

# Mechanics to Manifolds

## III. Quantum Mechanics

Andrew Stacey

June 7, 1999

### 1 Introduction

This is the third seminar in our mini series detailing how Physics feeds into Geometry. In the first seminar I gave an introduction to Mechanics from a Mathematician's point of view. In the second, Paul Cooper showed how some specific problems in Classical Mechanics give rise to problems in geometry. In this talk, I intend to do the same for Quantum Mechanics.

What I would like to do is concentrate on two particular problems which I know something about. These are the electron and the construction of path integration. I shall remind you of the basics of Quantum Mechanics and then show how considering the electron and similar particles leads to the necessity of supermanifolds. Then I shall introduce the path integral formulation of Quantum Mechanics, much beloved of Feynman.

### 2 Schrödinger's Equation and Fermions

#### 2.1 Quantum Theory Revisited

We recall the basic postulates of Quantum Theory from the Mathematical point of view.

1. The instantaneous pure states of the universe are rays in a Hilbert space,  $\mathcal{H}$  (that is, unit vectors of arbitrary phase).
2. The instantaneous observables in a Quantum system are Hermitian operators on this space. The expectation value of an observable  $B$  of a state  $\psi$  is  $\langle \psi | B \psi \rangle = \langle \psi | B | \psi \rangle$ . This is the mean of many observations of  $B$  all carried out on the same state  $\psi$ .
3. The Hamiltonian  $H$  is the infinitesimal generator of the unitary time evolution group. The momentum  $P$  is the infinitesimal generator of the unitary space translation group. The angular momentum  $J$  is the infinitesimal generator of the unitary spatial rotation group. What this means is that

there is a group homomorphism  $\mathbb{R} \mapsto U(\mathcal{H})$  given by  $t \mapsto U(t) = e^{-itH/\hbar}$  which controls the dynamics, and similarly for spatial translation and rotation.

Now there is a difference between actually making a measurement of an operator  $B$  on the system and simply calculating the expectation value  $\langle \psi | B | \psi \rangle$ . This latter is a statement of probability and says that the mean value of doing experiment  $B$  on state  $\psi$  a large number of times should be about  $\langle \psi | B | \psi \rangle$ . One can do this calculation without disturbing the system<sup>1</sup>.

Thus if we are in the prediction game and never actually make a measurement (so never disturb the system), all we can deal with and all we want to deal with is quantities of the form  $\langle \psi | B | \psi \rangle$ . All we really want to know is how this evolves in time. Postulate 3 gives us the answer:

$$\langle \psi | B | \psi \rangle (t) = \langle U(t)\psi | B | U(t)\psi \rangle \quad (1)$$

$$= \langle \psi | U(t)^* B U(t) | \psi \rangle \quad (2)$$

Equation 1 refers to the Schrödinger picture of Quantum Mechanics in which the evolution in time is assigned to the states and we define  $\psi(t) = U(t)\psi_0$ , where  $\psi_0$  is the initial state. Then states satisfy Schrödinger's equation:

$$i\hbar \frac{d\psi}{dt} = H\psi(t)$$

Equation 2 refers to the Heisenberg picture of Quantum Mechanics in which the evolution in time is assigned to the observables (and operators in general). We define  $B(t) = U(t)^* B U(t)$ . Then observables satisfy the equation:

$$\hbar \frac{dB(t)}{dt} = [iH, B(t)]$$

These equations have formal solutions in terms of  $H$  which we shall use to indicate certain important results. The actual results require a little more analysis than we shall use here.

The formal solution of the Schrödinger Equation is:

$$\begin{aligned} \psi(t) &= e^{-itH/\hbar}\psi \\ &= \sum_{n=0}^{\infty} \frac{(-it/\hbar)^n}{n!} H^n \psi \end{aligned} \quad (3)$$

and of the Heisenberg Equation is:

$$B(t) = \sum_{n=0}^{\infty} \frac{(-it/\hbar)^n}{n!} [H, [H, \dots [H, B] \dots]] \quad (4)$$

Now if we have a self-adjoint operator  $A$  acting on the Hilbert space,  $\mathcal{H}$ , then using the formal solution of the Heisenberg picture it is easy to see that if

---

<sup>1</sup>Unless the system consists of the computer you are using to do the calculation.

$[H, A] = 0$  - i.e. that  $A$  commutes with  $H$  - then  $A(t) = A$  for all time. This is particularly true for the Hamiltonian itself.

Transferring to the Schrödinger picture, this translates into the fact that if the universe is in an eigenstate of  $A$  at one time then it must always be in an eigenstate of  $A$  with the same eigenvalue. Of course, the actual state may vary within the eigenspace corresponding to that that eigenvalue.

## 2.2 The Exchange Operator

I now wish to consider a particular type of operator. Suppose we have a two particle system, say the Hydrogen atom. Then our overall state must encode the information of the individual states of the electron and proton. In fact, one would hope that this encoding was reversible since electrons and protons are easily distinguishable. Thus whenever we do an experiment and, say, find a particle in a particular place then it ought to be possible to work out which particle it is.

The way it works is by using the tensor product. If  $\mathcal{H}_e$  refers to the states of the electron and  $\mathcal{H}_p$  refers to the states of the proton then we find that the total Hilbert Space of the pair is  $\mathcal{H}_e \otimes \mathcal{H}_p$ .

So if we know that the proton is in state  $\psi_p$  and the electron is in state  $\psi_e$  then the overall state is  $\psi_e \otimes \psi_p$ .

Protons and electrons are easily distinguished. Suppose we had two identical particles and for simplicity's sake suppose they don't interact, for example two neutrons. Then one would expect:

$$\begin{aligned} \mathcal{H}_{nn} &= \mathcal{H}_n \otimes \mathcal{H}_n \\ H_{nn} &= H_n \otimes I + I \otimes H_n \end{aligned} \tag{5}$$

Consider the exchange operator acting on the Hilbert Space  $\mathcal{H}_{nn}$ . Since this space is the tensor product of two identical spaces, we can define an operator  $R_{12}$  which swaps the factors. This is clearly well defined and moreover it clearly commutes with  $H_{nn}$ . Thus from the above, if we start in an eigenstate of  $R_{12}$  we must stay in an eigenstate of  $R_{12}$ . Since  $R_{12}^2 = I$  then it is clearly self-adjoint and has spectrum  $\{-1, 1\}$ . Thus it divides  $\mathcal{H}_{nn}$  into two eigenspaces and if a system starts in an eigenspace of  $R_{12}$  then it stays in this eigenspace.

This would not be half as interesting as it is were it not for the fact that for nearly all elementary particles, the system does always start in one or other eigenspace and the eigenspace depends solely on the particle concerned. Particles whose states start in the +1-eigenspace are called bosons, those whose states start in the -1-eigenspace are called fermions.

As an indication of why this should be, suppose we were to do an experiment on our system and measure a neutron at a particular place, then we have no way of knowing which neutron it was. Even if we knew the initial positions of each neutron, the indeterminacy of Quantum Theory means that we cannot trace their paths and so we don't know which is which. Thus if we know that

one neutron is in state  $\psi_1$  and the other in state  $\psi_2$  then the overall state could be either  $\psi_1 \otimes \psi_2$  or  $\psi_2 \otimes \psi_1$ .

Thus in passing from one neutron to two, we have lost information and so our Hilbert Space of states should be smaller to reflect this loss. Smaller in what way is determined by whether the particles are bosons or fermions.

Typical bosons are the photon, the pion and the Helium nucleus. Typical fermions are the electron, the neutron and the proton. The deciding factor is a quantity called intrinsic spin. Bosons have integral spin, fermions have half-integral spin.

### 3 Fermions and Supermanifolds

#### 3.1 Quantization

From the Physicists' point of view, the Hilbert Spaces of states and the spaces of observables are constructed via a process known as Quantization. It should be emphasized again that this process is not well-defined and does depend on choices, but even so it is well understood.

The simplest form of this process starts with the classical problem. This consists of the configuration space, which is a Riemannian manifold,  $M$ . It then forms the cotangent space  $T^*M$ . This is the classical phase space. The classical space of observables is the algebra  $C^\infty(T^*M, \mathbb{R})$ .

We then consider the Hilbert space  $L^2(M, \mathbb{C})$  with a certain connection. We can then construct an almost homomorphism from  $C^\infty(T^*M, \mathbb{R})$  into the self-adjoint operators on this Hilbert Space.

By an "almost" homomorphism we mean that we have a well-defined scale on the universe (given by our Riemannian metric) and with respect to this scale then the images of two functions commute up to  $O(\hbar)$ .

We also want a certain amount of irreducibility in our Hilbert Space as a representation of the image of  $C^\infty(T^*M, \mathbb{C})$ . The technical term for this is the *unique vacuum vector postulate*. A vacuum vector is an eigenstate of  $H$  with zero eigenvalue.

#### 3.2 Quantization of Fermions

In the case of our two neutrons, that they are identical implies that their configuration spaces are the same. Then Equation 5 corresponds to the fact that:

$$L^2(M \times M, \mathbb{C}) \cong L^2(M, \mathbb{C}) \otimes L^2(M, \mathbb{C})$$

Now we have a problem because the unique vacuum vector postulate combined with the fact that neutrons are fermions means that we have to take at most the  $-1$ -eigenstate of  $R_{12}$  as our Hilbert Space. Ideally we would like a construction which takes us straight from the manifold, via quantization, to the correct Hilbert Space.

It isn't too hard to see that the  $-1$ -eigenspace is spanned by states of the form  $\psi_1 \otimes \psi_2 - \psi_2 \otimes \psi_1$  and the operator space of this eigenspace is spanned by observables of the form  $B_1 \otimes B_2 + B_2 \otimes B_1$ .

There is an easy way to define the process which takes us from  $\mathcal{H}_n$  to this space and that is to use the graded tensor product. We define a grading on  $\mathcal{H}_n$  by insisting that all states are odd. Then all observables become even as they map an odd space to an odd space. Thus  $\mathcal{H}_n$  becomes a  $\mathbb{Z}_2$ -graded space and for such spaces then there is a natural graded tensor product given by:

$$a \widehat{\otimes} b = (-1)^{ab} \widehat{\otimes} a$$

where if an element appears as a superscript then it is understood that we mean the degree of that element. Note that this only makes sense if we are tensoring the same object. This can also be expressed as a quotient space of the normal tensor product.

This construction can evidently be extended to cases where the Hilbert Spaces are not purely odd or even but a sum of odd and even parts.

What we now require is some extra structure on  $M$  which gives the required parity on the Hilbert Spaces  $L^2(M, \mathbb{C})$ . This means that the following identities must hold:

$$\begin{aligned} L^2(M \times M, \mathbb{C}) &\cong L^2(M, \mathbb{C}) \widehat{\otimes} L^2(M, \mathbb{C}) \\ C^\infty(M \times M, \mathbb{R}) &\cong C^\infty(M, \mathbb{R}) \widehat{\otimes} C^\infty(M, \mathbb{R}) \end{aligned}$$

This clearly means that we have to redefine our concepts of differentiable.

This leads to the notion of a supermanifold. Basically a supermanifold is like a manifold with a local  $\mathbb{Z}_2$ -grading and then everything else is defined to take this into account. I shan't say much more on this here except to say that I may give a seminar on this subject later on if we don't get enough volunteers for other talks and that supermanifolds are the right place to start if you want to know why people bother with spin manifolds.

## 4 Feynman's Path Integrals

### 4.1 The Action Integral

As has been said before, one of the most elegant methods of stating the laws of Classical Mechanics involves the *action integral*. This involves a quantity assigned to each possible path a system could take through phase space called the action of that path. Classical Mechanics is based on the statement that the path followed by the system is that which is an extremum of the action.

We shall consider the case of a system which is given by configuration data<sup>2</sup>  $x_a$  at time  $t_a$  and at  $x_b$  at time  $t_b$  with  $t_b > t_a$ . In other words, we know the

---

<sup>2</sup>For ease of notation, we shall refer to a point in configuration space as a "position" although the configuration coordinates can include data other than physical position.

initial and final positions of the system and the time it took to get from one to the other. The action is defined by:

$$S(\gamma) = \int_{t_a}^{t_b} L(\dot{\gamma}, \gamma, t) dt$$

where  $L$  is the Lagrangian of the system and  $\gamma$  is a path from  $x_a$  at time  $t_a$  and  $x_b$  at time  $t_b$ . Assuming sufficient differentiability of  $L$  then we find that the extrema of  $S$  are precisely those paths which obey the Lagrangian laws of motion.

In Quantum Mechanics, there is not just one path which the system can travel along. Somehow we have to include the fact into the equation that all paths contribute to the outcome. However, in Quantum Theory we are no longer looking for the path which the system travelled along as it doesn't exist, we are looking at the probability that the system which is at  $x_a$  at time  $t_a$  now is at  $x_b$  at time  $t_b$ .

This probability is then a function of all the possible paths between  $x_a$  and  $x_b$ . In a very basic form, we can write this probability as  $P(b, a)$  where  $P(b, a) = |K(b, a)|^2$ . This doesn't say much at present, but does at least state the problem in probabilistic form in a manner reminiscent of the postulates of Quantum Theory.

Our attention thus turns to  $K(b, a)$ . This is called the *kernel*. This must involve contributions from all possible paths:

$$K(b, a) = \sum_{\gamma} \phi[\gamma]$$

where the sum runs over all paths from  $x_a$  to  $x_b$ .

How we find these functions  $\phi$  is dictated to a certain degree by the Classical Limit. We need something which says that if the action of paths varies slowly<sup>3</sup> as the path varies, then the system must be most likely to travel along the path corresponding to an extremum of the action. This leads to the idea of the phase of a path:

$$\phi[\gamma] = \text{const} e^{(i/\hbar)S(\gamma)}$$

When we consider a classical problem then everything is large with comparison with  $\hbar$ . Thus  $S/\hbar$  is some large angle. If we change the path slightly on the classical scale,  $S$  changes very slightly on the classical scale but  $S/\hbar$  changes by a large amount. Thus a small change in path corresponds to a large change in phase and so the phase factor will vary wildly around the circle. Thus the contributions from nearby paths will all add up to zero. The only time this won't happen is if a small classical change in path leads to no change in action and so nearby paths are nearly in phase. Thus for this special path, the contributions all add together and so we find that  $K(b, a)$  depends only on the action of the classical path and so we do recover the classical case in the limit of the Quantum system.

---

<sup>3</sup>Slowly compared with the size of  $\hbar$

## 4.2 The Path Integral

The only problem with the above is the idea of taking sums over all possible paths. This is of course not rigorously defined. It is to path integration what counting squares is to Lebesgue integration. Fortunately one can put it on a slightly<sup>4</sup> more rigorous footing in a similar way to standard integration.

WARNING: The following method is not intended to be accepted as definitive. Path Integration is an extremely messy subject and dealing flippantly with the infinities involved should only be undertaken by experienced Quantum Mechanics.

We consider a system with Lagrangian given by:

$$L(\dot{\gamma}, \gamma, t) = \frac{1}{2}m\dot{\gamma}(t)^2 - V(\gamma(t), t)$$

As above, we consider paths in which our system starts at  $x_a$  at time  $t_a$  and ends at  $x_b$  at time  $t_b$ .

With this setup it is possible to define the concept of a piecewise classical path. We divide the time interval into  $N$  equal length parts, of length  $\epsilon = \frac{1}{N}(t_b - t_a)$ . This gives us  $N - 1$  times  $t_i$ . For each of these we select a point  $x_i$  and then construct our path by joining these  $x_i$  together into trajectories by taking the classical path (i.e. the geodesic) between them. Thus we take the classical path which the system would take if we forced it to go through each  $x_i$ .

For a given  $N$  we sum over all such paths by integrating with respect to the  $x_i$ . Thus:

$$K(b, a) \sim \int \int \dots \int \phi[\gamma(t)] dx_1 dx_2 \dots dx_{N-1}$$

The proportional sign is because if we let  $N$  tend to infinity, so does the integral. In general, finding a normalising factor is fiendishly difficult, but for our example we use  $A^{-N}$ , where

$$A = \left( \frac{2\pi i \hbar \epsilon}{m} \right)^{1/2}$$

With this factor then we have:

$$K(b, a) = \lim_{\epsilon \rightarrow 0} \frac{1}{A} \int \int \dots \int e^{(i/\hbar)S[b,a]} \frac{dx_1}{A} \frac{dx_2}{A} \dots \frac{dx_{N-1}}{A}$$

where:

$$S[b, a] = \int_{t_a}^{t_b} L(\dot{\gamma}, \gamma, t)$$

This method of definition mirrors that of the Riemann integral. There are other methods, just as there are other definitions of the ordinary integral. All

---

<sup>4</sup>This word is used advisedly.

are beset with the same renormalisation problems. Whichever method used to define it, it is written:

$$K(b, a) = \int_a^b e^{(i/\hbar)S[b,a]} \mathcal{D}\gamma$$

It is important to note that there is no general solution to the renormalisation problem and sometimes we even have to forget renormalisation and pretend that we know what infinity looks like<sup>5</sup>. However, it is extremely useful even if it doesn't technically exist.

### 4.3 Properties of the Path Integral

Suppose we start at point  $x_a$ , go through point  $x_c$  at time  $t_c$  before ending up at point  $x_b$ . As the action of a path depends only on the position, speed and time of the system at the points of that path, it is easy to see that the actions add. Thus:

$$S[b, a] = S[b, c] + S[c, a]$$

We can use this in our formula for  $K[b, a]$  by splitting the paths at time  $t_c$ . We consider the contributions for all paths which pass through a fixed  $x_c$  at this time and then integrate over all such points. However, given a path which goes through  $x_c$  we can further split this integral up by integrating first over all paths from  $x_a$  to  $x_c$  and then from  $x_c$  to  $x_b$ . Thus:

$$\begin{aligned} K[b, a] &= \int_c \int \int e^{(i/\hbar)(S[b,c]+S[c,a])} \mathcal{D}\gamma_{bc} \mathcal{D}\gamma_{ca} dx_c \\ &= \int_c \int e^{(i/\hbar)S[b,c]} \mathcal{D}\gamma_{bc} \int e^{(i/\hbar)S[c,a]} \mathcal{D}\gamma_{ca} dx_c \\ &= \int_c K[b, c] K[c, a] dx_c \end{aligned}$$

Thus the kernels for successive events multiply. This can of course be extended to multiple events. In fact, one can start from this point and work backwards to recover the definition of the kernel.

### 4.4 Wave Functions

A wave function is like a one-sided kernel. A wave function gives probabilities of the system now without considering where it came from. Thus in a sense it is the sum over  $a$  of  $K[b, a]$ , but renormalised. This is not useful for derivations, however, as we have yet more renormalisations to do. It is easier to start over with wave functions and rederive everything. However, we do arrive at a remarkably similar result.

---

<sup>5</sup>According to Terry Pratchett then it's blue. According to Douglas Adams then it's dull.

Recall that the kernel for arriving in state  $b$  having started in state  $a$  is given in terms of a point  $c$  in the middle by:

$$K[b, a] = \int_c K[b, c]K[c, a]dx_c$$

and the wave function,  $\psi[b]$ , at  $x_b, t_b$  given that the wave function at  $x_c, t_c$  was  $\psi[c]$  is given by:

$$\psi[b] = \int_c K[b, c]\psi[c]dx_c$$

Another way of stating the difference between kernels and wave functions is by reference to history independence. A wave function has no sense of history whereas a kernel does.

## 4.5 The Integral Formulation of Quantum Theory

What we have here is an integral formulation of Quantum Theory. Integral formulations have a lot going for them because annoying things like continuity of solutions are no longer necessary. Also, integration is better for getting estimates on approximate solutions.

However, path integrals have a major failing. There is no general method for renormalising the kernels to ensure that they exist. In certain cases then they have particularly nice forms and particularly nice renormalisations. This is a mathematical failing. It doesn't seem to upset the Physicists too much. Despite not always existing, path integration is a powerful tool because of the *form* of the integral. One can apply manipulations to the integral using the fact that it looks a lot like an exponential and then interpret the results. This interpretation may seem like a magic trick, but it does work. This is where Feynman's celebrated diagrams come into their own. These are an aid to the correct formulation, simplification and interpretation of path integrals.

## 4.6 The Weiner Measure

All is not quite lost, mathematically speaking. Although path integrals do not, strictly speaking, exist, they do have a close relative which does. The problem with the path integral is that the action integral is real and so the path integrand is just a phase factor. This is then integrated over an infinite dimensional space and can lead to problems. If our action integral could be replaced by something which was real and negative then we might be able to make some more sense of it.

Consider the Schrödinger Equation:

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi$$

This has formal solution:

$$\psi(t) = e^{-itH/\hbar}\psi$$

$$= \sum_{n=0}^{\infty} \frac{(-it/\hbar)^n}{n!} H^n \psi$$

Feynman's path integral formulation can then be expressed by the following horrendous equation:

$$\left( e^{-itH/\hbar} \psi \right) (y) = \int_x \int_{\gamma} \exp \left( (i/\hbar) \int_0^t L(\dot{\gamma}, \gamma, t) dt \right) \mathcal{D}\gamma \psi(x) dx$$

In which the path integral is over all paths from  $x$  to  $y$  taking time  $t$ .

In this equation,  $t$  is a real positive number and so this is some sort of function from  $\mathbb{R}^+$  to some space. However, at the moment then this has certain convergence problems. One trick which often gets round convergence problems, or at least highlights them in a new and useful way, is that of analytic continuation to the whole complex plane.

Doing this results in the Heat Equation:

$$-\hbar \frac{\partial \psi}{\partial t} = H \psi$$

This has formal solution:

$$\begin{aligned} \psi(t) &= e^{-tH/\hbar} \psi \\ &= \sum_{n=0}^{\infty} \frac{(-t/\hbar)^n}{n!} H^n \psi \end{aligned}$$

We can then define an analogue of the path integral equation to get:

$$\left( e^{-tH/\hbar} \psi \right) (y) = \int_x \int_{\gamma} \exp \left( \frac{1}{\hbar} \int_0^t L(\dot{\gamma}, \gamma, t) dt \right) \mathcal{D}\gamma \psi(x) dx$$

Given certain conditions on  $H$  then we can define this object. The most usual starting point is when  $H = H_0$ , the Hamiltonian for a free system.

With this Hamiltonian, we find that the path space  $\mathcal{W}(x, y, t)$  has a measure called the *Weiner Measure*. With respect to this measure then we have:

$$\int_{\mathcal{W}(x,y,t)} dW_{x,y}^t = \mathcal{K}_t^0(x, y)$$

Where we write  $\mathcal{K}$  for the Heat Kernel to distinguish it from the Schrödinger Kernel,  $K$ .

For  $H = H_0 + V$  we get the Feynman-Kac formula:

$$\mathcal{K}_t(x, y) = \int \exp \left( - \int_0^t V(\gamma(s)) ds \right) dW_{x,y}^t$$

This is an extremely useful formula, for example in Index Theory, and many of the path integral methods derived for Feynman's dodgy integrals work for this object as well.

## 5 Conclusion

What I have shown here is the emergence of two mathematical objects from Quantum Theory. These are the concept of a supermanifolds, which also feeds into the theory of spin manifolds, and the Weiner measure on path space, which also feeds into index theory and hence to elliptic genera. There are of course many other ways in which Quantum Theory has either suggested interesting mathematical objects or stimulated research into objects already known about but largely ignored. These two are ones which I know a little about already and so feel able to say something about without getting hopelessly lost in the detail.