On quantum mechanics

Warren D. Smith[*]
wds@research.nj.nec.com

date = October 13, 1998

Abstract — **Part I of 2 part lecture notes. Nonrigorous survey.**

There are several things wrong with orthodox quantum mechanics. First, there is the matter of the nature of "measurement," which was always the weak point in the Copenhagen interpretation. It's too vague. This is related to the usual philosophical difficulties which arise from the fact that QM is nonlocal but physics is "supposed" to be local. Second, there is the problem that the usual formulation of quantum mechanics does not allow the treatment of nonHamiltonian systems, e.g. friction. Third, there is the problem of computation. Theoretical physics is supposed to be a computational method, but conventional quantum mechanics seems to be a very bad one. Recent work by Peter Shor shows just how bad. One would like for theoretical physics to be simulatable with a Turing machine with at most polynomially bounded relative slowdown ("strong form of Church's thesis") but at present nobody knows how to do this for most scenarios in quantum mechanics.

I also claim that enhancements of quantum mechanics, developed over the last 30 years by a number of physicists, have solved all of the above problems – except for the computational one, and light may be about to be shed on that.

We survey all of these developments and contribute a few new ideas and observations.

Keywords — Dissipative quantum mechanics, measurement, Schrödinger's cat, stochastic mechanics, path integrals, hidden variable theories, Church's thesis, computational physics.

Part I

Summary of classical and quantum mechanics and orthodox world view

1 Introduction

We will present the usual formalism of classical and quantum mechanics, rarely saying anything new. Things get more interesting in §4, where we point out the flaws in the usual QM picture, and §5, where we list a lot of interesting unconventional ways to view quantum mechanics. §?? sketches QED in a vague and incomplete manner, merely mentioning the facts which we will need later.

2 Summary of classical mechanics

In the Beginning, there were Newton's laws. Mechanical systems were made of bodies linked together via springs, cams, levers, and whatnot, through which forces could be transmitted. In principle, Newton's laws allowed you to write down a system of second order differential equations describing the motion of any such system, but for a complicated system, this could be a painful and error-prone process.

Then, along came the Lagrangian and Hamiltonian reformulations. In the **Lagrange** formulation, there was a single real-valued function L of the state of the system, called the "Lagrangian." It was defined (nonrelativistically, and in the absence of a vector potential[1]) via

$$L = \text{Kinetic Energy} - \text{Potential Energy}. \tag{2}$$

The "state of the system" was defined by the "generalized position" coordinates q_r, which could be anything at all, and their time derivatives (and the time). The time evolution of the system was

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q_r}}\right) = \frac{\partial L}{\partial q_r}. \tag{3}$$

[1] The Lagrangian formulation still works (Lanczos 1966) in special relativity and with a vector potential $\vec{A}$, except that (EQ 2) has to be replaced by

$$L = \sum_j m_j c^2 (1 - \sqrt{1 - \vec{v}_j^2/c^2}) + \vec{A}\cdot\vec{v}_j - \text{Potential Energy}. \tag{1}$$

where j ranges over the particles in the system, which have rest masses m_j and velocities v_j. The function $f(v) = mc^2(1 - \sqrt{1 - v^2/c^2})$ was chosen so that $f'(v)$ is the momentum of a particle of rest mass m and velocity v; in the nonrelativistic limit $f(v) \approx mv^2/2$.

[*] NECI, Princeton NJ

This formulation was superior because one could use any coordinates whatsoever as the q_r, without having to worry about making errors while doing complicated coordinate transformations, and the only step at all tricky (since it alone encapsulates Newton's laws) is in writing down the Lagrangian function; the rest is automatic. Also this has the nice aesthetic feature that the "action"

$$S = \int_{t_1}^{t_2} L(q_r, \dot{q_r}) dt \tag{4}$$

is minimized (more precisely, extremized) by the actual trajectory $q_r(t)$.

Hamilton came up with a different formulation by defining the "generalized momenta" $p_r = \partial L / \partial \dot{q}_r$. The $2d$-dimensional space with coordinates q_r, p_r, $r \in \{1..d\}$, is called "phase space." The "Hamiltonian"

$$H(q_r, p_r, t) = \text{Kinetic Energy} + \text{Potential Energy} \tag{5}$$

is a real valued function on phase space. The time evolution of the system is then determined via the first order system

$$\dot{q_r} = \frac{\partial H}{\partial p_r}, \quad \dot{p_r} = -\frac{\partial H}{\partial q_r}. \tag{6}$$

In the Hamilton formulation, an important notational device is the "Poisson bracket" $\{A, B\}$, which means $\sum_r \frac{\partial A}{\partial p_r}\frac{\partial B}{\partial q_r} - \frac{\partial A}{\partial q_r}\frac{\partial B}{\partial p_r}$. The time evolution of any quantity Q which depends on q_r, p_r, and t under the Hamiltonian flow is governed by $\dot{Q} = \{H, Q\} + \frac{\partial Q}{\partial t}$, and Hamilton's equations are merely the special cases of this which arise when $Q = q_k$ and $Q = p_k$. Constants of the motion C obey $\{H, C\} + \frac{\partial C}{\partial t} = 0$ and thus for example $\{H, H\} = 0$ exposes the conservation of energy in systems in which H does not depend explicitly on t. The Hamiltonian flow preserves phase-space volume since any flow governed by $\dot{\vec{x}} = \vec{f}(\vec{x})$ has $d(\ln \text{vol})/dt = \vec{\nabla} \cdot \vec{f}$, which in our case is zero just due to equality of mixed partials.

From the modern computational complexity point of view, though, the Lagrange and Hamilton formulations do not seem to be any improvement on plain Newton's laws. But, they aren't worse.

Hamilton was still unsatisfied, because one can only solve for the time evolution of a system starting from some specified initial state. He wanted a formulation which would allow one to solve at once for all trajectories starting from all possible initial states. **Jacobi** invented a formulation that allows this: the "Hamilton Jacobi equation" is

$$H(q_r, \frac{\partial S}{\partial q_r}, t) + \frac{\partial S}{\partial t} = 0. \tag{7}$$

This is a first order *partial* differential equation, which lives in $\mathbf{R}^d$, where d is the number of positional degrees of freedom in the system. It is satisfied by the action function $S(q_r, t)$. If one could solve the HJ equation for $S(q_r, t)$ in closed form (there would be certain free parameters, to be sure, in this solution) one would then have a "master solution" from which all sorts of trajectory information could be extracted. Specifically, let the solution be $S(q_r, t) = f(q_r, t, \alpha_r) + A$ where the α_r are d constant free parameters, and A is one additional constant. Then to find out about the specific trajectory arising from some starting state, you would first choose A so that $S(q_r^{\text{initial}}, 0) = 0$, and then the generalized momenta at any point q_r on the trajectory would be

$$p_r = \frac{\partial S}{\partial q_r}. \tag{8}$$

(These would have to be forced to have the right values at $t = 0$ by choice of the free parameters α_r.) The trajectory itself could then be determined by noting that $\beta_r = \partial f / \partial \alpha_r$ are also constants of the motion; and these equations implicitly determine $q_r(t)$.

The Hamilton-Jacobi approach is definitely (from the computational complexity point of view) far *worse* than the old methods, because it requires the solution of a partial differential equation in a high dimensional space. For example, in a system with 100 degrees of freedom, and using a naive numerical method for PDEs based on a 10-point grid... that would be 10^{100} grid points. For this reason, the HJ technique is usually only regarded as a useful computational tool only when the HJ partial differential equations separate (so that their solution is reduced to solving ordinary differential equations). On the other hand, the HJ approach gives you more, since in some sense it is solving an infinite number of the sort of problems that the older approaches solve.

3 Summary of quantum mechanics

Schrödinger postulated the existence of a complex-valued "wave function" $\Psi(q_r)$ associated with any mechanical system, which obeys

$$H(q_r, -i\hbar\frac{\partial}{\partial q_r}, t)\Psi = i\hbar\frac{\partial \Psi}{\partial t}. \tag{9}$$

Here $\hbar$ is Planck's constant $h = 6.6260755(40) \times 10^{-34}$Joule-sec, divided by 2π, and the Hamiltonian H has now (thanks to its second argument) become a partial differential operator.

Since any complex constant times Ψ is also a solution, one generally requires Ψ to be "normalized," that is, $\int |\Psi|^2 d^d q_r = 1$. (This integral is a constant of the motion.) A more interesting boundary condition is the requirement that under the exchange of two particles which are identical fermions, Ψ should change sign, while under the exchange of two bosons, Ψ is unaffected.

How did Schrödinger decide on this equation? He wanted a partial differential equation, to allow the manifestation of full-dimensional wavelike behavior. He wanted to make Ψ *complex* to allow the possibility of an overall phase. The replacement of the generalized momentum p_r by the operator $-i\hbar\frac{\partial}{\partial q_r}$ was motivated by De Broglie's formula $h = \lambda p$ for the wavelength λ of a lump of matter with momentum p (which explains the Bohr atom) – notice that differentiating a complex exponential wave has the effect of multiplying it by $2\pi i/\lambda$.

The replacement of the energy by $i\hbar\frac{\partial}{\partial t}$ was similarly motivated by the formula $E = hf$ for the energy of a system with periodic time dependence at frequency f (needed to explain the photoelectric effect, Planck blackbody law, etc.). More precisely, you just verify by direct substitution that this choice made plane complex exponential waves are solutions if the potential is zero (free particle) and, in the case when the time dependence factors out of the wavefunction, time-oscillatory space-stationary solutions are exact solutions if and only if $E = hf$.

If the complex number Ψ is expressed via modulus and phase $\Psi = |\Psi|e^{iS/\hbar}$, Schrödinger's equation is rewritten (at least, under the assumption that

$$H(q_r, p_r, t) = U(q_r) + \sum_k \frac{(p_k - V_k(q_r))^2}{2m_k} \tag{10}$$

as arises in orthogonal coordinates q_r with a scalar potential U and a vector potential V_k, where[2] the generalized momenta of the masses m_r with velocities v_r are $p_r = m_r v_r + V_r$)

$$\frac{\partial S}{\partial t} + H(q_r, \frac{\partial S}{\partial q_r}, t) = \frac{\hbar^2}{2|\Psi|}\sum_k \frac{\partial^2 |\Psi|}{m_k \partial q_k^2} \tag{11}$$

$$\frac{\partial \rho}{\partial t} + \sum_k \frac{\partial}{\partial q_k}\left(\rho \frac{\frac{\partial S}{\partial q_k} - V_k}{m_k}\right) = 0, \text{ where } \rho = |\Psi|^2 \tag{12}$$

so that in some sense in the limit as $\hbar \to 0$, we recover classical mechanics: S obeys the Hamilton-Jacobi equation (EQ 7), and $\rho = |\Psi|^2$ obeys a mass conservation law. Another way in which classical mechanics is related to quantum mechanics is Ehrenfest's theorem: if $\hat{Q}$ is some linear differential operator (symbols with hats will represent operators from now on) then define its "expectation value" $\bar{Q} = \int \Psi^*(q_r)\hat{Q}\Psi(q_r)d^d q_r$; and then the time evolution of $\bar{Q}$ will be governed by an expectation value version of Hamilton's classical equation $\dot{\bar{Q}} = \overline{\{\widehat{H,Q}\}}$.

One can dispute whether this all makes sense, but most importantly, there is experiment. Schrödinger knew that his equation yielded the right energy levels for the hydrogen atom (within errors due to relativistic effects). Since then it has been found to yield agreement with experiment for many other systems. For example, Pekeris (1958-1962) solved the full 6-dimensional Schrödinger equation for the Helium atom numerically, and then applied small relativistic, Lamb, and "mass polarization" perturbative corrections. The theoretically predicted ionization energy corresponded to a $19831066.1 \pm 2.3\text{m}^{-1}$ line. This arises from Pekeris's nonrelativistic unperturbed value (but using the most recent value of the Rydberg $R_\infty = 10973731.534(13)\text{m}^{-1}$) of $19831737.0 \pm 2.2\text{m}^{-1}$; the nuclear mass and relativistic adjustments total $-534.8 \pm .05$ (about -478 of this is nuclear and -57 relativistic) and the Lamb adjustment of Saltpeter and Zaidi 1962 is -136.1 ± 2.1. Meanwhile the measured value (Herzberg 1958, with improved analysis by Seaton 1966) was $19831076 \pm 6\text{m}^{-1}$.

According to the "**Copenhagen interpretation**," Ψ (up to a complex phase factor) is the most that one can ever know about the system. $|\Psi(q_r, t)|^2$ is supposed to represent the probability density that, at time t, the system is in the classical state q_r. To each measurable classical quantity Q there corresponds a Hermitian linear differential[3] operator $\hat{Q}$; namely $\hat{p_r} = -i\hbar\frac{\partial}{\partial q_r}$ and $\hat{q_r} = q_r$, and if classically $C = \{A, B\}$, then $i\hbar\hat{C} = \hat{A}\hat{B} - \hat{B}\hat{A} = [\hat{A}, \hat{B}]$, and if classically $C = AB$ then $\hat{C} = \frac{1}{2}(\hat{A}\hat{B} + \hat{B}\hat{A})$. These operators $\hat{Q}$ have real eigenvalues λ and corresponding eigenfunctions ϕ forming a complete orthonormal basis: $\hat{Q}\phi = \lambda\phi$. Suppose $\Psi = \sum_k a_k \phi_k$ where the ϕ_k are the eigenfunctions of $\hat{Q}$ and the a_k are complex coefficients (the "sum" over k can be an integral). When one makes a "measurement" of a the classical quantity Q for a system with wavefunction Ψ, one will get as the result of the measurement λ_k, with probability $|a_k|^2$, and the wavefunction will instantly change from whatever it was before to ϕ_k and will then continue to evolve according to Schrödinger's equation.

In other words (in the abstract Hilbert space and operator viewpoint), there are two things you can do with quantum systems: you can let them **evolve**, or you can perform a **measurement** on them.

The time evolution of Ψ according to the Schrödinger equation $i\hbar\frac{\partial}{\partial t}\Psi = \hat{H}\Psi$ abstractly has the solution $e^{-it\hat{H}/\hbar}\Psi$, which is, for any t, a *unitary* transformation[4].

Performing a measurement effects a *sampling* from a probability space λ_k, followed by a *projection* of the state vector Ψ onto the ray of whatever eigenvector ϕ_k it is that the measurement gave (followed by a renormalization so Ψ remains a unit Hilbert vector). Although this projection operation is not a unitary transformation, it is still linear. The sampling operation may be thought of as in fact a nonlinear operation, although that arises from a naive view of "mixtures," and in the more sophisticated "random phase" view this operation still is linear.

To understand that, we have to introduce the notion of a **pure state** and a **mixture**. We said earlier that the most one could ever know about a quantum system was its wavefunction Ψ. It is possible to know less. If we are ignorant of the true wavefunction, our ignorance would be expressable as a probability distribution over which

[2]The mass m has subscripts because we are dealing with a multiparticle system. One is supposed to define m_r so that if the coordinates q_r refer to the jth particle, then m_r is the jth mass. Actually if we were to do this in fully general coordinates, m would have to be a two indexed tensor function of position. As long as we are talking about subscripts, note that we have often abused notation by using "q_r" to mean "$q_1, q_2, ..., q_d$" and "$\frac{\partial S}{\partial q_r}$" to mean "$\frac{\partial S}{\partial q_1} ... \frac{\partial S}{\partial q_d}$." It should always be clear what is meant.

[3]In cases such as spin, where the wave function is a function of a discrete rather than a continuous variable, the operators are no longer differential, but they are still linear and Hermitian.

[4]An operator $\hat{Q}$ is "Hermitian" if its transpose is its complex conjugate. It is "unitary" if its inverse is its conjugate transpose. If H is Hermitian then e^{iH} is unitary. If A and B are Hermitian so are $A + B$, $AB + BA$, $(AB - BA)i$, and A^n. Products of unitaries are unitary.

wavefunctions it could be. If we have Ψ_k with probability c_k^2, a useful quantity to introduce is $M = \sum_k c_k \Psi_k$, a wavefunction "mixture." It has many of the properties of a true wavefunction ("pure state") namely it also evolves according to the Schrödinger equation. However, it is no longer true that $|M|^2$ is the right probability density (or necessary even *a* probability density). The true density would be $\sum_k |c_k \Psi_k|^2$. If we regard the c_k's, instead of being real numbers, as being multiplied by independent and random complex phase factors $e^{i\theta_k}$, then $|M|^2$ would indeed be a probability density and give the right results, provided one always takes the expectation over the random phases at the end of the calculation (because the phase averaging kills the cross terms). If we have a pure state and we know that at some time in the future we are going to perform a measurement, then our ignorance of what the state of the quantum system after the measurement will be, is expressible as a mixture of eigenfunctions ϕ_k of the measurement operator with probabilistic weights $|a_k|^2$ where the expression of Ψ before the measurement was $\Psi = \sum a_k \phi_k$. Hence measurement may be thought of as the conversion of as pure state to a mixture, mathematically expressed by replacing a_k by $|a_k|$ in the above sum. This replacement is a nonlinear operation. On the other hand, in the random phase (as opposed to absolute value) view, measurement is not doing this nonlinear operation, it is instead merely multiplying the a_k's by random phases, which is a unitary operation. The random phase view of mixtures is in fact equivalent to the probability view if the probability distribution is in fact a mixture of random phases (of course this is a continuous distribution rather than a discrete one).

Mixtures are useful, not only to express our ignorance about what will happen after a measurement, but also to express our ignorance of the initial state of a system. For example, if all we know is the location and spin of one atom, but there are 10^{20} other atoms about which we know nothing, then we could express this as a mixture of all the eigenfunctions of the atom 1 (location and spin) operators, assigning a thermal probability weight to each one.

Incidentally, note that two operators $\hat{A}$ and $\hat{B}$ will have common eigenfunctions if and only if they commute. Hence if they do not commute with each other, then it is impossible to measure A and B simultaneously with certainty... indeed Heisenberg's **uncertainty theorem** states that the standard deviations $\sigma_A = \sqrt{\overline{A^2} - \bar{A}^2}$ and σ_B in measuring A and B will obey the inequality

$$\sigma_A \sigma_B \geq \frac{1}{2} |\overline{[\hat{A}, \hat{B}]}|. \tag{13}$$

In particular $\sigma_{p_a} \sigma_{q_b} \geq \delta_{ab} \hbar/2$, and equality is only achieved in the case of a unique "**minimum uncertainty wave packet.**" In the **Harmonic oscillator**[5] Hamiltonian $H(p,q) = \frac{1}{2m} p^2 + \frac{m}{2} \omega^2 q^2$, this wavepacket is

$$\Psi_{\text{mu}}(x,t) = (\frac{m\omega}{\pi\hbar})^{1/4} \exp\left(\frac{i\bar{p}x}{\hbar} - \frac{m\omega(x-\bar{x})^2}{2\hbar} - \frac{i\bar{p}\bar{x}}{2\hbar} - \frac{\omega t}{2} \right) \tag{14}$$

where $\bar{q}(t)$ and $\bar{p}(t)$ are the classical values, proportional to a circular function of ωt. The expectation value of the energy of this solution is

$$\bar{E}_{\text{mu}} = \int \Psi^* H \Psi dq = \frac{\overline{p^2}}{2m} + \frac{m\omega}{2} \overline{q^2}. \tag{15}$$

Meanwhile the orthonormal energy eigenstates of this Hamiltonian are

$$\Psi_n(x,t) = e^{-iE_n t/\hbar} 2^n n! (\frac{m\omega}{\pi\hbar})^{1/4} H_n(q\sqrt{m\omega/\hbar}) \exp \frac{-m\omega q^2}{2\hbar} \tag{16}$$

with corresponding eigenvalues $E_n = (n+\frac{1}{2})\hbar\omega$, $n = 0, 1, 2, \ldots$. Here the H_n are the Hermite polynomials

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}. \tag{17}$$

The minimum uncertainty wave packet expressed in this energy eigenbasis is

$$\Psi_{\text{mu}}(x,t) = \sum_{n \geq 0} c_n \Psi_n(x,t) \tag{18}$$

where

$$c_n = \int \Psi_n^* \Psi_{\text{mu}} dq|_{\mathrm{t=0}} = \sqrt{\mathrm{N^n e^{-N}/n!}} \tag{19}$$

and N is defined by $\bar{E} = (N + \frac{1}{2})\hbar\omega$.

4 What is wrong?

So? What is wrong with quantum mechanics? We have just given a complete prescription of how to set up a PDE corresponding to any classical Hamiltonian system, there is a clear interpretation of what it means and of exactly what the probability distribution of any measurement will be, and one merely has to solve it.

In fact, there are at least four things wrong, which we will now list. (Each one of these four problems will be discussed later in a seperate section devoted to it.)

1. Firstly, there is the matter of **nonlocality.** Notice that the system described by Ψ may consist of many particles, possibly far apart. But, there is only one wavefunction Ψ, and during the measurement process, it is supposed to collapse instantly. This might be thought of as somehow transmitting information faster than light. (And this still happens even in Dirac's relativistic replacement of Schrödinger's equations!)

2. Next, there is the matter of "**measurement**." Just what is a "measurement?" The whole lack of determinism, unitarity, and time reversal symmetry in quantum mechanics is happening precisely when a measurement is made, so this notion is a very important one, but we

[5] The limiting case when $\omega \to 0$ is the free particle.

haven't even got a definition. Furthermore, if one tries to look at any measurement process in detail, things get elusive. Just when did that projection/collapse process happen? You can follow the chain of system interacting with measuring device further and further along until one is actually applying Schrödinger's equation to the brain of the human supervising the measurement, and at no point in this process is anything nonunitary going on... Finally, there is the paradox of "Schrödinger's cat." You have a quantum system which is in a superposition of two states, with squared amplitudes $\approx 1/2$ for each. A device measures it to find out which state it is in, and depending on the result, does or does not drop a 10-ton weight on a cat. Now due to the linearity of Schrödinger's equation, the world afterward will be in a quantum linear superposition of two sorts of states, one in which the weight fell, the other in which it didn't. This contradicts the experimental evidence of our daily lives, in which we never see such superpositions!

3. Also, there is the matter of **nonHamiltonian systems**. It is often convenient, in classical mechanics, to put nonHamiltonian terms, such as "frictional dissipation" or "externally applied forcing," into the system. Suppose that the force of friction depends linearly on the generalized velocities:

$$f_a^{\text{friction}} = -\sum \mu_{ab}\dot{q}_b \tag{20}$$

where the friction coefficient is $\mu_{ab} = \mu_{ba}$. Define the nonnegative quadratic "dissipation function"

$$\Phi = \sum_a \sum_b \dot{q}_a \mu_{ab} \dot{q}_b \tag{21}$$

and Lagrange's equation with friction and with externally applied forces becomes

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_r}\right) = \frac{\partial L}{\partial q_r} - \frac{1}{2}\frac{\partial \Phi}{\partial \dot{q}_r} + f_r^{\text{extern}}(t). \tag{22}$$

The rate of increase in the system's energy is

$$\dot{H} = -\Phi + \sum_r \dot{q}_r f_r^{\text{extern}}(t), \tag{23}$$

and of phase space volume is

$$\frac{d \ln \text{vol}}{dt} = -\sum_k \frac{\mu_{kk}}{m_k}. \tag{24}$$

In classical mechanics with friction, we no longer have time reversal symmetry or conservation of phase space volume, energy, or momentum, but we still do have a differential equation (EQ 22) describing the motion[6]. However, quantum mechanics as described so far has given us no way to deal with such nonHamiltonian systems[7]. Ways to do this will be discussed in §**??**.

4. Finally, there is the matter of **computation**. Schrödinger's equation is a partial differential equation in $\mathbf{R}^d$ where there are d positional degrees of freedom. Again, the use of naive techniques to solve it (a grid with 10^d points?) would seem to be computationally infeasible – exponential growth of computational effort with polynomial growth of particle number. This is unacceptably poor performance, for any but the simplest problems. In fact, recent work by Peter Shor proves that (under reasonable assumptions) there *cannot* be a way to simulate orthodox quantum mechanics efficiently. This will be discussed in §**??**.

5 Other ways to look at (linear nonrelativistic) quantum mechanics

I have decided to stay away from relativistic quantum mechanics because it only seems to add complexity to the formulation without any great increase in understanding or scope.

Time independence. In practice, the two most common scenarios in which one uses the Schrödinger equation are not the full time evolution equation above.

First, one is often interested in the energy levels E (energy eigenvalues) of a bound system in which the Hamiltonian does not depend explicitly on t. A wave function that is an energy ($= E$) eigenvalue has known time dependence $e^{-iEt/\hbar}$ which factor may be removed from the definition of Ψ, yielding the time independent Schrödinger equation

$$H(q_r, i\hbar\frac{\partial}{\partial q_r})\Psi = E\Psi. \tag{26}$$

Second, one is often interested in scattering problems, in which there is a continuously impinging incoming plane wave and a fixed potential... these may also be phrased as a time independent problem, but with appropriate "radiation" boundary conditions.

A **variational principle** obeyed by Schrödinger's time independent equation is that $E = \bar{H} = \int \Psi^* H \Psi d^d q_r$ is minimized by choice of eigenfunction Ψ_E, subject to the constraint that Ψ has to be orthogonal to Ψ's with lesser values of E. This is simply Courant's variational principle for eigenvalues.

A variational principle which is less well known is obeyed by the full time dependent Schrödinger equation with the Hamiltonian (EQ 10): with the complex lagrangian density

$$L(\Psi, \dot{\Psi}, t) = i\hbar\Psi^*\dot{\Psi} - \Psi^* U \Psi \tag{27}$$
$$+ \sum_k \frac{1}{2m_k}(i\hbar\frac{\partial}{\partial q_k} + V_k)\Psi^* \cdot (i\hbar\frac{\partial}{\partial q_k} + V_k)\Psi,$$

[6]There is even a least action principle, obtained by adding the history-dependent term

$$\frac{1}{2}\int^t \Phi(t)dt + \sum_r q_r f_r^{\text{extern}}(t) \tag{25}$$

to the Lagrangian.

[7]Admittedly, one could in principle proceed by including all (external) systems, that are responsible for the friction or the driving force, inside "the" system. But from the standpoint of computation, this is unacceptable because it will make us include the entire universe, or at least an enormous amount of little-understood details.

$\delta \int \int L d^d q_r dt = 0$. Note that in this formulation, if the "generalized position" variables (of which there are an uncountably infinite number) are Ψ at every point of space, then the "generalized momenta" are $i\hbar\Psi^*$. Since this Lagrangian density is complex valued, this is really two real variational principles which must both be satisfied.

Von Neumann prefered the hermitian "density matrix"

$$\rho(q_r, u_r; t) = \Psi(q_r, t)\Psi^*(u_r, t) \tag{28}$$

to the wave function $\Psi(q_r, t)$; this is a function of *two* sets of position coordinates (or in the more abstract Hilbert space view, of two general Hilbertian indices). Of course when these two coordinate sets are the same ("diagonal entries of the density matrix"), this is just the probability density function $\Psi^*\Psi$. If you know ρ, you can always recover $\Psi(q_r, t)$ by integrating over one of the "matrix indices" and renormalizing the resulting function of one index. But this method of recovering Ψ will only work if ρ is a "pure" density function *of a wavefunction*, i.e. a projection operator, i.e. of rank 1. More generally, ρ could represent a wavefunction mixture, in which case $\rho = \sum_n p_n \rho_n$, where p_n is the probability that the true pure density matrix is ρ_n. In that case, the individual components Ψ_n of the mixture are recoverable (except for overall phase factors!) since they are the eigenvectors of ρ, and[8] the p_n are the corresponding eigenvalues. All the eigenvalues λ of either pure or mixed density matrices ρ obey $0 \le \lambda \le 1$. If ρ is pure $\rho^2 = \rho$ and all the λ are 0 except for one, which is equal to $\mathrm{Trace}(\rho) = 1$. If ρ is mixed, more than one eigenvalue is nonzero, and $\mathrm{Trace}(\rho^2) < 1 = \mathrm{Trace}(\rho)$ and $\rho^2 \neq \rho$. Schrödinger's equation rewritten in terms of the density matrix instead of Ψ is **Von Neumann's evolution equation**

$$\frac{\partial \rho}{\partial t} = -\frac{i}{\hbar}(\hat{H}\rho - \rho\hat{H}). \tag{29}$$

This time evolution equation works for either pure or mixed ρ's and preserves purity. Time evolution is abstractly a unitary similarity transformation[9]. The expectation value $\bar{Q}$ corresponding to a Hermitian operator $\hat{Q}$ is $\mathrm{Trace}(\rho(t)\hat{Q}) = \mathrm{Trace}(\hat{Q}\rho(t))$. The measurement (of A) operation corresponds to picking one of the diagonal elements of ρ – expressed in the eigenbasis of $\hat{A}$ – with probability proportional to itself, and then zeroing all the other entries of ρ. In more detail: Zeroing the off-diagonal entries of ρ reduces matters to a classical probability distribution, i.e. a wavefunction "mixture." (Note, in the random phase view of Ψ mixtures, any off-diagonal element of the density matrix would be, in expectation over the random phases, zero. Although any Hermitian matrix may be reduced to the diagonal matrix of its eigenvalues by a unitary similarity transformation, the reduction to *this* diagonal matrix, whose elements are not in general ρ's eigenvalues, is not a similarity transformation.) Picking one of the diagonal entries is sampling from this distribution. We then have to normalize by replacing the selected diagonal entry by 1 so that $\mathrm{Trace}\rho = \int |\Psi|^2 \mathrm{d}x = 1$ is maintained.

The greatest virtue of Von Neumann's representation is that it is a natural way to represent mixtures. In the Ψ representation, one had to know all the wavefunctions that were being mixed, and keep track of them as the mixture evolved, so that one could later assign them random phase factors. Since there might have been an infinite number of wavefunctions in the mixture, this is a lot to remember, and in fact, we are really remembering not merely a function (Ψ) of one position variable (q) but in fact also of a second variable which tells us which element of the mixture it is. But the ρ representation, which is explicitly a function of two variables, no random phases and no remembering extra quantities are needed and conversion to a mixture merely consists of zeroing the off-diagonal elements. Evolution might later cause the mixture's ρ to become nondiagonal in the basis you are using, but it is still a mixture, and one does not need to keep track of its individual components.

The greatest flaw of von Neumann's representation is the simple fact that dealing with infinite matrices is something that most physicists are unfamiliar with. Therefore, very few calculations have actually been carried through using ρ. In cases in which Ψ only depends on a discrete variable, so that ρ is a finite dimensional matrix, one is much more comfortable.

Heisenberg prefered to view quantum mechanics in such a way that the states Ψ (or ρ) remain fixed, but the operators corresponding to the observables (which, in the usual picture, were time invariant) evolve. In that case, to make things equivalent to the usual picture, the time evolution of an observable $\hat{Q}$ is abstractly $\hat{Q}(t) = e^{itH/\hbar}\hat{Q}(0)e^{-itH/\hbar}$, or equivalently, $\hat{Q}(t)$ obeys the "Heisenberg equation" (actually due to Dirac!)

$$\frac{\partial \hat{Q}}{\partial t} = \frac{i}{\hbar}(\hat{H}\hat{Q} - \hat{Q}\hat{H}). \tag{30}$$

which is the same as the von Neumann equation (EQ 29) with $\hat{Q}$ in place of ρ and the sign of the right hand sign reversed.

Shannon introduced the concept of **information theoretic entropy**, according to which the entropy of a system in a mixture of pure states, state n occuring with probability p_n, is

$$S_{\mathrm{info}} = -\sum_n p_n \log_2 p_n \tag{31}$$

bits. The **thermodynamic entropy** is $k_B \ln 2$ times this, where $k_B = 1.380658(12) \times 10^{-23}$Joules/Kelvin is

[8] Actually, this will only work if the Ψ_n are orthogonal. Otherwise, a set of orthogonal Ψ_n indeed will be recovered via this prescription, such that their p-mixture would have yielded ρ – but it won't be the Ψ's you thought. This distinction is immaterial for most purposes, since both the new and the old mixtures will always yield the same probability for any event. One is often concerned with the "thermal mixture" case where Ψ_n are the energy eigenfunctions, which *are* orthogonal, and $p_n = \exp(-\beta E_n)$ where $\beta = 1/(k_B T)$ and E_n is the energy eigenvalue corresponding to Ψ_n.

[9] A "similarity transformation" is a mapping from a matrix A to $Q^{-1}AQ$. It is "unitary" if Q is.

Boltzmann's constant. The information theoretic entropy of a wave function mixture with density matrix ρ is therefore expressible as above in terms of the eigenvalues p_n of ρ, or abstractly as Trace($\rho \ln \rho$). (Indeed all the thermodynamic quantities are expressible in terms of ρ.) Entropy measures our ignorance of the true state of the system, and this ignorance can only increase, leading to the law that entropy is nondecreasing.

Shannon entropy is also useful for measuring information transfer between different parts of a system. Suppose that in an appropriate basis, some variables pertain to part A of a two part physical system, and the rest to part B. Then the trace over the part of ρ's diagonal indexed by basis variables of type A, describes the part of the entropy attributable to part A of the system, and the rest of the trace describes the part of the entropy in part B of the system. As ρ evolves with time, one can track these two entropy values.

Wigner (1932) prefered the so called "Wigner transform" of ρ:

$$W(q_r, p_r; t) = \frac{1}{h^d} \int e^{i \sum_r p_r y_r / \hbar} \rho(q_r - y_r/2, q_r + y_r/2; t)\ d^d y_r \tag{32}$$

Since this is a Fourier transform along the antidiagonals of the density matrix, it is invertible and ρ is recoverable from W. Wigner's W is a *real* valued function of both position and momentum variables (this reality is a direct consequence of the Hermitian symmetry of ρ causing cancellation of the imaginary part) and $\int W(q_r, p_r) d^d p_r = |\Psi|^2$ is the probability density in position space, and $\int W(q_r, p_r) d^d q_r$ is the probability density in momentum space. Thus if one is so lucky as to have a W which happens to everywhere be nonnegative, then $W(q_r, p_r; t)$ may then be interpreted as a probability density in phase space. In the event that W is somewhere negative, one cannot make this interpretation, but W will still behave in many ways like a probability density, e.g. $\int\int W d^d p d^d q = 1$, and W is often highly oscillatory in regions where it goes negative, such that the mean value nearby is still nonnegative.

Wigner's formulation was further studied by Moyal 1949. The time evolution of Wigner's W is governed by Moyal's equation

$$\frac{\partial}{\partial t} W(p_r, q_r; t) = \tag{33}$$
$$\frac{2}{\hbar} \sin\left(\frac{\hbar}{2} \sum_r \frac{\partial}{\partial p_r^{(W)}} \frac{\partial}{\partial q_r^{(H)}} - \frac{\partial}{\partial q_r^{(W)}} \frac{\partial}{\partial p_r^{(H)}} \right)$$
$$H(p_r, q_r) W(p, q; t).$$

Here the sine of a bounded differential operator may be defined via the Taylor series for sine. Moyal's equation is only meaningful if the appropriate sums converge and if H is infinitely differentiable. Notation: the differentiations with respect to "$p^{(H)}$," for example, are only supposed to be applied to H.

Notice that in the classical limit $\hbar \to 0$ that this is the classical equation $\frac{\partial W}{\partial t} = \{H, W\}$ satisfied by a probability density function evolving under the Hamiltonian flow, up to additive error terms which vanish like $O(\hbar^2)$. Moyal's equation is in general a linear partial differential equation of infinite order, but if $H(p_r, q_r)$ is a polynomial in its arguments, then the Taylor series terminates and the PDE has finite order. In particular, in the event that the Hamiltonian $H(p_r, q_r)$ is quadratic in its arguments, as arises in "linear systems," then all partial derivatives of order ≥ 3 vanish and in fact the classical equation is obeyed *exactly*, leading to vast simplifications.

Wigner noted that the quantum expectation value of any function f of the p_r's only, or of any function of the q_r's only, or (since expectation values are linear) any sum of two such functions (for example, the total energy) was $\int\int W(p_r, q_r) f d^d p d^d q$. This also works for functions of the p_r's and q_r's, such as angular momentum, which correspond unambiguously to quantum operators because the p_a's are only combined with q_b's that commute with them. But is does not necessarily work for general functions of both the p_r's and the q_r's. Moyal, however, found out how to find the expectation value of such observables. You write the classical observable $f_{\text{classical}}$ in terms of the p_r and q_r combined via Poisson brackets, and then to convert it to a form suitable for Wignerian quantum mechanics, you replace each Poisson bracket $\{R, G\}$ by

$$\frac{2}{\hbar} \sin\left(\frac{\hbar}{2} \sum_r \frac{\partial}{\partial p_r^{(G)}} \frac{\partial}{\partial q_r^{(R)}} - \frac{\partial}{\partial p_r^{(R)}} \frac{\partial}{\partial q_r^{(G)}} \right) R(p, q) G(p, q) \tag{34}$$

and then the expectation value of f at time t will be given by $\bar{f} = \int\int W(p_r, q_r; t) f_{\text{quantum}} d^d p_r d^d q_r$.

The operation of measurement of A in Wignerian quantum mechanics can of course be expressed by transforming back to ρ land, but there ought to be a slick statement of what is going on, expressed totally in W space. No such statement has ever been worked out, but I conjecture the correct result is as follows:

Conjecture: Consider the region $R_{A=x}$ of phase space in which the A-function (of p_r, q_r, and W and its derivatives) assumes value x. With probability $\int_{R_{A=x}} W d^d p_r d^d q_r$, the result of the measurement will be x and the effect upon the quantum state will be to zero W outside $R_{A=x}$ and leave it invariant in $R_{A=x}$. Finally one has to renormalize W so its integral remains 1. (Note, this may entail making W a "generalized function.")

This is actually correct for many A's but I haven't proven it in general, could be wrong.

Because Wigner's W is just a linear transformation of von Neumann's ρ, mixtures are treatable just as naturally using W as by using ρ. Indeed, when you mix large numbers of uncorrelated W's, the oscillatory behaviors tend to smooth out and W gets more like a "classical" smooth, nonnegative, probability density.

Since the Wigner equation is linear, it may be solved via Green's functions. For example, the solution of the harmonic oscillator is simply a bodily rotation in the phase plane of $W(x, p)$, and of the free particle is simply a continuous shear. More generally, motion will depend locally on derivatives of W and will not be so simple to describe.

Presumably Feynman's path integral formulation may also be rewritten in Wigner land. ???Due to the all-real nature of W, this looks very promising???

Finally, Wigner proved (see Takabayashi 1954) that W obeys the inequality, and equality

$$|W| \leq (2/h)^d, \quad \int\int W^2\, d^d p d^d q = h^{-d}; \tag{35}$$

the former may be thought of as a new form of Planck's uncertainty principle. Assumptions behind this??? This proves oscillations are not too bad, very important, better re-examine???

Bohm (1952) suggested an interesting alternative to the Copenhagen interpretation of the Schrödinger equation, involving "hidden variables." Specifically, in Bohm's formulation, in addition to the wavefunction Ψ, which obeys Schrödinger's equation, the particles also have unique locations, given by the point $q_r(t)$ in $\mathbf{R}^d$. This point moves according to the law

$$\dot{q_r} = \frac{i\hbar}{2m_r} \frac{\Psi \frac{\partial}{\partial q_r}\Psi^* - \Psi^* \frac{\partial}{\partial q_r}\Psi}{|\Psi|^2} \tag{36}$$

i.e. $\dot{q_r}$ is the probability current divided by the probability density. Equivalently, under the assumption that H has the usual form (EQ 10), Bohm showed (see (EQ 7) and (EQ 11) for why) that the trajectory $q_r(t)$ is in fact the classical trajectory given by Hamilton's equations in the classical potential *plus* an extra "quantum potential"

$$U_{\text{quantum}}(q_r) = \frac{-\hbar^2}{2} \frac{\sum_k \frac{1}{m_k}\frac{\partial^2}{\partial q_k^2}|\Psi|}{|\Psi|}. \tag{37}$$

Bohm regards $\Psi(q_r)$ as a "quantum field," analagous to the electromagnetic field, and Schrödinger's equation as a field equation analagous to Maxwell's laws, and then the particles move totally classically except for the fact that the quantum field exerts "quantum forces" on the particles. In the classical limit $\hbar \to 0$, these quantum forces vanish proportionally to $\hbar^2$.

Bohm's mechanics contain the usual formulation, because if the particles's locations are initially assumed to be unknown but distributed according to the probability density $|\Psi|^2$, then, thanks to (EQ 36), this density will evolve via the correct probability current and will remain $|\Psi|^2$. However, it is also possible to start with a different probability density and watch it evolve according to Bohm's law, in which case Bohm's mechanics will be different from Schrödinger's. Bohm speculated that this extra capability might later be useful in attempts to extend quantum mechanics into regimes of ultrashort lengths ($< 10^{-15}$ m)... but that never happened.

The trajectories of Bohm particles moving nearly entirely freely are nice smooth curves, but during interactions these trajectories can be exceedingly complicated and the quantum forces can be quite violently fluctuating.

Interesting theorems in Bohm's mechanics include: When Ψ is standing waves in a box with infinite walls: the particles stay still. When Ψ is a solution of the time independent Schrödinger equation which is an eigenfunction of some observable, *every trajectory* has the same value of that observable (which is a constant of the motion).

One may ask: what about measurement in Bohm's mechanics? Bohm takes the cheap way out by

1. denying the nonunitary collapse process in the usual quantum mechanics and assuming that somehow, a full quantum mechanical description of the measuring apparatus and the test system would evolve in the correct way

2. Noting that his mechanics (if it also is applied to the measuring apparatus) is then equivalent to the usual one

and hence the results of any measurement will be the same in Bohm's and the usual formulations. Although particles have precise hidden values of momentum and position in Bohm's theory, and *everything* in Bohm's theory is totally deterministic, these variables will remain eternally hidden! For example in any attempt to measure both position and momentum very accurately, the position measurement will cause violent and unpredictable (if the hidden state of the measurement apparatus were not precisely known) quantum forces, which will alter the momentum enough to save the uncertainty principle[10].

Nobody has ever generalized Bohm's mechanics to relativistic quantum mechanics, according to Albert 1994. This is not suprising since I claim that Kochen and Specker's 1967 proof (see §**??**) shows that it is impossible to do so, at least while still satisfying Bohm's desiderata for what a hidden variable theory should do[11] Hence, Bohm's mechanics is purely of pedagogical interest and presumably cannot have fundamental significance.

Gutzwiller (1982) found an interesting relationship between quantum and classical mechanics. This is a formula expressing the energy eigenvalues of the time independent Schrödinger equation for a particle in an arbitrarily shaped box in terms of a sum over all periodic classical trajectories. It appears to be the same as "Selberg's trace formula" (Selberg 1959) expressing the eigenvalues of the laplacian on a compact Riemannian manifold of constant negative curvature in terms of its closed geodesics, and leading to various amazing zeta function identities. For particles not living in such a universe, or in a potential which is not everywhere either 0 or ∞, Gutzwiller's formula is only approximate.

[10] Why do the violent motions of Bohm's charged particles not produce radiation? Bohm would similarly take the cheap way out here and argue that were his formalism also applied to the electromagnetic field modes, it would be equivalent to the usual one.

[11] More precisely, any theory of spin having the algebraic behavior obeyed by genuine quantum spin operators, cannot be a deterministic hidden variable theory... This proof does not need to make any "locality" assumptions about the hidden variables.

This whole development seems to be of no practical use, since finding even one periodic orbit (much less all of them) is usually exceedingly difficult. And so long as Gutzwiller's formula remains only approximate, it doesn't seem to have much mathematical interest either.

Nelson (1966) found a way to view quantum mechanics as a stochastic process analagous to Brownian motion[12]. (Incidentally, this development *is* generalizable to relativistic quantum mechanics by considering Brownian motion on higher dimensional manifolds also incorporating "spin" degrees of freedom, see papers by Dohrn, Guerra et al., and Dankel in the bibliography.) The process has to be described in position space only, since the momentum and velocity of a particle undergoing Brownian motion is not defined. For simplicity we will specialize here to Cartesian position coordinates x_r, although Nelson 1985 carries everything through using generalized coordinates, two indexed mass and diffusion tensors, etc, on an arbitrary Riemannian manifold.

Consider the Markoff process

$$dx_r(t) = b_r^+(x_k(t), t)dt + \theta_r dw_r^+(t) \tag{38}$$

where θ_r is a positive "diffusion coefficient" and where w_r^+ is a d-dimensional Wiener process with $\overline{dw_j^+(t)} = 0$ and $\overline{dw_j^+(t_1)\, dw_k^+(t_2)} = 2\delta_{jk}\delta(t_1 - t_2)$ which is independent of the $x_k(s)$ with $s \le t$. This process is asymmetrical in time. Here b is a "bias" or "drift" term, and dw is a "white noise" stochastic term. The same process under time reversal is

$$dx_r(t) = b_r^-(x_k(t), t)dt + \theta_r dw_r^-(t) \tag{39}$$

where w_r^- is a d-dimensional Wiener process with the same properties as w_r^+ except that it is independent of the $x_k(s)$ with $s \ge t$. Here b_r^+ is the "mean forward velocity" and b_r^- the "mean backward velocity." The "drift velocity" v_r is $v_r = (b_r^+ + b_r^-)/2$, and the "osmotic velocity" is $u_r = (b_r^+ - b_r^-)/2$. The importance of these latter two "velocities" is that the probability density $\rho(x_r)$ obeys the continuity equation

$$\frac{\partial \rho}{\partial t} = -\sum_k \frac{\partial}{\partial x_k}(v_k \rho) \tag{40}$$

while

$$u_r = \theta_r \frac{\partial}{\partial x_r} \ln \rho \tag{41}$$

is a velocity corresponding to mean diffusive flux. Nelson also defines a notion of "mean acceleration" $a(t)$, namely $a_r(t) = \frac{1}{2}(\mathrm{D}^+\mathrm{D}^- + \mathrm{D}^-\mathrm{D}^+)\mathrm{x}_r(t)$, where the operators D^+ and D^- are defined by $D^\pm x(t) = \lim_{\epsilon\to 0\pm} \overline{(x(t+\epsilon) - x(t))/\epsilon}$.

[12] The first such stochastic view was advanced by Fürth 1933, who used his analogy to provide a different than usual proof of Heisenberg's uncertainty theorem, and to find a similar uncertainty theorem for ordinary diffusion processes. Other pre-Nelson developments, such as by Imre Fenyes and Friedrich Bopp, are discussed in chapter 9 of Jammer 1974; see also Nelson 1985.

Nelson's theorem is that this Markoff process is equivalent to Schrödinger's equation under the following transformation. In the modulus-phase decomposition (EQ 11) (EQ 12) of Schrödinger's equation, use $\rho = |\Psi|^2$, $\Psi = \sqrt{\rho}e^{iS/\hbar}$, and with the usual Hamiltonian (EQ 10) in orthogonal coordinates, define v_r by

$$m_r v_r + V_r = \partial S/\partial x_r, \tag{42}$$

u_r by (EQ 41), and let $\theta_r = \hbar/(2m_r)$. The fact that mean forward and backward velocities in the Markoff process are $b_r^\pm = v_r \pm u_r$ causes the mean acceleration to be identical with the classical acceleration in that potential.

To recapitulate, we have a point $x_r(t)$ in $\mathbf{R}^d$ executing a frictionless Brownian motion with diffusion coefficient $\theta_r = \hbar/(2m_r)$ and a certain mean forward velocity function b_r^+. In Nelson's words

> The picture which emerges is the following. If we have... a hydrogen atom in the ground state, the electron is in dynamical equilibrium between the random force causing the Brownian motion and the attractive Coulomb force of the nucleus. Its trajectory is very irregular. Most of the time the electron is near the nucleus, sometimes it goes farther away, but it always shows a general tendency to move toward the nucleus, and this is true no matter which direction we take for time. [There are also...] other states of dynamical equilibrium, at the usual discrete energy levels of the atom.

However, the time evolution of b_r^+ has to be found by solving the Schrödinger equation, or equivalently by solving Nelson's equations

$$\frac{\partial u_r}{\partial t} = -\theta_r \frac{\partial}{\partial x_r} \sum_k \frac{\partial}{\partial x_k} v_k - \frac{\partial}{\partial x_r} \sum_k v_k u_k \tag{43}$$

$$\frac{\partial v_r}{\partial t} = a_r - \sum_k v_k \frac{\partial}{\partial x_k} v_k + \sum_k u_k \frac{\partial}{\partial x_k} u_k + \theta_r \sum_k \frac{\partial^2}{\partial x_r^2} u_r \tag{44}$$

for the time evolution of the u_r and v_r velocity fields (in terms of the classical acceleration a_r of a particle with position x_r and velocity v_r) and then $b_r^+ = v_r + u_r$. In other words, Nelson's theory, just like Bohm's, is still inherently nonlocal and depends on "quantum fields" evolving according to their own peculiar "field equations;" the only difference is that while Bohm's particles move according to Newton's laws with an extra "quantum force" supplied by the Ψ field, Nelson's particles move according to Newton's laws with an added Brownian motion whose mean forward velocity is determined by the Ψ field. In fact, there is a more general, if mathematically uglier, formulation due to Davidson 1979, in which Schrödinger's equation is shown to arise from a Brownian-like process with a diffusion coefficient κ times as large as Nelson's, for any positive real number κ. When $\kappa = 1$ Davidson's formulation simplifies considerably and reduces to Nelson's. In the limit $\kappa \to 0$, Davidson's formulation (at least, if one ignores issues of mathematical rigor) turns into Bohm's. Note that Nelson's formulation is stochastic (thanks to the Brownian motion) but Bohm's is purely deterministic.

But, Nelson's formulation seemingly has a tremendous advantage over the usual formulation, which Nelson never remarked upon, although Yasue 1978 did. This is that it is easy (simply alter a_r appropriately) to put non-Hamiltonian forces into the picture and then see what happens. Indeed, we can ignore the particles and their Brownian motions entirely and forget all about stochastic processes, and simply solve Nelson's equations (EQ 43 and 44) for the time evolution of the $u_k(x_r)$ and $v_k(x_r)$, now regarded as the fundamental objects instead of $\Psi(x_r)$. And that[13] is a formulation of quantum mechanics which depends only on *forces*, and not on *potentials*. Unfortunately, as we will later see, this nonHamiltonian capability is more a mathematical illusion than a reality.

Yasue (1978) applied this procedure to derive a version of Schrödinger's equations[14] allowing linear friction and conservative external drive forces:

$$i\hbar\frac{\partial\Psi}{\partial t} = (H(x_r, -i\hbar\frac{\partial}{\partial x_r}, t) + \sum_k \frac{\mu_{kk}\hbar}{m_k}\arg\Psi - U_k^{\mathrm{extern}}(x_r,t))\Psi \quad (45)$$

Note that in the presence of friction the Yasue equation is nonlinear and does not exhibit a time reversal symmetry, unlike the Schrödinger equation. $|\Psi|^2$ is still a probability density, though. This equation will be discussed later as "Kostin's equation" – and our critique will make it clear this is not the right way to treat friction.

The modulus and phase version of Yasue's equation is

$$\frac{\partial S}{\partial t}+H(q_r,\frac{\partial S}{\partial q_r},t)+\sum_k\frac{\mu_{kk}}{m_k}S-U_k^{\mathrm{extern}} = \frac{\hbar^2}{2m|\Psi|}\sum_r\frac{\partial^2|\Psi|}{\partial q_r^2} \quad (46)$$

(and the modulus equation is the same as (EQ 12)).

In the classical limit $\hbar\to 0$, the right hand side vanishes and we obtain a version of the Hamilton-Jacobi equation that has external forces and friction.

Nelson (1985, starting page 124) recently pronounced himself to be philosophically unsatisfied with his own creation, because of nonlocality. Specifically, if you perturb the Hamiltonian affecting one particle, then some other particle, independent of the first one, will instantly have its trajectory affected. Of course, Schrödinger's equations are also nonlocal in the sense that the multiparticle wavefunction will be affected everywhere instantly, *but* no measurable effect will happen to the second particle and in particular its density $|\Psi|^2$ will be unaffected. In Nelson's mechanics presumably also the effect on the second particle is not measurable, but that is not enough to satisfy Nelson, who wants the trajectory to be unaffected whether you know it or not. Nelson speculates that an answer may lie in the use of nonMarkoffian stochastic processes, or by using a relativistic version of his mechanics with quantized (that is, Nelsonized) fields.

Feynman (1948; see also Feynman-Hibbs 1965) reformulated Schrödinger's equation using "propagators" and "path integrals." Since Schrödinger's equation is linear, it may be solved with the aid of Green's functions:

$$\Psi(q^{(2)},t^{(2)}) = \int K_H(q^{(2)},t^{(2)},q^{(1)},t^{(1)})\Psi(q^{(1)},t^{(1)})d^dq^{(1)} \quad (47)$$

where the Green's function K, also called the "propagator," depends on the particular Hamiltonian H in our problem, hence the subscript H. For example for the free particle Hamiltonian $H=\hat{p}^2/2m$ the propagator is

$$K_{\mathrm{free}}(q^{(2)},t^{(2)},q^{(1)},t^{(1)}) = \sqrt{\frac{m}{2\pi i\hbar(t^{(2)}-t^{(1)})}}\exp\frac{im\sum_r(q_r^{(2)}-q_r^{(1)})^2}{2\hbar(t^{(2)}-t^{(1)})}. \quad (48)$$

Although certainly K_H exists, it is not generally easy or even possible to find it in closed form. That is the whole problem. In the particular case of a free particle in one Cartesian coordinate x, the propagator is (with $t=t_2-t_1$)

$$K_{\mathrm{free}}(2,1) = (\frac{m}{ith})^{1/2}\exp\frac{i(x_2-x_1)^2m}{2t\hbar} \quad (49)$$

In the particular case of a one dimensional harmonic oscillator where $2H(x,p)=p^2/m+m\omega^2x^2$ it is

$$K_{\mathrm{harmonic}}(2,1) = (\frac{m\omega}{ih\sin\omega t})^{1/2} \times\exp\left\{\frac{im\omega}{2\hbar\sin\omega t}[(x_2^2+x_1^2)\cos\omega t-2x_1x_2]\right\} \quad (50)$$

In the general case of a particle in a general scalar potential $V(x)$ in cartesian coordinates $x\in\mathbf{R}^d$, Feynman was able to express the propagator K_V in terms of the free-particle propagator K_0 via an infinite series (which he called the "Born approximation," although it was known to Volterra)

$$K_H(B,A) = K_0(B,A)+\sum_{n\geq 1}(\frac{-i}{\hbar})^nK^{(n)}(B,A) \quad (51)$$

where

$$K^{(n)}(B,A) = \int\int...\int K_0(B,x_n)V(x_n) K_0(x_n,x_{n-1})V(x_{n-1})...K_0(x_2,x_1)V(x_1)K_0(x_1,A)d\tau_n...d\tau_1 \quad (52)$$

[13]Note: if the forces are Hamiltonian, then an examination of Nelson's equations shows that if the "drift momentum" $m_rv_r+V_r$ is initially a gradient (of S) then it will stay a gradient forever. (Also, a different derivation, by Guerra, of Nelson's picture, from a stochastic variational principle, shows that it should be a gradient.) However, with nonHamiltonian forces, the drift momentum will generally not stay a gradient, and thus $\arg\Psi$ will not be defined for $t>0$, and hence we are in fact *forced* to stay in the realm of u_r and v_r, since no Ψ will generally be definable. But note u_r remains a gradient for all time, regardless of Hamiltonicity. So we can get halfway there by regarding v_r and ρ as the fundamental objects.

[14]Yasue was only able to get back to the Ψ realm because the force field of *linear friction* is a gradient, if $\vec{v}$ is, so that $\vec{v}$ remains a gradient for all time.

where $d\tau_j$ means $dx_j dt_j$ and the t is integrated over the interval $t_A \le t \le t_B$, and where the x_j are each integrated throughout $\mathbf{R}^d$. This is thus a $(d+1)n$-dimensional integral with a known integrand. The series (EQ 51) will converge if the $K^{(n)}$ grow sufficiently more slowly than $\hbar^n$, i.e. if $t_B - t_A$ times the "typical energy" $V(x)$ is smaller than $\hbar$. (In particular, convergence will occur for all sufficiently short times.) Due to cancellations in the integrals, the magnitude of this "typical" energy is fortunately less than one might at first think, so that, heuristically, we ought to be able to go forward in time an amount substantially *greater* than the order given by the energy-time uncertainty principle.

What was more interesting was Feynman's expression of the propagator as a "path integral:"

$$K(2,1) = \int e^{iS/\hbar} \mathcal{D}\frac{x}{A} \tag{53}$$

Here the $\mathcal{D}$ denotes integration over "all possible paths" $x(t)$, $t_1 \le t \le t_2$, from x_1 to x_2, and S is the classical action. What does this mean? Feynman defined such path integrals via a limiting process:

$$\int f[x(t), \dot{x}(t)] \mathcal{D}\frac{x}{A} = \lim_{N\to\infty} \int\int \dots \int \tag{54}$$
$$f[x_{\text{lin}}^{(N)}(t), \dot{x}_{\text{lin}}^{(N)}(t)] \frac{dx_1}{A_N}\frac{dx_2}{A_N}\frac{dx_N}{A_N}$$

where f is a functional depending on the entire path $x(t)$ for $t_1 \le t \le t_2$ and where $x_{\text{lin}}^{(n)}(t)$ is the piecewise linear function that interpolates the points $x_{\text{lin}}^{(n)}(t_1 + \frac{(t_2-t_1)j}{N+1}) = x_j$ for $j = 1, 2, ..., N$ and where x_0 is the starting point and x_{N+1} is the endpoint. In fact, Feynman's particular path integral involves an action function S in the integrand which is a function, not of x and $\dot{x}$, but in fact of these quantities on the entire subpath previous to the current moment in time, but the same limiting process applies. If f depends also on kth order time derivatives of x with $k \ge 2$ then kth order interpolation needs to be used instead of linear. With Feynman's definition of path integrals, the magic normalizing constant A needs to be specially chosen to make it all work; the value in quantum mechanics is

$$A_N = \prod_{j=1}^{d} \sqrt{\frac{2(t_2 - t_1)hi}{N m_j}}. \tag{55}$$

Crudely speaking, quantum particles traverse all possible paths from source to destination, and each path contributes a complex amplitude proportional to $e^{iS/\hbar}$ to the final wavefunction value at that destination. Feynman argues that in the classical limit $\hbar \to 0$, only paths that are "near" the classical path (i.e. of extremized action) matter, because since a slight variation of the path will change the phase enormously everywhere except near a (locally flat!) extremum of the action, hence the complex contributions of all nonclassical paths will tend to cancel out.

Another application of the path integral is to quantum statistical mechanics; as was pointed out by Von Neumann the thermodynamic mean value of C arises from the operator Trace$(\hat{C}e^{\hat{H}/(k_B T)})$. This is exactly analagous to the usual expectation value only with $1/(k_B T)$ substituted for $it/\hbar$, thus mathematically the two problems are essentially the same.

Path integrals are integrals over (infinite dimensional) function space. Only in a few cases is it known how to evaluate them in closed form; in particular for quantum "linear systems" they may always be evaluated because all the integrals that arise in the limiting process are always Gaussian. (The gaussians have parameters that are complex and which depend upon t, but this is no problem.) It is more satisfactory (since magic normalizing factors will not be needed)[15] to define such integrals by defining a probability measure on the space of all functions by means of a stochastic process, and then the path integral is an expectation value under this probability measure. For example, one could use the unit "Wiener process" (Doob 1953) to generate a probability distribution over all paths and then define "$\int_W f[x(t)]\mathcal{D}x(t)$" to be the expectation value of some functional $f[x(t)]$. Now instead of the Wiener process, I point out that one could use a Langevin process

$$dx = b(x,t)dt + \theta dw \tag{56}$$

where w is a Wiener process, θ is a diffusion coefficient, and b is the "forward velocity." If this is done, then the integrand needs to be multiplied by a correction factor ???. (Also changing Wiener processes introduces a correction factor??? Also explain how to convert from Feynman to Wiener???). This leads to a whole bunch of Feynman integral formulas instead of just one. One interesting possibility in quantum mechanics is to use a drift velocity equal to the classical velocity, and choose the diffusion coefficient in the Wiener process to be equal to Nelson's choice above (i.e. $\hbar/(2m)$, in one dimension). This causes the important near-classical paths to lie explicitly at the center of things and should improve numerical properties.

One problem with Feynman's path integral formulation of QM is that it doesn't generalize in any simple manner to handle spin (although it will generalize to spinless relativistic quantum by using the relativistically invariant action). Fixing this is very messy and involves a noncommutative "ordered operator calculus." However, the more pedestrian propagator formalism works fine in relativistic QM.

[15] Convergence issues can remain murky and the notion of "changing variables" in a path integral is hard to justify... but both these situations are improved by this suggestion, and a Monte Carlo numerical method becomes obvious.