Physics 221A
Fall 2019
Notes 5
Time Evolution in Quantum Mechanics†

## 1. Introduction

In these notes we develop the formalism of time evolution in quantum mechanics, continuing the quasi-axiomatic approach that we have been following in earlier notes. First we introduce the time evolution operator and define the Hamiltonian in terms of it. Then we discuss the evolution of state vectors and the Schrödinger equation, the evolution of observables in the Heisenberg picture and the Heisenberg equations of motion, the evolution of the density operator, and the initial value problem. Next we introduce the Hamiltonian for potential motion and discuss aspects of it, including wave function representations of the Schrödinger equation, the probability density and current, and the Ehrenfest relations. Finally we discuss the Schrödinger equation for a charged particle in a magnetic field, and discuss its transformation properties under gauge transformations.

## 2. The Time-Evolution Operator

Let the pure state of a system at some initial time $t_0$ be described by the state vector $|\psi(t_0)\rangle$, and let the state at some final time $t$ be described by $|\psi(t)\rangle$. We postulate that these two state vectors are related by a linear operator $U(t, t_0)$, parameterized by the two times, that is,

$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle. \tag{1}$$

We write the final time first and the initial time second in $U(t, t_0)$. This is another (the seventh) postulate of quantum mechanics, which we can add to the list in Sec. 3.19.

The operator $U(t, t_0)$ must satisfy the following three properties. First, it reduces to the identity at $t = t_0$,

$$U(t_0, t_0) = 1. \tag{2}$$

Second, it must be unitary in order to preserve probabilities,

$$U(t, t_0)^{-1} = U(t, t_0)^\dagger. \tag{3}$$

For example, the probability of finding a particle somewhere in space must be 1, and, assuming particles are neither created or destroyed, this probability must be independent of time. In relativistic

---

† Links to the other sets of notes can be found at:
http://bohr.physics.berkeley.edu/classes/221/1920/221.html.

interactions particles are created and destroyed, but even here we require the unitarity of $U(t, t_0)$. The logic is that if we measure some observable, the probability of getting some answer must be 1, independent of $t$. As for measuring the position of a particle in relativity theory, we will see later in the course that the position operator is not well defined in relativistic quantum mechanics, precisely because of the possibility of particle creation or destruction.

The third property says that if we evolve a system from some initial time $t_0$ to some intermediate time $t_1$, and then from $t_1$ to some final time $t_2$, we must get the same answer as evolving directly from $t_0$ to $t_2$. In other words, we can regard the state at any intermediate time as the initial conditions for the second half of the time evolution. According to this *composition* property, we have

$$U(t_2, t_1)U(t_1, t_0) = U(t_2, t_0). \tag{4}$$

## 3. The Hamiltonian

Now consider an infinitesimal time evolution from time $t$ to $t + \epsilon$, where $\epsilon$ is small. We expand $U(t + \epsilon, t)$ out to first order in $\epsilon$ and write the correction term as follows:

$$U(t + \epsilon, t) = 1 - i\epsilon\Omega(t) + \dots, \tag{5}$$

where the ellipsis indicates higher order terms in $\epsilon$. Effectively, we have defined the operator $\Omega(t)$ as

$$\Omega(t) = i \left. \frac{\partial}{\partial t'} U(t', t) \right|_{t'=t}. \tag{6}$$

We introduce the factor of $i$ in Eq. (6) in order to make $\Omega$ Hermitian. This follows from the unitarity of $U(t, t_0)$, precisely as in the demonstration of the Hermiticity of $\hat{k}$ in Sec. 4.5. The operator $\Omega$ depends on $t$ in general because the right-hand side of Eq. (6) does so.

The operator $\Omega(t)$ can be regarded as the *generator* of time translations in quantum mechanics, because it gives the small correction needed to advance a state from time $t$ to $t + \epsilon$. Evolution over a finite time interval can be regarded as the result of a large number of small displacements in time, so knowing $\Omega(t)$ in principle allows us to evolve the system over any time interval. Since it is Hermitian, $\Omega(t)$ presumably corresponds to something we can observe.

The classical Hamiltonian $H$ is considered the generator of time translations in classical mechanics. This means that for small time increments we can calculate the correction in $q$ or $p$ or any other classical observable by forming the Poisson bracket with $H$. To see this explicitly, let $f(x, p)$ be a classical observable (for a particle moving in one dimension, for simplicity), which acquires a time dependence because $x$ and $p$ evolve in time as a particle moves along its orbit. Then, working to first order in a small time increment $\epsilon$, we have

$$\begin{aligned} f(t + \epsilon) &= f(t) + \epsilon \frac{df}{dt} = f(t) + \epsilon \left( \dot{x} \frac{\partial f}{\partial x} + \dot{p} \frac{\partial f}{\partial p} \right) \\ &= f(t) + \epsilon \left( \frac{\partial H}{\partial p} \frac{\partial f}{\partial x} - \frac{\partial H}{\partial x} \frac{\partial f}{\partial p} \right) = f(t) - \epsilon\{H, f\}, \end{aligned} \tag{7}$$

where we use Hamilton's equations, Eq. (B.78), and the definition of the Poisson bracket, Eq. (B.100). This should be compared to Eq. (4.64), which shows the analogous role of the momentum in classical mechanics as the generator of translations.

Because of the role of the classical Hamiltonian as the generator of time translations, we guess that $\Omega(t)$ is closely related to the operator that we should define as the *quantum Hamiltonian*, denoted by $\hat{H}$ (with a hat to distinguish it from the classical Hamiltonian $H$). We cannot have $\Omega = \hat{H}$ because $\Omega$ has units of inverse time, while $\hat{H}$ has units of energy. We fix this by defining

$$\hat{H}(t) = \hbar\Omega(t). \tag{8}$$

With this definition, the infinitesimal time evolution (5) becomes

$$U(t + \epsilon, t) = 1 - \frac{i\epsilon}{\hbar}\hat{H}(t) + \dots. \tag{9}$$

We remark that Hamiltonians in classical mechanics depend on time when there are time-dependent forces or fields present. See the discussion in Sec. B.20. For example, the motion of a charged particle in an electromagnetic wave is described by a time-dependent Hamiltonian. We expect the same to be true in quantum mechanics, so $\hat{H}$ depends on $t$, in general. As in classical mechanics, however, time-independent Hamiltonians are an important special case, and offer some simplifications.

Our presentation leads to the question of whether the $\hbar$ in Eq. (8) is the same as in Eq. (4.65), $\hat{\mathbf{p}} = \hbar\hat{\mathbf{k}}$. See Sec. 4.12 for an overview of the history of the relations $E = \hbar\omega$ and $\mathbf{p} = \hbar\mathbf{k}$. Here we simply note that these relations are obviously the space- and time-components of a single 4-vector, in which energy and momentum are generators of displacements in time and space, respectively. Relativistic covariance requires that the $\hbar$'s in these relations be the same. We are only concerning ourselves with the nonrelativistic theory at this point, but it is obvious that relativistic covariance will be impossible to establish if these $\hbar$'s are different.

## 4. Differential Equations for Time Evolution

To obtain a differential equation for $U(t, t_0)$ we take the derivative with respect to the final time $t$, using the definition of the derivative:

$$\frac{\partial U(t, t_0)}{\partial t} = \lim_{\epsilon \to 0} \frac{U(t + \epsilon, t_0) - U(t, t_0)}{\epsilon}. \tag{10}$$

But by the composition property (4), we can write

$$U(t + \epsilon, t_0) = U(t + \epsilon, t)U(t, t_0), \tag{11}$$

so the operator $U(t, t_0)$ can be factored out to the right in Eq. (10), giving

$$\frac{\partial U(t, t_0)}{\partial t} = \left[\lim_{\epsilon \to 0} \frac{U(t + \epsilon, t) - 1}{\epsilon}\right] U(t, t_0). \tag{12}$$

The remaining limit is $\partial U(t',t)/\partial t'|_{t'=t}$, which by Eqs. (6) and (8) is $-i\Omega(t) = -i\hat{H}(t)/\hbar$. Altogether, Eq. (12) becomes

$$i\hbar\frac{\partial U(t,t_0)}{\partial t} = \hat{H}(t)U(t,t_0). \tag{13}$$

This differential equation is subject to the initial conditions (2). In practice $\hat{H}(t)$ is usually a given operator as a function of time and we wish to solve for $U(t,t_0)$. With the given initial conditions, the solution is unique.

Now differentiating Eq. (1) with respect to $t$ and using Eq. (13), we obtain a differential equation for the state vector $|\psi(t)\rangle$,

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}(t)|\psi(t)\rangle, \tag{14}$$

which is subject to the given initial condition $|\psi(t_0)\rangle$ for the state vector. This is the *Schrödinger equation*. This is a ket equation that differs only trivially from the operator equation (13).

The case that the Hamiltonian is independent of time, $\partial\hat{H}(t)/\partial t = 0$, is important in practice. In this case, Eq. (13) has the solution,

$$U(t,t_0) = \exp[-i\hat{H}(t-t_0)/\hbar], \tag{15}$$

showing that $U(t,t_0)$ only depends on the elapsed time $t - t_0$ when $\hat{H}$ is independent of $t$. Writing simply $t$ for the elapsed time (or choosing $t_0 = 0$), we have

$$U(t) = e^{-i\hat{H}t/\hbar}. \tag{16}$$

In this case, $\hat{H}$ commutes with $U(t)$.

In the general case in which $\hat{H}$ depends on time, there is no simple solution for $U(t,t_0)$. In this case $\hat{H}(t)$ does not commute with $U(t,t_0)$, nor for that matter does it usually commute with itself at different times.

## 5. The Heisenberg Picture

The time evolution of a classical observable along an orbit is given by Eq. (B.98) or (B.101). In quantum mechanics we see a closer analogy with classical mechanics if we allow observables to evolve in time. Up to this point in our discussion of time evolution in quantum mechanics we have used the *Schrödinger picture*, in which kets evolve according to the Schrödinger equation (14), and operators do not evolve (unless they have an explicit time dependence, in which case their only evolution is due to this time dependence). We shall continue to use the Schrödinger picture for most of this course, but there are certain problems and topics that are best handled in the *Heisenberg picture*, in which kets do not evolve in time but operators do. We shall now explain these two pictures and the relation between them.

In the following we drop the hats on $H$, it being understood that we are speaking of the quantum Hamiltonian.

The Schrödinger and Heisenberg pictures differ by a time-dependent, unitary transformation. In the following we shall put an $S$ subscript on kets and operators in the Schrödinger picture and an $H$ subscript on them in the Heisenberg picture. If a ket or an operator appears without a subscript, the Schrödinger picture is assumed. If $|\psi_S(t)\rangle = |\psi(t)\rangle$ is the time-dependent state vector in the Schrödinger picture, then the state vector in the Heisenberg picture is defined by

$$|\psi_H\rangle = U(t, t_0)^\dagger |\psi_S(t)\rangle. \tag{17}$$

By Eqs. (1) and (3), this is equivalent to

$$|\psi_H\rangle = |\psi_S(t_0)\rangle, \tag{18}$$

where $t_0$ is the initial time. The state vector in the Heisenberg picture does not evolve because it is equal to the ket in the Schrödinger picture at the initial time $t_0$. As for operators, if $A_S(t) = A(t)$ is an operator in the Schrödinger picture, which for generality we allow to have an explicit time dependence, then the corresponding operator in the Heisenberg picture is

$$A_H(t) = U(t, t_0)^\dagger A_S(t) U(t, t_0). \tag{19}$$

The Heisenberg operator $A_H(t)$ has a time dependence for two reasons, first because the Schrödinger operator $A_S(t)$ may have an (explicit) time dependence, and second because the operators $U(t, t_0)$ and $U(t, t_0)^\dagger$ depend on time, what we regard as the implicit time dependence. Notice that the Heisenberg operator depends on time even if the Schrödinger operator does not (because of the implicit time dependence). See Sec. B.20 for a discussion of explicit and implicit time dependence in classical observables.

There is no evolution equation for $|\psi_H\rangle$ in the Heisenberg picture, but the operators obey an evolution equation obtained by taking the derivative of the definition (19) and using Eq. (13). This gives

$$i\hbar \frac{dA_H}{dt} = -U^\dagger H A_S U + U^\dagger A_S H U + i\hbar U^\dagger \frac{\partial A_S}{\partial t} U, \tag{20}$$

where $H = H_S$ is the Hamiltonian in the Schrödinger picture, which may be time dependent. But by inserting the product $UU^\dagger$ between $A_S$ and $H$ in the first two terms above, the result can be written purely in terms of operators in the Heisenberg picture,

$$i\hbar \frac{dA_H}{dt} = [A_H, H_H] + i\hbar \left( \frac{\partial A}{\partial t} \right)_H, \tag{21}$$

where the last term means, differentiate $A$ with respect to its explicit time dependence in the Schrödinger picture, then convert to the Heisenberg picture. This may also be written

$$\frac{dA_H}{dt} = -\frac{i}{\hbar} [A_H, H_H] + \left( \frac{\partial A}{\partial t} \right)_H, \tag{22}$$

whereupon the analogy with the classical equation (B.101) is more striking. Equations (21) or (22) are the Heisenberg equations of motion for operators.

## 6. Algebraic Relations Same in Both Pictures

Suppose we have an operator $A_S$ in the Schrödinger picture, so that in the Heisenberg picture it is $A_H = U^\dagger A_S U$, where $U = U(t, t_0)$. Then if $B_S = f(A_S)$ for some function $f$, it turns out that

$$B_H = U^\dagger f(A_S) U = f(U^\dagger A_S U) = f(A_H). \tag{23}$$

This is easy to see in the case that $f$ is a polynomial or power series, for example, if $B_S = A_S^2$ we have

$$B_H = U^\dagger A_S A_S U = U^\dagger A_S U U^\dagger A_S U = A_H^2, \tag{24}$$

but the result can be proved more generally by using the definition of a function of an operator given in Sec. 1.25. Likewise, suppose we have an algebraic relation in the Schrödinger picture such as $C_S = A_S B_S$, where $A_S$ and $B_S$ need not commute. Then it is easy to show that $C_H = A_H B_H$. In other words, algebraic relations among operators in the Schrödinger picture are converted into the corresponding algebraic relations in the Heisenberg picture simply by replacing the $S$ subscripts with $H$'s.

In particular, commutation relations have the same form in both pictures. If $C_S = [A_S, B_S]$, then $C_H = [A_H, B_H]$. Also, if $V(\mathbf{x}_S)$ is a potential energy in the Schrödinger picture, then

$$U^\dagger V(\mathbf{x}_S) U = V(\mathbf{x}_H) \tag{25}$$

is the potential energy in the Heisenberg picture. This means that the Hamiltonian has the same form in both pictures, for example, if

$$H_S = \frac{\mathbf{p}_S^2}{2m} + V(\mathbf{x}_S), \tag{26}$$

then

$$H_H = \frac{\mathbf{p}_H^2}{2m} + V(\mathbf{x}_H). \tag{27}$$

## 7. The Heisenberg and Schrödinger Pictures are Physically Equivalent

All measurable quantities in quantum mechanics can be expressed in terms of matrix elements of the form $\langle \phi | A | \psi \rangle$. These, however, are the same in both the Heisenberg and Schrödinger pictures,

$$\langle \phi_S(t) | A_S | \psi_S(t) \rangle = \langle \phi_H | A_H(t) | \psi_H \rangle, \tag{28}$$

since the unitary operators cancel when Eqs. (17) and (19) are substituted. Thus, if we can find the time evolution of operators in the Heisenberg picture, we can predict the outcome of all experiments equally well as by finding the time evolution of the state vector in the Schrödinger picture. For some problems, the Heisenberg picture is easier to use (the harmonic oscillator, for example). In addition, the Heisenberg picture is needed in relativistic quantum field theory to establish manifest Lorentz covariance. Although most of nonrelativistic quantum mechanics is usually carried out in

the Schrödinger picture (and also in this course), it can be argued that the Heisenberg picture is more fundamental.

In the special case that the Schrödinger Hamiltonian is independent of time, the operator $U(t, t_0) = U(t - t_0)$ is a function of $H$ and so commutes with it, according to Eq. (16). Thus, in this case, the Schrödinger and Heisenberg Hamiltonians are the same,

$$H_H = U(t)^\dagger H_S U(t) = H_S, \tag{29}$$

where we have set $t_0 = 0$ as in Eq. (16).

There are other pictures besides the Schrödinger and Heisenberg pictures. In the interaction picture, which we will take up in Notes 33, the time dependence due to some unperturbed Hamiltonian $H_0$ is stripped off from the kets and transferred to the operators. But the time dependence remaining from some perturbation $H_1$ is left on the kets. Thus, the interaction picture is intermediate between the Schrödinger and Heisenberg pictures. It is especially useful in time-dependent perturbation theory and scattering theory.

## 8. Time Evolution of the Density Operator

Equation (14), the Schrödinger equation, gives the time evolution of the state vector in quantum mechanics. As we have seen in Notes 3, however, the state of a quantum system is described in general by a density operator, not a state vector. Therefore we require an equation of evolution for the density operator. We work here in the Schrödinger picture, and consider the density operator for a discrete ensemble of pure states $|\psi_i\rangle$ with statistical weights $f_i$, as in Eq. (3.16). The density operator $\hat{\rho}$ is a function of time because the states $|\psi_i\rangle$ are functions of time. We put a hat on $\hat{\rho}$ to distinguish it from the classical density $\rho$ discussed in Sec. B.24.

At the initial time $t_0$ the density operator is given by

$$\hat{\rho}(t_0) = \sum_i f_i |\psi_i(t_0)\rangle\langle\psi_i(t_0)|, \tag{30}$$

while at the final time it is

$$\hat{\rho}(t) = \sum_i f_i |\psi_i(t)\rangle\langle\psi_i(t)| = U(t, t_0) \left( \sum_i f_i |\psi_i(t_0)\rangle\langle\psi_i(t_0)| \right) U(t, t_0)^\dagger$$

$$= U(t, t_0)\hat{\rho}(t_0)U(t, t_0)^\dagger. \tag{31}$$

Differentiating this last expression with respect to time and using Eq. (13), we find

$$i\hbar \frac{\partial \hat{\rho}}{\partial t} = [\hat{H}, \hat{\rho}]. \tag{32}$$

This equation replaces the Schrödinger equation when we have mixed states. By a slight change Eq. (32) can be written

$$\frac{\partial \hat{\rho}}{\partial t} - \frac{i}{\hbar}[\hat{\rho}, \hat{H}] = 0, \tag{33}$$

whereupon the analogy with the classical Liouville equation (B.112) is more clear. Equation (32) looks similar to the Heisenberg equation of motion for some observable, Eq. (22), but they are not the same since we are working here in the Schrödinger picture.

## 9. Commutators and Poisson Brackets

Comparing the quantum equations (22) and (33) with their classical counterparts (B.101) and (B.112), respectively, we see that one goes into the other if we map commutators into Poisson brackets according to the rule,

$$[A, B] \to i\hbar\{A, B\}. \tag{34}$$

This association was first noticed by Dirac, and it helped him to establish the formal structure of quantum mechanics on the classical model. In important cases it is an exact correspondence, that is, quantum commutators are the same as classical Poisson brackets, apart from the factor $i\hbar$ and a reinterpretation of the symbols as operators (quantum observables) instead of classical observables. Compare, for example, the classical canonical Poisson bracket relations (B.108) with the Heisenberg-Born commutation relations (4.69). Another example is the classical Poisson bracket relations for the components of orbital angular momentum, $\mathbf{L} = \mathbf{x}\times\mathbf{p}$,

$$\{L_i, L_j\} = \epsilon_{ijk} L_k, \tag{35}$$

which may be compared to the quantum commutation relations for the components of the angular momentum operator,

$$[L_i, L_j] = i\hbar\, \epsilon_{ijk} L_k. \tag{36}$$

In other cases, Dirac's correspondence (34) is only approximate.

## 10. Constants of Motion

In classical mechanics, an observable $A = A(q, p)$ is a *constant of motion* if its value is constant as $q$ and $p$ evolve along an orbit. For simplicity in this section we only speak of observables without an explicit time dependence, and we assume that the Hamiltonian, too, has no explicit time dependence. But according to Eq. (B.101), this means that $A$ has vanishing Poisson bracket with the Hamiltonian, $\{A, H\} = 0$.

Similarly, in quantum mechanics, we say that an observable $A$ without an explicit time dependence is a *constant of motion* if it commutes with the Hamiltonian, $[A, H] = 0$. Then the Heisenberg equations of motion (21) imply

$$\frac{dA_H}{dt} = 0. \tag{37}$$

This in turn implies that the expectation value of $A$ with respect to any state is constant in time,

$$\frac{d}{dt}\langle A\rangle = 0, \tag{38}$$

where, as we have seen, the expectation value can be computed in either picture (Schrödinger or Heisenberg).

Energy eigenstates in many practical problems are expressed as simultaneous eigenstates of a complete set of commuting observables, of which the Hamiltonian is one. For example, in central force motion we often use $(H, L^2, L_z)$. But since the other observables in the list, besides the Hamiltonian, all commute with the Hamiltonian, they are all constants of motion, and their expectation values with respect to arbitrary states are independent of time. And since the Hamiltonian commutes with itself, it is also a constant of motion.

## 11. The Initial Value Problem for Time-Independent Hamiltonians

The initial value problem refers to the problem of finding the final state $|\psi(t)\rangle$ at some time $t$, given the initial state $|\psi(t_0)\rangle$ at some initial time $t_0$. When the Hamiltonian is time-independent, it is possible to give an explicit solution to the initial value problem in terms of the eigenstates and eigenvalues of the Hamiltonian.

Suppose $H$ in Eq. (14) is time-independent. Let us denote its eigenstates by $|n\rangle$ and corresponding eigenvalues by $E_n$, so that

$$H|n\rangle = E_n|n\rangle, \tag{39}$$

where $n$ can stand for one or more quantum numbers needed to specify an energy eigenstate. For simplicity we assume the index $n$ is discrete, but the formulas are easily extended to handle a continuous or mixed spectrum. Let us expand the time-dependent state $|\psi(t)\rangle$ in terms of the energy eigenstates,

$$|\psi(t)\rangle = \sum_n c_n(t)|n\rangle, \tag{40}$$

where the expansion coefficients $c_n$ depend on time. Then by substituting Eq. (40) into Eq. (14), we find

$$i\hbar \dot{c}_n = E_n c_n, \tag{41}$$

or,

$$c_n(t) = e^{-iE_n t/\hbar} c_n(0). \tag{42}$$

Thus we have

$$|\psi(t)\rangle = \sum_n e^{-iE_n t/\hbar} c_n(0)|n\rangle, \tag{43}$$

where

$$c_n(0) = \langle n|\psi(0)\rangle. \tag{44}$$

These results are equivalent to expanding the time-evolution operator in energy eigenstates,

$$U(t) = e^{-iHt/\hbar} = \sum_n e^{-iE_n t/\hbar} |n\rangle\langle n|. \tag{45}$$

This is a special case of a function of an observable expressed in terms of the projectors of that observable, as described in Sec. 1.25.

Often the easiest way to solve the initial value problem in quantum mechanics is first to find the eigenfunctions of the Hamiltonian, then to compute the expansion coefficients of $|\psi\rangle$ at $t = 0$ by Eq. (44), and finally to sum the series (43). This only works for time-independent Hamiltonians, however.

In particular, if the system is known to be in an energy eigenstate at $t = 0$, then it stays in this state forever. The only thing that changes in time is the overall phase of the state, $e^{-iE_n t/\hbar}$. For this reason an energy eigenstate is also called a *stationary state*.

## 12. Hamiltonian for Potential Motion

So far we have defined the Hamiltonian operator as the generator of time evolution, but we have not said what the Hamiltonian actually is in any specific case. The simplest case is that of a particle of mass $m$ moving in a potential $V(\mathbf{x}, t)$, which for generality we allow to depend on time. See Sec. B.8 for the physical meaning of time-dependent potentials, including an example. See also See. B.20 for a discussion of time-dependent Hamiltonians. In practice, the potential is often time-independent. The classical equations of motion are

$$m\mathbf{a} = -\nabla V(\mathbf{x}, t), \tag{46}$$

where $\mathbf{a} = \ddot{\mathbf{x}}$. The classical Hamiltonian is a single-particle version of Eq. (B.83),

$$H = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x}, t). \tag{47}$$

To obtain the quantum Hamiltonian, we take this formula and reinterpret the classical observables $\mathbf{x}$ and $\mathbf{p}$ as operators acting on the ket space spanned by the position eigenkets $|\mathbf{x}\rangle$, as in Notes 4. This space is isomorphic to the space of wave functions $\psi(\mathbf{x})$. The operators $\mathbf{x}$ and $\mathbf{p}$ satisfy the Heisenberg-Born commutation relations (4.69). As in classical mechanics, the Hamiltonian (47) is the sum of the kinetic and the potential energies.

Quantum mechanics must go over to classical mechanics in the limit in which $\hbar$ is small, that is, when all actions of a problem are large in comparison to $\hbar$. So it is an educated guess that we can adopt a classical Hamiltonian in this manner, reinterpret it, and get the correct quantum Hamiltonian. Ultimately, the success of our guess must be verified by comparison with experiment, or comparison with a more fundamental theory that we trust. And of course we trust fundamental theories when they have been verified by experiment. There is no process of logical deduction to obtain quantum mechanics from classical mechanics, since new physics is involved.

The potentials $V$ that are used in practice are often phenomenological models that represent guesses or hopes that some complicated physics can be represented by a potential. For example, we may try to represent the interaction of two atoms, or of a proton and a neutron, by a potential.

Actually, atoms and nucleons are composite particles with internal structure, and their interactions can only be represented by a potential in an approximate manner.

A large part of atomic, molecular and condensed matter physics can be understood in terms of the electrostatic interactions of charged particles. These are described by the potential (B.24), which can be regarded as "fundamental." But effects involving spin, magnetic fields, relativistic corrections, the emission and absorption of radiation and other phenomena require other types of Hamiltonians. We will consider magnetic fields below in Sec. 17, and the other generalizations will be taken up later in the course.

The multiparticle generalization of Eq. (47) is straightforward. Let there be $N$ particles with masses $m_i$, positions $\mathbf{x}_i$ and momenta $\mathbf{p}_i$, and assume that they interact by means of a potential. Then the classical Hamiltonian is

$$H = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{x}_1, \ldots, \mathbf{x}_N). \tag{48}$$

In a multiparticle system, the wave function is $\psi(\mathbf{x}_1, \ldots, \mathbf{x}_N, t)$. It is a function defined over the multidimensional configuration space (see Sec. B.3).

## 13. Representations of the Schrödinger Equation

The Hamiltonian (47) is presented in terms of operators $\mathbf{x}$ and $\mathbf{p}$. We will not put hats on these operators as we did in Notes 4, since it will usually be clear from context whether a quantum operator is intended and or a $c$-number. These operators are represented in different ways in different representations, as discussed in Secs. 4.13 and 4.14 (see Table 4.1). We now translate the ket version of the Schrödinger equation, Eq. (14), into its configuration space representation, using the Hamiltonian (47). We write $\psi(\mathbf{x}, t) = \langle \mathbf{x} | \psi(t) \rangle$. Then using Eqs. (4.20) and (4.71) or Table 4.1 for the actions of $\mathbf{x}$ and $\mathbf{p}$ on configuration space wave functions, we obtain

$$-\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{x}, t) + V(\mathbf{x}, t) \psi(\mathbf{x}, t) = i\hbar \frac{\partial \psi(\mathbf{x}, t)}{\partial t}. \tag{49}$$

If the potential is independent of time, then we often wish to find the energy eigenfunctions of the Hamiltonian, call them $\psi_n(\mathbf{x}) = \langle \mathbf{x} | n \rangle$ where $|n\rangle$ are the energy eigenkets as in Eq. (39). These satisfy

$$\left[ -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{x}) \right] \psi_n(\mathbf{x}) = E_n \psi_n(\mathbf{x}), \tag{50}$$

which is the configuration space representation of Eq. (39). Equations (49) and (50) are the time-dependent and time-independent versions, respectively, of the Schrödinger equation in configuration space, for a single particle moving in a given potential.

We can also write down the Schrödinger equation in momentum space. In the time-independent version, it is

$$\frac{p^2}{2m} \phi_n(\mathbf{p}) + \int d^3 \mathbf{p}' \, \tilde{V}(\mathbf{p} - \mathbf{p}') \phi_n(\mathbf{p}') = E_n \phi_n(\mathbf{p}), \tag{51}$$

where $\phi_n(\mathbf{p}) = \langle \mathbf{p}|n\rangle$ and where

$$\tilde{V}(\mathbf{p}) = \int \frac{d^3\mathbf{x}}{(2\pi\hbar)^3}\, e^{-i\mathbf{p}\cdot\mathbf{x}/\hbar}\, V(\mathbf{x}). \tag{52}$$

The first term on the left of Eq. (51) (the kinetic energy) is a multiplicative operator in momentum space, which is simpler than the differential operator (the Laplacian) that occurs in configuration space. The potential, on the other hand, has become an integral operator. The momentum space version of the Schrödinger equation is useful for some problems, notably the hydrogen atom.

## 14. The Probability Density and Current

Let us consider a single particle moving in three-dimensional space under the influence of a potential $V$, as described by the Hamiltonian (40). The potential may be time-dependent. Let us write

$$\rho(\mathbf{x}, t) = |\psi(\mathbf{x}, t)|^2 \tag{53}$$

for the probability density on configuration space (not to be confused with the density operator). The probability density changes in time, in general, but the integral over all space,

$$\int d^3\mathbf{x}\,\rho(\mathbf{x}, t) = \int d^3\mathbf{x}\,|\psi(\mathbf{x}, t)|^2 = \int d^3\mathbf{x}\,\langle\psi(t)|\mathbf{x}\rangle\langle\mathbf{x}|\psi(t)\rangle = \langle\psi(t)|\psi(t)\rangle = 1, \tag{54}$$

does not, since it is just the square norm of the wave function which is constant in time due to the unitarity of $U(t, t_0)$.

Associated with the probability density $\rho(\mathbf{x}, t)$ is a probability current $\mathbf{J}(\mathbf{x}, t)$ satisfying

$$\nabla \cdot \mathbf{J} + \frac{\partial\rho}{\partial t} = 0. \tag{55}$$

The current is conveniently expressed in terms of the velocity operator, which in the configuration representation is given by

$$\mathbf{v} = \frac{1}{m}\mathbf{p} = -\frac{i\hbar}{m}\nabla. \tag{56}$$

The current itself is

$$\mathbf{J}(\mathbf{x}, t) = \frac{1}{2}[\psi^*(\mathbf{x}, t)\mathbf{v}\psi(\mathbf{x}, t) + \psi(\mathbf{x}, t)\mathbf{v}\psi^*(\mathbf{x}, t)] = \mathrm{Re}[\psi^*(\mathbf{x}, t)\mathbf{v}\psi(\mathbf{x}, t)] \tag{57}$$

The current looks like the real part of the expectation value of $\mathbf{v}$, but it is not since we do not integrate over all space. It is however the real part of the integrand of that expectation value. It is straightforward to prove the continuity equation (55) with the aid of the Schrödinger equation (49).

Notice that if $\psi(\mathbf{x}, t)$ is an energy eigenfunction $\psi_n(\mathbf{x})e^{-iE_n t/\hbar}$, then $\rho$ is independent of time and the continuity equation reduces to $\nabla \cdot \mathbf{J} = 0$. Also, if the Hamiltonian is of the kinetic-plus-potential form, then time-reversal invariance (see Notes 6 and 21) implies that the energy eigenfunctions can be chosen to be real. In that case, $\mathbf{J} = 0$.

The definition of the probability current depends on the system. For the case considered here, a single particle moving in 3-dimensional space, the probability current $\mathbf{J}$ is a 3-vector. In an $N$-particle system the probability current is a $3N$-vector on the $3N$-dimensional configuration space. Also, even for a single particle, the definition is modified in the presence of magnetic fields, as we shall explain in Sec. 17 below, or when there is spin.

## 15. Physical Significance of the Probability Density and Current

The existence of a probability density $\rho$ and current $\mathbf{J}$ that satisfy the continuity equation (55) is usually regarded as an essential part of the probabilistic interpretation of the wave function $\psi$. There is, however, the question of the physical interpretation of $\rho$ and $\mathbf{J}$.

The physical interpretation of the density $\rho$ is clear on the basis of the postulates of quantum mechanics: It gives the distribution of measured positions of the particle when measurements are made across an ensemble of identically prepared systems. We should not think that the particle in an individual system has been somehow smeared around in space. If we make a measurement of the position of the particle, we find some definite value, not some fraction of a particle in one region and another fraction in another. In particular, if the particle has charge, we never measure a fraction of the charge. Thus, fundamentally, the density $\rho$ describes the statistics of the ensemble, not some kind of smearing of an individual system. Nevertheless, there are circumstances in which the latter interpretation also appears, as we shall explain. Similar considerations apply to the current $\mathbf{J}$.

In the case of a charged particle, it is suggestive to multiply both the probability density $\rho$ and current $\mathbf{J}$ by the charge of the particle $q$, and to write

$$\rho_c = q\rho, \qquad \mathbf{J}_c = q\mathbf{J}, \tag{58}$$

since then the continuity equation for $\rho_c$ and $\mathbf{J}_c$ looks like the law of conservation of charge,

$$\nabla \cdot \mathbf{J}_c + \frac{\partial \rho_c}{\partial t} = 0. \tag{59}$$

But now the question arises, do this $\rho_c$ and $\mathbf{J}_c$ serve as sources for electromagnetic fields? If so, are these fields measurable? That is, are electric and magnetic fields produced by the average positions and motions of the charged particle across the ensemble, or by their "actual" positions and velocities in an individual system? (We put "actual" in quotes because position and velocity do not commute, and so cannot simultaneously have a precise value.)

Consider, for example, a hydrogen atom in the ground state. The wave function for the electron is rotationally symmetric and dies off exponentially with a scale length which is the Bohr radius, so the probability density $\rho = |\psi|^2$ and effective charge density $\rho_c = -e\rho$ do the same. At a distance large compared to the Bohr radius, the electron wave function and hence probability density are essentially zero. Since the probability density is rotationally symmetric, that is, it is a function only of the radius $r$, the electric field produced by $\rho_c$ is easy to calculate by Gauss's law. The electric field is purely in the radial direction, and, far from the proton, it is the same as that of an electron

at the origin (the position of the proton). If we add the electric field of the proton, the total electric field far from the hydrogen atom is zero.

On the other hand, in a classical model of the hydrogen atom, in which the electron has a definite position, the electric field produced by the electron and proton is a dipole field, which falls off as $1/r^3$. Moreover, the dipole is time-dependent, as the electron moves around in its orbit. This is quite different from the electric field produced by the average charge density $\rho_c$.

If we take an actual hydrogen atom in the ground state and measure the electric field, which do we get? A dipole field, or a field that falls off exponentially to zero within a short distance?

This is a somewhat subtle question. In the first place, the electron is in orbit around the proton, so if we wish to see a dipole field we would need to make the measurement on a time scale shorter than the orbital period of the electron, which is something like $10^{-16}$ seconds. It would be hard to make such a measurement in practice. On longer time scales we will see an average field. In fact, taking an expectation value of an observable with respect to an energy eigenfunction is similar to making a time average in classical mechanics. (The one goes over into the other in the classical limit.) This can be seen crudely by the time-energy uncertainty principle, $\Delta t \Delta E >\approx \hbar$. In an energy eigenstate, $\Delta E = 0$ since the energy is exactly known. Thus the time is completely unknown.

But if it were practical to make a measurement of the electric field on a very short time scale, what would we find? To answer this question, we must first remember that according to the measurement postulates of quantum mechanics, all observables are represented by operators that act on some ket space. This must also apply to the measurement of the electric field that we are contemplating. According to classical electromagnetic theory, the electric field at field point $\mathbf{r}$ due to an electron of charge $-e$ at position $\mathbf{x}$ is

$$\mathbf{E}(\mathbf{r}) = -\frac{e(\mathbf{r} - \mathbf{x})}{|\mathbf{r} - \mathbf{x}|^3}, \tag{60}$$

in the electrostatic approximation. There are two measurable quantities that appear in this formula, the position $\mathbf{x}$ of the electron and the value $\mathbf{E}$ of the electric field at observation point $\mathbf{r}$. But the observation point $\mathbf{r}$ is not an observable, since we are not measuring it; it is just the place where we make the measurement of $\mathbf{E}$. With this understanding, we see that $\mathbf{E}$ is a function of $\mathbf{x}$, that is, $\mathbf{E}(\mathbf{r})$ is an operator (really, a vector of operators) that act on the ket space for the electron. Since the components of $\mathbf{x}$ commute with one another, so also do the components of $\mathbf{E}$, and $\mathbf{E}$ at one point $\mathbf{r}_1$ commutes with $\mathbf{E}$ at another point $\mathbf{r}_2$. In fact, by measuring $\mathbf{E}$ at any point $\mathbf{r}$ and knowing the charge of the electron, we can solve Eq. (60) for $\mathbf{x}$, the position of the electron. Measuring the electric field at a point is thus equivalent to measuring the position of the electron. The value measured in any individual system is thus a dipole field, as in the classical case.

We emphasize that this discussion is based on the electrostatic approximation, in which we ignore the retardation in the communication between the source of the field (the electron, at position $\mathbf{x}$) and the observer (at position $\mathbf{r}$). We also ignore magnetic fields and other effects that lie outside the electrostatic approximation.

It is possible to make a variation on this situation that is not hampered by short time scales. Consider a double slit (really, a double hole) experiment, in which a plane wave (a pure state of a given energy) of electrons is directed against a screen $S$. See Fig. 1. There are two holes in the screen that are large compared to the wavelength, so that two parallel beams emerge on the other side. We assume the beams are wide enough to ignore diffraction effects (the spreading of the beam) over the extent of the experiment. Then probability density $\rho$ is nonzero inside two cylinders (the two beams) downstream from the holes. This is a variation in the double slit experiment, in which one of the questions is whether a single electron passes somehow through both holes, or how we can know which hole it passed through.
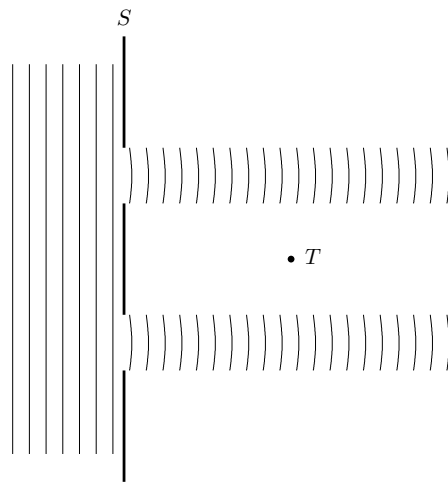


**Fig. 1.** A variation of the double slit experiment, in which we attempt to measure which hole in the screen $S$ the electron passed through by measuring the electric field produced on test charge $T$ as the electron passes by.

Let us place a test charge $T$ half way between the two beams, as in the figure. By symmetry, the electric field produced by $\rho_c$ is zero on the axis between the two beams, so if $T$ is affected by the electric field of the particle smeared out according to $\rho_c$, then it should not move. But if we think of an individual system in a classical sense, each electron will pass by the test charge on one side or the other. If we imagine that $T$ is massive enough that it does not move much while the electron is passing by, then the main effect of the electron is to deliver an impulse in the transverse direction (up or down in the figure). If we see which direction the test charge moves, we can detect which of the two beams the electron was in. So now the question is, what does the test particle really do in a double hole experiment of this type? Does it respond to the average electric field, or to the field of an individual electron, as in a classical model?

To answer this question we must realize that the test particle is a quantum mechanical system in its own right, so we must represent it by a wave function, say a wave packet centered initially on the symmetry axis between the two beams. In fact, because the test particle interacts with the electron in the beam, we are really dealing with a two-particle system. Suppose for simplicity that

initially the wave function has the product form, $\Psi(\mathbf{x}_e, \mathbf{x}_T) = \psi_e(\mathbf{x}_e)\psi_T(\mathbf{x}_T)$, where $e$ refers to the electron and $T$ to the test particle, and where $\psi_T(\mathbf{x}_T)$ is a wave packet centered on position $T$ in the figure. There is of course the extra assumption that the entire system is in a pure state.

Then we are inclined to ask, does the wave packet of the test particle move up or down, as if the electron were in one beam or the other, or does it stay put, in response to the electric field produced by $\rho_c$? As it turns out, this is not exactly the right question, because although the test particle has a wave function (a wave packet) at the initial time, once the interaction begins the two-particle wave function can no longer be represented as a product of two single-particle wave functions. As we say, the two particles become "entangled." That is, the concept of the wave function of the test particle, by itself, loses meaning. See Prob. 3.3 for more on this point. It does make sense, however, to talk about the probability density of the test particle in space, and we can ask whether that moves in one direction or the other, or whether it stays put. The answer is that it does neither; it splits into two parts, with one going up and the other down. The probability distribution of the test particle responds to the probability distribution of the electron, but that tells us nothing about what an actual test particle will do. That is, if we measure the velocity of the test particle after some time, we will find that it has some probability of pointing upward, and some of pointing downward (50% each, in the symmetrical arrangement of Fig. 1).

What we are seeing here is the first step in the measurement process, in this case, measuring which of the two beams the electron is in. In general, in a measurement process, the system interacts with the measuring apparatus, causing the combined quantum mechanical system of system plus measuring apparatus to become entangled. This entanglement is roughly similar to correlations of random variables in ordinary statistics, but more subtle because it is quantum statistics, not classical. The combined system remains entangled even if the interaction ceases after some time (for example, when the measuring apparatus is removed).

To pursue line of thinking further would take us deeper into the quantum theory of measurement. Let us instead return to the probability density $\rho$ and current $\mathbf{J}$, and just mention that similar considerations apply to $\mathbf{J}$ as to $\rho$. That is, a moving charge produces a magnetic field, which is in principle measurable. If the time scale for the measurement is long compared to the time scale of the system being measured, then the magnetic field is the one produced by the ensemble average of the system, that is, by $\mathbf{J}_c$. On shorter time scales, it will be the field produced by an individual member of the ensemble, with probabilities determined by the rules of quantum mechanics.

An example we will see later in the course in which one electron is influenced by the average field of other electrons is the Hartree approximation in many-electron atoms. See Notes 31.

To summarize with a simple rule, it sometimes leads to physical insight to view $\rho_c$ and $\mathbf{J}_c$ as the sources of electric and magnetic fields in individual systems. A detailed understanding of this fact must take into account adiabatic considerations (that is, time scales) and the measurement postulates of quantum mechanics.

## 16. The Ehrenfest Relations

Let us continue with a single particle moving in 3-dimensional space under the influence of the potential $V(\mathbf{x}, t)$. The Hamiltonian is Eq. (47). Let us work out the Heisenberg equations of motion (22) for the position $\mathbf{x}$ and momentum $\mathbf{p}$. Neither of these operators has any explicit time dependence, so the equations of motion are

$$\dot{\mathbf{x}} = -\frac{i}{\hbar}[\mathbf{x}, H],$$
$$\dot{\mathbf{p}} = -\frac{i}{\hbar}[\mathbf{p}, H].$$

(61)

We omit the $H$-subscripts, but these operators are in the Heisenberg picture. The commutators can be evaluated with the help of the results of Prob. 4.4.

The result is

$$\dot{\mathbf{x}} = \frac{\mathbf{p}}{m},$$

(62a)

$$\dot{\mathbf{p}} = -\nabla V(\mathbf{x}, t).$$

(62b)

These are operator versions of Newton's laws (46). Now we take the expectation value of Eq. (62a) with respect to some state $|\psi\rangle$, remembering that expectation values are the same in the Heisenberg and Schrödinger pictures. Here we attach subscripts $H$ and $S$ to indicate the picture. We have

$$\langle\psi_H|\frac{d\mathbf{x}_H}{dt}|\psi_H\rangle = \frac{d}{dt}\langle\psi_H|\mathbf{x}_H(t)|\psi_H\rangle = \frac{d}{dt}\langle\psi_S(t)|\mathbf{x}_S|\psi_S(t)\rangle$$
$$= \frac{1}{m}\langle\psi_H|\mathbf{p}_H(t)|\psi_H\rangle = \frac{1}{m}\langle\psi_S(t)|\mathbf{p}_S|\psi_S(t)\rangle,$$

(63)

where we use the fact that the ket $|\psi_H\rangle$ and the operators $\mathbf{x}_S$ and $\mathbf{p}_S$ are independent of time. A similar calculation applies to Eq. (62b). Then the expectation values of Eqs. (62) become

$$\frac{d\langle\mathbf{x}\rangle}{dt} = \frac{1}{m}\langle\mathbf{p}\rangle,$$
$$\frac{d\langle\mathbf{p}\rangle}{dt} = -\langle\nabla V(\mathbf{x}, t)\rangle,$$

(64)

where we are now in the Schrödinger picture and have dropped the $S$ subscripts. These are versions of the classical Hamilton's equations, applied to expectation values. Eliminating $\langle\mathbf{p}\rangle$ from them, we have

$$m\frac{d^2\langle\mathbf{x}\rangle}{dt^2} = -\langle\nabla V(\mathbf{x}, t)\rangle,$$

(65)

an expectation value version of the classical Newton's equations (46). Equations (64) or (65) are the *Ehrenfest relations* for a particle moving in a potential in 3-dimensional space.

The Ehrenfest relations are exact, and they make no assumption about the state $|\psi\rangle$ that is used for the expectation values. In particular, it need not be a wave packet. But these equations do not say that the expectation value of $\mathbf{x}$ follows the classical motion. That would be true if the right hand side of Eq. (65) were $-\nabla V(\langle\mathbf{x}\rangle, t)$ instead of $-\langle\nabla V(\mathbf{x}, t)\rangle$.

On the other hand, suppose the wave function $\psi(\mathbf{x})$ is localized about its expectation value $\mathbf{x}_0 = \langle \mathbf{x} \rangle$, that is, suppose it is a wave packet, and suppose $f(\mathbf{x})$ is a function that is slowly varying on the scale length of the width of $\psi$. Both $\psi$ and $f$ may depend on time, but we suppress the time dependence. Then

$$\langle f(\mathbf{x}) \rangle = \int d^3\mathbf{x} \, |\psi(\mathbf{x})|^2 f(\mathbf{x})$$

$$\approx \int d^3\mathbf{x} \, |\psi(\mathbf{x})|^2 \Big[ f(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0) \cdot \nabla f(\mathbf{x}_0) + \dots \Big] = f(\mathbf{x}_0) = f(\langle \mathbf{x} \rangle), \tag{66}$$

since the first order term in the Taylor series vanishes upon integration. Identifying $f$ with one of the components of $\nabla V$, we see that $-\langle \nabla V(\mathbf{x}, t) \rangle$ is approximately equal to $-\nabla V(\langle \mathbf{x} \rangle, t)$, under the conditions stated. In other words, the expectation value $\langle \mathbf{x} \rangle$ of a wave packet moving in a slowly varying potential does indeed approximately follow the classical motion. The corrections to Eq. (66) that have been omitted involve the second derivatives of $f$.

It should be noted that the width of the wave packet is not constant in time, and that in general, wave packets spread, so that a wave function that is initially localized in configuration space will not stay that way. When the width of the wave function becomes comparable to the scale length of the potential, the expectation values no longer follow the classical motion even approximately.

In the special case that $f$ in Eq. (66) is a linear function of $\mathbf{x}$, the correction terms all vanish, and Eq. (66) is exact. Note that the force $-\nabla V$ is linear in $\mathbf{x}$ when the potential $V$ itself is quadratic in $\mathbf{x}$. We see that the expectation value of the position of the particle $\langle \mathbf{x} \rangle$ follows the classical motion in potentials that are quadratic functions of the coordinates. Such potentials include the free particle, the particle in a uniform electric or gravitational field, the harmonic oscillator, and other examples. For such potentials, the result holds regardless of the nature of the wave function (it need not be a wave packet).

One can show that the expectation values follow the classical motion more generally when the Hamiltonian is any quadratic function of $\mathbf{x}$ and $\mathbf{p}$. In addition to the systems mentioned, this includes a charged particle moving in a uniform magnetic field.

The Ehrenfest relations provide one way of understanding how the classical limit emerges from quantum mechanics. Another way involves WKB theory, which we will take up in Notes 7.

## 17. Particles in Magnetic Fields

Let us now consider a single particle of charge $q$ moving in an electromagnetic field, which for generality we allow to be time-dependent. The fields are given in terms of the scalar potential $\Phi$ and the vector potential $\mathbf{A}$ by

$$\mathbf{E} = -\nabla\Phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \tag{67a}$$

$$\mathbf{B} = \nabla \times \mathbf{A}. \tag{67b}$$

The classical equations of motion are expressed in terms of $\mathbf{E}$ and $\mathbf{B}$,

$$m\mathbf{a} = q\Big[\mathbf{E}(\mathbf{x},t) + \frac{1}{c}\mathbf{v}{\times}\mathbf{B}(\mathbf{x},t)\Big], \tag{68}$$

where $\mathbf{v} = \dot{\mathbf{x}}$ and $\mathbf{a} = \ddot{\mathbf{x}}$. But the classical Hamiltonian requires the use of the potentials $\Phi$ and $\mathbf{A}$,

$$H = \frac{1}{2m}\Big[\mathbf{p} - \frac{q}{c}\mathbf{A}(\mathbf{x},t)\Big]^2 + q\Phi(\mathbf{x},t), \tag{69}$$

which is a single-particle version of Eq. (B.86). The momentum $\mathbf{p}$ appearing here is the *canonical momentum*, related to the velocity by

$$\mathbf{p} = m\mathbf{v} + \frac{q}{c}\mathbf{A}(\mathbf{x},t), \tag{70}$$

which is not to be confused with the *kinetic momentum* $\mathbf{p}_{\mathrm{kin}} = m\mathbf{v}$. In the presence of a magnetic field, these two momenta are not the same. See the discussion in Sec. B.11. Because of the relationship (70), the first major term of the Hamiltonian (69) is just the kinetic energy, $(1/2)mv^2$, albeit written in a complicated way.

It is strange that the classical Hamiltonian requires the use of potentials and the canonical momentum, none of which is gauge-invariant. This is, however, a fact. See the discussion in Sec. 15.

To obtain the Hamiltonian for a charged particle moving in an electromagnetic field in quantum mechanics we take over the classical Hamiltonian (68) and reinterpret $\mathbf{x}$ and $\mathbf{p}$ as operators. We call this the *quantization* of the classical Hamiltonian. This makes $\mathbf{p}$ the quantum analog of the canonical momentum in classical mechanics, but, as discussed in Notes 4, that is the momentum that should be interpreted as the generator of translations and the one that corresponds to the operator $-i\hbar\nabla$ in the configuration representation (not the kinetic momentum).

There is one issue that arises in this process, namely the fact that the product $\mathbf{p} \cdot \mathbf{A}$ occurs in the expansion of the kinetic energy. Classically, $\mathbf{p} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{p}$, but in quantum mechanics these are not the same because of the commutation relations $[x_i, p_j] = i\hbar\,\delta_{ij}$. Thus there is an ambiguity in the quantization of the classical Hamiltonian, because the results depend on the order in which the classical expression is written (before the classical $\mathbf{x}$'s and $\mathbf{p}$'s are replaced by quantum operators). We resolve this ambiguity by interpreting the kinetic energy operator as

$$\frac{1}{2m}\Big[\mathbf{p} - \frac{q}{c}\mathbf{A}(\mathbf{x},t)\Big]^2 = \frac{1}{2m}\Big[\mathbf{p} - \frac{q}{c}\mathbf{A}(\mathbf{x},t)\Big] \cdot \Big[\mathbf{p} - \frac{q}{c}\mathbf{A}(\mathbf{x},t)\Big]$$
$$= \frac{p^2}{2m} - \frac{q}{2mc}(\mathbf{p} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{p}) + \frac{q^2}{2mc^2}A^2. \tag{71}$$

By this interpretation, $H$ is Hermitian, as it must be. Notice that in the presence of a magnetic field, $p^2/2m$ is not the kinetic energy, rather the kinetic energy is the entire expression (71).

As in our previous exercise in quantization, it is a guess that the Hamiltonian (69), reinterpreted as a quantum operator, is correct physically. In fact, this Hamiltonian neglects several effects that we shall incorporate later.

Now translating the time-dependent Schrödinger equation into the configuration representation, we obtain the differential equation,

$$\frac{1}{2m}\left[-i\hbar\nabla - \frac{q}{c}\mathbf{A}(\mathbf{x},t)\right]^2\psi(\mathbf{x},t) + q\Phi(\mathbf{x},t)\psi(\mathbf{x},t) = i\hbar\frac{\partial\psi(\mathbf{x},t)}{\partial t}, \tag{72}$$

which generalizes Eq. (49). Explicitly, the differential operator on the left of Eq. (72) is given by

$$\left[-i\hbar\nabla - \frac{q}{c}\mathbf{A}(\mathbf{x},t)\right]^2\psi = -\hbar^2\nabla^2\psi + \frac{i\hbar q}{c}\nabla\cdot(\mathbf{A}\psi) + \frac{i\hbar q}{c}\mathbf{A}\cdot\nabla\psi + \frac{q^2}{c^2}A^2\psi. \tag{73}$$

Similarly, when $\Phi$ and $\mathbf{A}$ are time-independent, we can write down the time-independent Schrödinger equation,

$$\frac{1}{2m}\left[-i\hbar\nabla - \frac{q}{c}\mathbf{A}(\mathbf{x})\right]^2\psi(\mathbf{x}) + q\Phi(\mathbf{x})\psi(\mathbf{x}) = E\psi(\mathbf{x}), \tag{74}$$

an equation for the energy eigenfunction $\psi(\mathbf{x})$ and eigenvalue $E$.

The Schrödinger equation (72) possesses a conserved probability current $\mathbf{J}$. The definition is Eq. (57), the same as in the absence of a magnetic field, but now the velocity operator must be interpreted as

$$\mathbf{v} = \frac{1}{m}\left[\mathbf{p} - \frac{q}{c}\mathbf{A}(\mathbf{x},t)\right] = \frac{1}{m}\left(-i\hbar\nabla - \frac{q}{c}\mathbf{A}(\mathbf{x},t)\right). \tag{75}$$

It will be left as an exercise to show that $\rho = |\psi|^2$ and $\mathbf{J}$ with this definition satisfy the continuity equation (55).

## 18. Gauge Transformations in Quantum Mechanics

The potentials $\Phi$ and $\mathbf{A}$ that give $\mathbf{E}$ and $\mathbf{B}$ through Eqs. (67) are not unique, rather they can be subjected to a *gauge transformation*, specified by a scalar field $f$ that we call the *gauge scalar*. The new (primed, or gauge-transformed) potentials are given in terms of the old ones by

$$\Phi' = \Phi - \frac{1}{c}\frac{\partial f}{\partial t}, \tag{76a}$$

$$\mathbf{A}' = \mathbf{A} + \nabla f. \tag{76b}$$

The electric and magnetic fields, however, do not change under the gauge transformation, that is, $\mathbf{E}' = \mathbf{E}$ and $\mathbf{B}' = \mathbf{B}$. Thus we might say that the potentials $\Phi$ and $\mathbf{A}$ contain a physical part (because the physical field $\mathbf{E}$ and $\mathbf{B}$ can be computed from them), and a nonphysical part (the part that changes under a gauge transformation). Since the Hamiltonian (69) is expressed in terms of $\Phi$ and $\mathbf{A}$, we must understand how it and the Schrödinger equation change under a gauge transformation.

It turns out that when we change the gauge of $\Phi$ and $\mathbf{A}$, we must also change the wave function $\psi(\mathbf{x},t)$. In fact, $\psi$ changes by a space- and time-dependent phase factor, related to the gauge scalar $f$,

$$\psi' = e^{iqf/\hbar c}\psi. \tag{77}$$

With this substitution we have

$$\left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)\psi = \left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)e^{-iqf/\hbar c}\,\psi' = e^{-iqf/\hbar c}\left(\mathbf{p} - \frac{q}{c}\mathbf{A} - \frac{q}{c}\nabla f\right)\psi'$$

$$= e^{-iqf/\hbar c}\left(\mathbf{p} - \frac{q}{c}\mathbf{A}'\right)\psi', \tag{78}$$

so

$$\frac{1}{2m}\left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)^2\psi = e^{-iqf/\hbar c}\frac{1}{2m}\left(\mathbf{p} - \frac{q}{c}\mathbf{A}'\right)^2\psi'. \tag{79}$$

Similarly,

$$\left(i\hbar\frac{\partial}{\partial t} - q\Phi\right)\psi = \left(i\hbar\frac{\partial}{\partial t} - q\Phi\right)e^{-iqf/\hbar c}\psi' = e^{-iqf/\hbar c}\left(i\hbar\frac{\partial}{\partial t} - q\Phi + \frac{q}{c}\frac{\partial f}{\partial t}\right)\psi'$$

$$= e^{-iqf/\hbar c}\left(i\hbar\frac{\partial}{\partial t} - q\Phi'\right)\psi', \tag{80}$$

so when we transform the Schrödinger equation (72) to the new gauge, a common factor of $e^{-iqf/\hbar c}$ cancels, and we have

$$\frac{1}{2m}\left(-i\hbar\nabla - \frac{q}{c}\mathbf{A}'\right)^2\psi' + q\Phi'\psi' = i\hbar\frac{\partial\psi'}{\partial t}. \tag{81}$$

This is a primed version of Eq. (72). We see that if the wave function transforms according to Eq. (77), then the Schrödinger equation maintains its form under a gauge transformation.

The Schrödinger wave function $\psi(\mathbf{x}, t)$ is not like a classical field, such as the pressure in a fluid. Such classical fields can be measured at a given spatial point at a given time, in principle without difficulty. Likewise, the measurement of the probability density $|\psi(\mathbf{x}, t)|^2$ in quantum mechanics presents in principle no difficulty, since one just accumulates the statistics of position measurements. But the discussion above shows that one cannot measure the phase of $\psi(\mathbf{x}, t)$ unless one is committed to a certain gauge convention (a choice of $\mathbf{A}$) for whatever magnetic field is present. In principle this applies even in the case $\mathbf{B} = 0$, although here the usual gauge $\mathbf{A} = 0$ would be the obvious choice. As for $\Phi$, it affects the manner in which the phase of $\psi$ depends on time.

**Problems**

**1.** Consider the Hamiltonian for a particle of charge $q$ in an electromagnetic field:

$$H = \frac{1}{2m}\left[\mathbf{p} - \frac{q}{c}\mathbf{A}(\mathbf{x}, t)\right]^2 + q\Phi(\mathbf{x}, t), \tag{82}$$

where we allow all fields to be space- and time-dependent.

**(a)** Define the kinetic momentum operator $\boldsymbol{\pi}$ by

$$\boldsymbol{\pi} = \mathbf{p} - \frac{q}{c}\mathbf{A}(\mathbf{x}, t). \tag{83}$$

Notice that $\boldsymbol{\pi}$ has an explicit time dependence, due to the $\mathbf{A}$ term. Work out the commutation relations, $[x_i, \pi_j]$ and $[\pi_i, \pi_j]$. Use these to work out the Heisenberg equations of motion for the

operators $\mathbf{x}$, $\boldsymbol{\pi}$. Then eliminate $\boldsymbol{\pi}$, and find an expression for $\ddot{\mathbf{x}}$ (the Heisenberg analog of the Newton-Lorentz equations).

**(b)** By taking expectation values, show that if $\mathbf{B}$ is uniform in space and $\mathbf{E}$ has the form,

$$E_i(\mathbf{x}) = a_i + \sum_j b_{ij}\, x_j, \tag{84}$$

where $a_i$ and $b_{ij}$ are constants, then the expectation value of $\mathbf{x}$ follows the classical orbit.

**2.** Consider a particle of charge $q$ moving in an electric and magnetic field. The Hamiltonian is (69). Let $\rho(\mathbf{x}, t) = |\psi(\mathbf{x}, t)|^2$, and let $\mathbf{J}$ be given by Eq. (57), where the velocity operator $\mathbf{v}$ is defined by Eq. (75). Show that $\rho$ and $\mathbf{J}$ satisfy the continuity equation.

**3.** In this problem we consider the spreading of wave packets of a free particle in one dimension. The Hamiltonian is

$$H = \frac{p^2}{2m}. \tag{85}$$

**(a)** Let the wave function at $t = 0$ be

$$\psi_0(x) = \psi(x, 0) = Ce^{-x^2/4L^2}, \tag{86}$$

where $C$ and $L$ are constants. Assuming $C$ is real and positive, normalize the wave function and find $C$. See Appendix C. Then determine $\langle x \rangle$, $\langle p \rangle$ and $\Delta x$, the last one being the dispersion (or standard deviation), see Eq. (2.40).

**(b)** Compute the initial momentum space wave function $\phi_0(p)$, and from that compute $\Delta p$ (do the latter calculation in momentum space). Compare $\Delta x \Delta p$ to the minimum value allowed by the uncertainty principle.

**(c)** Now compute $\phi(p, t)$ for any time $t$. Use it to find $\langle p \rangle$ and $\Delta p$ for any time $t$.

**(d)** Now compute $\psi(x, t)$ for any time $t$. Express your answer in terms of the quantity,

$$D = L + \frac{it\hbar}{2mL} \tag{87}$$

(a useful abbreviation). Then find $\langle x \rangle$ and $\Delta x$ for any time $t$. Notice what happens to the product $\Delta x \Delta p$ for times $t > 0$.

**(e)** If you had not done a detailed calculation you could estimate $\Delta x$ as a function of $t$ by using the uncertainty principle. Do this and compare to the results of the detailed calculation.

**4.** Show that if $A$ and $B$ are constants of motion with no explicit time dependence, then so is $[A, B]$. See Sec. 10.