_n - t_{n-1}) \quad (108)$$

Here each integral stands for a triple integral. $P(\vec{\rho}|\vec{r}, t)$ is now the Gaussian distribution in three-dimensions:

$$P(\vec{\rho}|\vec{r},t) = \frac{e^{-|\vec{r}-\vec{\rho}|^2/2t}}{(2\pi t)^{3/2}} \tag{109}$$

Now you have a complete analogy. We can build up a measure theory in the same way as before, starting from these sets whose measure we know. As a matter of fact, I may as well mention that this is not really necessary. We don't have to go through this process a second time. We can take the measure already introduced and simply build up what's called product measure. Both ways get you there, so take your choice.

We now have a measure and hence also the possibility of integration in the space of all continuous paths. In this case, the paths emanate from the origin. With these tools we can build up very neatly the potential theory in three dimensions. The problem we are going to consider is the following: Take a region Ω and a point $\vec{y}$. In Figure 4 the point is drawn outside the region, but it can be anywhere.

Figure 4.

Now consider $\vec{y} + \vec{r}(t)$. This is a Brownian motion path which starts from $\vec{y}$ rather than from the origin. That is the only difference. The problem is to find out how much time is spent in the region Ω.

In order to get a nice symbolism going I will introduce the following function:

$$V(\vec{x}) = \begin{cases} 1 & \text{if } \vec{x} \in \Omega \\ 0 & \text{if } \vec{x} \notin \Omega \end{cases} \tag{110}$$

This is what is known as the characteristic function of the region Ω. Let me also consider the following functional:

$$T_\Omega(\vec{y}, \vec{r}(\tau)) = \int_0^\infty V(\vec{y} + \vec{r}(\tau))\, d\tau \tag{111}$$

This integral may be infinite, but let me first of all see what its meaning is. Whenever $\vec{y} + \vec{r}(\tau)$ is in the region, I am integrating simply one. Otherwise I am integrating zero. So it clearly represents the total time the curve spends in the region Ω. It depends, of course, upon which curve $\vec{r}(\tau)$ the Brownian particle follows.

Now let me find the average of this quantity $T_\Omega(\vec{y}, \vec{r}(\tau))$ The average means the integral with respect to the measure we introduced. This is sometimes called the mathematical expectation and we have used the symbol E for it. What we want then is

$$E\{T_\Omega(\vec{y}, \vec{r}(\tau))\} = E\left\{ \int_0^\infty V(\vec{y} + \vec{r}(\tau))\, d\tau \right\} \tag{112}$$

Now one does the usual trick. One interchanges these two operations (remember that this operator E is also an integration, an integration in function space) so you get:

$$E\left\{T_\Omega(\vec{y},\vec{r}(\tau))\right\} = \int_0^\infty d\tau\, E\left\{V(\vec{y}+\vec{r}(\tau))d\tau\right\} \tag{113}$$

Now this thing has a very vivid significance, because the function V is either zero or one depending on whether $\vec{y}+\vec{r}(\tau)$ *is in* Ω *or not.* A moment of thought will convince you that this just gives:

$$E\left\{T_\Omega(\vec{y},\vec{r}(\tau))\right\} = \int_0^\infty d\tau \left[\text{Prob}\left\{\vec{y}+\vec{r}(\tau)\in\Omega\right\}\right] \tag{114}$$

But this probability, that $\vec{y}+\vec{r}(\tau)\in\Omega$, I can immediately decipher from my basic assumption. It is the simplest case, where we have only one time involved. This gives us:

$$E\left\{T_\Omega(\vec{y},\vec{r}(\tau))\right\} = \int_0^\infty d\tau \int_\Omega d\vec{r}\, \frac{e^{-(\vec{r}-\vec{y})^2/2\tau}}{(2\pi\tau)^{3/2}} \tag{115}$$

Now we interchange the order of integration again:

$$E\left\{T_\Omega(\vec{y},\vec{r}(\tau))\right\} = \int_\Omega d\vec{r} \int_0^\infty d\tau\, \frac{e^{-(\vec{r}-\vec{y})^2/2\tau}}{(2\pi\tau)^{3/2}} \tag{116}$$

This last integral is well-known. You will find it in almost all texts on Laplace transform. So I will just put it in and obtain the final answer.

$$E\left\{T_\Omega(\vec{y},\vec{r}(\tau))\right\} = \frac{1}{2\pi}\int_\Omega \frac{d\vec{r}}{|\vec{r}-\vec{y}|} \tag{117}$$

This is finite in three or more dimensions. And this is true even if $\vec{y}$ would happen to be inside the region. It would still be finite, as you can see, by introducing polar coordinates. The integrand is essentially $1/r$ and the surface element will be of order r^2. So there is plenty to spare to make it convergent.

Now what is the conclusion? So far, it is that the average of $T_\Omega(\vec{y},\vec{r}(\tau))$ is finite. But the quantity itself is non-negative. And if the average of a non-negative quantity is finite then with probability one the quantity itself must be finite. If it were infinite with a probability bigger than zero then of course the average would also have to be infinite. So you discover a very interesting, very simple fact: Almost every Brownian motion curve spends a finite amount of time in a region of finite volume. (I use "almost every" in a technical sense -- it means except for a set of curves of measure zero.) That is not so in the plane. For this reason the plane has to be treated separately. If you were to go through the whole argument you would come out with

$$\frac{e^{-\frac{|\vec{r}-\vec{\rho}|^2}{2\tau}}}{2\pi\tau} \tag{118}$$

as the integrand in (116). The t in the denominator means that this is not integrable at infinity. You end up with an infinite integral. This doesn't necessarily imply that almost every path spends an infinite time in the region. If the integral of a function is infinite that doesn't mean that the function is infinite. However, it can be shown by a somewhat more refined argument that in two dimensions almost every curve indeed spends an infinite amount of time in any region.

This distinction between the plane and space is very interesting. In fact, as you will see a little later, this is the beginning of a probabilistic explanation of the great difference between the two- and three-dimensional spaces in potential theory. Everybody knows there is a tremendous difference. In one case the basic Green's function is $1/r$ and in the other is $\log r$. That causes the great difference. From our point of view, you discover that the real difference is that in three dimensions almost every Brownian motion curve spends a finite amount of time in a given region. Whereas in two dimensions there is not enough room. The poor curve wanders around and comes back and back to the same region. In fact, one can refine the argument and prove the following: Almost every curve in two-dimensional space is everywhere dense. That means it comes arbitrarily close to every point. In three dimensions almost every curve is nowhere dense. Then you have the remarkable theorem that while every three-dimensional curve is nowhere dense the projections, the plane projections, are almost everywhere dense. Of course, one should not take this literally. This is not a statement about curves but a statement about the measures. It's not quite easy, actually, to imagine even one such curve. A curve which occupies hardly any space, which is kind of sparse,

but one whose projections on any plane go arbitrarily near to every point in that plane. Our particular argument shows that there are such curves.

Now this fact, this strange difference between two and three dimensions, was already noticed in a much simpler situation long, long ago by Polya. It was in 1923, I think, that Polya considered the following problem: Take a lattice, a square lattice, in the plane and consider a random walk, an ordinary discrete random walk. From every point you go to one of its four neighbors with probability one-fourth. Then, wherever you are, you go again to one of the neighbors with probability one-fourth. And so you continue. What is the probability that starting, say, from the origin you will eventually pass through a given point in the lattice? And the answer, in the plane, is one. That means that with probability one the curve will eventually pass through any pre-assigned point. You pick yourself a point and you are almost sure to hit it. In space, however, that isn't so. If you consider the corresponding problem in space you have six possible ways to go, six neighbors, and you make the probability one-sixth. If you now ask the same question for this cubic lattice, the answer is distinctly less than one. The probability that you will eventually hit a chosen point is no longer one. It can actually be calculated. The answer is expressible in terms of a certain integral.

It is really quite interesting, this distinction between the plane and space. In the plane the Brownian motion path tends to come arbitrarily close to every point and in space it does not. That's a very important point for people like Stan Natanson who don't know the city, and drives at random. Because, assuming that Dallas is a plane there is a chance of one that he will eventually be at the Melrose Hotel if at each intersection he will flip

a coin. It may be by way of Ithaca, of course, or Los Angeles. And, in fact, if you would calculate the average number of steps he takes, it is infinite. And indeed that was experimentally verified when he drove me to the Melrose.

These results can be extended to higher dimensions. The higher the dimension, the less is the probability that it will hit a given point. For every compact creature, like a sphere the probability will always be one, because then you really have no room. The only problem is for a non-compact thing, where you have definitely a possibility of escape. The phenomenon has not been investigated with any care on surfaces other than the plane, but the method would apply. As a matter of fact, in principle it is easy to tell you what the answer is on any surface, although nobody has ever proved this with complete rigor. You take the heat conduction equation on that surface and find its fundamental solution. Then you take the integral of this from zero to infinity in time as in equation (116). If the integral is finite then you know the probability is less than one. If the integral is infinite then you have the same case as in the plane. It's purely a property of the kind of singularity the heat conduction solution has.

This actually is a very superficial theorem. It is extremely elegant and very appealing but there is really no depth at all. There is no dynamical content to it. It is essentially a theorem which combines dimensionality of the space you are considering with some very, very rough idea of compactness. If the dimension is high enough and if it isn't compact then you will have the possibility of escaping. However, these facts form a beginning of an entirely new, pictorial way of looking at potential theory.

To go to potential theory I must do something more refined than simply calculating the average time (117). I must attempt to calculate the distribution of the times. This is really the crucial step. And this means that we must calculate all the moments, not simply the average (117). Now let me just perform for you the calculation of the second moment:

$$E\left\{\left[T_\Omega(\vec{y},\vec{r}(\tau))\right]^2\right\} = E\left\{\left[\int_0^\infty V(\vec{y}+\vec{r}(\tau))\,d\tau\right]^2\right\} \tag{119}$$

Our procedure will be almost identical in spirit to the one we used with the telegrapher's equation. Let me again write the square of an integral as a double integral:

$$E\left\{\left[T_\Omega(\vec{y},\vec{r}(\tau))\right]^2\right\} = E\left\{\int_0^\infty\int_0^\infty V\left(\vec{y}+\vec{r}(\tau_1)\right)V\left(\vec{y}+\vec{r}(\tau_2)\right)d\tau_1 d\tau_2\right\} \tag{119'}$$

Now this integral is over a square, an infinite square to be sure. Just as before I can integrate over only half of the square and then multiply by two:

$$2!\,E\left\{\int_0^\infty d\tau_2\int_0^{\tau_2} d\tau_1 V(\vec{y}+\vec{r}(\tau_1))V\left(\vec{y}+\vec{r}(\tau_2)\right)\right\} \tag{120}$$

Now you interchange expectation and integration and write it so:

$$2!\int_0^\infty d\tau_2 \int_0^{\tau_2} d\tau_1\, E\left\{V(\vec{y}+\vec{r}(\tau_1))\,V(\vec{y}+\vec{r}(\tau_2))\right\} \tag{121}$$

This is perfectly justifiable. Now $V(\vec{y}+\vec{r}(\tau_1))$ is either zero or one depending on whether $\vec{y}+\vec{r}(\tau_1)$ is in Ω at time τ_1 or not. And the same goes also for $V(\vec{y}+\vec{r}(\tau_2))$. So the expectation of the product of these two is again included in formula (108), except that I start from $\vec{y}$ rather than from zero. So I simply apply formula (108) with Ω_1 and Ω_2 both equal to Ω:

$$2!\int_0^\infty d\tau_2 \int_0^{\tau_2} d\tau_1 \int_\Omega\int_\Omega P(\vec{y}|\vec{r}_1,\tau_1)\,P(\vec{r}_1|\vec{r}_2,\tau_2-\tau_1)\,d\vec{r}_1\,d\vec{r}_2 \tag{122}$$

You can see this easily. I must simply make a transition from $\vec{y}$ to some point $\vec{r}_1$ in Ω and then from that point I must again make a transition to a point $\vec{r}_2$ in Ω. Now you can see that it would be much more convenient to take the integration over the spatial variables lastly:

$$2!\int_\Omega\int_\Omega d\vec{r}_1\,d\vec{r}_2 \int_0^\infty d\tau_2 \int_0^{\tau_2} d\tau_1\, P(\vec{y}|\vec{r}_1,\tau_1)\,P(\vec{r}_1|\vec{r}_2,\tau_2-\tau_1) \tag{123}$$

The integrations over time are again related to convolution. That is simply the Markovian property making an appearance, together with the temporal homogeneity. Temporal homogeneity means that everything depends on the difference of times. You see $\tau_2 - \tau_1$ rather than τ_2 , τ_1 separately. Now you can easily see what happens with the time integrals. Because suppose I insert $e^{-s\tau_2}$:

$$\int_0^\infty d\tau_2 e^{-s\tau_2} \int_0^{\tau_2} d\tau_1 P(\vec{y}|\vec{r}_1,\tau_1) P(\vec{r}_1|\vec{r}_2,\tau_2-\tau_1) \qquad (124)$$

With this factor there, (124) is simply the Laplace transform of the convolution. Of course, I don't really have the factor there. But this means I simply put $s=0$. Now the Laplace transform of a convolution is simply the product of Laplace transforms. But these Laplace transforms are known for $s=0$ (refer back to formulas (116) and (117)). It immediately follows from this argument that

$$E\left\{\left[T_\Omega(\vec{y},\vec{r}(\tau))\right]^2\right\} = \frac{2!}{2\pi} \int_\Omega \int_\Omega \frac{1}{|\vec{r}_1-\vec{y}|} \frac{1}{|\vec{r}_2-\vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \qquad (125)$$

This is the second moment. Now it's perfectly clear how the moments will come out in general. You can use exactly the same argument, the only difference is that a little more writing will be required. You have the following general formula for the kth moment μ_k :

$$\frac{\mu_k}{k!} = \frac{1}{k!} E\left\{T_\Omega^k(\vec{y}, \vec{r}(\tau))\right\} = \tag{126}$$

$$= \frac{1}{(2\pi)^k} \int_\Omega \cdots \int_\Omega d\vec{r}_1 \cdots d\vec{r}_k \frac{1}{|\vec{r}_1 - \vec{y}|} \frac{1}{|\vec{r}_2 - \vec{r}_1|} \cdots \frac{1}{|\vec{r}_k - \vec{r}_{k-1}|}$$

Now everybody who has ever dealt with integral equations will recognize that in (126) we have something which looks very much like an iterated kernel. It is like raising a matrix to a power. To see this simply write the integrand of (126) in the form:

$$a(\vec{y}, \vec{r}_1)\, a(\vec{r}_1, \vec{r}_2)\, a(\vec{r}_2, \vec{r}_3) \cdots a(\vec{r}_{k-1}, \vec{r}_k) \tag{127}$$

Now I will have to assume a little bit of knowledge of integral equations. If you have this knowledge, then it is very natural to associate with this problem the following integral equation:

$$\frac{1}{2\pi} \int_\Omega \frac{1}{|\vec{r} - \vec{\rho}|} \varphi(\vec{\rho})\, d\vec{\rho} = \lambda \varphi(\vec{r}), \quad \vec{r} \in \Omega \tag{128}$$

I want to call your attention to the fact that this is a three-dimensional integral equation. The integral is not taken over the surface of Ω. It is taken over the volume. It is an honest-to-goodness three-dimensional integral equation. Now this kernel $1/|\vec{r} - \vec{\rho}|$ which may look singular

to you, is not really very bad. In fact, the integral:

$$\int_{\Omega} \frac{d\vec{\rho}}{|\vec{r}-\vec{\rho}|^2} \tag{129}$$

is finite. It is what is known as a completely continuous kernel. So all the theory of integral equations is applicable -- in particular, the whole theory of Hilbert and Schmidt. You can then show that the integral equation (128) has eigenvalues. In fact, the eigenvalues are all positive. It also has a complete set of normalized eigenfunctions.

In terms of the eigenfunctions and eigenvalues of the integral equation one can very easily express the moments. That is really the only thing you have to know from the theory of integral equations. The only difficulty is that the equation (128) holds only if $\vec{r} \in \Omega$. Now the only variable which may refer to a point outside Ω is $\vec{y}$. So $\vec{y}$ has to be treated a little bit separately. You get the formula

$$\frac{\mu_k}{k!} = \frac{1}{2\pi} \sum_{j=1}^{\infty} \lambda_j^{k-1} \int_{\Omega} \varphi_j(\vec{\rho})\, d\vec{\rho} \int_{\Omega} \frac{\varphi_j(\vec{\rho})\, d\vec{\rho}}{|\vec{\rho}-\vec{y}|} \tag{130}$$

This is very straight-forward. It is a formula for the iterated kernel that you'll find in any text book on integral equations. It is completely analogous to expressing the kth power of a matrix in terms of the eigenvalues and eigenvectors of the original matrix. Now you might jump at me and ask why I don't use the integral equation to replace

$$\int_{\Omega} \frac{\varphi_j(\vec{\rho})\, d\vec{\rho}}{|\vec{\rho} - \vec{y}|} \tag{131}$$

by $2\pi\lambda_j \varphi_j(\vec{y})$. I can do this provided $\vec{y}$ is in the region Ω . But if $\vec{y}$ is outside then I cannot. So, since I don't know where $\vec{y}$ is I must leave it in this form. Now we will calculate the moment generating function:

$$E\left\{e^{-u T_{\Omega}(\vec{y}, \vec{r}(\tau))}\right\} \tag{132}$$

It is easy to see how to do it. You simply expand in a power series in u . You need a little bit of a discussion of convergence and you finally get the following:

$$E\left\{e^{-u T_{\Omega}(\vec{y}, \vec{r}(\tau))}\right\} = 1 - \sum_{j=1}^{\infty} \frac{u}{1+\lambda_j u} \int_{\Omega} \varphi_j(\vec{\rho})\, d\vec{\rho} \frac{1}{2\pi} \int_{\Omega} \frac{\varphi_j(\vec{\rho})\, d\vec{\rho}}{|\vec{\rho} - \vec{y}|} \tag{133}$$

Let me call your attention to the following feature: On the right hand side there is a perfectly classical expression. And on the left hand side we have an average, an integral, over a function space. We will play the game of deciphering the properties of the classical expression by looking at the average over the function space.

Notice that the right hand side, the whole thing, satisfies Laplace's equation for $\vec{y}$ outside the region. That is perfectly obvious. (131) is a sum of potentials of spatial mass distributions. The mass distributions are given by $\varphi_j(\vec{\rho})$. Of course, you have an infinite series. But it is easy to show that the series converges uniformly in every finite region.

Now let me see what happens as $u \to \infty$. This will be very interesting. To see what happens, just look at $\exp -u T_\Omega(\vec{y}, \vec{r}(\tau))$. Now as u goes to infinity this can go to two limits. It goes to zero if $T_\Omega(\vec{y}, \vec{r}) > 0$. And it goes to one if $T_\Omega(\vec{y}, \vec{r}) = 0$. These are the only possibilities because the time $T_\Omega(\vec{y}, \vec{r})$ is certainly non-negative. Now let's try to understand what it means. Don't forget that $T_\Omega(\vec{y}, \vec{r})$ depends upon the path $\vec{r}(\tau)$. If this path penetrates the region then you get the limit zero. Because then the particle spends a finite time in the region. For every path which does not penetrate the region the limit is one. That is the limit of the function, but I still have to take the expectation. That is, I still have to integrate over all the paths. For those of you who studied measure theory in your youth, this is one of the few places one needs to use the more refined properties of measure, namely the complete additivity property. Because now you need the theorem that the limit of the integral is the integral of the limit. In this case it is very easy because you are dealing with a set of decreasing functions, as u goes to infinity. So dominated convergence is operable. So in the limit the left hand side of (133) is simply the integral of a function which is zero for those paths which penetrate the region and is one for those paths which do not. But what is the integral of

a function which is one on some set and zero outside that set? That is just the probability attached to that set. So in the limit we get:

$$\lim_{u\uparrow\infty} E\left\{e^{-u T_\Omega(\vec{y},\vec{F}(\tau))}\right\} = \text{Prob}\left\{T_\Omega(\vec{y},\vec{F}(\tau))=0\right\}$$

$$= 1 - \lim_{u\uparrow\infty} \sum_{j=1}^{\infty} \frac{1}{\frac{1}{u}+\lambda_j} \int_\Omega \varphi_j(\vec{\rho})\,d\vec{\rho}\,\frac{1}{2\pi}\int_\Omega \frac{\varphi_j(\vec{\rho})\,d\vec{\rho}}{|\vec{\rho}-\vec{y}|} \qquad (134)$$

Now it's actually more convenient to write this in the form

$$\text{Prob}\left\{T_\Omega(\vec{y},\vec{F}(\tau))>0\right\} = \lim_{u\uparrow\infty} \sum_{j=1}^{\infty} \frac{1}{\frac{1}{u}+\lambda_j} \int_\Omega \varphi_j(\vec{\rho})\,d\vec{\rho}\cdot\frac{1}{2\pi}\int_\Omega \frac{\varphi_j(\vec{\rho})\,d\vec{\rho}}{|\vec{\rho}-\vec{y}|} \qquad (135)$$

This says the same thing but it's much nicer. This gives the probability that you will penetrate the region in the form of a limit as u goes to infinity of a series. So far, this is true for every $\vec{y}$, on the boundary, inside or out.

I want to call your attention to one thing immediately. The limit (135) is also a harmonic function. To prove that I must appeal to a very well-known theorem of Harnack. It states that the limit of a bounded decreasing sequence of harmonic functions is harmonic. This is certainly the case we have. You can see this most easily from the left-hand side of (134). So if you believe this theorem, which you can find in any standard textbook on potential

theory, then (135) is also a harmonic function. I am going to call it $U(\vec{y})$. And I know it satisfies $\Delta U = 0$ when $\vec{y} \notin \Omega$.

Not only is $U(\vec{y})$ a harmonic function, it is nothing but the capacitory potential. That, by definition, is the harmonic function which vanishes at infinity and which assumes the boundary value one on the boundary of the region Ω. Incidentally, we produce as a by-product the formula:

$$U(\vec{y}) = \lim_{u \uparrow \infty} \sum_{j=1}^{\infty} \frac{1}{\frac{1}{u} + \lambda_j} \int_{\Omega} \varphi_j(\vec{\rho})\, d\vec{\rho} \cdot \frac{1}{2\pi} \int_{\Omega} \frac{\varphi_j(\vec{\rho})\, d\vec{\rho}}{|\vec{\rho} - \vec{y}|} \tag{136}$$

for the capacitory potential. It is a remarkable thing that despite the advanced age of potential theory no one had ever discovered this formula. It comes, you see, very naturally out of this argument, which I have originally made around 1949. The key to the answer is not hard to find. You might say, that if u goes to infinity, and I have $1/u$ in (136), then why bother with it? Simply cross away the limit and put zero for $1/u$. Unfortunately, then, the series makes no sense whatsoever. It diverges. That was the reason why the approach based on the integral equation (128) was originally abandoned. Because it led directly to the divergent series:

$$\sum_{j=1}^{\infty} \frac{1}{\lambda_j} \int_{\Omega} \varphi_j(\vec{\rho})\, d\vec{\rho} \, \frac{1}{2\pi} \int_{\Omega} \frac{\varphi_j(\vec{\rho})\, d\vec{\rho}}{|\vec{\rho} - \vec{y}|} \tag{137}$$

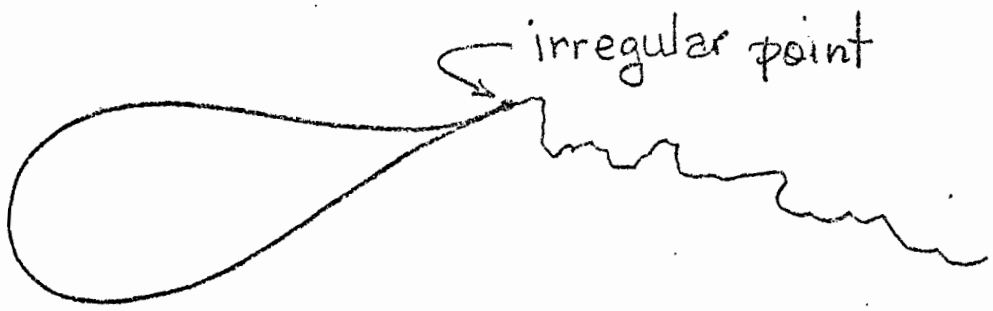
This was perhaps fortunate for the history of science, because it forced a different approach. The double layer idea was introduced which led to a very nice development of certain mathematical techniques and ideas. But (136) is a formula of a type that Karl Neumann, and others, tried to arrive at in 1870 but couldn't succeed. It simply introduces what we call in mathematics a summability method. The serious trouble with (137) is divergence. In (136) we have done a little repair work and then it makes perfect sense.

Let's have a look at what happens if $\vec{y}$ is inside the region. The integral equation then becomes operative and (135) can be simplified. The whole thing simplifies dramatically and becomes:

$$\text{Prob}\left\{T_\Omega(\vec{y},\vec{r}(\tau))>0\right\} = \lim_{u\to\infty} \sum_1^\infty \frac{\lambda'_j}{\frac{1}{u}+\lambda_j} \cdot \varphi_j(\vec{y}) \int_\Omega \varphi_j(\vec{\rho})\, d\vec{\rho} \qquad (138)$$

It's very easy to tell what this is. If you start inside the region you've got to spend some positive time there. So the probability on the left hand side is simply one. The Brownian path, being continuous, can't simply jump out of the region. Notice how easy it was to see that (138) converged to one. It is almost trivial. You just use the fact that the particle can't jump out from inside the region.

Now what is really interesting is what happens on the boundary. Suppose that $\vec{y}$ is on the boundary. Then you can't say anymore that the probability is one that you spend sometime within the region Ω. In fact, there are points, called irregular points, for which this is not true. The boundary has to be very, very sharp at such points:



An example was first given by Lebesque in 1913, or so. It corresponds to a fact which is well-known in electrostatics. If you have an extremely sharp point on your conductor you cannot maintain a constant potential. The charge leaks off. This leak is simply interpretable. If you start the Brownian motion at such a point it is possible with non-zero probability to go out without ever coming into the region. However, the limit (135) still exists. It exists for all values of $\vec{y}$ -- there is no problem at all there. But it need not be equal to one. If a point is a regular point, if it is not a point where the boundary is very sharp, then the limit clearly is going to be one again. Intuitively it is perfectly obvious that if you have a round point then you have a pretty good chance of almost immediately entering the region. The point has got to be extremely sharp to prevent it.

Now let's take a look at this formula (136) and see what happens when $\vec{y}$ goes to infinity. It's perfectly clear that it goes to zero because each term goes to zero. This can be made precise without any trouble. Now what happens as $\vec{y}$ approaches the boundary is a little bit more complicated. It depends on what kind of point you have. If the point is an irregular point then the potential will not approach one. But if you have a reasonable

boundary then you can show that $U(\vec{y})$ approaches one as $\vec{y}$ approaches a boundary point. So, if you believe my intuitive argument, you see that $U(\vec{y})$ is indeed what is called the capacitory potential. It is that potential which is zero at infinity and which approaches one on the boundary wherever it can. If there are bad points, then of course it can't.

TENTH LECTURE

Now I would like to make a few remarks about this method of looking at things from a probabilistic point of view. If we look at the problem of the capacitory potential classically then what one wants is the following. One wants a harmonic function which vanishes at infinity and approaches one on the boundary. Now you might take a very naive point of view. I'll try to find a mass distribution $\psi(\vec{\rho})$ over the region Ω so that the corresponding potential:

$$U(\vec{y}) = \frac{1}{2\pi} \int_{\Omega} \frac{\psi(\vec{\rho})\, d\vec{\rho}}{|\vec{\rho} - \vec{y}|} \tag{139}$$

has the desired properties. Well, certainly it's a potential all right. It's a harmonic function. It is also zero at infinity. But what is the best way to make it one on the boundary? Well, you say, I'll try to make it one on the boundary by making it one inside; everywhere inside. In fact, you know from electrostatics that that's how it's going to be. The potential inside is going to be uniform. This gives the integral equation:

$$\frac{1}{2\pi} \int_{\Omega} \frac{\psi(\vec{\rho})\, d\vec{\rho}}{|\vec{\rho} - \vec{y}|} = 1 \quad , \quad \vec{y} \in \Omega \tag{140}$$

This is what peopls did when they first tried to solve problems of this sort. They tried to solve this integral equation. Now how do we solve such an

integral equation? One of the simplest ways is to suppose that the function $\psi(\vec{\rho})$ can be expanded in the eigenfunctions φ_j of the integral equation (128). Lo and behold, you get

$$\psi(\vec{\rho}) = \sum_{j=1}^{\infty} \frac{1}{\lambda_j} \int_{\Omega} \varphi_j(\vec{r})\, d\vec{r}\, \varphi_j(\vec{\rho}) \tag{141}$$

This is simply the expression I would get formally if I put $1/u$ equal to zero in (138). Then, if I substitute this in to find the potential, it comes out

$$U(\vec{\rho}) = \sum_{j=1}^{\infty} \frac{1}{\lambda_j} \int_{\Omega} \varphi_j(\vec{r})\, d\vec{r}\, \frac{1}{2\pi} \int_{\Omega} \frac{\varphi_j(\vec{\rho})\, d\vec{\rho}}{|\vec{\rho} - \vec{\rho}|} \tag{142}$$

Again, this compares with my formula (136) if I put $1/u$ equal to zero.

Now, why is this bad? This is, to be sure, the most natural way to solve the problem. But it is bad because the series (142) makes no sense. That is already obvious on physical grounds. Because you know that there is no mass distribution inside Ω which will give such a potential. All the charge is concentrated on the boundary. So you could not expect (142) to give you anything sensible. Since everybody knew that there was no mass distribution inside which would give a reasonable result this approach was abandoned.

The interesting thing is that the simple change made in (136) makes possible a perfectly sensible interpretation. And, moreover, this interpretation is so closely related to the probability viewpoint. Now I might add,

in this connection, the following observation: One can very easily prove that the series

$$\sum_{j=1}^{\infty} \frac{1}{\frac{1}{u}+\lambda_j} \int_{\Omega} \varphi_j(\vec{\rho})\,d\vec{\rho}\; \frac{1}{2\pi} \int_{\Omega} \frac{\varphi_j(\vec{\rho})\,d\vec{\rho}}{|\vec{\rho}-\vec{y}|} \tag{143}$$

for a finite u , is the potential of a mass distribution. It can be written

$$\frac{1}{2\pi} \int_{\Omega} \frac{\psi_u(\vec{\rho})\,d\vec{\rho}}{|\vec{\rho}-\vec{y}|} \tag{143'}$$

Moreover, $\psi(\vec{\rho})$, the mass density, is non-negative. For every finite u you get a perfectly good mass distribution. But as u goes to infinity what happens to this function is that it gets smaller and smaller everywhere inside. The mass gets more and more concentrated near the boundary. So the process $u \to \infty$ is a sweeping-out process. The mass gets swept out from the inside and in the limit you collect it all on the boundary.

Now let's take another look at relation (136). From it one can very easily see what the capacity is. There are very many definitions of capacity, but one of them is:

$$U(\vec{y}) \sim \frac{c}{|\vec{y}|} \quad , \; |\vec{y}| \to \infty \tag{144}$$

This says that at infinity the capacitory potential behaves like a certain constant divided by the distance of the point from the origin. The constant is known as the capacity. Now we can find this constant very easily. Of course, there is a matter of interchanging various limiting processes. But this can be easily justified. Now for large $\vec{y}$ (136) becomes:

$$U(\vec{y}) \sim \lim_{u \uparrow \infty} \sum_{j=1}^{\infty} \frac{1}{\frac{1}{u} + \lambda_j} \frac{\left[\int_\Omega \varphi_j(\vec{\rho})\, d\vec{\rho}\right]^2}{2\pi |\vec{y}|} \tag{145}$$

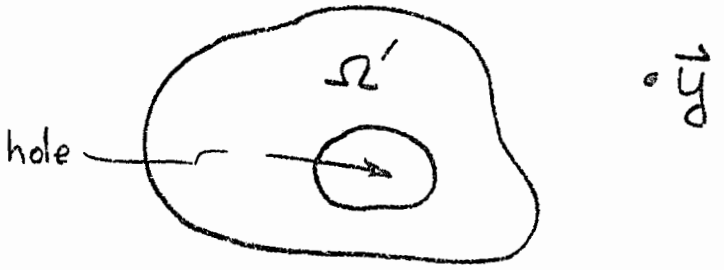
But now all the terms in the series are positive so you ought to be able to let u go to infinity inside the sum. Then you get the following formula for the capacity:

$$C = \frac{1}{2\pi} \sum_{j=1}^{\infty} \frac{\left[\int_\Omega \varphi_j(\vec{\rho})\, d\vec{\rho}\right]^2}{\lambda_j} \tag{146}$$

This gives you an expression for the capacity in terms of the eigenvalues and eigenfunctions of the integral equation (128). It is equivalent to the formula obtained from a variational principle.

I would like to make one final observation in this connection. Throughout the whole theory here, we worked strictly with the volume integral equation (128). The surface, as such, never entered the consideration. From the purely mathematical point of view this was a great convenience. Surface

considerations are always very tedious. You might wonder if we really got something for nothing. Well, we didn't, really. And again, this point of view makes it very clear. To see it, suppose I scoop a hole out of Ω. Let me call the region I have left Ω'.



Now, if I start a Brownian particle from the point $\vec{y}$ then what is the probability that it will spend a positive time in Ω'? Why, it's clearly the same as the probability it will spend a positive time in Ω. Because when it enters Ω' it automatically enters Ω, also. On the other hand, the eigenfunctions and the eigenvalues for Ω' will be vastly different. But, nevertheless, the strange combination (135) must be the same. So you see that you can cut out an arbitrary chunk and still get the same answer. From this it is obvious that we are really dealing with a surface phenomenon, because you can badly butcher up the volume Ω and it makes no difference. For a reasonable region, one can say it as follows: The probability of spending positive time in the region Ω is the same as the probability of crossing the boundary. But if you try to build the theory from this viewpoint, that of crossing the boundary, then you're in trouble. Things would become very tedious and messy. So you simply say that this is the same as spending positive time in the interior. And then one is led to all these formulas in terms of volume integrals. It's much nicer.

Now what about two dimensions? There is no complication in going to higher dimensions. And that is, after all, only an academic exercise. But in two dimensions all these fundamental difficulties we mentioned come in. Almost every Brownian path spends an infinite amount of time in any region Ω. So one cannot carry through the same theory. It has to be modified. The modifications lead to more involved computations which I will spare you. I will just give you the results; at least, the probabilistic part of the results. These are the following: We again take a region Ω and a point $\vec{y}$. Let us start a Brownian particle from $\vec{y}$.

Now I know, with probability one, that the curve will eventually enter the region Ω. But I can ask for the probability that it will not enter the region up to time t. In other words, I can ask for:

$$\text{Prob}\left\{\vec{y}+\vec{r}(\tau) \notin \Omega \text{ for } 0 \leq \tau \leq t\right\} \tag{147}$$

In the plane, as t approaches infinity, this probability must approach zero. But the remarkable thing is that it approaches zero, asymptotically, as

$$\frac{R(\vec{y})}{\log \sqrt{t}} \tag{148}$$

$R(\vec{y})$ is the two-dimensional analogue of the capacitory potential. It is, namely, that harmonic function which is zero on the boundary of the region and behaves like

$$C + \log|\vec{y}| \tag{149}$$

at infinity. In other words, it has a logarithmic singularity at infinity.

So even here you have a probabilitistic interpretation of the capacitory potential, a much more unpleasant one. It would be useless for Monte Carlo purposes, because you would really have to wait a long, long time. Convergence would be woefully slow.

To prove these things the methods have to be modified. The proofs become reasonably unpleasant. Some of it has been published. This theorem was conjectured by me and then proved by one of my colleagues. I have now, under somewhat stringent conditions, a reasonably simple proof. It will eventually appear in print.

Now, to finish up this thing I will make some remarks. You remember that all this measure theory, starting from a one-dimensional case, was built on this function:

$$\frac{1}{\sqrt{2\pi t}}\, e^{-\frac{(x-x_0)^2}{2t}} \tag{150}$$

The reason we needed such a function was that it was a solution of the Chapman-Kolmogoroff equation. It is of interest to see what would happen if we were to take other solutions of this equation. For instance, the one which I showed you:

$$\frac{t}{\pi}\,\frac{1}{t^2+(x-x_0)^2} \tag{151}$$

The second moment of this one is infinite. It has to be. If the second moment is finite then you can only get (150).

So suppose we try to build a measure theory based on (151). We introduce the windows (see page 136) and consider the set of paths passing through them, and so forth. But then you discover to your great surprise that with probability one the paths are discontinuous. This is one of the most interesting facts in this business. I think it was first discovered by Paul Levy. You cannot build a measure in the set of continuous paths based on this function (151). So the best you can do is to build the measure in a certain set of discontinuous functions.

One of the very interesting outcomes is the change this makes in our potential theory. The integral equation (128) has to be changed. But you can get a formula similar to (136). But now, as u goes to infinity, the mass becomes distributed all over the region. In the other case it got concentrated on the boundary. The difference is a direct consequence of the discontinuity of the paths. That is very interesting and very charming. Before, the only reason that I could use the argument that crossing the boundary was the same as entering the interior was that the paths were continuous. That's not true anymore, because the paths can jump. You might enter the interior without ever crossing the boundary. So you see, the fact

that in ordinary potential theory the paths are continuous is equivalent to the statement that the charges giving rise to the capacitory potential are distributed on the boundary.

The measure in the space of all these discontinuous paths is a very pathological affair. In fact, considerable caution has to be exercised. Nevertheless, it is very intimately connected with purely analytic questions. You might think that such a measure would only be of interest in itself, to see how bad things can get. But one finds that it can be applied with reasonable success to analytical problems.

I may as well finish up with a famous statement due either to Russell or to Whitehead. I don't remember which. One of them gave a lecture and the other was presiding. Let us say that Russell was presiding and Whitehead gave the lecture. The lecture was on the foundations of quantum mechanics. It was apparently not only too much for the audience, but also too much for the presiding officer. The whole thing was extraordinarily difficult, abstruse, and unclear. Yet, when the lecture was over the chairman felt that he must make some comment. He made one which was both polite and true. He simply said, "We must be thankful to the speaker for not further darkening this vastly obscure subject."

TABLE OF CONTENTS

PART I

FOUNDATIONS OF STATISTICAL MECHANICS

	Page
The Classical Paradoxes	2
The H-Theorem	3
Reversability Paradox	3
Liouville's Theorem	4
Poincare's Recurrence Theorem	7
A Simple Model with the Same Difficulties	10
The H-Theorem	13
Reversability Paradox	13
Recurrence Paradox	14
Probabilistic Analysis	15
The Paradoxes Resolved	23
A Second Model, but Easier	23
The Liouville Equation	28
The Master Equation	30
The Two Basic Approaches	32
They are different	41
But not very	42
The Boltzmann Equation for a Gas	47
Statistical Approach	51
The Master Equation	56
A Simpler Model of a Gas	59
The Master Equation	59
Contracted Distributions	60
Lo, the Boltzmann Equation	63
Chaos -- Chaotic Distributions	64
The H-Theorem	69
The Maxwell Distribution	71
A Class of Chaotic Distributions	74
The Linear Boltzmann Equation	80
The Linearized Boltzmann Equation	84
Hilbert's Procedure	90
Relation to the Master Equation Approach	93

TABLE OF CONTENTS
(Continued)

PART II

STOCHASTIC MODELS

	Page
Stochastic Model Related to Telegrapher's Equation	102
A Discrete Random-Walk	103
Limiting Case	108
Monte Carlo	109
A Continuous Model	110
The Poisson Process	110
Solution of the Telegrapher's Equation	113
Related Equations in Higher Dimensions	114
Asymptotic Behavior of Eigenvalues of the Laplacian	123
Relation to Diffusion Equation	124
Stochastic Interpretation	125
The Principle of Not Feeling the Boundary for Early Times	126
Use of a Tauberian Theorem	127
Advantages of this Approach	130
Brownian Motion	131
The Chapman-Kolmogoroff Equation	132
Solutions of this Equation	133
Wiener Measure	134
A Functional, Its Distribution, and a Related Differential Equation	140
Stochastic Interpretation	142
The Fundamental Solution	147
Eigenvalues of the Schroedinger Equation	148
Monte Carlo	155
Potential Theory	158
Average Time a Brownian Particle Spends in a Region Ω	161
Difference Between Space and the Plane	163
Distribution of Sojourn Times	166
A Related Integral Equation	169
Probabilistic Expression for the Capacitory Potential	173
The Capacity	181
Two Dimensions	183
Another Measure Based on the Chapman-Kolmogoroff Equation	185