2

Reminder: classical and quantum theory

Summary Action principle. Functional derivatives. Lagrangian and Hamiltonian formalisms. Canonical quantization. Operators and vectors in Hilbert space. Schrödinger equation.

2.1 Lagrangian formalism

Quantum theory is built by applying a quantization procedure to a classical theory. The starting point of a classical theory is the action principle.

The action principle The state of a classical system is described by a (set of) generalized coordinate(s) $q(t)$. The variable q is a "symbolic combined notation," which includes all degrees of freedom of the system. For a system with N degrees of freedom we have $q \equiv \{q_1, q_2, \ldots q_N\}$. A field theory describes systems with an infinite number of degrees of freedom. For example, the configuration of a classical real scalar field at a given moment of time is characterized by a function $\phi(\mathbf{x})$. In this case $q(t) \equiv \{\phi_{\mathbf{x}}(t)\}$, where the spatial coordinate $\mathbf{x}$ can be considered as an "index" enumerating the infinite number of degrees of freedom of the scalar field.

The classical trajectory $q(t)$ connecting the states at two moments of time t_1 and t_2 is an extremum of an action functional[1]

$$S[q(t)] = \int_{t_1}^{t_2} L(t, q(t), \dot{q}(t), \ldots)\, dt. \tag{2.1}$$

The function $L(t, q, \dot{q}, \ldots)$ is called the *Lagrangian*. In the case of a field theory, the Lagrangian is a functional too. Different Lagrangians describe different systems. For a harmonic oscillator with unit mass and frequency ω we have

$$L(q, \dot{q}) = \frac{1}{2}\left(\dot{q}^2 - \omega^2 q^2\right). \tag{2.2}$$

[1] See Appendix A1.1 for more details concerning functionals.

If ω is time independent this Lagrangian has no explicit time dependence.

Equations of motion The requirement that the function $q(t)$ extremizes the action leads to a differential equation for $q(t)$. Let us derive this equation for the action

$$S[q] = \int_{t_1}^{t_2} L(t, q, \dot{q})\, dt. \tag{2.3}$$

If the function $q(t)$ is an extremum of the action functional (2.3), then a small variation $\delta q(t)$ of the trajectory $q(t)$ changes the value of $S[q]$ only by terms which are quadratic in $\delta q(t)$. In other words, the variation

$$\delta S[q, \delta q] \equiv S[q + \delta q] - S[q]$$

does not contain any first-order terms in δq. To obtain the differential equation for $q(t)$, let us compute δS:

$$\begin{aligned}
\delta S[q; \delta q] &= S[q(t) + \delta q(t)] - S[q(t)] \\
&= \int_{t_1}^{t_2} \left[\frac{\partial L(t, q, \dot{q})}{\partial q} \delta q(t) + \frac{\partial L(t, q, \dot{q})}{\partial \dot{q}} \delta \dot{q}(t) \right] dt + O\left(\delta q^2\right) \\
&= \delta q(t) \frac{\partial L}{\partial \dot{q}} \Big|_{t_1}^{t_2} + \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right] \delta q(t) dt + O\left(\delta q^2\right).
\end{aligned} \tag{2.4}$$

If we impose the boundary conditions $q(t_1) = q_1$ and $q(t_2) = q_2$ and require that the perturbed trajectory also satisfy them, then $\delta q(t_1) = \delta q(t_2) = 0$. Hence the boundary terms in equation (2.4) vanish and we obtain

$$\delta S = \int_{t_1}^{t_2} \left[\frac{\partial L(t, q, \dot{q})}{\partial q} - \frac{d}{dt} \frac{\partial L(t, q, \dot{q})}{\partial \dot{q}} \right] \delta q(t) dt + O\left(\delta q^2\right). \tag{2.5}$$

The requirement that δS is second order in δq means that the first-order terms should vanish for any $\delta q(t)$. This is possible only if the expression inside the square brackets in equation (2.5) vanishes. Thus we obtain the *Euler–Lagrange equation*

$$\frac{\partial L(t, q, \dot{q})}{\partial q} - \frac{d}{dt} \frac{\partial L(t, q, \dot{q})}{\partial \dot{q}} = 0, \tag{2.6}$$

which is the classical equation of motion for a system with the Lagrangian $L(t, q, \dot{q})$. Note that in a field theory, where $L(t, q, \dot{q})$ is also a functional, the derivatives with respect to q and $\dot{q}$ in equation (2.6) are replaced by the functional derivatives (see below).

Example 2.1 For the harmonic oscillator with Lagrangian (2.2), the Euler–Lagrange equation reduces to

$$\ddot{q} + \omega^2 q = 0. \tag{2.7}$$

Generally the path $q(t)$ that extremizes the action and satisfies the boundary conditions is unique. However, there are cases when the extremum is not unique, and sometimes it does not even exist.

Exercise 2.1
Find the solution of equation (2.7) satisfying the boundary conditions $q(t_1) = q_1$ and $q(t_2) = q_2$. Determine when this solution is unique.

2.1.1 Functional derivatives

The variation of a functional can always be written in the form

$$\delta S = \int \frac{\delta S}{\delta q(t)} \delta q(t) dt + O\left(\delta q^2\right). \tag{2.8}$$

The expression denoted by $\delta S/\delta q(t)$ in (2.8) is called the *functional derivative* (or the *variational derivative*) of $S[q]$ with respect to $q(t)$.

Comparing the definition in (2.8) with (2.5), we find that the functional derivative of the functional $S[q]$ given in (2.3) is

$$\frac{\delta S}{\delta q(t)} = \frac{\partial L(t, q, \dot{q})}{\partial q} - \frac{d}{dt} \frac{\partial L(t, q, \dot{q})}{\partial \dot{q}}. \tag{2.9}$$

Example 2.2 For a harmonic oscillator with the Lagrangian (2.2) we have

$$\frac{\delta S}{\delta q(t)} = -\omega^2 q(t) - \ddot{q}(t). \tag{2.10}$$

It is important to keep track of the argument t in the functional derivative $\delta S/\delta q(t)$. A functional $S[q]$ generally depends on the values of q at all moments of time t. Discretizing the time interval $t_2 > t > t_1$ as $t_k = t_1 + \varepsilon k$, where k is integer, we may approximate the function $q(t)$ by its values at points t_i. Then the functional $S[q(t)]$ can be visualized as a function of many variables $q_k \equiv q(t_k)$,

$$S[q(t)] = \text{“}S(q_1, q_2, q_3, \ldots)\text{”}.$$

In the limit $\varepsilon \to 0$, the properly normalized partial derivative of this “function” with respect to one of its arguments, say $q_1 \equiv q(t_1)$, becomes the functional derivative $\delta S/\delta q(t_1)$. Clearly the derivative $\delta S/\delta q(t_1)$ is in general different from $\delta S/\delta q(t_2)$ and therefore $\delta S/\delta q(t)$ is a function of t.

For a functional $S[\phi]$ of a field $\phi(\mathbf{x}, t)$, the functional derivative with respect to $\phi(\mathbf{x}, t)$ retains both the arguments $\mathbf{x}$ and t and is written as $\delta S/\delta \phi(\mathbf{x}, t)$.

To calculate the functional derivatives, one has to convert the functionals to an integral form. When the original functional does not contain any integration, the Dirac δ function must be used. (See Appendix A1.1 to recall the definition

and the properties of the δ function.) Below we demonstrate how the functional derivative is calculated in a few simple cases.

Example 2.3 For the functional

$$A[q] \equiv \int q^3 dt$$

the functional derivative is

$$\frac{\delta A[q]}{\delta q(t_1)} = 3q^2(t_1).$$

Example 2.4 The functional

$$\begin{aligned} B[q] &\equiv 3\sqrt{q(1)} + \sin[q(2)] \\ &= \int \left[3\delta(t-1)\sqrt{q(t)} + \delta(t-2)\sin q(t)\right] dt \end{aligned}$$

has the functional derivative

$$\frac{\delta B[q]}{\delta q(t)} = \frac{3\delta(t-1)}{2\sqrt{q(1)}} + \delta(t-2)\cos[q(2)].$$

Example 2.5 For the functional

$$S[\phi] = \frac{1}{2}\int d^3\mathbf{x}\, dt (\nabla\phi)^2,$$

which depends on a field $\phi(\mathbf{x}, t)$, the functional derivative is found after an integration by parts:

$$\frac{\delta S[\phi]}{\delta \phi(\mathbf{x}, t)} = -\Delta\phi(\mathbf{x}, t).$$

The boundary terms have been omitted because the integration in $S[\phi]$ is performed over the entire spacetime and the field ϕ is assumed to decay sufficiently rapidly at infinity.

Remark: alternative definition The functional derivative may equivalently be defined with the help of the δ function:

$$\frac{\delta A[q]}{\delta q(t_1)} = \frac{d}{ds}\bigg|_{s=0} A[q(t) + s\delta(t-t_1)].$$

One can prove that this definition is equivalent to (2.8).

Because the δ function is a distribution, the definition above can be understood in a more rigorous way as

$$\frac{\delta A[q]}{\delta q(t_1)} = \lim_{n\to\infty} \frac{d}{ds}\bigg|_{s=0} A[q_n(t)],$$

where $q_n(t)$, $n = 1, 2, \ldots$ is a sequence of functions that converges to $q(t) + s\delta(t - t_1)$ in the sense of distributions.

Second functional derivative A derivative of a functional is a functional again. Therefore we can define the second functional derivative via

$$\frac{\delta^2 S}{\delta q(t_1)\,\delta q(t_2)} \equiv \frac{\delta}{\delta q(t_2)} \left\{ \frac{\delta S}{\delta q(t_1)} \right\}.$$

Exercise 2.2
Calculate the second functional derivative

$$\frac{\delta^2 S[q]}{\delta q(t_1)\,\delta q(t_2)}$$

for a harmonic oscillator with the action

$$S[q] = \int \frac{1}{2}\left(\dot{q}^2 - \omega^2 q^2\right) dt.$$

2.2 Hamiltonian formalism

The starting point of a canonical quantum theory is a classical theory in the Hamiltonian formulation. The Hamiltonian formalism is based on the Legendre transform of the Lagrangian $L(t, q, \dot{q})$ with respect to the velocity $\dot{q}$.

Legendre transform Given a function $f(x)$, one can introduce a new variable p instead of x,

$$p \equiv \frac{df}{dx}, \tag{2.11}$$

and map the function $f(x)$ to a new function

$$g(p) \equiv px(p) - f(x(p)).$$

Here we imply that x has been expressed through p using (2.11); the function $f(x)$ must be such that p, which is the slope of $f(x)$, is uniquely related to x. The function $g(p)$ is called the *Legendre transform* of $f(x)$. A nice property of the Legendre transform is that when applied to the function $g(p)$, it restores the old variable $x = dg(p)/dp$ and the old function $f(x)$.

The Hamiltonian To define the Hamiltonian, one performs the Legendre transform of the Lagrangian $L(t, q, \dot{q})$ replacing $\dot{q}$ by a new variable

$$p = \frac{\partial L(t, q, \dot{q})}{\partial \dot{q}}, \tag{2.12}$$

called the *canonical momentum*. The variables t and q remain as parameters and the ubiquitously used notation $\partial/\partial\dot{q}$ means simply the partial derivative of $L(t, q, \dot{q})$ with respect to its third argument.

Remark If the coordinate q is a multi-dimensional vector, $q \equiv q_j$, the Legendre transform is performed with respect to each velocity $\dot{q}_j$ and the momentum vector p_j is introduced. In field theory there is a continuous set of "coordinates," so we need to use a functional derivative when defining the momenta.

Assuming that equation (2.12) can be solved for the velocity $\dot{q}$ as a function of t, q and p,

$$\dot{q} = v(p; q, t), \tag{2.13}$$

one defines the *Hamiltonian* $H(p, q, t)$ by

$$H(p, q, t) \equiv [p\dot{q} - L(t, q, \dot{q})]_{\dot{q}=v(p;q,t)}. \tag{2.14}$$

In the above expression, $\dot{q}$ is replaced by the function $v(p; q, t)$.

Remark: the existence of the Legendre transform The possibility of performing the Legendre transform hinges on the invertibility of equation (2.12) which requires that the Lagrangian $L(t, q, \dot{q})$ should be a suitably nondegenerate function of the velocity $\dot{q}$. Many physically important theories with constraints, such as gauge theories or Einstein's general relativity, are described by Lagrangians that do not admit an immediate Legendre transform in the velocities. In those cases (not considered in this book) a more complicated formalism is needed to obtain an adequate Hamiltonian formulation of the theory.

The Hamilton equations of motion The Euler–Lagrange equations (2.6) are second-order differential equations for $q(t)$. In the Hamiltonian formalism they are replaced by first-order differential equations for the variables $q(t)$ and $p(t)$. Due to the definition (2.12), we can recast (2.6) as

$$\frac{dp}{dt} = \left.\frac{\partial L(t, q, \dot{q})}{\partial q}\right|_{\dot{q}=v(p;q,t)}, \tag{2.15}$$

where the substitution $\dot{q} = v$ must be carried out after the differentiation $\partial L/\partial q$. The other equation is (2.13),

$$\frac{dq}{dt} = v(p; q, t). \tag{2.16}$$

Using the Hamiltonian $H(p, q, t)$, defined in (2.14), we can rewrite the above equations in a more symmetrical form. As a result of straightforward algebra, one has

$$\frac{\partial H}{\partial q} = \frac{\partial}{\partial q}(pv - L) = p\frac{\partial v}{\partial q} - \frac{\partial L}{\partial q} - \frac{\partial L}{\partial \dot{q}}\frac{\partial v}{\partial q} = -\frac{\partial L}{\partial q},$$

$$\frac{\partial H}{\partial p} = \frac{\partial}{\partial p}(pv - L) = v + p\frac{\partial v}{\partial p} - \frac{\partial L}{\partial \dot{q}}\frac{\partial v}{\partial p} = v,$$

and hence equations (2.15)–(2.16) become

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q}. \tag{2.17}$$

These are the well-known Hamilton equations of motion.

Example 2.6 For a harmonic oscillator described by the Lagrangian (2.2), the canonical momentum is equal to $p = \dot{q}$ and the corresponding Hamiltonian is

$$H(p, q) = p\dot{q} - L = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2. \tag{2.18}$$

The Hamilton equations are then

$$\dot{q} = p, \quad \dot{p} = -\omega^2 q.$$

The action principle The Hamilton equations follow from the action principle

$$S_H[q(t), p(t)] = \int [p\dot{q} - H(p, q, t)]\, dt, \tag{2.19}$$

where the *Hamiltonian action* S_H is a functional of two functions $q(t)$ and $p(t)$ which must be varied independently to extremize S_H.

Exercise 2.3

(a) Derive equations (2.17) by extremizing the action (2.19). Determine the boundary conditions which must be imposed on $p(t)$ and $q(t)$.
(b) Verify that the Hamilton equations imply $dH/dt = 0$ when $H(p, q)$ does not depend explicitly on the time t.
(c) Show that the expression $p\dot{q} - H$ evaluated for the classical trajectory (on-shell) is equal to the Lagrangian $L(q, \dot{q}, t)$.

2.3 Quantization of Hamiltonian systems

In quantum theory we replace the canonical variables $q(t)$, $p(t)$ by noncommuting operators $\hat{q}(t)$, $\hat{p}(t)$ for which one postulates the equal-time commutation relation

$$[\hat{q}(t), \hat{p}(t)] \equiv \hat{q}(t)\hat{p}(t) - \hat{p}(t)\hat{q}(t) = i\hbar\hat{1}. \tag{2.20}$$

(We shall frequently omit the identity operator $\hat{1}$ in formulae below.) Relation (2.20) reflects the impossibility of simultaneously measuring the coordinate and the momentum with unlimited accuracy (Heisenberg's uncertainty relation). Note that the commutation relations for unequal times, for instance $[\hat{q}(t_1), \hat{p}(t_2)]$, are not postulated but are *derived* using the equations of motion.

Exercise 2.4

Simplify the expression $\hat{q}\hat{p}^2\hat{q} - \hat{p}^2\hat{q}^2$ using the commutation relation (2.20).

The problem of operator ordering Consider the following classical Hamiltonian $H(p, q) = 2p^2q$. Since $\hat{p}\hat{q} \neq \hat{q}\hat{p}$, it is not a priori clear whether the corresponding quantum Hamiltonian should be $\hat{p}^2\hat{q} + \hat{q}\hat{p}^2$, or perhaps $2\hat{p}\hat{q}\hat{p}$, or some other combination of the noncommuting operators $\hat{p}$ and $\hat{q}$. The difference between the possible Hamiltonians is of order $\hbar$ or higher, so the classical limit $\hbar \to 0$ is the same for any choice of the operator ordering. The ambiguity in the choice of the quantum Hamiltonian is called the *operator ordering problem.*

The operator ordering needs to be chosen by hand in each case when it is not unique. In principle, only a precise measurement of quantum effects could unambiguously determine the correct operator ordering in such cases.

Remark For frequently used Hamiltonians of the form

$$H(p, q) = \frac{1}{2m}p^2 + U(q)$$

there is no operator ordering problem.

2.4 Hilbert spaces and Dirac notation

The non commuting operators $\hat{q}$ and $\hat{p}$ can be represented as linear transformations ("matrices") in a suitable vector space (the space of *quantum states*). Since relation (2.20) cannot be satisfied by any finite-dimensional matrices,[2] the corresponding space of quantum states is necessarily infinite-dimensional and must be defined over the field $\mathbb{C}$ of complex numbers.

Infinite-dimensional vector spaces A vector in a *finite*-dimensional space can be visualized as a collection of components, e.g. $\vec{a} \equiv (a_1, a_2, a_3, a_4)$, where each a_k is a (complex) number. A vector in infinite-dimensional space has infinitely many components. An important example of an infinite-dimensional complex vector space is the space L^2 of square-integrable complex-valued functions $\psi(q)$ for which the integral

$$\int_{-\infty}^{+\infty} |\psi(q)|^2 \, dq$$

converges. One can check that a linear combination of two such functions, $\lambda_1\psi_1(q) + \lambda_2\psi_2(q)$, with constant coefficients λ_1 and $\lambda_2 \in \mathbb{C}$, is again an element of L^2. A function $\psi \in L^2$ can be thought of as a set of infinitely many "components" $\psi_q \equiv \psi(q)$, where the "index" q is continuous.

It turns out that the space of quantum states of a system with one degree of freedom is exactly the space of square-integrable functions $\psi(q)$, where q is a

[2] This is easy to prove by considering the trace of a commutator. If $\hat{A}$ and $\hat{B}$ are arbitrary finite-dimensional matrices, then $\mathrm{Tr}\,[\hat{A}, \hat{B}] = \mathrm{Tr}\hat{A}\hat{B} - \mathrm{Tr}\hat{B}\hat{A} = 0$ which contradicts equation (2.20). In an infinite-dimensional space, this reasoning no longer holds because the trace is not well-defined for an arbitrary operator.

generalized coordinate of a system (for example, position of a particle moving in one dimensional space). In this case the function $\psi(q)$ is called the *wave function.* In the case of two degrees of freedom the wave function depends on both coordinates q_1 and q_2 characterizing the state of the corresponding classical system, $\psi = \psi(q_1, q_2)$. In quantum field theory, the "coordinates" are field configurations $\phi(\mathbf{x}) \equiv \phi_{\mathbf{x}}$ and the wave function depends on infinitely many "coordinates" $\phi_{\mathbf{x}}$; in other words, it is a functional, $\psi[\phi(\mathbf{x})]$.

The Dirac notation Linear algebra is used in many areas of physics, and the Dirac notation is a convenient shorthand for calculations in both finite- and infinite-dimensional vector spaces.

To denote vectors in abstract linear space, Dirac proposed to use symbols such as $|a\rangle$, $|x\rangle$, $|\lambda\rangle$ ..., which he called "ket"-vectors. Then the symbol $2|v\rangle - 3i|w\rangle$, for example, denotes a linear combination of the vectors $|v\rangle$ and $|w\rangle$.

Dual space In a vector space V one can define linear forms, which act on a vector to produce a (complex) number; $f: V \to \mathbb{C}$. A linear form is called *covector* or "bra"-vector and denoted by $\langle f|$. A complex number produced by linear form $\langle f|$ as a result of acting on a vector $|v\rangle$ is denoted by $\langle f|v\rangle$ (the mnemonic rule is: "bra"-vector acting on "ket"-vector makes a "bracket", which is a complex number). The fact that a form is linear means that

$$\langle f|(\alpha|v\rangle + \beta|w\rangle) = \alpha\langle f|v\rangle + \beta\langle f|w\rangle,$$

where α and β are arbitrary complex numbers. In the space of all linear forms one can define the multiplication by a complex number and a sum of two linear forms in a natural way as

$$(\alpha\langle f| + \beta\langle g|)|v\rangle \equiv \alpha\langle f|v\rangle + \beta\langle g|v\rangle,$$

valid for any vector $|v\rangle$. Then the space of all linear forms becomes a vector space called the *dual space.* At this stage one still has to distinguish the dual space from the original linear space because they have a different mathematical origin. For the reader familiar with differential geometry we point out that the dual space is an analog of the space of one-forms.

Hilbert space A linear space of quantum states must possess an extra structure, namely, a *Hermitian* scalar product. A scalar product maps any two vectors $|v\rangle$ and $|w\rangle$ into a complex number $(|v\rangle, |w\rangle)$. A complete, separable vector space[3] with a Hermitian scalar product is called a Hilbert space. The Hermitian

[3] A normed vector space is *complete* if all Cauchy sequences converge in it; then all norm-convergent infinite sums always have a unique vector as their limit. The space is *separable* if there exists a countable set of vectors $\{|e_n\rangle\}$ which is dense everywhere in the space. Separability ensures that all vectors can be approximated arbitrarily well by *finite* combinations of the basis vectors.

scalar product satisfies the usual axioms, while the hermiticity means that

$$(|v\rangle, |w\rangle) = (|w\rangle, |v\rangle)^*,$$

where the asterisk * denotes the complex conjugate.

The scalar product allows us to establish a one-to-one correspondence between "bra"- and "ket"-vectors. We say that the "bra"-vector $\langle v|$ corresponds to a "ket" vector $|v\rangle$ if

$$\langle v|w\rangle = (|v\rangle, |w\rangle)$$

for *any* vector $|w\rangle$. The scalar product of two "bra"-vectors $\langle w|$ and $\langle v|$ can then be defined as

$$(\langle w|, \langle v|) \equiv (|v\rangle, |w\rangle),$$

and the dual space also becomes a Hilbert space. Because the original Hilbert space and its dual space are isomorphic, we can "identify" them from now on and consider $|v\rangle$ and $\langle v|$ as simply different symbols designating the same quantum state. The scalar product $(|v\rangle, |w\rangle)$ can then be always written in somewhat more concise form as $\langle v|w\rangle$.

Exercise 2.5
Verify that the "bra"-vector $\alpha^* \langle v| + \beta^* \langle w|$ corresponds to a "ket"-vector $\alpha |v\rangle + \beta |w\rangle$.

2.5 Operators, eigenvalue problem and basis in a Hilbert space

Operators An operator $\hat{A}$ maps a vector space to itself; it transforms a vector $|v\rangle$ to the vector $|w\rangle \equiv \hat{A}|v\rangle$. For example, the identity operator $\hat{1}$ does not change any vectors: $\hat{1}|v\rangle = |v\rangle$. In quantum theory it is enough to consider only linear operators for which

$$\hat{A}(\alpha |v\rangle + \beta |w\rangle) = \alpha \hat{A}|v\rangle + \beta \hat{A}|w\rangle.$$

The product of two operators $\hat{A}$ and $\hat{B}$ is defined in a natural way as

$$\left(\hat{A} \cdot \hat{B}\right)|v\rangle \equiv \hat{A}\left(\hat{B}|v\rangle\right)$$

for any vector $|v\rangle$. Generically, operators do not commute with each other, that is,

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} \neq 0.$$

The notation $\langle v|\hat{A}|w\rangle$ denotes the scalar product of the vectors $|v\rangle$ and $\hat{A}|w\rangle$ and the quantity $\langle v|\hat{A}|w\rangle$ is also called the *matrix element* of the operator $\hat{A}$

with respect to the states $|v\rangle$ and $|w\rangle$. Acting on a "bra"-vector $\langle v|$ the operator $\hat{A}$ produces the "bra"-vector $\langle v|\hat{A}$, such that

$$\left(\langle v|\hat{A}\right)|w\rangle = \langle v|\hat{A}|w\rangle$$

for any $|w\rangle$.

Let us consider a "ket"-vector

$$|g\rangle = \hat{A}|v\rangle .$$

To the "ket"-vectors $|g\rangle$ and $|v\rangle$ there correspond the "bra"-vectors $\langle g|$ and $\langle v|$ respectively. Generically, $\langle g| \neq \langle v|\hat{A}$, and the vectors $\langle g|$ and $\langle v|$ are related by another operator $\hat{A}^\dagger$,

$$\langle g| = \langle v|\hat{A}^\dagger,$$

which is called the *Hermitian conjugate* of operator $\hat{A}$. Because $\langle g|w\rangle = \langle w|g\rangle^*$ we have

$$\langle v|\hat{A}^\dagger|w\rangle = \left(\langle w|\hat{A}|v\rangle\right)^* . \tag{2.21}$$

It is easy to see that the operation of Hermitian conjugation has the following properties:

$$(\hat{A}+\hat{B})^\dagger = \hat{A}^\dagger + \hat{B}^\dagger; \quad (\lambda\hat{A})^\dagger = \lambda^*\hat{A}^\dagger; \quad (\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger.$$

The following subsets of all operators play a particularly important role in a quantum theory: *Hermitian* operators for which $\hat{A}^\dagger = \hat{A}$, *skew-Hermitian* operators satisfying $\hat{B}^\dagger = -\hat{B}$, and *unitary* operators: $\hat{U}^\dagger\hat{U} = \hat{U}\hat{U}^\dagger = \hat{1}$.

Eigenvalue problem Given an operator $\hat{A}$ one can consider the equation

$$\hat{A}|v\rangle = v|v\rangle . \tag{2.22}$$

The vector $|v\rangle \neq 0$ satisfying this equation is called an eigenvector of the operator $\hat{A}$ corresponding to the eigenvalue v.

According to quantum theory the result of any measurement of some quantity corresponding to an operator $\hat{A}$ is always an eigenvalue of this operator. Therefore, observables must be described by operators with real eigenvalues. Setting $|w\rangle = |v\rangle$ in equation (2.21) and assuming that $|v\rangle$ is an eigenvector of an operator $\hat{A}$ we find that if this operator is Hermitian, $\hat{A}^\dagger = \hat{A}$, then its eigenvalues are always real. This motivates an important assumption made in quantum theory: the operators corresponding to all observables are Hermitian.

Remark The operators of position $\hat{q}$ and momentum $\hat{p}$ must be Hermitian, $\hat{q}^\dagger = \hat{q}$ and $\hat{p}^\dagger = \hat{p}$. The commutator of two Hermitian operators $\hat{A}$, $\hat{B}$ is anti-Hermitian: $[\hat{A},\hat{B}]^\dagger = -[\hat{A},\hat{B}]$. Accordingly, the commutator of $\hat{q}$ and $\hat{p}$ contains the imaginary unit i. The operator $\hat{p}\hat{q}$ is neither Hermitian nor skew-Hermitian: $(\hat{p}\hat{q})^\dagger = \hat{q}\hat{p} = \hat{p}\hat{q} + i\hbar\hat{1} \neq \pm\hat{p}\hat{q}$.

Eigenvectors of an Hermitian operator corresponding to different eigenvalues are always orthogonal. This is easy to see: if $|v_1\rangle$ and $|v_2\rangle$ are eigenvectors of an Hermitian operator $\hat{A}$ with eigenvalues v_1 and v_2, then $\langle v_1|\hat{A} = v_1\langle v_1|$ because v_1 is real. Consequently,

$$\langle v_1|\hat{A}|v_2\rangle = v_2\langle v_1|v_2\rangle = v_1\langle v_1|v_2\rangle,$$

and $\langle v_1|v_2\rangle = 0$ if $v_1 \neq v_2$.

Basis in Hilbert space In an N-dimensional vector space, one can find a finite set of linearly independent vectors $|e_1\rangle, \ldots, |e_N\rangle$ and uniquely express any vector $|v\rangle$ as a linear combination of these vectors,

$$|v\rangle = \sum_{n=1}^{N} v_n |e_n\rangle.$$

The coefficients v_n are called the *components* of the vector $|v\rangle$ in the basis $\{|e_n\rangle\}$. In an orthonormal basis satisfying $\langle e_m|e_n\rangle = \delta_{mn}$, the scalar product of two vectors $|v\rangle$, $|w\rangle$ is expressed through their components v_n, w_n as

$$\langle v|w\rangle = \sum_{n=1}^{N} v_n^* w_n.$$

By definition, a vector space is infinite-dimensional if no finite set of vectors can serve as a basis. In that case, one might expect to have an infinite countable basis $|e_1\rangle$, $|e_2\rangle, \ldots$, such that any vector $|v\rangle$ is uniquely expressible as

$$|v\rangle = \sum_{n=1}^{\infty} v_n |e_n\rangle. \tag{2.23}$$

However, the convergence of this infinite series is a nontrivial issue. For instance, if the basis vectors $|e_n\rangle$ are orthonormal, then the norm of the vector $|v\rangle$ is

$$\langle v|v\rangle = \left(\sum_{m=1}^{\infty} v_m^* \langle e_m|\right)\left(\sum_{n=1}^{\infty} v_n |e_n\rangle\right) = \sum_{n=1}^{\infty} |v_n|^2. \tag{2.24}$$

This series must converge if the vector $|v\rangle$ has a finite norm. Therefore, for example, the sum $\sum_{n=1}^{\infty} n^2 |e_n\rangle$ does not correspond to a vector that belongs to a Hilbert space. The coefficients v_n must fall off sufficiently rapidly so that the series (2.24) is finite and only in this case is it plausible that the infinite linear combination (2.23) converges and uniquely specifies the vector $|v\rangle$. This does not hold in all infinite-dimensional spaces. However, as we have already mentioned, the required properties, known in functional analysis as completeness and separability, are fulfilled in a Hilbert space. When defining a quantum theory, one always chooses the space of quantum states as a separable Hilbert space. In some instances we have to "enclose" the system inside a large finite box and only

then will a countable basis $\{|e_n\rangle\}$ exist. Once an orthonormal basis is chosen, any vector $|v\rangle$ is unambiguously represented by a collection of its components $(v_1, v_2, \ldots)$. Therefore a separable Hilbert space can be visualized as the space of infinite sets of complex numbers, $|v\rangle \equiv (v_1, v_2, \ldots)$, such that the sums $\sum_{n=1}^{\infty} |v_n|^2$ converge. The convergence guarantees that all scalar products $\langle v|w\rangle = \sum_{n=1}^{\infty} v_n^* w_n$ are finite.

Example 2.7 The space L^2 of square-integrable wave functions $\psi(q)$ defined on an interval $a < q < b$ is a separable Hilbert space, although it may appear to be "much larger" than the space of infinite rows of numbers. The scalar product of two wave functions $\psi_1(q)$ and $\psi_2(q)$ is defined by

$$\langle \psi_1 | \psi_2 \rangle = \int_a^b \psi_1^*(q) \psi_2(q) dq,$$

and the canonical operators $\hat{p}$, $\hat{q}$ can be represented as linear operators in the space L^2 that act on functions $\psi(q)$ as

$$\hat{p} : \psi(q) \to -i\hbar \frac{\partial \psi}{\partial q}, \quad \hat{q} : \psi(q) \to q\psi(q). \tag{2.25}$$

It is straightforward to verify that the commutation relation (2.20) holds.

Remark When one considers a field $\phi(\mathbf{x})$ in an infinite three-dimensional space, the corresponding space of quantum states is not separable. Therefore, to obtain a mathematically consistent theory, we need to enclose the field inside a finite box and impose appropriate boundary conditions.

Decomposition of unity Let $\{|e_n\rangle\}$ be a complete orthonormal basis in a separable Hilbert space. Then any vector $|v\rangle$ can be written as

$$|v\rangle = \sum_{n=1}^{\infty} v_n |e_n\rangle = \sum_{n=1}^{\infty} \langle e_n | v \rangle |e_n\rangle = \left(\sum_{n=1}^{\infty} |e_n\rangle \langle e_n| \right) |v\rangle .$$

Hence, the identity operator $\hat{1}$ can be decomposed as

$$\hat{1} = \sum_{n=1}^{\infty} |e_n\rangle \langle e_n| .$$

The symbol $|e_n\rangle \langle e_n|$ must be interpreted as the operator which acts on a vector $|v\rangle$ according to the rule

$$|v\rangle \to (|e_n\rangle \langle e_n|) |v\rangle \equiv \langle e_n | v \rangle |e_n\rangle .$$

It projects a vector $|v\rangle$ onto the one-dimensional subspace spanned by $|e_n\rangle$. Thus, the identity operator $\hat{1}$ can be written as a sum of projectors onto orthonormal basis vectors.

The choice of basis in Hilbert space corresponds to a choice of a certain representation in quantum theory (for example, coordinate or momentum representation). The decomposition of unity is very useful for establishing the relation between different representations ("coordinate systems").

Remark One must carefully distinguish the symbols $\langle w|v\rangle$ and $|w\rangle\langle v|$. The first one is a number, while the second is the operator which acts on the "ket"- and "bra"-vectors as $(|w\rangle\langle v|)|g\rangle = \langle v|g\rangle|w\rangle$ and $\langle g|(|w\rangle\langle v|) = \langle g|w\rangle\langle v|$ respectively.

2.6 Generalized eigenvectors and basic matrix elements

We can build a *basis* in a Hilbert space if we take all eigenvectors of a suitable Hermitian operator. This operator must have a discrete spectrum because its eigenvectors must form a countable set.

In calculations, however, it is often more convenient to use as a basis in Hilbert space the eigenbasis of the major operators $\hat{q}$ and $\hat{p}$. The position operator has a continuous spectrum and its eigenvalues are all possible generalized coordinates q. Therefore, it turns out that the operator $\hat{q}$ has no eigenvectors which belong to a separable Hilbert space. It is possible, however, to extend a Hilbert space to a larger vector space and introduce the "generalized vectors" $|q\rangle$ that are the eigenvectors of operator $\hat{q}$. Assuming the completeness of the basis $\{|q\rangle\}$ we can expand a vector $|\psi\rangle$ as

$$|\psi\rangle = \int dq\, \psi(q)\,|q\rangle\,. \tag{2.26}$$

Note that $|\psi\rangle$ belongs to the Hilbert space while the generalized vectors $|q\rangle$ do not. This is very similar to the situation in the theory of distributions (generalized functions), where for example the delta-function, $\delta(x-y)$, is well-defined only when applied to some function $f(x)$ from the space of base functions.

Since the operator $\hat{q}$ is Hermitian, its different eigenvectors $|q_1\rangle$ and $|q_2\rangle$ are orthogonal:

$$\langle q_1|q_2\rangle = 0 \text{ for } q_1 \neq q_2.$$

The generalization of the decomposition of unity for the case of continuous q must be

$$\hat{1} = \int dq\,|q\rangle\langle q|\,. \tag{2.27}$$

Substituting this decomposition into (2.26) we obtain

$$|\psi\rangle = \int dq\,\psi(q)\hat{1}\,|q\rangle = \int dqdq'\,\psi(q)\langle q'|q\rangle\,|q'\rangle = \int dq'\psi(q')\,|q'\rangle$$

and hence the identity

$$\psi(q') = \int dq \psi(q) \langle q'|q\rangle$$

must be satisfied for an arbitrary $\psi(q)$. This is possible only if

$$\langle q'|q\rangle = \delta(q'-q).$$

Thus, the basis $\{|q\rangle\}$ is normalized on the *delta-function* and in particular $\langle q|q\rangle = \delta(0)$ is undefined. Generally, the matrix elements such as $\langle q|\hat{A}|q'\rangle$ are also distributions.

The basis $\{|p\rangle\}$ of generalized eigenvectors of the momentum operator $\hat{p}$ is constructed in a similar way and has the same properties as the basis $\{|q\rangle\}$.

Below we will use the commutation relation (2.20) to calculate the basic matrix elements for the position and momentum operators.

The matrix elements $\langle q_1|\hat{p}|q_2\rangle$ and $\langle p_1|\hat{q}|p_2\rangle$ To determine $\langle q_1|\hat{p}|q_2\rangle$ let us consider the matrix element of the commutator $[\hat{q}, \hat{p}]$:

$$\langle q_1|[\hat{q}, \hat{p}]|q_2\rangle = \langle q_1|\hat{q}\hat{p} - \hat{p}\hat{q}|q_2\rangle = (q_1 - q_2)\langle q_1|\hat{p}|q_2\rangle.$$

Taking into account the commutation relation (2.20) we find that on the other hand

$$\langle q_1|[\hat{q}, \hat{p}]|q_2\rangle = \langle q_1|i\hbar|q_2\rangle = i\hbar\delta(q_1 - q_2).$$

Therefore the matrix element $\langle q_1|\hat{p}|q_2\rangle \equiv F(q_1, q_2)$ must satisfy the equation

$$i\hbar\delta(q_1 - q_2) = (q_1 - q_2)F(q_1, q_2). \tag{2.28}$$

To solve this equation for $F(q_1, q_2)$ we cannot simply divide both sides by $q_1 - q_2$ because the expression obtained, $x^{-1}\delta(x)$, is undefined. Therefore we first apply the Fourier transform to (2.28). Introducing the variable $q \equiv q_1 - q_2$ and taking into account that

$$\int \delta(q) e^{-ipq} dq = 1,$$

we obtain

$$i\hbar = \int qF(q_1, q_1 - q) e^{-ipq} dq = i\frac{\partial}{\partial p}\int F(q_1, q_1 - q) e^{-ipq} dq.$$

Integration over p gives

$$\hbar p + C(q_1) = \int F(q_1, q_1 - q) e^{-ipq} dq,$$

where $C(q_1)$ is an undetermined function. The inverse Fourier transform yields

$$F(q_1, q_2) = \frac{1}{2\pi}\int (\hbar p + C) e^{ipq} dp = \left[-i\hbar\frac{\partial}{\partial q_1} + C(q_1)\right]\delta(q_1 - q_2),$$

and hence

$$\langle q_1|\hat{p}|q_2\rangle = -i\hbar\delta'(q_1-q_2)+C(q_1)\delta(q_1-q_2)\,, \tag{2.29}$$

where the prime denotes the derivative of $\delta(q)$ with respect to $q = q_1 - q_2$. The function $C(q_1)$ cannot be determined from the commutation relation alone because the replacement of the operator $\hat{p}$ by $\hat{p}+c(\hat{q})$, where c is an arbitrary function, does not destroy the commutation relation. The above transformation changes the matrix element $\langle q_1|\hat{p}|q_2\rangle$ by the term $c(q_1)\delta(q_1-q_2)$ and therefore the term proportional to $\delta(q_1-q_2)$ in (2.29) can always be removed by redefinition of the operator $\hat{p}$. Thus the final result for $\langle q_1|\hat{p}|q_2\rangle$ is

$$\langle q_1|\hat{p}|q_2\rangle = -i\hbar\delta'(q_1-q_2)\,. \tag{2.30}$$

Remark Note that for a given $\hat{p}$ the term $C(q_1)\delta(q_1-q_2)$ in (2.29) can be removed by redefining the basis vectors $|q\rangle$ themselves. Multiplying each $|q\rangle$ by a q-dependent phase,

$$|\tilde{q}\rangle \equiv e^{-ic(q)}|q\rangle\,,$$

we obtain

$$\langle\tilde{q}_1|\hat{p}|\tilde{q}_2\rangle = \hbar c'(q)\delta(q_1-q_2) - i\hbar\delta'(q_1-q_2)+C(q_1)\delta(q_1-q_2)\,.$$

The function $c(q)$ can always be chosen such that $C(q_1)\delta(q_1-q_2)$ is canceled.

Because $\hat{q}$ and $\hat{p}$ enter the commutation relation on the same "footing" (up to the sign) the matrix element $\langle p_1|\hat{q}|p_2\rangle$ can be immediately inferred by interchanging $q \longleftrightarrow p$ in (2.30) and changing the sign:

$$\langle p_1|\hat{q}|p_2\rangle = i\hbar\delta'(p_1-p_2)\,. \tag{2.31}$$

The matrix elements $\langle p|q\rangle$ To calculate $\langle p|q\rangle$, we consider the matrix element

$$\langle p|\hat{p}|q\rangle = p\langle p|q\rangle\,. \tag{2.32}$$

On the other hand, we have

$$\langle p|\hat{p}|q\rangle = \langle p|\left[\int dq_1\,|q_1\rangle\langle q_1|\right]\hat{p}|q\rangle = \int dq_1\,\langle p|q_1\rangle\langle q_1|\hat{p}|q\rangle = i\hbar\frac{\partial}{\partial q}\langle p|q\rangle\,, \tag{2.33}$$

where we have substituted the expression in (2.30) for $\langle q_1|\hat{p}|q\rangle$. Comparing (2.32) and (2.33) we obtain

$$p\langle p|q\rangle = i\hbar\frac{\partial}{\partial q}\langle p|q\rangle\,.$$

Similarly, considering the matrix element $\langle p|\hat{q}|q\rangle$, one derives

$$q\langle p|q\rangle = i\hbar\frac{\partial}{\partial p}\langle p|q\rangle\,.$$

Integrating these equations, we find

$$\langle p|q\rangle = C_1(p)\exp\left[-\frac{ipq}{\hbar}\right], \quad \langle p|q\rangle = C_2(q)\exp\left[-\frac{ipq}{\hbar}\right],$$

respectively. The above solutions are compatible only if $C_1(p) = C_2(q) = \text{const}$, and thus

$$\langle p|q\rangle = C\exp\left[-\frac{ipq}{\hbar}\right], \tag{2.34}$$

where the constant of integration C is determined (up to an irrelevant phase factor) by the normalization condition $C = (2\pi\hbar)^{-1/2}$ (see Exercise 2.6). Thus, the final result is

$$(\langle q|p\rangle)^* = \langle p|q\rangle = \frac{1}{\sqrt{2\pi\hbar}}\exp\left[-\frac{ipq}{\hbar}\right]. \tag{2.35}$$

Exercise 2.6
Let $|p\rangle$, $|q\rangle$ be the δ-normalized eigenvectors of the momentum and the position operators, i.e.

$$\hat{p}\,|p_1\rangle = p_1\,|p_1\rangle\,, \quad \langle p_1|p_2\rangle = \delta\,(p_1 - p_2)\,,$$

and the same for $\hat{q}$. Show that the coefficient C in equation (2.34) satisfies $|C| = (2\pi\hbar)^{-1/2}$.

2.7 Evolution in quantum theory

Heisenberg picture In the limit $\hbar \to 0$ the expectation values of the position and momentum operators must satisfy the classical equations of motion. Therefore the simplest way to implement the time evolution in quantum theory is to postulate that the state vector of the system is time-independent, while the operators $\hat{q}\,(t)$ and $\hat{p}\,(t)$ satisfy the "Hamilton equations of motion":

$$\frac{d\hat{q}}{dt} = \frac{\partial H}{\partial p}\,(\hat{p}, \hat{q}, t) + O(\hbar), \quad \frac{d\hat{p}}{dt} = -\frac{\partial H}{\partial q}\,(\hat{p}, \hat{q}, t) + O(\hbar), \tag{2.36}$$

where $\hat{p}$, $\hat{q}$ are substituted into $\partial H/\partial q$, $\partial H/\partial p$ *after* taking the derivatives. The expressions on the right-hand side are well-defined if the Hamiltonian H is a polynomial function in p and q. The operator ordering induces the terms which are proportional to $\hbar$ and hence does not influence the classical limit. Most non-polynomial functions can be approximated by polynomials and therefore below we shall not dwell on the mathematical details of defining the operator $\hat{H} = H\,(\hat{p}, \hat{q}, t)$ for a general case.

Of course, ultimately the correct form of the quantum equations of motion is decided by their agreement with experimental data. Presently, in most cases a

theory based on the "classical equations of motion" for the $\hat{p}$, $\hat{q}$ operators is in excellent agreement with experiments.

It is convenient to rewrite the equations

$$\frac{d\hat{q}}{dt} = \frac{\partial H}{\partial p}(\hat{p}, \hat{q}, t), \quad \frac{d\hat{p}}{dt} = -\frac{\partial H}{\partial q}(\hat{p}, \hat{q}, t) \tag{2.37}$$

in a purely algebraic form. Using the identities (see Exercise 2.7):

$$[\hat{q}, f(\hat{p}, \hat{q})] = i\hbar \frac{\partial f}{\partial p}(\hat{p}, \hat{q}), \quad [\hat{p}, f(\hat{p}, \hat{q})] = -i\hbar \frac{\partial f}{\partial q}(\hat{p}, \hat{q}),$$

the equations (2.37) become:

$$\frac{d\hat{q}}{dt} = -\frac{i}{\hbar}\left[\hat{q}, \hat{H}\right], \quad \frac{d\hat{p}}{dt} = -\frac{i}{\hbar}\left[\hat{p}, \hat{H}\right]. \tag{2.38}$$

They are called the Heisenberg equations of motion.

Exercise 2.7

(a) Using the canonical commutation relation, prove that

$$[\hat{q}, \hat{q}^m \hat{p}^n] = i\hbar n \hat{q}^m \hat{p}^{n-1}.$$

This relation can symbolically be written as

$$[\hat{q}, \hat{q}^m \hat{p}^n] = i\hbar \frac{\partial}{\partial \hat{p}}(\hat{q}^m \hat{p}^n).$$

Derive the similar relation for $\hat{p}$,

$$[\hat{p}, \hat{p}^m \hat{q}^n] = -i\hbar \frac{\partial}{\partial \hat{q}}(\hat{p}^m \hat{q}^n).$$

(b) Suppose that $f(p, q)$ is an analytic function in p, q given by a series expansion that converges for all p and q. The operator $f(\hat{p}, \hat{q})$ is defined by substituting the operators $\hat{p}$, $\hat{q}$ into that expansion (the ordering of $\hat{q}$ and $\hat{p}$ must be somehow fixed). Show that

$$[\hat{q}, f(\hat{p}, \hat{q})] = i\hbar \frac{\partial}{\partial \hat{p}} f(\hat{p}, \hat{q}). \tag{2.39}$$

Here it is implied that the derivative $\partial/\partial\hat{p}$ acts on each $\hat{p}$ with no change to the operator ordering, e.g.

$$\frac{\partial}{\partial \hat{p}}(\hat{p}^3 \hat{q} \hat{p}^2 \hat{q}) = 3\hat{p}^2 \hat{q} \hat{p}^2 \hat{q} + 2\hat{p}^3 \hat{q} \hat{p} \hat{q}.$$

Exercise 2.8

Show that an operator $\hat{A} = f(\hat{p}, \hat{q}, t)$, where $f(p, q, t)$ is an analytic function, satisfies the equation

$$\frac{d}{dt}\hat{A} = -\frac{i}{\hbar}\left[\hat{A}, \hat{H}\right] + \frac{\partial \hat{A}}{\partial t}. \tag{2.40}$$

So far we have considered the time-dependent operators $\hat{q}(t)$, $\hat{p}(t)$ that act on fixed state vectors $|\psi\rangle$; this description of quantized systems is called the *Heisenberg picture.*

Schrödinger picture An alternative way to describe the time evolution in quantum theory is to refer the time dependence to the state vector assuming that the operators are time-independent. In fact, what is relevant for the measurements is not the operators themselves, but only their eigenvalues and the expectation values. It turns out that the time evolution of the expectation values can be entirely encoded in $|\psi(t)\rangle$. Let us consider for simplicity an operator which does not depend on time explicitly, that is, $\hat{A} = f(\hat{p}, \hat{q})$. The general solution of equation (2.40) is then

$$\hat{A}(t) = \exp\left[\frac{i}{\hbar}(t-t_0)\hat{H}\right]\hat{A}_0 \exp\left[-\frac{i}{\hbar}(t-t_0)\hat{H}\right], \tag{2.41}$$

where $\hat{A}_0 \equiv A(t_0) = \text{const}$, and for an arbitrary quantum state $|\psi_0\rangle$ the time-dependent expectation value of $\hat{A}(t)$ is

$$\langle A(t)\rangle \equiv \langle\psi_0|\,\hat{A}(t)\,|\psi_0\rangle = \langle\psi_0|\,e^{\frac{i}{\hbar}\hat{H}(t-t_0)}\hat{A}_0 e^{-\frac{i}{\hbar}\hat{H}(t-t_0)}\,|\psi_0\rangle\,.$$

This relation can be rewritten using a time-dependent state

$$|\psi(t)\rangle \equiv e^{-\frac{i}{\hbar}\hat{H}(t-t_0)}\,|\psi_0\rangle \tag{2.42}$$

and the time-independent operator $\hat{A}_0$ as

$$\langle A(t)\rangle = \langle\psi(t)|\,\hat{A}_0\,|\psi(t)\rangle\,.$$

The description of dynamics using evolving state vectors and time-independent operators is called the *Schrödinger picture.*

Taking the time derivative of (2.42), we find that the state vector $|\psi(t)\rangle$ satisfies the *Schrödinger equation,*

$$i\hbar\frac{\partial\,|\psi(t)\rangle}{\partial t} = \hat{H}\,|\psi(t)\rangle\,. \tag{2.43}$$

The above quantization procedure is equally well applicable to nonrelativistic mechanics, to solid state physics (a very large but finite number of degrees of freedom), and to relativistic field theory (infinitely many degrees of freedom). In the case of a system with local symmetries, some complications (mainly of technical nature) arise, but the general idea of quantization remains the same. It is clear that the Schrödinger equation (2.43) is simply a way to implement the Hamiltonian dynamics in quantum theory and it can be either relativistically invariant or not, depending on the Hamiltonian of the physical system. If the Hamiltonian $\hat{H}$ describes a relativistic system, then the corresponding Schrödinger equation

is also relativistically invariant. For example, in field theory the Hamiltonian depends on infinitely many degrees of freedom and, as we will see later, equation (2.43) becomes a functional differential equation. The relativistic invariance of this equation is not manifest and is only revealed with extra effort. This is related to the fact that the quantization procedure is based on the commutation relations which are naturally implemented only in the Hamiltonian approach, where Lorentz invariance is also not manifest.

Remark: Schrödinger equations The use of a Schrödinger equation does *not* necessarily imply nonrelativistic physics. There is a widespread confusion about the role of the Schrödinger equation vs. that of the basic relativistic field equations: the Klein–Gordon equation, the Dirac equation, or the Maxwell equations. It would be a mistake to think that the Dirac equation and the Klein–Gordon equation are "relativistic forms" of the Schrödinger equation (although some textbooks say that). This was how the Dirac and the Klein–Gordon equations were discovered, but their actual place in quantum theory is quite different. The three field equations named above describe *classical* relativistic fields of spin 0, 1/2 and 1 respectively. These equations need to be quantized to obtain a quantum field theory. Their role is analogous to that of the harmonic oscillator equation: they provide a classical Hamiltonian for quantization. The Schrödinger equations corresponding to the Klein–Gordon, the Dirac and the Maxwell equations describe quantum theories of these classical fields. (In practice, Schrödinger equations are rarely used in quantum field theory because in most cases it is much easier to work in the Heisenberg picture.)

Remark: second quantization The term "second quantization" is frequently used to refer to quantum field theory, whereas "first quantization" means ordinary quantum mechanics. However, this is obsolete terminology originating from the historical development of QFT as a relativistic extension of quantum mechanics. In fact, a quantization procedure can only be applied to a *classical* theory and yields the corresponding quantum theory. One does not quantize a *quantum* theory for a second time. It is more logical to say "quantization of fields" instead of "second quantization."

Historically it was not immediately realized that relativistic particles can be described only by quantized fields. At first, fields were regarded as wave functions of point particles. Old QFT textbooks present the picture of (1) "quantizing" a relativistic point particle to obtain the Klein–Gordon or Dirac equations, which are sometimes mistakenly identified with the relativistic generalization of the Schrödinger equation; and (2) "second-quantizing" the "relativistic Schrödinger wave function" to obtain a quantum field theory. The confusion between Schrödinger equations and relativistic wave equations has been cleared, but the old illogical terminology of "first" and "second" quantization persists. It is unnecessary to talk about a "second-quantized Dirac equation" if the Dirac equation is actually quantized only once.

The modern view is that one must describe relativistic particles by fields. Therefore one starts right away with a classical relativistic field equation, such as the Dirac equation (for the electron field) and the Maxwell equations (for the photon field), and applies the quantization procedure (only once) to obtain the relativistic quantum theory of photons and electrons.