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 $0<t_1<t_2<\cdots<t_n$.)

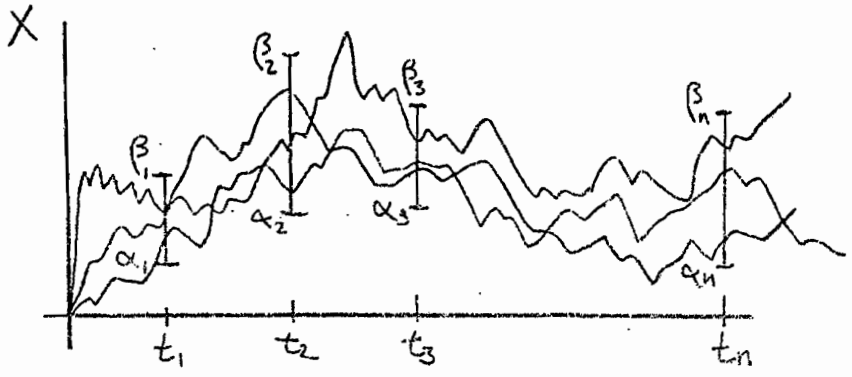
In the integrand I have a product of probabilities. That I can always separate the probability into such a product is referred to as the Markovian property. In our case, it's actually simpler than the Markovian property. It's simply the independence of the increments during non-overlapping time intervals.

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

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

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

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

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

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

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:

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

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

$$E\left\{e^{-u\int_0^\infty V(x_0+x(\tau))d\tau}\right\} = \int_{-\infty}^{\infty} Q(x_0|x,t)\,dx \tag{84}$$

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:

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

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.