

# Quantum Theories

Frederick J. Ernst

February 21, 2005

## Abstract

When I was an assistant professor of physics, first at Boston University and subsequently at the Illinois Institute of Technology in Chicago, I specialized in the core courses of theoretical physics; i.e., classical mechanics, electromagnetism and quantum theory. The following lecture is based upon a 1965 lecture I gave to second semester seniors, not long after I joined the faculty at I.I.T. The purpose of the lecture was to draw students' attention to some important features of quantum theory that are often misunderstood.

It is probably safe to say that 80% of students who have completed one or more years of quantum theory do not have a working knowledge of the essential ideas of the theory. I should be overly optimistic to imagine that a one-hour lecture is going to modify significantly the statistics. However, during this hour I should like to make an effort to outline the structure of quantum theory and to emphasize certain points which tend to elude students.

I should like to begin by proposing that quantum "theory" is not really a theory at all, but can be more properly described as a way of life, or at least as an "approach" rather than as a "theory." Certain, but by no means all, aspects of the *quantum approach* were implicit in the 19th century approach to mechanics and electromagnetism based upon the variational principle of Hamilton.

## 1 Hamilton's Principle of Least Action

When beginning physics students learn to solve simple mechanical problems, they employ Newton's laws directly, starting with the elusive concept of a

*free body diagram*. Most students are delighted when they learn in a more advanced class that such mechanical problems can be solved in a very different and much more elegant way using the concept of a Lagrangian

$$L := T - V, \tag{1}$$

where  $V$  is the potential energy of the mechanical system expressed in terms of *generalized coordinates*, and  $T$  is the kinetic energy of the mechanical system expressed in terms of generalized coordinates and the corresponding *generalized velocities*. In Lagrange's formulation, the equations of motion of the mechanical system (which govern the temporal behavior of the generalized coordinates) were derived using a *variational principle*, Hamilton's principle of least action,

$$\delta \int L dt = 0. \tag{2}$$

This equation says, in effect, that the action  $\int L dt$  should be a minimum. More specifically, it should not change when one varies the generalized coordinates slightly as functions of the time  $t$ .

By focussing attention upon the Lagrangian  $L$ , it was found possible to describe not only mechanical systems, but also continuous media and even electromagnetic fields. In the case of field theories, the Lagrangian was expressed in terms of a *Lagrangian density*  $\mathcal{L}$ , in terms of which Hamilton's principle became

$$\delta \int \mathcal{L} dx dy dz dt = 0. \tag{3}$$

In extending Hamilton's principle from pure mechanics to cover also field theories such as electromagnetism, the original origin of the Lagrangian as  $L := T - V$  was suppressed. The formulation of a new theory came to mean the choosing of a Lagrangian or Lagrangian density. Hamilton's principle thus became elevated to the level of an approach, within the framework of which new theories of diverse phenomena could be formulated. Severe restrictions were placed upon the choice of Lagrangians by the invariance properties which you demanded of the resulting theories. In this connection the advent of special relativity, with its requirement of Lorentz covariance, gave rise to extremely severe restrictions upon the choice of Lagrangians in formulating new theories.

The development of the action principle, with the concomitant centering of attention upon the choice of the correct Lagrangian, or equivalently, the

choice of the correct Hamiltonian,

$$H(p, q) := p\dot{q} - L(q, \dot{q}), \quad (4)$$

where the generalized momentum variable  $p$  associated with the generalized coordinate  $q$  is defined by

$$p := \frac{\partial L}{\partial \dot{q}}, \quad (5)$$

marks a philosophical transition, which has played a crucial role in the development of the quantum approach. However, before launching into a discussion of the postulates of the quantum approach, I should like to give a brief outline of the historical development of quantum theories:

### **PERIOD I: The Heroic Era**

Planck's explanation of black-body radiation	1904
Einstein's explanation of photo-electric effect	1904
Rutherford's scattering experiments	1911
Bohr model of H atom	1913
Quantum mechanics & Wave mechanics (Heisenberg & Schrödinger)	1926
Quantum theory of fields (Dirac, Heisenberg, Pauli)	1928
Discovery of neutron & positron	1931
Hitler's rise to power	1933

### **PERIOD II: The Dark Ages**

The Thousand Year Reich 1933-1945

### **PERIOD III: The Renaissance of Modern Physics**

End of world war two	1945
Lorentz covariant QED (Feynman, Schwinger, Tomonaga)	1945
Renormalization (Bethe, Dyson)	1947
Discovery of pion	1947
Strange particles	
Axiomatic field theory	1955
Dispersion relations	1955

## 2 Features of all Quantum Theories

I should now like to discuss those features which all quantum theories share, whether they be the quantum mechanics of 1926 taught in Physics 405-6 and in Physics 526-27, the quantum field theories of 1928, those of 1945 or of 1955:

All quantum theories involve entities called *state vectors* and entities called *observables*. The former are vectors in a linear vector space over a complex field, in which space a scalar product is defined. It is sometimes described as a Hilbert space, which it usually isn't. The observables are self adjoint linear transformations (operators) defined on this "Hilbert" space of state vectors. That is, when such an operator is applied to a vector belonging to the Hilbert space, it gives a vector belonging to the Hilbert space.

There are vectors in the space whose direction is not altered by the application of a given operator. These vectors are called the *eigenvectors* of the observable in question. The magnification factor to which the application of the operator gives rise is called the *eigenvalue*.

In any quantum theory the observables correspond to the things which are measurable. Furthermore, it is postulated that the only possible results of the measurement of an observable  $A$  are eigenvalues of the operator  $A$ . No other result may be obtained. In any specific case the eigenvalue spectrum may be discrete or continuous, or part discrete and part continuous. In any event, if an observable  $A$  is measured and the result of the measurement is  $\mathbf{a}$ , then, immediately after the measurement is performed, the state vector of the system must be an eigenvector of  $A$  belonging to the eigenvalue  $\mathbf{a}$ . There may be a unique such eigenvector of  $A$ , or there may be many, even infinitely many. If there are many eigenvectors of  $A$  belonging to the eigenvalue  $\mathbf{a}$ , it indicates that some other observable will have to be measured to specify completely the state vector of the system.

Suppose we measure a second observable  $B$ . If  $A$  and  $B$  have no eigenvectors in common, it is impossible to measure both  $A$  and  $B$  simultaneously, for the resulting state of the system cannot be characterized as a simultaneous eigenstate of  $A$  and  $B$ . An operator  $A$  may be measured simultaneously with an operator  $B$  if  $A$  and  $B$  commute. They then have simultaneous eigenvectors.

A complete set of commuting observables is a set of commuting observables, the measurement of which suffices to specify completely the state vector of the system. The simultaneous eigenvectors of a complete set of

commuting observables can usually be shown to constitute a basis system of vectors in the Hilbert space. An arbitrary state vector can be expanded in terms of such a basis system.

Popular basis systems in quantum mechanics (1926 variety) are energy eigenvectors (the Heisenberg representation) and the position eigenvectors (Schrödinger representation) and also the momentum eigenvectors (Momentum representation). The popular basis system in quantum field theory consists of eigenvectors of particle-number operators.

We have spoken about what happens to the system when you make a measurement, and we have delimited the possible results of such a measurement. However, we may still inquire as to the probability of getting one or another eigenvalue as the result of a measurement of the observable  $A$ . The postulate is that the probability that as a result of measuring  $A$  you will cast the system into the state with eigenvector  $\phi_a$  of  $A$ , if immediately prior to the measurement the state of the system is  $\psi$ , is

$$P_a = |(\phi_a, \psi)|^2. \quad (6)$$

In writing this I have tacitly assumed that all state vectors are normalized, which may or may not be the case actually.

All the preceding considerations take place at a given instant of time. Between observations, the system developed in a deterministic fashion. Depending upon whether you adopt the Heisenberg picture or the Schrödinger picture of the temporal evolution, either the operators will evolve in a deterministic fashion, or the state vectors will so evolve. In the latter case the law of motion of the state vector is the time-dependent Schrödinger equation. This picture is only popular in non-relativistic quantum mechanics.

### 3 The Simplest Atom

Consider the lowly hydrogen atom. It consists of one proton and one electron bound together by an electric field. But let's look a little more closely. The proton is a rather complicated structure in itself, as has been demonstrated by the electron scattering experiments of Hofstadter. Furthermore the structure of the proton is more than a matter of a smeared out charge and magnetic moment distribution. Strongly interacting particles like protons, neutrons, pi-mesons, etc., constantly come into being and disappear again, so that the proton is actually a seething mass. The electron is less complicated, for it

doesn't seem to participate in strong interactions. However, it does interact with photons (light). In fact, this is why an atom can radiate light.

A proper description of the hydrogen atom would involve strong interaction theory to handle the proton, and quantum electrodynamics to handle the interplay between light and the orbital electron. In fact, strictly speaking quantum electrodynamics is necessary to describe correctly the force between the proton and the electron.

Although quantum electrodynamics is well understood and reasonably manageable mathematically, strong interaction physics is still in pretty sad shape. Fortunately, one can describe certain aspects of atomic structure using a rather more naive quantum theory, in which the subtle interaction between proton and electron is approximated by a fixed classical Coulomb field. In this case, one can completely describe the state of the system by measuring, for example, the total momentum of the atom, the center-of-mass energy, and certain angular momentum components. The proton and electron spins are sometimes neglected too.

Using this rather naive model of the hydrogen atom, one is able to calculate energy levels quite accurately, but I believe many students draw the mistaken conclusion that the theory so formulated actually accounts for the emission of radiation by the hydrogen atom. As a matter of fact, nowhere in the theory is the photo-emission process described. Actually, there is no reason why excited atomic states should not be absolutely stable, for the theory embodies no mechanism for the emission of light. Quantum mechanics such as it is taught for example in Physics 405-6 is designed to answer only certain types of questions. This fact is usually obscured by appealing to semi-classical arguments to bolster up the quantum theory. Of course, there are pragmatic reasons for doing this, but I believe the student should be aware that there does exist a more or less correct and self-contained way of attacking the problem via quantum field theory.

## 4 The Postulates of the Quantum Approach

Finally, I should like to illustrate the postulates of the quantum approach with a trivial example; namely, the one-dimensional simple harmonic oscillator, a physicist's best friend.

The Hamiltonian of such an oscillator may be written

$$H(p, x) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2, \quad (7)$$

where  $p$  and  $x$  are operators satisfying the commutation relations

$$[p, x] = \frac{\hbar}{i}. \quad (8)$$

Many of you, I am sure, know that the above Hamiltonian has a discrete eigenvalue spectrum; namely, the eigenvalues of  $H$  are

$$E_n = (n + 1/2)\hbar\omega \quad (n \geq 0). \quad (9)$$

The corresponding eigenvectors  $\Phi_n$ , which are solutions of

$$H\Phi_n = E_n\Phi_n, \quad (10)$$

form an orthonormal basis in the Hilbert space of state vectors.

Any vector can be expanded as

$$\Psi = \sum_n C_n \Phi_n, \quad (11)$$

where the  $C_n$  are complex numbers rendered by the scalar product

$$C_n = (\Phi_n, \Psi). \quad (12)$$

## A The Measurement Postulate

If the state vector of the system is known to be  $\Psi$  at a given instant, and the energy of the oscillator is measured immediately thereafter, the probability of getting  $(n + 1/2)\hbar\omega$  and casting the system into the state  $\Phi_n$  is given by

$$P_n = |C_n|^2, \quad (13)$$

providing all the vectors have been normalized to unity. It is important to realize that no matter what the state vector of the system may be before the measurement is performed, no value of the energy can be obtained other than the values  $(n + 1/2)\hbar\omega$ .

For some reason, even after having studied quantum mechanics for one or two years, *most* students seem to retain a mistaken notion concerning the

simple measurement postulate described above. The “expectation value” of a quantum mechanical operator does not tell you anything about the probability of getting one or another result for a single measurement of the observable in question! It only tells you what to expect for the mean value of many measurements of the observable, starting off each time with the same state before the measurement.

One can measure other observables besides the total energy of the simple harmonic oscillator. For example, there is no reason why you can’t measure the *kinetic* energy of the oscillator. Unlike the total energy, the kinetic energy operator,

$$T = \frac{p^2}{2m}, \quad (14)$$

has a continuous eigenvalue spectrum. The kinetic energy can be anything between 0 and  $\infty$ . Suppose the state of the system before any measurement is made is  $\Psi$ . The probability of obtaining a specific value of the kinetic energy immediately thereafter is zero. However, it is reasonable to inquire as to the probability of getting  $T$  between certain limits. One must expand the given state  $\Psi$  in terms of the eigenstates of  $T$ . The latter are essentially identical to the eigenstates of the momentum  $p$ . If these are denoted by  $\varphi_p$ , we can expand  $\Psi$  as an integral

$$\Psi = \int_{-\infty}^{\infty} \psi(p)\varphi_p dp. \quad (15)$$

The expansion “coefficients”  $\psi(p)$  here are a function of the continuous momentum variable  $p$ . The function  $\psi(p)$  is nothing but the “wave function in momentum representation.” The probability of getting momentum between  $p$  and  $p + dp$  is

$$P(p)dp = |\psi(p)|^2 dp. \quad (16)$$

From this the probability of finding  $T$  between certain limits can be easily ascertained.

Similarly, you might be interested in predicting the probability of getting one or another result for the *position* of a simple harmonic oscillator immediately after the state vector is  $\Psi$ . The position operator has a continuous eigenvalue spectrum. It can assume any value between  $-\infty$  and  $+\infty$ . Denoting the eigenvectors of  $X$  by  $\varphi_x$ , we can expand  $\Psi$  as an integral

$$\Psi = \int_{-\infty}^{\infty} \psi(x)\varphi_x dx, \quad (17)$$

where  $\psi(x)$  is the “wave function in position representation.” The probability of getting position between  $x$  and  $x + dx$  is

$$P(x)dx = |\psi(x)|^2 dx. \quad (18)$$

It is not difficult to show from the commutation relation

$$[p, x] = \frac{\hbar}{i} \quad (19)$$

that the wave functions  $\psi(x)$  and  $\psi(p)$  are related by the Fourier transformation

$$\psi(x) = h^{-1/2} \int_{-\infty}^{\infty} \psi(p) e^{ipx/\hbar} dp. \quad (20)$$

For completeness I should mention that the “wave function in Heisenberg representation” is nothing but the infinite sequence of numbers  $C_n$  in the expansion

$$\Psi = \sum_n C_n \Phi_n. \quad (21)$$

## B The Time-development Postulate

There are two popular ways of describing the time-development of a quantum mechanical system. In the *Heisenberg picture* the operators  $A$  develop in a prescribed fashion; namely,

$$A(t) = e^{iHt/\hbar} A(0) e^{-iHt/\hbar}, \quad (22)$$

while in the *Schrödinger picture* the state vector  $\Psi$  develops in a prescribed fashion,

$$\Psi(t) = e^{-iHt/\hbar} \Psi(0). \quad (23)$$

Here  $e^{-iHt/\hbar}$  is a unitary operator defined by its power series expansion.

For some reason many students assume that all quantum mechanical state vectors and wave functions have a simple exponential time dependence. Perhaps this fallacy arise because often their first exposure to quantum mechanics was seeing the separation of the time-dependent Schrödinger equation. In any event, only in the case of energy eigenstates  $\Phi_n$  does the state vector evolve in a simple exponential fashion,

$$\Phi_n(t) = \Phi_n(0) e^{-iE_n t/\hbar}. \quad (24)$$

For a more general state vector  $\Psi$  we have instead

$$\Psi(t) = \sum_n C_n \Phi_n(0) e^{iE_n t/\hbar}, \quad (25)$$

where  $C_n = (\Phi_n(0), \Psi(0))$ .

It is easy to show that the Schrödinger representation wave function  $\psi(x, t)$ , defined by the expansion

$$\Psi(t) = \int \psi(x, t) \varphi_x dx, \quad (26)$$

satisfies Schrödinger's equation,

$$H\left(\frac{\hbar}{i} \frac{\partial}{\partial x}, x\right) \psi(x, t) = -\frac{\hbar}{i} \frac{\partial}{\partial t} \psi(x, t). \quad (27)$$

This follows directly from the basic commutation relation  $[p, x] = \frac{\hbar}{i}$ .

If at time  $t = 0$  you perform sufficiently many measurements to ascertain the state vector of the system, you may use the time-development hypothesis as here formulated to determine what the state vector will be at all subsequent times, until a new measurement is made. From the evolved state vector you can predict the probability for obtaining one or another result for an experiment. When the experiment is actually performed, however, the system is cast into some definite eigenstate of the observable measured. The new state vector of the system proceeds to evolve deterministically until a further measurement is performed.

## Recommendation

I hope that this discussion has served some useful purpose. I should suggest that during your first course in quantum mechanics you supplement your assigned reading with a perusal of Dirac's lovely book *Quantum Mechanics*, which clearly describes the foundations of the theory, and at the same time introduces the reader to a very useful language, the so-called Dirac notation.