Ay, call it holy ground,
The soil where first they trod!
F. D. HEMANS (1793-1835), Landing of the Pilgrim Fathers

1

Fundamentals

Path integrals deal with fluctuating line-like structures. These appear in nature in a variety of ways, for instance, as particle orbits in spacetime continua, as polymers in solutions, as vortex lines in superfluids, as defect lines in crystals and liquid crystals. Their fluctuations can be of quantum-mechanical, thermodynamic, or statistical origin. Path integrals are an ideal tool to describe these fluctuating line-like structures, thereby leading to a unified understanding of many quite different physical phenomena. In developing the formalism we shall repeatedly invoke well-known concepts of classical mechanics, quantum mechanics, and statistical mechanics, to be summarized in this chapter. In Section 1.13, we emphasize some important problems of operator quantum mechanics in spaces with curvature and torsion. These problems will be solved in Chapters 10 and 8 by means of path integrals.[1]

1.1 Classical Mechanics

The orbits of a classical-mechanical system are described by a set of time-dependent generalized coordinates $q_1(t), \ldots, q_N(t)$. A Lagrangian

$$L(q_i, \dot{q}_i, t) \tag{1.1}$$

depending on $q_1, \ldots, q_N$ and the associated velocities $\dot{q}_1, \ldots, \dot{q}_N$ governs the dynamics of the system. The dots denote the time derivative d/dt. The Lagrangian is at most a quadratic function of $\dot{q}_i$. The time integral

$$\mathcal{A}[q_i] = \int_{t_a}^{t_b} dt\, L(q_i(t), \dot{q}_i(t), t) \tag{1.2}$$

of the Lagrangian along an arbitrary path $q_i(t)$ is called the *action* of this path. The path being actually chosen by the system as a function of time is called the *classical path* or the *classical orbit* $q_i^{\rm cl}(t)$. It has the property of extremizing the action in comparison with all neighboring paths

$$q_i(t) = q_i^{\rm cl}(t) + \delta q_i(t) \tag{1.3}$$

[1]Readers familiar with the foundations may start directly with Section 1.13.

having the same endpoints $q(t_b)$, $q(t_a)$. To express this property formally, one introduces the *variation* of the action as the linear term in the Taylor expansion of $\mathcal{A}[q_i]$ in powers of $\delta q_i(t)$:

$$\delta\mathcal{A}[q_i] \equiv \{\mathcal{A}[q_i + \delta q_i] - \mathcal{A}[q_i]\}_{\text{lin term in }\delta q_i} \quad . \tag{1.4}$$

The extremal principle for the classical path is then

$$\delta\mathcal{A}[q_i]\Big|_{q_i(t)=q_i^{\text{cl}}(t)} = 0 \tag{1.5}$$

for all variations of the path around the classical path, $\delta q_i(t) \equiv q_i(t) - q_i^{\text{cl}}(t)$, which vanish at the endpoints, i.e., which satisfy

$$\delta q_i(t_a) = \delta q_i(t_b) = 0. \tag{1.6}$$

Since the action is a time integral of a Lagrangian, the extremality property can be phrased in terms of differential equations. Let us calculate the variation of $\mathcal{A}[q_i]$ explicitly:

$$\begin{aligned}
\delta\mathcal{A}[q_i] &= \{\mathcal{A}[q_i + \delta q_i] - \mathcal{A}[q_i]\}_{\text{lin}} \\
&= \int_{t_a}^{t_b} dt\ \{L\left(q_i(t) + \delta q_i(t), \dot{q}_i(t) + \delta\dot{q}_i(t), t\right) - L\left(q_i(t), \dot{q}_i(t), t\right)\}_{\text{lin}} \\
&= \int_{t_a}^{t_b} dt\ \left\{\frac{\partial L}{\partial q_i}\delta q_i(t) + \frac{\partial L}{\partial \dot{q}_i}\delta\dot{q}_i(t)\right\} \\
&= \int_{t_a}^{t_b} dt\ \left\{\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right\}\delta q_i(t) + \frac{\partial L}{\partial \dot{q}_i}\delta q_i(t)\Big|_{t_a}^{t_b}.
\end{aligned} \tag{1.7}$$

The last expression arises from a partial integration of the $\delta\dot{q}_i$ term. Here, as in the entire text, repeated indices are understood to be summed (*Einstein's summation convention*). The endpoint terms (*surface* or *boundary terms*) with the time t equal to t_a and t_b may be dropped, due to (1.6). Thus we find for the classical orbit $q_i^{\text{cl}}(t)$ the *Euler-Lagrange equations*:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i}. \tag{1.8}$$

There is an alternative formulation of classical dynamics which is based on a Legendre-transformed function of the Lagrangian called the *Hamiltonian*

$$H \equiv \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i - L(q_i, \dot{q}_i, t). \tag{1.9}$$

Its value at any time is equal to the energy of the system. According to the general theory of Legendre transformations [1], the natural variables which H depends on are no longer q_i and $\dot{q}_i$, but q_i and the generalized momenta p_i, the latter being defined by the N equations

$$p_i \equiv \frac{\partial}{\partial \dot{q}_i}L(q_i, \dot{q}_i, t). \tag{1.10}$$

In order to express the Hamiltonian $H(p_i, q_i, t)$ in terms of its proper variables p_i, q_i, the equations (1.10) have to be solved for $\dot{q}_i$ by a velocity function

$$\dot{q}_i = v_i(p_i, q_i, t). \tag{1.11}$$

This is possible provided the *Hessian metric*

$$h_{ij}(q_i, \dot{q}_i, t) \equiv \frac{\partial^2}{\partial \dot{q}_i \partial \dot{q}_j} L(q_i, \dot{q}_i, t) \tag{1.12}$$

is nonsingular. The result is inserted into (1.9), leading to the Hamiltonian as a function of p_i and q_i:

$$H(p_i, q_i, t) = p_i v_i(p_i, q_i, t) - L(q_i, v_i(p_i, q_i, t), t). \tag{1.13}$$

In terms of this Hamiltonian, the action is the following functional of $p_i(t)$ and $q_i(t)$:

$$\mathcal{A}[p_i, q_i] = \int_{t_a}^{t_b} dt \left[p_i(t)\dot{q}_i(t) - H(p_i(t), q_i(t), t) \right]. \tag{1.14}$$

This is the so-called *canonical form* of the action. The classical orbits are now specified by $p_i^{\rm cl}(t)$, $q_i^{\rm cl}(t)$. They extremize the action in comparison with all neighboring orbits in which the coordinates $q_i(t)$ are varied at fixed endpoints [see (1.3), (1.6)] whereas the momenta $p_i(t)$ are varied without restriction:

$$\begin{aligned} q_i(t) &= q_i^{\rm cl}(t) + \delta q_i(t), \qquad \delta q_i(t_a) = \delta q_i(t_b) = 0, \\ p_i(t) &= p_i^{\rm cl}(t) + \delta p_i(t). \end{aligned} \tag{1.15}$$

In general, the variation is

$$\begin{aligned} \delta\mathcal{A}[p_i, q_i] &= \int_{t_a}^{t_b} dt \left[\delta p_i(t)\dot{q}_i(t) + p_i(t)\delta\dot{q}_i(t) - \frac{\partial H}{\partial p_i}\delta p_i - \frac{\partial H}{\partial q_i}\delta q_i \right] \\ &= \int_{ta}^{t_b} dt \left\{ \left[\dot{q}_i(t) - \frac{\partial H}{\partial p_i} \right] \delta p_i - \left[\dot{p}_i(t) + \frac{\partial H}{\partial q_i} \right] \delta q_i \right\} \\ &\qquad + p_i(t)\delta q_i(t) \Big|_{t_a}^{t_b}. \end{aligned} \tag{1.16}$$

Since this variation has to vanish for the classical orbits, we find that $p_i^{\rm cl}(t), q_i^{\rm cl}(t)$ must be solutions of the *Hamilton equations* of motion

$$\begin{aligned} \dot{p}_i &= -\frac{\partial H}{\partial q_i}, \\ \dot{q}_i &= \frac{\partial H}{\partial p_i}. \end{aligned} \tag{1.17}$$

These agree with the Euler-Lagrange equations (1.8) via (1.9) and (1.10), as can easily be verified. The $2N$-dimensional space of all p_i and q_i is called the *phase space*.

An arbitrary function $O(p_i(t), q_i(t), t)$ changes along an arbitrary path as follows:

$$\frac{d}{dt}O\left(p_i(t), q_i(t), t\right) = \frac{\partial O}{\partial p_i}\dot{p}_i + \frac{\partial O}{\partial q_i}\dot{q}_i + \frac{\partial O}{\partial t}. \tag{1.18}$$

If the path coincides with a classical orbit, we may insert (1.17) and find

$$\begin{aligned} \frac{dO}{dt} &= \frac{\partial H}{\partial p_i}\frac{\partial O}{\partial q_i} - \frac{\partial O}{\partial p_i}\frac{\partial H}{\partial q_i} + \frac{\partial O}{\partial t} \\ &\equiv \{H, O\} + \frac{\partial O}{\partial t}. \end{aligned} \tag{1.19}$$

Here we have introduced the symbol $\{\ldots,\ldots\}$ called *Poisson brackets*:

$$\{A, B\} \equiv \frac{\partial A}{\partial p_i}\frac{\partial B}{\partial q_i} - \frac{\partial B}{\partial p_i}\frac{\partial A}{\partial q_i}, \tag{1.20}$$

again with the Einstein summation convention for the repeated index i. The Poisson brackets have the obvious properties

$$\{A, B\} = -\{B, A\} \qquad \text{antisymmetry}, \tag{1.21}$$

$$\{A, \{B, C\}\} + \{B, \{C, A\}\} + \{C, \{A, B\}\} = 0 \qquad \text{Jacobi identity}. \tag{1.22}$$

If two quantities have vanishing Poisson brackets, they are said to *commute*.

The original Hamilton equations are a special case of (1.19):

$$\begin{aligned} \frac{d}{dt}p_i &= \{H, p_i\} = \frac{\partial H}{\partial p_j}\frac{\partial p_i}{\partial q_j} - \frac{\partial p_i}{\partial p_j}\frac{\partial H}{\partial q_j} = -\frac{\partial H}{\partial q_i}, \\ \frac{d}{dt}q_i &= \{H, q_i\} = \frac{\partial H}{\partial p_j}\frac{\partial q_i}{\partial q_j} - \frac{\partial q_i}{\partial p_j}\frac{\partial H}{\partial q_j} = \frac{\partial H}{\partial p_i}. \end{aligned} \tag{1.23}$$

By definition, the phase space variables p_i, q_i satisfy the Poisson brackets

$$\begin{aligned} \{p_i, q_j\} &= \delta_{ij}, \\ \{p_i, p_j\} &= 0, \\ \{q_i, q_j\} &= 0. \end{aligned} \tag{1.24}$$

A function $O(p_i, q_i)$ which has no *explicit* dependence on time and which, moreover, commutes with H (i.e., $\{O, H\} = 0$), is a *constant of motion* along the classical path, due to (1.19). In particular, H itself is often time-independent, i.e., of the form

$$H = H(p_i, q_i). \tag{1.25}$$

Then, since H commutes with itself, the energy is a constant of motion.

The Lagrangian formalism has the virtue of being independent of the particular choice of the coordinates q_i. Let Q_i be any other set of coordinates describing the system which is connected with q_i by what is called a *local*[2] or *point transformation*

$$q_i = f_i(Q_j, t). \tag{1.26}$$

Certainly, to be of use, this relation must be invertible, at least in some neighborhood of the classical path,

$$Q_i = {f^{-1}}_i(q_j, t). \tag{1.27}$$

Otherwise Q_i and q_i could not both parametrize the same system. Therefore, f_i must have a nonvanishing Jacobi determinant:

$$\det\left(\frac{\partial f_i}{\partial Q_j}\right) \neq 0. \tag{1.28}$$

In terms of Q_i, the initial Lagrangian takes the form

$$L'\left(Q_j, \dot{Q}_j, t\right) \equiv L\left(f_i\left(Q_j, t\right), \dot{f}_i\left(Q_j, t\right), t\right) \tag{1.29}$$

and the action reads

$$\begin{aligned} \mathcal{A} &= \int_{t_a}^{t_b} dt\, L'\left(Q_j(t), \dot{Q}_j(t), t\right) \\ &= \int_{t_a}^{t_b} dt\, L\left(f_i\left(Q_j(t), t\right), \dot{f}_i\left(Q_j(t), t\right), t\right). \end{aligned} \tag{1.30}$$

By performing variations $\delta Q_j(t)$, $\delta\dot{Q}_j(t)$ in the first expression while keeping $\delta Q_j(t_a) = \delta Q_j(t_b) = 0$, we find the equations of motion

$$\frac{d}{dt}\frac{\partial L'}{\partial \dot{Q}_j} - \frac{\partial L'}{\partial Q_j} = 0. \tag{1.31}$$

The variation of the lower expression, on the other hand, gives

$$\begin{aligned} \delta\mathcal{A} &= \int_{t_a}^{t_b} dt \left(\frac{\partial L}{\partial q_i}\delta f_i + \frac{\partial L}{\partial \dot{q}_i}\delta\dot{f}_i\right) \\ &= \int_{t_a}^{t_b} dt \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right)\delta f_i + \frac{\partial L}{\partial \dot{q}_i}\delta f_i\Big|_{t_a}^{t_b}. \end{aligned} \tag{1.32}$$

If δq_i is arbitrary, then so is δf_i. Moreover, with $\delta q_i(t_a) = \delta q_i(t_b) = 0$, also δf_i vanishes at the endpoints. Hence the extremum of the action is determined equally well by the Euler-Lagrange equations for $Q_j(t)$ [as it was by those for $q_i(t)$].

[2]The word *local* means here *at a specific time*. This terminology is of common use in field theory where *local* means, more generally, *at a specific spacetime point*.

Note that the locality property is quite restrictive for the transformation of the generalized velocities $\dot{q}_i(t)$. They will necessarily be linear in $\dot{Q}_j$:

$$\dot{q}_i = \dot{f}_i(Q_j, t) = \frac{\partial f_i}{\partial Q_j}\dot{Q}_j + \frac{\partial f_i}{\partial t}. \tag{1.33}$$

In phase space, there exists also the possibility of performing local changes of the canonical coordinates p_i, q_i to new ones P_j, Q_j. Let them be related by

$$\begin{aligned} p_i &= p_i(P_j, Q_j, t), \\ q_i &= q_i(P_j, Q_j, t), \end{aligned} \tag{1.34}$$

with the inverse relations

$$\begin{aligned} P_j &= P_j(p_i, q_i, t), \\ Q_j &= Q_j(p_i, q_i, t). \end{aligned} \tag{1.35}$$

However, while the Euler-Lagrange equations maintain their form under *any* local change of coordinates, the Hamilton equations do not hold, in general, for any transformed coordinates $P_j(t)$, $Q_j(t)$. The local transformations $p_i(t), q_i(t) \to P_j(t), Q_j(t)$ for which they hold, are referred to as *canonical*. They are characterized by the form invariance of the action, up to an arbitrary surface term,

$$\begin{aligned} \int_{t_a}^{t_b} dt\, [p_i \dot{q}_i - H(p_i, q_i, t)] &= \int_{t_a}^{t_b} dt\, \left[P_j \dot{Q}_j - H'(P_j, Q_j, t)\right] \\ &\qquad + F(P_j, Q_j, t)\Big|_{t_a}^{t_b}, \end{aligned} \tag{1.36}$$

where $H'(P_j, Q_j, t)$ is some new Hamiltonian. Its relation with $H(p_i, q_i, t)$ must be chosen in such a way that the equality of the action holds for *any* path $p_i(t), q_i(t)$ connecting the same endpoints (at least any in some neighborhood of the classical orbits). If such an invariance exists then a variation of this action yields for $P_j(t)$ and $Q_j(t)$ the Hamilton equations of motion governed by H':

$$\begin{aligned} \dot{P}_i &= -\frac{\partial H'}{\partial Q_i}, \\ \dot{Q}_i &= \frac{\partial H'}{\partial P_i}. \end{aligned} \tag{1.37}$$

The invariance (1.36) can be expressed differently by rewriting the integral on the left-hand side in terms of the new variables $P_j(t), Q_j(t)$,

$$\int_{t_a}^{t_b} dt \left\{ p_i \left(\frac{\partial q_i}{\partial P_j}\dot{P}_j + \frac{\partial q_i}{\partial Q_j}\dot{Q}_j + \frac{\partial q_i}{\partial t} \right) - H(p_i(P_j, Q_j, t), q_i(P_j, Q_j, t), t) \right\}, \tag{1.38}$$

and subtracting it from the right-hand side, leading to

$$\begin{aligned} \int_{t_a}^{t_b} \Bigg\{ \left(P_j - p_i \frac{\partial q_i}{\partial Q_j} \right) dQ_j - p_i \frac{\partial q_i}{\partial P_j} dP_j \\ - \left(H' + p_i \frac{\partial q_i}{\partial t} - H \right) dt \Bigg\} &= -F(P_j, Q_j, t)\Big|_{t_a}^{t_b}. \end{aligned} \tag{1.39}$$

The integral is now a line integral along a curve in the $(2N+1)$-dimensional space, consisting of the $2N$-dimensional phase space variables p_i, q_i and of the time t. The right-hand side depends only on the endpoints. Thus we conclude that the integrand on the left-hand side must be a total differential. As such it has to satisfy the standard Schwarz integrability conditions [2], according to which all second derivatives have to be independent of the sequence of differentiation. Explicitly, these conditions are

$$\begin{aligned} \frac{\partial p_i}{\partial P_k}\frac{\partial q_i}{\partial Q_l} - \frac{\partial q_i}{\partial P_k}\frac{\partial p_i}{\partial Q_l} &= \delta_{kl}, \\ \frac{\partial p_i}{\partial P_k}\frac{\partial q_i}{\partial P_l} - \frac{\partial q_i}{\partial P_k}\frac{\partial p_i}{\partial P_l} &= 0, \\ \frac{\partial p_i}{\partial Q_k}\frac{\partial q_i}{\partial Q_l} - \frac{\partial q_i}{\partial Q_k}\frac{\partial p_i}{\partial Q_l} &= 0, \end{aligned} \tag{1.40}$$

and

$$\begin{aligned} \frac{\partial p_i}{\partial t}\frac{\partial q_i}{\partial P_l} - \frac{\partial q_i}{\partial t}\frac{\partial p_i}{\partial P_l} &= \frac{\partial (H' - H)}{\partial P_l}, \\ \frac{\partial p_i}{\partial t}\frac{\partial q_i}{\partial Q_l} - \frac{\partial q_i}{\partial t}\frac{\partial p_i}{\partial Q_l} &= \frac{\partial (H' - H)}{\partial Q_l}. \end{aligned} \tag{1.41}$$

The first three equations define the so-called *Lagrange brackets* in terms of which they are written as

$$\begin{aligned} (P_k, Q_l) &= \delta_{kl}, \\ (P_k, P_l) &= 0, \\ (Q_k, Q_l) &= 0. \end{aligned} \tag{1.42}$$

Time-dependent coordinate transformations satisfying these equations are called *symplectic*. After a little algebra involving the matrix of derivatives

$$J = \begin{pmatrix} \partial P_i/\partial p_j & \partial P_i/\partial q_j \\ \partial Q_i/\partial p_j & \partial Q_i/\partial q_j \end{pmatrix}, \tag{1.43}$$

its inverse

$$J^{-1} = \begin{pmatrix} \partial p_i/\partial P_j & \partial p_i/\partial Q_j \\ \partial q_i/\partial P_j & \partial q_i/\partial Q_j \end{pmatrix}, \tag{1.44}$$

and the symplectic unit matrix

$$E = \begin{pmatrix} 0 & \delta_{ij} \\ -\delta_{ij} & 0 \end{pmatrix}, \tag{1.45}$$

we find that the Lagrange brackets (1.42) are equivalent to the Poisson brackets

$$\begin{aligned} \{P_k, Q_l\} &= \delta_{kl}, \\ \{P_k, P_l\} &= 0, \\ \{Q_k, Q_l\} &= 0. \end{aligned} \tag{1.46}$$

This follows from the fact that the $2N \times 2N$ matrix formed from the Lagrange brackets

$$\mathcal{L} \equiv \begin{pmatrix} -(Q_i, P_j) & -(Q_i, Q_j) \\ (P_i, P_j) & (P_i, Q_j) \end{pmatrix} \tag{1.47}$$

can be written as $(E^{-1}J^{-1}E)^T J^{-1}$, while an analogous matrix formed from the Poisson brackets

$$\mathcal{P} \equiv \begin{pmatrix} \{P_i, Q_j\} & -\{P_i, P_j\} \\ \{Q_i, Q_j\} & -\{Q_i, P_j\} \end{pmatrix} \tag{1.48}$$

is equal to $J(E^{-1}JE)^T$. Hence $\mathcal{L} = \mathcal{P}^{-1}$, so that (1.42) and (1.46) are equivalent to each other. Note that the Lagrange brackets (1.42) [and thus the Poisson brackets (1.46)] ensure $p_i\dot{q}_i - P_j\dot{Q}_j$ to be a total differential of some function of P_j and Q_j in the $2N$-dimensional phase space:

$$p_i\dot{q}_i - P_j\dot{Q}_j = \frac{d}{dt}G(P_j, Q_j, t). \tag{1.49}$$

The Poisson brackets (1.46) for P_i, Q_i have the same form as those in Eqs. (1.24) for the original phase space variables p_i, q_i.

The other two equations (1.41) relate the new Hamiltonian to the old one. They can always be used to construct $H'(P_j, Q_j, t)$ from $H(p_i, q_i, t)$. The Lagrange brackets (1.42) or Poisson brackets (1.46) are therefore both necessary and sufficient for the transformation $p_i, q_i \to P_j, Q_j$ to be canonical.

A canonical transformation preserves the volume in phase space. This follows from the fact that the matrix product $J(E^{-1}JE)^T$ is equal to the $2N \times 2N$ unit matrix (1.48). Hence $\det(J) = \pm 1$ and

$$\prod_i \int [dp_i\, dq_i] = \prod_j \int [dP_j\, dQ_j] \,. \tag{1.50}$$

It is obvious that the process of canonical transformations is reflexive. It may be viewed just as well from the opposite side, with the roles of p_i, q_i and P_j, Q_j exchanged [we could just as well have considered the integrand (1.39) as a complete differential in P_j, Q_j, t space].

Once a system is described in terms of new canonical coordinates P_j, Q_j, we introduce the new Poisson brackets

$$\{A, B\}' \equiv \frac{\partial A}{\partial P_j}\frac{\partial B}{\partial Q_j} - \frac{\partial B}{\partial P_j}\frac{\partial A}{\partial Q_j}, \tag{1.51}$$

and the equation of motion for an arbitrary observable quantity $O\left(P_j(t), Q_j(t), t\right)$ becomes with (1.37)

$$\frac{dO}{dt} = \{H', O\}' + \frac{\partial O}{\partial t}, \tag{1.52}$$

by complete analogy with (1.19). The new Poisson brackets automatically guarantee the canonical commutation rules

$$\begin{aligned}\{P_i, Q_j\}' &= \delta_{ij}, \\ \{P_i, P_j\}' &= 0, \\ \{Q_i, Q_j\}' &= 0.\end{aligned} \tag{1.53}$$

A standard class of canonical transformations can be constructed by introducing a *generating function* F satisfying a relation of the type (1.36), but depending explicitly on half an old and half a new set of canonical coordinates, for instance

$$F = F(q_i, Q_j, t). \tag{1.54}$$

One now considers the equation

$$\int_{t_a}^{t_b} dt\, [p_i\dot{q}_i - H(p_i, q_i, t)] = \int_{t_a}^{t_b} dt \left[P_j\dot{Q}_j - H'(P_j, Q_j, t) + \frac{d}{dt}F(q_i, Q_j, t)\right], \tag{1.55}$$

replaces $P_j\dot{Q}_j$ by $-\dot{P}_jQ_j + \frac{d}{dt}P_jQ_j$, defines

$$F(q_i, P_j, t) \equiv F(q_i, Q_j, t) + P_jQ_j,$$

and works out the derivatives. This yields

$$\begin{aligned}&\int_{t_a}^{t_b} dt \left\{p_i\dot{q}_i + \dot{P}_jQ_j - [H(p_i, q_i, t) - H'(P_j, Q_j, t)]\right\} \\ &= \int_{t_a}^{t_b} dt \left\{\frac{\partial F}{\partial q_i}(q_i, P_j, t)\dot{q}_i + \frac{\partial F}{\partial P_j}(q_i, P_j, t)\dot{P}_j + \frac{\partial F}{\partial t}(q_i, P_j, t)\right\}.\end{aligned} \tag{1.56}$$

A comparison of the two sides yields the equations for the canonical transformation

$$\begin{aligned}p_i &= \frac{\partial}{\partial q_i}F(q_i, P_j, t), \\ Q_j &= \frac{\partial}{\partial P_j}F(q_i, P_j, t).\end{aligned} \tag{1.57}$$

The second equation shows that the above relation between $F(q_i, P_j, t)$ and $F(q_i, Q_j, t)$ amounts to a Legendre transformation.

The new Hamiltonian is

$$H'(P_j, Q_j, t) = H(p_i, q_i, t) + \frac{\partial}{\partial t}F(q_i, P_j, t). \tag{1.58}$$

Instead of (1.54) we could, of course, also have chosen functions with other mixtures of arguments such as $F(q_i, P_j, t)$, $F(p_i, Q_j, t)$, $F(p_i, P_j, t)$ to generate simple canonical transformations.

A particularly important canonical transformation arises by choosing a generating function $F(q_i, P_j)$ in such a way that it leads to time-independent momenta $P_j \equiv \alpha_j$. Coordinates Q_j with this property are called *cyclic*. To find cyclic coordinates we must search for a generating function $F(q_j, P_j, t)$ which makes the transformed H' in (1.58) vanish identically. Then all derivatives with respect to the coordinates vanish and the new momenta P_j are trivially constant. Thus we seek a solution of the equation

$$\frac{\partial}{\partial t} F(q_i, P_j, t) = -H(p_i, q_i, t), \tag{1.59}$$

where the momentum variables in the Hamiltonian obey the first equation of (1.57). This leads to the following partial differential equation for $F(q_i, P_j, t)$:

$$\partial_t F(q_i, P_j, t) = -H(\partial_{q_i} F(q_i, P_j, t), q_i, t), \tag{1.60}$$

called the *Hamilton-Jacobi equation*. Here and in the sequel we shall often use the short notations for partial derivatives $\partial_t \equiv \partial/\partial t$, $\partial_{q_i} \equiv \partial/\partial_{q_i}$.

A generating function which achieves this goal is supplied by the action functional (1.14). When following the classical solutions starting from a fixed initial point and running to all possible final points q_i at a time t, the associated actions of these solutions form a function $A(q_i, t)$. Expression (1.14) show that if a particle moves along a classical trajectory, and the path is varied *without* keeping the endpoints fixed, the action changes as a function of the end positions (1.16) by

$$\delta \mathcal{A}[p_i, q_i] = p_i(t_b)\delta q_i(t_b) - p_i(t_a)\delta q_i(t_a). \tag{1.61}$$

From this we deduce immediately the first of the equations (1.57), now for the generating function $A(q_i, t)$:

$$p_i = \frac{\partial}{\partial q_i} A(q_i, t). \tag{1.62}$$

Moreover, the function $A(q_i, t)$ has the time derivative

$$\frac{d}{dt} A(q_i(t), t) = p_i(t)\dot{q}_i(t) - H(p_i(t), q_i(t), t). \tag{1.63}$$

Together with (1.62) this implies

$$\partial_t A(q_i, t) = -H(p_i, q_i, t). \tag{1.64}$$

If the momenta p_i on the right-hand side are replaced according to (1.62), $A(q_i, t)$ is indeed seen to be a solution of the Hamilton-acobi differential equation:

$$\partial_t A(q_i, t) = -H(\partial_{q_i} A(q_i, t), q_i, t). \tag{1.65}$$

1.2 Relativistic Mechanics in Curved Spacetime

The classical action of a relativistic spinless point particle in a curved four-dimensional spacetime is usually written as an integral

$$\mathcal{A} = -Mc^2 \int d\tau L(q,\dot{q}) = -Mc^2 \int d\tau \sqrt{g_{\mu\nu}\dot{q}^\mu(\tau)\dot{q}^\nu(\tau)}, \tag{1.66}$$

where τ is an arbitrary parameter of the trajectory. It can be chosen in the final trajectory to make $L(q,\dot{q}) \equiv 1$, in which case it coincides with the *proper time* of the particle. For arbitrary τ, the Euler-Lagrange equation (1.8) reads

$$\frac{d}{dt}\left[\frac{1}{L(q,\dot{q})} g_{\mu\nu}\dot{q}^\nu\right] = \frac{1}{2L(q,\dot{q})}\left(\partial_\mu g_{\kappa\lambda}\right)\dot{q}^\kappa \dot{q}^\lambda. \tag{1.67}$$

If τ is the proper time where $L(q,\dot{q}) \equiv 1$, this simplifies to

$$\frac{d}{dt}\left(g_{\mu\nu}\dot{q}^\nu\right) = \frac{1}{2}\left(\partial_\mu g_{\kappa\lambda}\right)\dot{q}^\kappa \dot{q}^\lambda, \tag{1.68}$$

or

$$g_{\mu\nu}\ddot{q}^\nu = \left(\frac{1}{2}\partial_\mu g_{\kappa\lambda} - \partial_\lambda g_{\mu\kappa}\right)\dot{q}^\kappa \dot{q}^\lambda. \tag{1.69}$$

For brevity, we have denoted partial derivatives $\partial/\partial q^\mu$ by ∂_μ. This partial derivative is supposed to apply only to the quantity right behind it. At this point one introduces the *Christoffel symbol*

$$\bar{\Gamma}_{\lambda\nu\mu} \equiv \frac{1}{2}(\partial_\lambda g_{\nu\mu} + \partial_\nu g_{\lambda\mu} - \partial_\mu g_{\lambda\nu}), \tag{1.70}$$

and the Christoffel symbol of the second kind[3]

$$\bar{\Gamma}_{\kappa\nu}{}^{\mu} \equiv g^{\mu\sigma}\bar{\Gamma}_{\kappa\nu\sigma}. \tag{1.71}$$

Then (1.69) can be written as

$$\ddot{q}^\mu + \bar{\Gamma}_{\kappa\lambda}{}^{\mu}\dot{q}^\kappa \dot{q}^\lambda = 0. \tag{1.72}$$

Since the solutions of this equation minimize the length of a curve in spacetime, they are called *geodesics*.

[3]In many textbooks, for instance S. Weinberg, *Gravitation and Cosmology*, Wiley, New York, 1972, the upper index and the third index in (1.70) stand at the first position. Our notation follows J.A. Schouten, *Ricci Calculus*, Springer, Berlin, 1954. It will allow for a closer analogy with gauge fields in the construction of the Riemann tensor as a covariant curl of the Christoffel symbol in Chapter 10. See H. Kleinert, *Gauge Fields in Condensed Matter*, Vol. II *Stresses and Defects*, World Scientific Publishing Co., Singapore 1989, pp. 744-1443 (`http://www.physik.fu-berlin.de/~kleinert/b2`).

1.3 Quantum Mechanics

Historically, the extension of classical mechanics to quantum mechanics became necessary in order to understand the stability of atomic orbits and the discrete nature of atomic spectra. It soon became clear that these phenomena reflect the fact that at a sufficiently short length scale, small material particles such as electrons behave like waves, called *material waves*. The fact that waves cannot be squeezed into an arbitrarily small volume without increasing indefinitely their frequency and thus their energy, prevents the collapse of the electrons into the nucleus, which would take place in classical mechanics. The discreteness of the atomic states of an electron are a manifestation of standing material waves in the atomic potential well, by analogy with the standing waves of electromagnetism in a cavity.

1.3.1 Bragg Reflections and Interference

The most direct manifestation of the wave nature of small particles is seen in diffraction experiments on periodic structures, for example of electrons diffracted by a crystal. If an electron beam of fixed momentum $\mathbf{p}$ passes through a crystal, it emerges along sharply peaked angles. These are the well-known *Bragg reflections*. They look very similar to the interference patterns of electromagnetic waves. In fact, it is possible to use the same mathematical framework to explain these patterns as in electromagnetism. A free particle moving with momentum

$$\mathbf{p} = (p^1, p^2, \ldots, p^D). \tag{1.73}$$

through a D-dimensional Euclidean space spanned by the Cartesian coordinate vectors

$$\mathbf{x} = (x^1, x^2, \ldots, x^D) \tag{1.74}$$

is associated with a *plane wave*, whose field strength or *wave function* has the form

$$\Psi_{\mathbf{p}}(\mathbf{x}, t) = e^{i\mathbf{k}\mathbf{x} - i\omega t}, \tag{1.75}$$

where $\mathbf{k}$ is the *wave vector* pointing into the direction of $\mathbf{p}$ and ω is the *wave frequency*. Each scattering center, say at $\mathbf{x}'$, becomes a source of a spherical wave with the spatial behavior e^{ikR}/R (with $R \equiv |\mathbf{x} - \mathbf{x}'|$ and $k \equiv |\mathbf{k}|$) and the wavelength $\lambda = 2\pi/k$. At the detector, all field strengths have to be added to the total field strength $\Psi(\mathbf{x}, t)$. The absolute square of the total field strength, $|\Psi(\mathbf{x}, t)|^2$, is proportional to the number of electrons arriving at the detector.

The standard experiment where these rules can most simply be applied consists of an electron beam impinging vertically upon a flat screen with two parallel slits a with spacing d. At a large distance R behind these, one observes the number of particles arriving per unit time (see Fig. 1.1)

$$\frac{dN}{dt} \propto |\Psi_1 + \Psi_2|^2 \approx \left| e^{ik(R + \frac{1}{2}d\sin\varphi)} + e^{ik(R - \frac{1}{2}d\sin\varphi)} \right|^2 \frac{1}{R^2}, \tag{1.76}$$

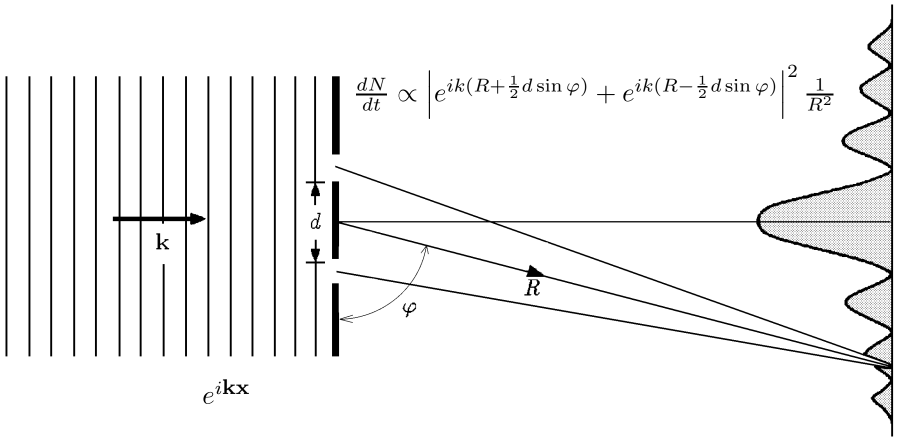

Figure 1.1 Probability distribution of particle behind double slit, being proportional to the absolute square of the sum of the two complex field strengths.

where φ is the angle of deflection from the normal.

Conventionally, the wave function $\Psi(\mathbf{x}, t)$ is normalized to describe a single particle. Its absolute square gives directly the probability density of the particle at the place $\mathbf{x}$ in space, i.e., $d^3x\, |\Psi(\mathbf{x}, t)|^2$ is the probability of finding the particle in the volume element d^3x around $\mathbf{x}$.

1.3.2 Matter Waves

From the experimentally observed relation between the momentum and the size of the angular deflection φ of the diffracted beam of the particles, one deduces the relation between momentum and wave vector

$$\mathbf{p} = \hbar\mathbf{k}, \tag{1.77}$$

where $\hbar$ is the universal *Planck constant* whose dimension is equal to that of an action,

$$\hbar \equiv \frac{h}{2\pi} = 1.0545919(80) \times 10^{-27} \text{erg}\,\text{sec} \tag{1.78}$$

(the number in parentheses indicating the experimental uncertainty of the last two digits before it). A similar relation holds between the energy and the frequency of the wave $\Psi(\mathbf{x}, t)$. It may be determined by an absorption process in which a light wave hits an electron and kicks it out of the surface of a metal, the well-known *photoelectric effect*. From the threshold property of this effect one learns that an electromagnetic wave oscillating in time as $e^{-i\omega t}$ can transfer to the electron the energy

$$E = \hbar\omega, \tag{1.79}$$

where the proportionality constant $\hbar$ is the same as in (1.77). The reason for this lies in the properties of electromagnetic waves. On the one hand, their frequency ω and the wave vector $\mathbf{k}$ satisfy the relation $\omega/c = |\mathbf{k}|$, where c is the light velocity defined to be $c \equiv 299\,792.458\,\mathrm{km/s}$. On the other hand, energy and momentum are related by $E/c = |\mathbf{p}|$. Thus, the quanta of electromagnetic waves, the *photons*, certainly satisfy (1.77) and the constant $\hbar$ must be the same as in Eq. (1.79).

With matter waves and photons sharing the same relations (1.77), it is suggestive to postulate also the relation (1.79) between energy and frequency to be universal for the waves of all particles, massive and massless ones. All free particles of momentum $\mathbf{p}$ are described by a *plane wave* of wavelength $\lambda = 2\pi/|\mathbf{k}| = 2\pi\hbar/|\mathbf{p}|$, with the explicit form

$$\Psi_{\mathbf{p}}(\mathbf{x}, t) = \mathcal{N} e^{i(\mathbf{p}\mathbf{x} - E_{\mathbf{p}} t)/\hbar}, \tag{1.80}$$

where $\mathcal{N}$ is some normalization constant. In a finite volume, the wave function is normalized to unity. In an infinite volume, this normalization makes the wave function vanish. To avoid this, the *current density* of the particle probability

$$\mathbf{j}(\mathbf{x}, t) \equiv -i\frac{\hbar}{2m}\psi^*(\mathbf{x}, t) \overleftrightarrow{\boldsymbol{\nabla}} \psi(\mathbf{x}, t) \tag{1.81}$$

is normalized in some convenient way, where $\overleftrightarrow{\boldsymbol{\nabla}}$ is a short notation for the difference between forward- and backward-derivatives

$$\begin{aligned} \psi^*(\mathbf{x}, t) \overleftrightarrow{\boldsymbol{\nabla}} \psi(\mathbf{x}, t) &\equiv \psi^*(\mathbf{x}, t) \overrightarrow{\boldsymbol{\nabla}} \psi(\mathbf{x}, t) - \psi^*(\mathbf{x}, t) \overleftarrow{\boldsymbol{\nabla}} \psi(\mathbf{x}, t) \\ &\equiv \psi^*(\mathbf{x}, t)\boldsymbol{\nabla}\psi(\mathbf{x}, t) - [\boldsymbol{\nabla}\psi^*(\mathbf{x}, t)]\,\psi(\mathbf{x}, t). \end{aligned} \tag{1.82}$$

The energy $E_{\mathbf{p}}$ depends on the momentum of the particle in the classical way, i.e., for nonrelativistic material particles of mass M it is $E_{\mathbf{p}} = \mathbf{p}^2/2M$, for relativistic ones $E_{\mathbf{p}} = c\sqrt{\mathbf{p}^2 + M^2c^2}$, and for massless particles such as photons $E_{\mathbf{p}} = c|\mathbf{p}|$. The common relation $E_{\mathbf{p}} = \hbar\omega$ for photons and matter waves is necessary to guarantee conservation of energy in quantum mechanics.

In general, both momentum and energy of a particle are not sharply defined as in the plane-wave function (1.80). Usually, a particle wave is some superposition of plane waves (1.80)

$$\Psi(\mathbf{x}, t) = \int \frac{d^3p}{(2\pi\hbar)^3}\, f(\mathbf{p}) e^{i(\mathbf{p}\mathbf{x} - E_{\mathbf{p}} t)/\hbar}. \tag{1.83}$$

By the Fourier inversion theorem, $f(\mathbf{p})$ can be calculated via the integral

$$f(\mathbf{p}) = \int d^3x\, e^{-i\mathbf{p}\mathbf{x}/\hbar}\Psi(\mathbf{x}, 0). \tag{1.84}$$

With an appropriate choice of $f(\mathbf{p})$ it is possible to prepare $\Psi(\mathbf{x}, t)$ in any desired form at some initial time, say at $t = 0$. For example, $\Psi(\mathbf{x}, 0)$ may be a function sharply centered around a space point $\bar{\mathbf{x}}$. Then $f(\mathbf{p})$ is approximately a pure phase $f(\mathbf{p}) \sim e^{-i\mathbf{p}\bar{\mathbf{x}}/\hbar}$, and the wave contains all momenta with equal probability. Conversely, if the particle amplitude is spread out in space, its momentum distribution is confined to a small region. The limiting $f(\mathbf{p})$ is concentrated at a specific momentum $\bar{\mathbf{p}}$. The particle is found at each point in space with equal probability, with the amplitude oscillating like $\Psi(\mathbf{x}, t) \sim e^{i(\bar{\mathbf{p}}\mathbf{x} - E_{\bar{\mathbf{p}}} t)/\hbar}$.

In general, the width of $\Psi(\mathbf{x}, 0)$ in space and of $f(\mathbf{p})$ in momentum space are inversely proportional to each other:

$$\Delta\mathbf{x}\,\Delta\mathbf{p} \sim \hbar. \tag{1.85}$$

This is the content of *Heisenberg's principle of uncertainty*. If the wave is localized in a finite region of space while having at the same time a fairly well-defined average momentum $\bar{\mathbf{p}}$, it is called a *wave packet*. The maximum in the associated probability density can be shown from (1.83) to move with a velocity

$$\bar{\mathbf{v}} = \partial E_{\bar{\mathbf{p}}}/\partial\bar{\mathbf{p}}. \tag{1.86}$$

This coincides with the velocity of a classical particle of momentum $\bar{\mathbf{p}}$.

1.3.3 Schrödinger Equation

Suppose now that the particle is nonrelativistic and has a mass M. The classical Hamiltonian, and thus the energy $E_{\mathbf{p}}$, are given by

$$H(\mathbf{p}) = E_{\mathbf{p}} = \frac{\mathbf{p}^2}{2M}. \tag{1.87}$$

We may therefore derive the following identity for the wave field $\Psi_{\mathbf{p}}(\mathbf{x}, t)$:

$$\int \frac{d^3p}{(2\pi\hbar)^3}\, f(\mathbf{p})\, [H(\mathbf{p}) - E_{\mathbf{p}}]\, e^{i(\mathbf{p}\mathbf{x} - E_{\mathbf{p}}t)/\hbar} = 0. \tag{1.88}$$

The arguments *inside* the brackets can be removed from the integral by observing that $\mathbf{p}$ and $E_{\mathbf{p}}$ inside the integral are equivalent to the differential operators

$$\begin{aligned} \hat{\mathbf{p}} &= -i\hbar\boldsymbol{\nabla}, \\ \hat{E} &= i\hbar\partial_t \end{aligned} \tag{1.89}$$

outside. Then, Eq. (1.88) may be written as the differential equation

$$[H(-i\hbar\boldsymbol{\nabla}) - i\hbar\partial_t)]\Psi(\mathbf{x}, t) = 0. \tag{1.90}$$

This is the *Schrödinger equation* for the wave function of a free particle. The equation suggests that the motion of a particle with an arbitrary Hamiltonian $H(\mathbf{p}, \mathbf{x}, t)$ follows the straightforward generalization of (1.90)

$$\left(\hat{H} - i\hbar\partial_t\right)\Psi(\mathbf{x}, t) = 0, \tag{1.91}$$

where $\hat{H}$ is the differential operator

$$\hat{H} \equiv H(-i\hbar\boldsymbol{\nabla}, \mathbf{x}, t). \tag{1.92}$$

The rule of obtaining $\hat{H}$ from the classical Hamiltonian $H(\mathbf{p}, \mathbf{x}, t)$ by the substitution $\mathbf{p} \to \hat{\mathbf{p}} = -i\hbar\boldsymbol{\nabla}$ will be referred to as the *correspondence principle*.[4] We shall see in Sections 1.13–1.15 that this simple correspondence principle holds only in Cartesian coordinates.

The Schrödinger operators (1.89) of momentum and energy satisfy with $\mathbf{x}$ and t the so-called canonical commutation relations

$$[\hat{p}_i, x_j] = -i\hbar, \qquad [\hat{E}, t] = 0 = i\hbar. \tag{1.93}$$

The linear combinations of the solutions of the Schrödinger equation (1.91) form, at each time t, a *Hilbert space*. If the Hamiltonian does not depend explicitly on time, If the Hamiltonian does not depend explicitly on time, the Hilbert space can be spanned by the energy eigenstates $\Psi_{E_n}(\mathbf{x}, t) = e^{-iE_n t/\hbar}\Psi_{E_n}(\mathbf{x})$, where $\Psi_{E_n}(\mathbf{x})$ are time-independent *stationary states*, which solve the time-independent Schrödinger equation

$$\hat{H}(\hat{\mathbf{p}}, \mathbf{x})\Psi_{E_n}(\mathbf{x}) = E_n\Psi_{E_n}(\mathbf{x}). \tag{1.94}$$

The validity of the Schrödinger theory (1.91) is confirmed by experiment, most notably for the Coulomb Hamiltonian

$$H(\mathbf{p}, \mathbf{x}) = \frac{\mathbf{p}^2}{2M} - \frac{e^2}{r}, \tag{1.95}$$

which governs the quantum mechanics of the hydrogen atom in the center-of-mass coordinate system of electron and proton, where M is the reduced mass of the two particles.

Since the square of the wave function, $|\Psi(\mathbf{x}, t)|^2$, is interpreted as the probability density of a single particle in a finite volume, the integral over the entire volume must be normalized to unity:

$$\int d^3x \; |\Psi(\mathbf{x}, t)|^2 = 1. \tag{1.96}$$

For a stable particle, this normalization must remain the same at all times. If $\Psi(\mathbf{x}, t)$ is to follow the Schrödinger equation (1.91), this is assured if and only if the Hamiltonian operator is Hermitian,[5] i.e., if it satisfies for arbitrary wave functions Ψ_1, Ψ_2 the equality

$$\int d^3x \, [\hat{H}\Psi_2(\mathbf{x}, t)]^*\Psi_1(\mathbf{x}, t) = \int d^3x \, \Psi_2^*(\mathbf{x}, t)\hat{H}\Psi_1(\mathbf{x}, t). \tag{1.97}$$

---

[4]Our formulation of this principle is slightly stronger than the historical one used in the initial phase of quantum mechanics, which gave certain translation rules between classical and quantum-mechanical relations. The substitution rule for the momentum runs also under the name *Jordan rule*.

[5]Problems arising from unboundedness or discontinuities of the Hamiltonian and other quantum-mechanical operators, such as restrictions of the domains of definition, are ignored here since they are well understood. Correspondingly we do not distinguish between Hermitian and self-adjoint operators (see J. v. Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin, 1932). Some quantum-mechanical operator subtleties will manifest themselves in this book as problems of path integration to be solved in Chapter 12. The precise relationship between the two calls for further detailed investigations.

The left-hand side defines the *Hermitian-adjoint* $\hat{H}^\dagger$ of the operator $\hat{H}$, which satisfies the identity

$$\int d^3x\, \Psi_2^*(\mathbf{x},t)\hat{H}^\dagger\Psi_1(\mathbf{x},t) \equiv \int d^3x\, [\hat{H}\Psi_2(\mathbf{x},t)]^*\Psi_1(\mathbf{x},t) \tag{1.98}$$

for all wave functions $\Psi_1(\mathbf{x},t), \Psi_2(\mathbf{x},t)$. An operator $\hat{H}$ is *Hermitian*, if it coincides with its Hermitian-adjoint $\hat{H}^\dagger$:

$$\hat{H} = \hat{H}^\dagger. \tag{1.99}$$

Let us calculate the time change of the integral over two arbitrary wave functions, $\int d^3x\, \Psi_2^*(\mathbf{x},t)\Psi_1(\mathbf{x},t)$. With the Schrödinger equation (1.91), this time change vanishes indeed as long as $\hat{H}$ is Hermitian:

$$\begin{aligned} i\hbar\frac{d}{dt}\int d^3x\, \Psi_2^*(\mathbf{x},t)\Psi_1(\mathbf{x},t) \\ = \int d^3x\, \Psi_2^*(\mathbf{x},t)\hat{H}\Psi_1(\mathbf{x},t) - \int d^3x\, [\hat{H}\Psi_2(\mathbf{x},t)]^*\Psi_1(\mathbf{x},t) = 0. \end{aligned} \tag{1.100}$$

This also implies the time independence of the normalization integral $\int d^3x\, |\Psi(\mathbf{x},t)|^2 = 1$.

Conversely, if $\hat{H}$ is not Hermitian, one can always find an eigenstate of $\hat{H}$ whose norm changes with time: any eigenstate of $(H - H^\dagger)/i$ has this property.

Since $\hat{\mathbf{p}} = -i\hbar\boldsymbol{\nabla}$ and $\mathbf{x}$ are themselves Hermitian operators, $\hat{H}$ will automatically be a Hermitian operator if it is a sum of a kinetic and a potential energy:

$$H(\mathbf{p},\mathbf{x},t) = T(\mathbf{p},t) + V(\mathbf{x},t). \tag{1.101}$$

This is always the case for nonrelativistic particles in Cartesian coordinates $\mathbf{x}$. If $\mathbf{p}$ and $\mathbf{x}$ appear in one and the same term of H, for instance as $\mathbf{p}^2\mathbf{x}^2$, the correspondence principle does not lead to a unique quantum-mechanical operator $\hat{H}$. Then there seem to be, in principle, several Hermitian operators which, in the above example, can be constructed from the product of two $\hat{\mathbf{p}}$ and two $\hat{\mathbf{x}}$ operators [for instance $\alpha\hat{\mathbf{p}}^2\hat{\mathbf{x}}^2+\beta\hat{\mathbf{x}}^2\hat{\mathbf{p}}^2+\gamma\hat{\mathbf{p}}\hat{\mathbf{x}}^2\hat{\mathbf{p}}$ with $\alpha+\beta+\gamma=1$]. They all correspond to the same classical $\mathbf{p}^2\mathbf{x}^2$. At first sight it appears as though only a comparison with experiment could select the correct operator ordering. This is referred to as the operator-ordering problem of *quantum mechanics* which has plagued many researchers in the past. If the ordering problem is caused by the geometry of the space in which the particle moves, there exists a surprisingly simple geometric principle which specifies the ordering in the physically correct way. Before presenting this in Chapter 10 we shall avoid ambiguities by assuming $H(\mathbf{p},\mathbf{x},t)$ to have the standard form (1.101), unless otherwise stated.

1.3.4 Particle Current Conservation

The conservation of the total probability (1.96) is a consequence of a more general *local conservation law* linking the *current density* of the particle probability

$$\mathbf{j}(\mathbf{x},t) \equiv -i\frac{\hbar}{2m}\psi(\mathbf{x},t)\stackrel{\leftrightarrow}{\boldsymbol{\nabla}}\psi(\mathbf{x},t) \tag{1.102}$$

with the *probability density*

$$\rho(\mathbf{x},t) = \psi^*(\mathbf{x},t)\psi(\mathbf{x},t) \tag{1.103}$$

via the relation

$$\partial_t\rho(\mathbf{x},t) = -\boldsymbol{\nabla}\cdot\mathbf{j}(\mathbf{x},t). \tag{1.104}$$

By integrating this *current conservation law* over a volume V enclosed by a surface S, and using Green's theorem, one finds

$$\int_V d^3x\,\partial_t\rho(\mathbf{x},t) = -\int_V d^3x\,\boldsymbol{\nabla}\cdot\mathbf{j}(\mathbf{x},t) = -\int_S d\mathbf{S}\cdot\mathbf{j}(\mathbf{x},t), \tag{1.105}$$

where $d\mathbf{S}$ are the directed infinitesimal surface elements. This equation states that the probability in a volume decreases by the same amount by which probability leaves the surface via the current $\mathbf{j}(\mathbf{x},t)$.

By extending the integral (1.105) over the entire space and assuming the currents to vanish at spatial infinity, we recover the conservation of the total probability (1.96).

More general dynamical systems with N particles in Euclidean space are parametrized in terms of $3N$ Cartesian coordinates $\mathbf{x}_\nu$ ($\nu = 1,\ldots,N$). The Hamiltonian has the form

$$H(\mathbf{p}_\nu,\mathbf{x}_\nu,t) = \sum_{\nu=1}^{N}\frac{\mathbf{p}_\nu^2}{2M_\nu} + V(\mathbf{x}_\nu,t), \tag{1.106}$$

where the arguments $\mathbf{p}_\nu,\mathbf{x}_\nu$ in H and V stand for all $\mathbf{p}_\nu$'s, $\mathbf{x}_\nu$ with $\nu = 1,2,3,\ldots,N$. The wave function $\Psi(\mathbf{x}_\nu,t)$ satisfies the N-particle Schrödinger equation

$$\left\{-\sum_{\nu=1}^{N}\left[\frac{\hbar^2}{2M_\nu}\partial_{\mathbf{x}_\nu}{}^2 + V(\mathbf{x}_\nu,t)\right]\right\}\Psi(\mathbf{x}_\nu,t) = i\hbar\partial_t\Psi(\mathbf{x}_\nu,t). \tag{1.107}$$

1.4 Dirac's Bra-Ket Formalism

Mathematically speaking, the wave function $\Psi(\mathbf{x},t)$ may be considered as a vector in an infinite-dimensional complex vector space called *Hilbert space*. The configuration space variable $\mathbf{x}$ plays the role of a continuous "index" of these vectors. An obvious contact with the usual vector notation may be established, in which a D-dimensional vector $\mathbf{v}$ is given in terms of its components v_i with a subscript $i = 1,\ldots D$, by writing the argument $\mathbf{x}$ of $\Psi(\mathbf{x},t)$ as a subscript:

$$\Psi(\mathbf{x},t) \equiv \Psi_{\mathbf{x}}(t). \tag{1.108}$$

The usual norm of a complex vector is defined by

$$|\mathbf{v}|^2 = \sum_i v_i^* v_i. \tag{1.109}$$

The continuous version of this is

$$|\Psi|^2 = \int d^3x \, \Psi_{\mathbf{x}}^*(t)\Psi_{\mathbf{x}}(t) = \int d^3x \, \Psi^*(\mathbf{x},t)\Psi(\mathbf{x},t). \tag{1.110}$$

The normalization condition (1.96) requires that the wave functions have the norm $|\Psi| = 1$, i.e., that they are unit vectors in the Hilbert space.

1.4.1 Basis Transformations

In a vector space, there are many possible choices of orthonormal basis vectors ${b_i}^a$ labeled by $a = 1, \dots, D$, in terms of which[6]

$$v_i = \sum_a {b_i}^a v_a, \tag{1.111}$$

with the components v_a given by the scalar products

$$v_a \equiv \sum_i {b_i}^{a*} v_i. \tag{1.112}$$

The latter equation is a consequence of the *orthogonality relation*[7]

$$\sum_i {b_i}^{a*} {b_i}^{a'} = \delta^{aa'}, \tag{1.113}$$

which in a finite-dimensional vector space implies the *completeness relation*

$$\sum_a {b_i}^{a*} {b_j}^a = \delta^{ij}. \tag{1.114}$$

In the space of wave functions (1.108) there exists a special set of basis functions called *local basis functions* of particular importance. It may be constructed in the following fashion: Imagine the continuum of space points to be coarse-grained into a cubic lattice of mesh size ϵ, at positions

$$\mathbf{x_n} = (n_1, n_2, n_3)\epsilon, \qquad n_{1,2,3} = 0, \pm 1, \pm 2, \dots \,. \tag{1.115}$$

Let $h^{\mathbf{n}}(\mathbf{x})$ be a function that vanishes everywhere in space, except in a cube of size ϵ^3 centered around $\mathbf{x_n}$, i.e., for each component x_i of $\mathbf{x}$,

$$h^{\mathbf{n}}(\mathbf{x}) = \begin{cases} 1/\sqrt{\epsilon^3} & |x_i - x_{\mathbf{n}\,i}| \le \epsilon/2, \quad i = 1,2,3. \\ 0 & \text{otherwise.} \end{cases} \tag{1.116}$$

[6] Mathematicians would expand more precisely $v_i = \sum_a {b_i}^a v_a^{(b)}$, but physicists prefer to shorten the notation by distinguishing the different components via different types of subscripts, using for the initial components $i, j, k, \dots$ and for the b-transformed components $a, b, c, \dots$.

[7] An orthogonality relation implies usually a unit norm and is thus really an *orthonormality relation* but this name is rarely used.

These functions are certainly orthonormal:

$$\int d^3x\, h^{\mathbf{n}}(\mathbf{x})^* h^{\mathbf{n}'}(\mathbf{x}) = \delta^{\mathbf{n}\mathbf{n}'}. \tag{1.117}$$

Consider now the expansion

$$\Psi(\mathbf{x},t) = \sum_{\mathbf{n}} h^{\mathbf{n}}(\mathbf{x})\Psi_{\mathbf{n}}(t) \tag{1.118}$$

with the coefficients

$$\Psi_{\mathbf{n}}(t) = \int d^3x\, h^{\mathbf{n}}(\mathbf{x})^* \Psi(\mathbf{x},t) \approx \sqrt{\epsilon^3}\Psi(\mathbf{x}_{\mathbf{n}},t). \tag{1.119}$$

It provides an excellent approximation to the true wave function $\Psi(\mathbf{x},t)$, as long as the mesh size ϵ is much smaller than the scale over which $\Psi(\mathbf{x},t)$ varies. In fact, if $\Psi(\mathbf{x},t)$ is integrable, the integral over the sum (1.118) will always converge to $\Psi(\mathbf{x},t)$. The same convergence of discrete approximations is found in any scalar product, and thus in any observable probability amplitudes. They can all be calculated with arbitrary accuracy knowing the discrete components of the type (1.119) in the limit $\epsilon \to 0$. The functions $h^{\mathbf{n}}(\mathbf{x})$ may therefore be used as an approximate basis in the same way as the previous basis functions $f^a(\mathbf{x}), g^b(\mathbf{x})$, with any desired accuracy depending on the choice of ϵ.

In general, there are many possible orthonormal basis functions $f^a(\mathbf{x})$ in the Hilbert space which satisfy the orthonormality relation

$$\int d^3x\, f^a(\mathbf{x})^* f^{a'}(\mathbf{x}) = \delta^{aa'}, \tag{1.120}$$

in terms of which we can expand

$$\Psi(\mathbf{x},t) = \sum_a f^a(\mathbf{x})\Psi_a(t), \tag{1.121}$$

with the coefficients

$$\Psi_a(t) = \int d^3x\, f^a(\mathbf{x})^*\, \Psi(\mathbf{x},t). \tag{1.122}$$

Suppose we use another orthonormal basis $\tilde{f}^b(\mathbf{x})$ with the orthonormality relation

$$\int d^3x\, \tilde{f}^b(\mathbf{x})^* \tilde{f}^{b'}(\mathbf{x}) = \delta^{bb'}, \qquad \sum_b \tilde{f}^b(\mathbf{x})\tilde{f}^b(\mathbf{x}')^* = \delta^{(3)}(\mathbf{x}-\mathbf{x}'), \tag{1.123}$$

to re-expand

$$\Psi(\mathbf{x},t) = \sum_b \tilde{f}^b(\mathbf{x})\tilde{\Psi}_b(t), \tag{1.124}$$

with the components

$$\tilde{\Psi}_b(t) = \int d^3x\, \tilde{f}^b(\mathbf{x})^*\, \Psi(\mathbf{x},t). \tag{1.125}$$

Inserting (1.121) shows that the components are related to each other by

$$\tilde{\Psi}_b(t) = \sum_a \left[\int d^3x\, \tilde{f}^b(\mathbf{x})^* f^a(\mathbf{x})\right] \Psi_a(t). \tag{1.126}$$

1.4.2 Bracket Notation

It is useful to write the scalar products between two wave functions occurring in the above basis transformations in the so-called *bracket notation* as

$$\langle \tilde{b}|a\rangle \equiv \int d^3x\, \tilde{f}^b(\mathbf{x})^* f^a(\mathbf{x}). \tag{1.127}$$

In this notation, the components of the state vector $\Psi(\mathbf{x},t)$ in (1.122), (1.125) are

$$\begin{aligned} \Psi_a(t) &= \langle a|\Psi(t)\rangle, \\ \tilde{\Psi}_b(t) &= \langle \tilde{b}|\Psi(t)\rangle. \end{aligned} \tag{1.128}$$

The transformation formula (1.126) takes the form

$$\langle \tilde{b}|\Psi(t)\rangle = \sum_a \langle \tilde{b}|a\rangle\langle a|\Psi(t)\rangle. \tag{1.129}$$

The right-hand side of this equation may be formally viewed as a result of inserting the abstract relation

$$\sum_a |a\rangle\langle a| = 1 \tag{1.130}$$

between $\langle \tilde{b}|$ and $|\Psi(t)\rangle$ on the left-hand side:

$$\langle \tilde{b}|\Psi(t)\rangle = \langle \tilde{b}|1|\Psi(t)\rangle = \sum_a \langle \tilde{b}|a\rangle\langle a|\Psi(t)\rangle. \tag{1.131}$$

Since this expansion is only possible if the functions $f^b(\mathbf{x})$ form a complete basis, the relation (1.130) is alternative, abstract way of stating the completeness of the basis functions. It may be referred to as completeness relation à la Dirac.

Since the scalar products are written in the form of brackets $\langle a|a'\rangle$, Dirac called the formal objects $\langle a|$ and $|a'\rangle$, from which the brackets are composed, *bra* and *ket*, respectively. In the bracket notation, the orthonormality of the basis $f^a(\mathbf{x})$ and $g^b(\mathbf{x})$ may be expressed as follows:

$$\begin{aligned} \langle a|a'\rangle &= \int d^3x\, f^a(\mathbf{x})^* f^{a'}(\mathbf{x}) = \delta^{aa'}, \\ \langle \tilde{b}|\tilde{b}'\rangle &= \int d^3x\, \tilde{f}^b(\mathbf{x})^* \tilde{f}^{b'}(\mathbf{x}) = \delta^{bb'}. \end{aligned} \tag{1.132}$$

In the same spirit we introduce abstract bra and ket vectors associated with the basis functions $h^{\mathbf{n}}(\mathbf{x})$ of Eq. (1.116), denoting them by $\langle \mathbf{x_n}|$ and $|\mathbf{x_n}\rangle$, respectively, and writing the orthogonality relation (1.117) in bracket notation as

$$\langle \mathbf{x_n}|\mathbf{x_{n'}}\rangle \equiv \int d^3x\, h^{\mathbf{n}}(\mathbf{x})^* h^{\mathbf{n}'}(\mathbf{x}) = \delta_{\mathbf{nn}'}. \tag{1.133}$$

The components $\Psi_{\mathbf{n}}(t)$ may be considered as the scalar products

$$\Psi_{\mathbf{n}}(t) \equiv \langle \mathbf{x_n}|\Psi(t)\rangle \approx \sqrt{\epsilon^3}\,\Psi(\mathbf{x_n},t). \tag{1.134}$$

Changes of basis vectors, for instance from $|\mathbf{x_n}\rangle$ to the states $|a\rangle$, can be performed according to the rules developed above by inserting a completeness relation à la Dirac of the type (1.130). Thus we may expand

$$\Psi_{\mathbf{n}}(t) = \langle \mathbf{x_n}|\Psi(t)\rangle = \sum_a \langle \mathbf{x_n}|a\rangle\langle a|\Psi(t)\rangle. \tag{1.135}$$

Also the inverse relation is true:

$$\langle a|\Psi(t)\rangle = \sum_{\mathbf{n}} \langle a|\mathbf{x_n}\rangle\langle \mathbf{x_n}|\Psi(t)\rangle. \tag{1.136}$$

This is, of course, just an approximation to the integral

$$\int d^3x\, h^{\mathbf{n}}(\mathbf{x})^* \langle \mathbf{x}|\Psi(t)\rangle. \tag{1.137}$$

The completeness of the basis $h^{\mathbf{n}}(\mathbf{x})$ may therefore be expressed via the abstract relation

$$\sum_{\mathbf{n}} |\mathbf{x_n}\rangle\langle \mathbf{x_n}| \approx 1. \tag{1.138}$$

The approximate sign turns into an equality sign in the limit of zero mesh size, $\epsilon \to 0$.

1.4.3 Continuum Limit

In ordinary calculus, finer and finer sums are eventually replaced by integrals. The same thing is done here. We define new continuous scalar products

$$\langle \mathbf{x}|\Psi(t)\rangle \approx \frac{1}{\sqrt{\epsilon^3}} \langle \mathbf{x_n}|\Psi(t)\rangle, \tag{1.139}$$

where $\mathbf{x_n}$ are the lattice points closest to $\mathbf{x}$. With (1.134), the right-hand side is equal to $\Psi(\mathbf{x_n}, t)$. In the limit $\epsilon \to 0$, $\mathbf{x}$ and $\mathbf{x_n}$ coincide and we have

$$\langle \mathbf{x}|\Psi(t)\rangle \equiv \Psi(\mathbf{x}, t). \tag{1.140}$$

The completeness relation can be used to write

$$\begin{aligned} \langle a|\Psi(t)\rangle &\approx \sum_{\mathbf{n}} \langle a|\mathbf{x_n}\rangle\langle \mathbf{x_n}|\Psi(t)\rangle \\ &\approx \sum_{\mathbf{n}} \epsilon^3 \langle a|\mathbf{x}\rangle\langle \mathbf{x}|\Psi(t)\rangle\Big|_{\mathbf{x}=\mathbf{x^n}} \;, \end{aligned} \tag{1.141}$$

which in the limit $\epsilon \to 0$ becomes

$$\langle a|\Psi(t)\rangle = \int d^3x\, \langle a|\mathbf{x}\rangle\langle \mathbf{x}|\Psi(t)\rangle. \tag{1.142}$$

This may be viewed as the result of inserting the formal completeness relation of the limiting local bra and ket basis vectors $\langle \mathbf{x}|$ and $|\mathbf{x}\rangle$,

$$\int d^3x\, |\mathbf{x}\rangle\langle \mathbf{x}| = 1, \tag{1.143}$$

evaluated between the vectors $\langle a|$ and $|\Psi(t)\rangle$.

With the limiting local basis, the wave functions can be treated as components of the state vectors $|\Psi(t)\rangle$ with respect to the local basis $|\mathbf{x}\rangle$ in the same way as any other set of components in an arbitrary basis $|a\rangle$. In fact, the expansion

$$\langle a|\Psi(t)\rangle = \int d^3x \, \langle a|\mathbf{x}\rangle\langle \mathbf{x}|\Psi(t)\rangle \tag{1.144}$$

may be viewed as a re-expansion of a component of $|\Psi(t)\rangle$ in one basis, $|a\rangle$, into those of another basis, $|\mathbf{x}\rangle$, just as in (1.129).

In order to express all these transformation properties in a most compact notation, it has become customary to deal with an arbitrary physical state vector in a *basis-independent* way and denote it by a ket vector $|\Psi(t)\rangle$. This vector may be specified in any convenient basis by multiplying it with the corresponding completeness relation

$$\sum_a |a\rangle\langle a| = 1, \tag{1.145}$$

resulting in the expansion

$$|\Psi(t)\rangle = \sum_a |a\rangle\langle a|\Psi(t)\rangle. \tag{1.146}$$

This can be multiplied with any bra vector, say $\langle b|$, from the left to obtain the expansion formula (1.131):

$$\langle b|\Psi(t)\rangle = \sum_a \langle b|a\rangle\langle a|\Psi(t)\rangle. \tag{1.147}$$

The continuum version of the completeness relation (1.138) reads

$$\int d^3x \, |\mathbf{x}\rangle\langle \mathbf{x}| = 1, \tag{1.148}$$

and leads to the expansion

$$|\Psi(t)\rangle = \int d^3x \, |\mathbf{x}\rangle\langle \mathbf{x}|\Psi(t)\rangle, \tag{1.149}$$

in which the wave function $\Psi(\mathbf{x},t) = \langle \mathbf{x}|\Psi(t)\rangle$ plays the role of an $\mathbf{x}$th component of the state vector $|\Psi(t)\rangle$ in the local basis $|\mathbf{x}\rangle$. This, in turn, is the limit of the discrete basis vectors $|\mathbf{x_n}\rangle$,

$$|\mathbf{x}\rangle \approx \frac{1}{\sqrt{\epsilon^3}} \, |\mathbf{x_n}\rangle \,, \tag{1.150}$$

with $\mathbf{x_n}$ being the lattice points closest to $\mathbf{x}$.

A vector can be described equally well in bra or in ket form. To apply the above formalism consistently, we observe that the scalar products

$$\begin{aligned} \langle a|\tilde{b}\rangle &= \int d^3x \, f^a(\mathbf{x})^* \tilde{f}^b(\mathbf{x}), \\ \langle \tilde{b}|a\rangle &= \int d^3x \, \tilde{f}^b(\mathbf{x})^* f^a(\mathbf{x}) \end{aligned} \tag{1.151}$$

satisfy the identity

$$\langle\tilde{b}|a\rangle \equiv \langle a|\tilde{b}\rangle^*. \tag{1.152}$$

Therefore, when expanding a ket vector as

$$|\Psi(t)\rangle = \sum_a |a\rangle\langle a|\Psi(t)\rangle, \tag{1.153}$$

or a bra vector as

$$\langle\Psi(t)| = \sum_a \langle\Psi(t)|a\rangle\langle a|, \tag{1.154}$$

a multiplication of the first equation with the bra $\langle\mathbf{x}|$ and of the second with the ket $|\mathbf{x}\rangle$ produces equations which are complex-conjugate to each other.

1.4.4 Generalized Functions

Dirac's bra-ket formalism is elegant and easy to handle. As far as the vectors $|\mathbf{x}\rangle$ are concerned there is, however, one inconsistency with some fundamental postulates of quantum mechanics: When introducing state vectors, the norm was required to be unity in order to permit a proper probability interpretation of single-particle states. The limiting states $|\mathbf{x}\rangle$ introduced above do not satisfy this requirement. In fact, the scalar product between two different states $\langle\mathbf{x}|$ and $|\mathbf{x}'\rangle$ is

$$\langle\mathbf{x}|\mathbf{x}'\rangle \approx \frac{1}{\epsilon^3}\langle\mathbf{x_n}|\mathbf{x_{n'}}\rangle = \frac{1}{\epsilon^3}\delta_{\mathbf{nn}'}, \tag{1.155}$$

where $\mathbf{x_n}$ and $\mathbf{x_{n'}}$ are the lattice points closest to $\mathbf{x}$ and $\mathbf{x}'$. For $\mathbf{x} \neq \mathbf{x}'$, the states are orthogonal. For $\mathbf{x} = \mathbf{x}'$, on the other hand, the limit $\epsilon \to 0$ is infinite, approached in such a way that

$$\epsilon^3 \sum_{\mathbf{n}'} \frac{1}{\epsilon^3}\delta_{\mathbf{nn}'} = 1. \tag{1.156}$$

Therefore, the limiting state $|\mathbf{x}\rangle$ is not a properly normalizable vector in the Hilbert space. For the sake of elegance, it is useful to weaken the requirement of normalizability (1.96) by admitting the limiting states $|\mathbf{x}\rangle$ to the physical Hilbert space. In fact, one admits all states which can be obtained by a limiting sequence from properly normalized state vectors.

The scalar product between states $\langle\mathbf{x}|\mathbf{x}'\rangle$ is not a proper function. It is denoted by the symbol $\delta^{(3)}(\mathbf{x}-\mathbf{x}')$ and called *Dirac δ-function*:

$$\langle\mathbf{x}|\mathbf{x}'\rangle \equiv \delta^{(3)}(\mathbf{x}-\mathbf{x}'). \tag{1.157}$$

The right-hand side vanishes everywhere, except in the infinitely small box of width ϵ around $\mathbf{x} \approx \mathbf{x}'$. Thus the δ-function satisfies

$$\delta^{(3)}(\mathbf{x}-\mathbf{x}') = 0 \quad \text{for} \quad \mathbf{x} \neq \mathbf{x}'. \tag{1.158}$$

At $\mathbf{x} = \mathbf{x}'$, it is so large that its volume integral is unity:

$$\int d^3x'\, \delta^{(3)}(\mathbf{x}-\mathbf{x}') = 1. \tag{1.159}$$

Obviously, there exists no proper function that can satisfy both requirements, (1.158) and (1.159). Only the finite-ϵ approximation in (1.155) to the δ-function are proper functions. In this respect, the scalar product $\langle \mathbf{x}|\mathbf{x}'\rangle$ behaves just like the states $|\mathbf{x}\rangle$ themselves: Both are $\epsilon \to 0$ -limits of properly defined mathematical objects.

Note that the integral Eq. (1.159) implies the following property of the δ -function:

$$\delta^{(3)}(a(\mathbf{x}-\mathbf{x}')) = \frac{1}{|a|}\delta^{(3)}(\mathbf{x}-\mathbf{x}'). \tag{1.160}$$

In one dimension, this leads to the more general relation

$$\delta(f(x)) = \sum_i \frac{1}{|f'(x_i)|}\delta(x-x_i), \tag{1.161}$$

where x_i are the simple zeros of $f(x)$.

In mathematics, one calls the δ-function a *generalized function* or a *distribution*. It defines a linear functional of arbitrary smooth *test functions* $f(\mathbf{x})$ which yields its value at any desired place $\mathbf{x}$:

$$\delta[f;\mathbf{x}] \equiv \int d^3x\, \delta^{(3)}(\mathbf{x}-\mathbf{x}')f(\mathbf{x}') = f(\mathbf{x}). \tag{1.162}$$

Test functions are arbitrarily often differentiable functions with a sufficiently fast falloff at spatial infinity.

There exist a rich body of mathematical literature on distributions [3]. They form a linear space. This space is restricted in an essential way in comparison with ordinary functions: products of δ-functions or any other distributions remain undefined. In Section 10.8.1 we shall find, however, that physics forces us to go beyond these rules. An important requirement of quantum mechanics is coordinate invariance. If we want to achieve this for the path integral formulation of quantum mechanics, we must set up a definite extension of the existing theory of distributions, which specifies uniquely integrals over products of distributions.

In quantum mechanics, the role of the test functions is played by the wave packets $\Psi(\mathbf{x},t)$. By admitting the generalized states $|\mathbf{x}\rangle$ to the Hilbert space, we also admit the scalar products $\langle \mathbf{x}|\mathbf{x}'\rangle$ to the space of wave functions, and thus all distributions, although they are not normalizable.

1.4.5 Schrödinger Equation in Dirac Notation

In terms of the bra-ket notation, the Schrödinger equation can be expressed in a basis-independent way as an operator equation

$$\hat{H}|\Psi(t)\rangle \equiv H(\hat{\mathbf{p}},\hat{\mathbf{x}},t)|\Psi(t)\rangle = i\hbar\partial_t|\Psi(t)\rangle, \tag{1.163}$$

to be supplemented by the following specifications of the canonical operators:

$$\langle \mathbf{x}|\hat{\mathbf{p}} \equiv -i\hbar\boldsymbol{\nabla}\langle \mathbf{x}|, \tag{1.164}$$

$$\langle \mathbf{x}|\hat{\mathbf{x}} \equiv \mathbf{x}\langle \mathbf{x}|. \tag{1.165}$$

Any matrix element can be obtained from these equations by multiplication from the right with an arbitrary ket vector; for instance with the local basis vector $|\mathbf{x}'\rangle$:

$$\begin{aligned}\langle\mathbf{x}|\hat{\mathbf{p}}|\mathbf{x}'\rangle &= -i\hbar\boldsymbol{\nabla}\langle\mathbf{x}|\mathbf{x}'\rangle = -i\hbar\boldsymbol{\nabla}\delta^{(3)}(\mathbf{x}-\mathbf{x}'), &(1.166)\\ \langle\mathbf{x}|\hat{\mathbf{x}}|\mathbf{x}'\rangle &= \mathbf{x}\langle\mathbf{x}|\mathbf{x}'\rangle = \mathbf{x}\delta^{(3)}(\mathbf{x}-\mathbf{x}'). &(1.167)\end{aligned}$$

The original differential form of the Schrödinger equation (1.91) follows by multiplying the basis-independent Schrödinger equation (1.163) with the bra vector $\langle\mathbf{x}|$ from the left:

$$\begin{aligned}\langle\mathbf{x}|H(\hat{\mathbf{p}},\hat{\mathbf{x}},t)|\Psi(t)\rangle &= H(-i\hbar\boldsymbol{\nabla},\mathbf{x},t)\langle\mathbf{x}|\Psi(t)\rangle\\ &= i\hbar\partial_t\langle\mathbf{x}|\Psi(t)\rangle.\end{aligned} \tag{1.168}$$

Obviously, $\hat{\mathbf{p}}$ and $\hat{\mathbf{x}}$ are Hermitian matrices in any basis,

$$\begin{aligned}\langle a|\hat{\mathbf{p}}|a'\rangle &= \langle a'|\hat{\mathbf{p}}|a\rangle^*, &(1.169)\\ \langle a|\hat{\mathbf{x}}|a'\rangle &= \langle a'|\hat{\mathbf{x}}|a\rangle^*, &(1.170)\end{aligned}$$

and so is the Hamiltonian

$$\langle a|\hat{H}|a'\rangle = \langle a'|\hat{H}|a\rangle^*, \tag{1.171}$$

as long as it has the form (1.101).

The most general basis-independent operator that can be constructed in the generalized Hilbert space spanned by the states $|\mathbf{x}\rangle$ is some function of $\hat{\mathbf{p}}, \hat{\mathbf{x}}, t$,

$$\hat{O}(t) \equiv O(\hat{\mathbf{p}},\hat{\mathbf{x}},t). \tag{1.172}$$

In general, such an operator is called Hermitian if all its matrix elements have this property. In the basis-independent Dirac notation, the definition (1.97) of a Hermitian-adjoint operator $\hat{O}^\dagger(t)$ implies the equality of the matrix elements

$$\langle a|\hat{O}^\dagger(t)|a'\rangle \equiv \langle a'|\hat{O}(t)|a\rangle^*. \tag{1.173}$$

Thus we can rephrase Eqs. (1.169)–(1.171) in the basis-independent form

$$\begin{aligned}\hat{\mathbf{p}} &= \hat{\mathbf{p}}^\dagger,\\ \hat{\mathbf{x}} &= \hat{\mathbf{x}}^\dagger,\\ \hat{H} &= \hat{H}^\dagger.\end{aligned} \tag{1.174}$$

The stationary states in Eq. (1.94) have a Dirac ket representation $|E_n\rangle$, and satisfy the time-independent operator equation

$$\hat{H}|E_n\rangle = E_n|E_n\rangle. \tag{1.175}$$

1.4.6 Momentum States

Let us now look at the momentum $\hat{\mathbf{p}}$. Its eigenstates are given by the eigenvalue equation

$$\hat{\mathbf{p}}|\mathbf{p}\rangle = \mathbf{p}|\mathbf{p}\rangle. \quad (1.176)$$

By multiplying this with $\langle\mathbf{x}|$ from the left and using (1.164), we find the differential equation

$$\langle\mathbf{x}|\hat{\mathbf{p}}|\mathbf{p}\rangle = -i\hbar\partial_{\mathbf{x}}\langle\mathbf{x}|\mathbf{p}\rangle = \mathbf{p}\langle\mathbf{x}|\mathbf{p}\rangle. \quad (1.177)$$

The solution is

$$\langle\mathbf{x}|\mathbf{p}\rangle \propto e^{i\mathbf{p}\mathbf{x}/\hbar}. \quad (1.178)$$

Up to a normalization factor, this is just a plane wave introduced before in Eq. (1.75) to describe free particles of momentum $\mathbf{p}$.

In order for the states $|\mathbf{p}\rangle$ to have a finite norm, the system must be confined to a finite volume, say a cubic box of length L and volume L^3. Assuming periodic boundary conditions, the momenta are discrete with values

$$\mathbf{p}^{\mathbf{m}} = \frac{2\pi\hbar}{L}(m_1, m_2, m_3), \quad m_i = 0, \pm1, \pm2, \dots \ . \quad (1.179)$$

Then we adjust the factor in front of $\exp(i\mathbf{p}^{\mathbf{m}}\mathbf{x}/\hbar)$ to achieve unit normalization

$$\langle\mathbf{x}|\mathbf{p}^{\mathbf{m}}\rangle = \frac{1}{\sqrt{L^3}}\exp\left(i\mathbf{p}^{\mathbf{m}}\mathbf{x}/\hbar\right), \quad (1.180)$$

and the discrete states $|\mathbf{p}^{\mathbf{m}}\rangle$ satisfy

$$\int d^3x\, |\langle\mathbf{x}|\mathbf{p}^{\mathbf{m}}\rangle|^2 = 1. \quad (1.181)$$

The states $|\mathbf{p}^{\mathbf{m}}\rangle$ are complete:

$$\sum_{\mathbf{m}} |\mathbf{p}^{\mathbf{m}}\rangle\langle\mathbf{p}^{\mathbf{m}}| = 1. \quad (1.182)$$

We may use this relation and the matrix elements $\langle\mathbf{x}|\mathbf{p}^{\mathbf{m}}\rangle$ to expand any wave function within the box as

$$\Psi(\mathbf{x}, t) = \langle\mathbf{x}|\Psi(t)\rangle = \sum_{\mathbf{m}}\langle\mathbf{x}|\mathbf{p}^{\mathbf{m}}\rangle\langle\mathbf{p}^{\mathbf{m}}|\Psi(t)\rangle. \quad (1.183)$$

If the box is very large, the sum over the discrete momenta $\mathbf{p}^{\mathbf{m}}$ can be approximated by an integral over the momentum space [4].

$$\sum_{\mathbf{m}} \approx \int \frac{d^3pL^3}{(2\pi\hbar)^3}. \quad (1.184)$$

In this limit, the states $|\mathbf{p}^{\mathbf{m}}\rangle$ may be used to define a continuum of basis vectors with an improper normalization

$$|\mathbf{p}\rangle \approx \sqrt{L^3}|\mathbf{p}^{\mathbf{m}}\rangle, \quad (1.185)$$

in the same way as $|\mathbf{x_n}\rangle$ was used in (1.150) to define $|\mathbf{x}\rangle \sim (1/\sqrt{\epsilon^3})|\mathbf{x^n}\rangle$. The momentum states $|\mathbf{p}\rangle$ satisfy the orthogonality relation

$$\langle \mathbf{p}|\mathbf{p}'\rangle = (2\pi\hbar)^3 \delta^{(3)}(\mathbf{p}-\mathbf{p}'), \tag{1.186}$$

with $\delta^{(3)}(\mathbf{p}-\mathbf{p}')$ being again the Dirac δ-function. Their completeness relation reads

$$\int \frac{d^3p}{(2\pi\hbar)^3} |\mathbf{p}\rangle\langle\mathbf{p}| = 1, \tag{1.187}$$

such that the expansion (1.183) becomes

$$\Psi(\mathbf{x},t) = \int \frac{d^3p}{(2\pi\hbar)^3} \langle\mathbf{x}|\mathbf{p}\rangle\langle\mathbf{p}|\Psi(t)\rangle, \tag{1.188}$$

with the momentum eigenfunctions

$$\langle\mathbf{x}|\mathbf{p}\rangle = e^{i\mathbf{px}/\hbar}. \tag{1.189}$$

This coincides precisely with the Fourier decomposition introduced above in the description of a general particle wave $\Psi(\mathbf{x},t)$ in (1.83), (1.84), with the identification

$$\langle\mathbf{p}|\Psi(t)\rangle = f(\mathbf{p})e^{-iE_{\mathbf{p}}t/\hbar}. \tag{1.190}$$

The bra-ket formalism accommodates naturally the technique of Fourier transforms. The Fourier inversion formula is found by simply inserting into $\langle\mathbf{p}|\Psi(t)\rangle$ a completeness relation $\int d^3x|\mathbf{x}\rangle\langle\mathbf{x}| = 1$ which yields

$$\begin{aligned} \langle\mathbf{p}|\Psi(t)\rangle &= \int d^3x\, \langle\mathbf{p}|\mathbf{x}\rangle\langle\mathbf{x}|\Psi(t)\rangle \\ &= \int d^3x\, e^{-i\mathbf{px}/\hbar}\Psi(\mathbf{x},t). \end{aligned} \tag{1.191}$$

The amplitudes $\langle\mathbf{p}|\Psi(t)\rangle$ are referred to as *momentum space wave functions*.

By inserting the completeness relation

$$\int d^3x|\mathbf{x}\rangle\langle\mathbf{x}| = 1 \tag{1.192}$$

between the momentum states on the left-hand side of the orthogonality relation (1.186), we obtain the Fourier representation of the δ-function (1.186):

$$\begin{aligned} \langle\mathbf{p}|\mathbf{p}'\rangle &= \int d^3x\, \langle\mathbf{p}|\mathbf{x}\rangle\langle\mathbf{x}|\mathbf{p}'\rangle \\ &= \int d^3x\, e^{-i(\mathbf{p}-\mathbf{p}')\mathbf{x}/\hbar} = (2\pi\hbar)^3\delta^{(3)}(\mathbf{p}-\mathbf{p}'). \end{aligned} \tag{1.193}$$

1.4.7 Incompleteness and Poisson's Summation Formula

For many physical applications it is important to find out what happens to the completeness relation (1.148) if one restrict the integral so a subset of positions. Most relevant will be the one-dimensional integral,

$$\int dx\,|x\rangle\langle x| = 1, \tag{1.194}$$

restricted to a sum over equally spaced points $x_n = na$:

$$\sum_{n=-N}^{N} |x_n\rangle\langle x_n|. \tag{1.195}$$

Taking this sum between momentum eigenstates $|p\rangle$, we obtain

$$\sum_{n=-N}^{N} \langle p|x_n\rangle\langle x_n|p'\rangle = \sum_{n=-N}^{N} \langle p|x_n\rangle\langle x_n|p'\rangle = \sum_{n=-N}^{N} e^{i(p-p')na/\hbar} \tag{1.196}$$

For $N \to \infty$ we can perform the sum with the help of *Poisson's summation formula*

$$\sum_{n=-\infty}^{\infty} e^{2\pi i\mu n} = \sum_{m=-\infty}^{\infty} \delta(\mu - m). \tag{1.197}$$

Identifying μ with $(p-p')a/2\pi\hbar$, we find using Eq. (1.160):

$$\sum_{n=-\infty}^{\infty} \langle p|x_n\rangle\langle x_n|p'\rangle = \sum_{m=-\infty}^{\infty} \delta\left(\frac{(p-p')a}{2\pi\hbar} - m\right) = \sum_{m=-\infty}^{\infty} \frac{2\pi\hbar}{a}\delta\left(p-p'-\frac{2\pi\hbar m}{a}\right). \tag{1.198}$$

In order to prove the Poisson formula (1.197), we observe that the sum $s(\mu) \equiv \sum_m \delta(\mu - m)$ on the right-hand side is periodic in μ with a unit period and has the Fourier series $s(\mu) = \sum_{n=-\infty}^{\infty} s_n e^{2\pi i\mu n}$. The Fourier coefficients are given by $s_n = \int_{-1/2}^{1/2} d\mu\, s(\mu) e^{-2\pi i\mu n} \equiv 1$. These are precisely the Fourier coefficients on the left-hand side.

For a finite N, the sum over n on the left-hand side of (1.197) yields

$$\begin{aligned}
\sum_{n=-N}^{N} e^{2\pi i\mu n} &= 1 + \left(e^{2\pi i\mu} + e^{2\cdot 2\pi i\mu} + \ldots + e^{N\cdot 2\pi i\mu} + \text{c.c.}\right) \\
&= -1 + \left(\frac{1 - e^{2\pi i\mu(N+1)}}{1 - e^{2\pi i\mu}} + \text{c.c.}\right) \\
&= 1 + \frac{e^{2\pi i\mu} - e^{2\pi i\mu(N+1)}}{1 - e^{2\pi i\mu}} + \text{c.c.} = \frac{\sin \pi\mu(2N+1)}{\sin \pi\mu}.
\end{aligned} \tag{1.199}$$

This function is well known in wave optics (see Fig. 2.4). It determines the diffraction pattern of light behind a grating with $2N+1$ slits. It has large peaks at

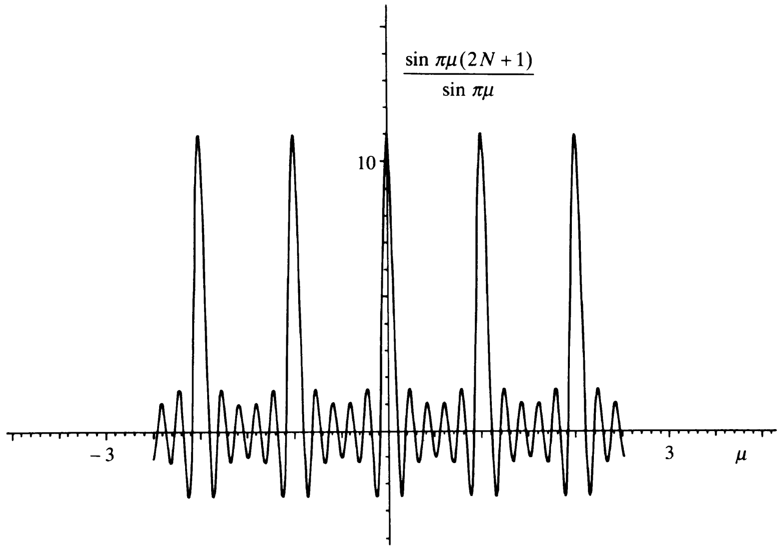

Figure 1.2 Relevant function $\sum_{n=-N}^{N} e^{2\pi i\mu n}$ in Poisson's summation formula. In the limit $N \to \infty$, μ is squeezed to the integer values.

$\mu = 0, \pm 1, \pm 2, \pm 3, \ldots$ and $N-1$ small maxima between each pair of neighboring peaks, at $\nu = (1+4k)/2(2N+1)$ for $k = 1, \ldots, N-1$. There are zeros at $\nu = (1+2k)/(2N+1)$ for $k = 1, \ldots, N-1$.

Inserting $\mu = (p-p')a/2\pi\hbar$ into (1.199), we obtain

$$\sum_{n=-N}^{N} \langle p|x_n\rangle\langle x_n|p'\rangle = \frac{\sin{(p-p')a(2N+1)/2\hbar}}{\sin{(p-p')a/2\hbar}}. \tag{1.200}$$

Let us see how the right-hand side of (1.199) turns into the right-hand side of (1.197) in the limit $N \to \infty$. In this limit, the area under each large peak can be calculated by an integral over the central large peak plus a number n of small maxima next to it:

$$\int_{-n/2N}^{n/2N} d\mu \frac{\sin \pi\mu(2N+1)}{\sin \pi\mu} = \int_{-n/2N}^{n/2N} d\mu \frac{\sin 2\pi\mu N \cos \pi\mu + \cos 2\pi\mu N \sin \pi\mu}{\sin \pi\mu}. \tag{1.201}$$

Keeping a fixed ratio $n/N \ll 1$, we may replace in the integrand $\sin \pi\mu$ by $\pi\mu$ and $\cos \pi\mu$ by 1. Then the integral becomes, for $N \to \infty$ at fixed n/N,

$$\int_{-n/2N}^{n/2N} d\mu \frac{\sin \pi\mu(2N+1)}{\sin \pi\mu} \xrightarrow{N\to\infty} \int_{-n/2N}^{n/2N} d\mu \frac{\sin 2\pi\mu N}{\pi\mu} + \int_{-n/2N}^{n/2N} d\mu \cos 2\pi\mu N$$
$$\xrightarrow{N\to\infty} \frac{1}{\pi}\int_{-\pi n}^{\pi n} dx \frac{\sin x}{x} + \frac{1}{2\pi N}\int_{-\pi n}^{\pi n} dx \cos x \xrightarrow{N\to\infty} 1, \tag{1.202}$$

where we have used the integral formula

$$\int_{-\infty}^{\infty} dx \frac{\sin x}{x} = \pi. \tag{1.203}$$

In the limit $N \to \infty$, we find indeed (1.197) and thus (1.205).

There exists another useful way of expressing Poisson's formula. Consider an arbitrary smooth function $f(\mu)$ which possesses a convergent sum

$$\sum_{m=-\infty}^{\infty} f(m). \tag{1.204}$$

Then Poisson's formula (1.197) implies that the sum can be rewritten as an integral and an auxiliary sum:

$$\sum_{m=-\infty}^{\infty} f(m) = \int_{-\infty}^{\infty} d\mu \sum_{n=-\infty}^{\infty} e^{2\pi i\mu n} f(\mu). \tag{1.205}$$

The auxiliary sum over n squeezes μ to the integer numbers.

1.5 Observables

Changes of basis vectors are an important tool in analyzing the physically observable content of a wave vector. Let $A = A(\mathbf{p}, \mathbf{x})$ be an arbitrary time-independent real function of the phase space variables $\mathbf{p}$ and $\mathbf{x}$. Important examples for such an A are $\mathbf{p}$ and $\mathbf{x}$ themselves, the Hamiltonian $H(\mathbf{p}, \mathbf{x})$, and the angular momentum $\mathbf{L} = \mathbf{x} \times \mathbf{p}$. Quantum-mechanically, there will be an observable operator associated with each such quantity. It is obtained by simply replacing the variables $\mathbf{p}$ and $\mathbf{x}$ in A by the corresponding operators $\hat{\mathbf{p}}$ and $\hat{\mathbf{x}}$:

$$\hat{A} \equiv A(\hat{\mathbf{p}}, \hat{\mathbf{x}}). \tag{1.206}$$

This replacement rule is the extension of the correspondence principle for the Hamiltonian operator (1.92) to more general functions in phase space, converting them into observable operators. It must be assumed that the replacement leads to a unique Hermitian operator, i.e., that there is no ordering problem of the type discussed in context with the Hamiltonian (1.101).[8] If there are ambiguities, the naive correspondence principle is insufficient to determine the observable operator. Then the correct ordering must be decided by comparison with experiment, unless it can be specified by means of simple geometric principles. This will be done for the Hamiltonian operator in Chapter 8.

Once an observable operator $\hat{A}$ is Hermitian, it has the useful property that the set of all eigenvectors $|a\rangle$ obtained by solving the equation

$$\hat{A}|a\rangle = a|a\rangle \tag{1.207}$$

can be used as a basis to span the Hilbert space. Among the eigenvectors, there is always a choice of orthonormal vectors $|a\rangle$ fulfilling the completeness relation

$$\sum_a |a\rangle\langle a| = 1. \tag{1.208}$$

[8] Note that this is true for the angular momentum $\mathbf{L} = \mathbf{x} \times \mathbf{p}$.

The vectors $|a\rangle$ can be used to extract physical information concerning the observable A from arbitrary state vector $|\Psi(t)\rangle$. For this we expand this vector in the basis $|a\rangle$:

$$|\Psi(t)\rangle = \sum_a |a\rangle\langle a|\Psi(t)\rangle. \tag{1.209}$$

The components

$$\langle a|\Psi(t)\rangle \tag{1.210}$$

yield the probability amplitude for measuring the eigenvalue a for the observable quantity A.

The wave function $\Psi(\mathbf{x}, t)$ itself is an example of this interpretation. If we write it as

$$\Psi(\mathbf{x}, t) = \langle \mathbf{x}|\Psi(t)\rangle, \tag{1.211}$$

it gives the probability amplitude for measuring the eigenvalues $\mathbf{x}$ of the position operator $\hat{\mathbf{x}}$, i.e., $|\Psi(\mathbf{x}, t)|^2$ is the probability density in $\mathbf{x}$-space.

The expectation value of the observable operator (1.206) in the state $|\Psi(t)\rangle$ is defined as the matrix element

$$\langle\Psi(t)|\hat{A}|\Psi(t)\rangle \equiv \int d^3x \langle\Psi(t)|\mathbf{x}\rangle A(-i\hbar\boldsymbol{\nabla}, \mathbf{x})\langle\mathbf{x}|\Psi(t)\rangle. \tag{1.212}$$

1.5.1 Uncertainty Relation

We have seen before [see the discussion after (1.83), (1.84)] that the amplitudes in real space and those in momentum space have widths inversely proportional to each other, due to the properties of Fourier analysis. If a wave packet is localized in real space with a width $\Delta\mathbf{x}$, its momentum space wave function has a width $\Delta\mathbf{p}$ given by

$$\Delta\mathbf{x}\,\Delta\mathbf{p} \sim \hbar. \tag{1.213}$$

From the Hilbert space point of view this uncertainty relation can be shown to be a consequence of the fact that the operators $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$ do not commute with each other, but the components satisfy the canonical commutation rules

$$\begin{aligned} [\hat{p}_i, \hat{x}_j] &= -i\hbar\delta_{ij}, \\ [\hat{x}_i, \hat{x}_j] &= 0, \\ [\hat{p}_i, \hat{p}_j] &= 0. \end{aligned} \tag{1.214}$$

In general, if an observable operator $\hat{A}$ is measured to have a sharp value a in one state, this state must be an eigenstate of $\hat{A}$ with an eigenvalue a:

$$\hat{A}|a\rangle = a|a\rangle. \tag{1.215}$$

This follows from the expansion

$$|\Psi(t)\rangle = \sum_a |a\rangle\langle a|\Psi(t)\rangle, \tag{1.216}$$

in which $|\langle a|\Psi(t)\rangle|^2$ is the probability to measure an arbitrary eigenvalue a. If this probability is sharply focused at a specific value of a, the state necessarily coincides with $|a\rangle$.

Given the set of all eigenstates $|a\rangle$ of $\hat{A}$, we may ask under what circumstances another observable, say $\hat{B}$, can be measured sharply in each of these states. The requirement implies that the states $|a\rangle$ are also eigenstates of $\hat{B}$,

$$\hat{B}|a\rangle = b_a|a\rangle, \tag{1.217}$$

with some a-dependent eigenvalue b_a. If this is true for all $|a\rangle$,

$$\hat{B}\hat{A}|a\rangle = b_a a|a\rangle = a b_a|a\rangle = \hat{A}\hat{B}|a\rangle, \tag{1.218}$$

the operators $\hat{A}$ and $\hat{B}$ necessarily commute:

$$[\hat{A}, \hat{B}] = 0. \tag{1.219}$$

Conversely, it can be shown that a vanishing commutator is also sufficient for two observable operators to be simultaneously diagonalizable and thus to allow for simultaneous sharp measurements.

1.5.2 Density Matrix and Wigner Function

An important object for calculating observable properties of a quantum-mechanical system is the quantum mechanical density operator associated with a pure state

$$\hat{\rho}(t) \equiv |\Psi(t)\rangle\langle\Psi(t)|, \tag{1.220}$$

and the associated density matrix associated with a pure state

$$\rho(\mathbf{x}_1, \mathbf{x}_2; t) = \langle\mathbf{x}_1|\Psi(t)\rangle\langle\Psi(t)|\mathbf{x}_2\rangle. \tag{1.221}$$

The expectation value of any function $f(\mathbf{x}, \hat{\mathbf{p}})$ can be calculated from the trace

$$\langle\Psi(t)|f(\mathbf{x}, \hat{\mathbf{p}})|\Psi(t)\rangle = \mathrm{tr}\,[f(\mathbf{x}, \hat{\mathbf{p}})\hat{\rho}(t)] = \int d^3x \langle\Psi(t)|\mathbf{x}\rangle f(\mathbf{x}, -i\hbar\boldsymbol{\nabla})\langle\mathbf{x}|\Psi(t)\rangle. \tag{1.222}$$

If we decompose the states $|\Psi(t)\rangle$ into stationary eigenstates $|E_n\rangle$ of the Hamiltonian operator $\hat{H}$ [recall (1.175)], $|\Psi(t)\rangle = \sum_n |E_n\rangle\langle E_n|\Psi(t)\rangle$, then the density matrix has the expansion

$$\hat{\rho}(t) \equiv \sum_{n,m} |E_n\rangle\rho_{nm}(t)\langle E_m| = \sum_{n,m} |E_n\rangle\langle E_n|\Psi(t)\rangle\langle\Psi(t)|E_m\rangle\langle E_m|. \tag{1.223}$$

Wigner showed that the Fourier transform of the density matrix, the *Wigner function*

$$W(\mathbf{X}, \mathbf{p}; t) \equiv \int \frac{d^3\Delta x}{(2\pi\hbar)^3} e^{i\mathbf{p}\Delta\mathbf{x}/\hbar}\rho(\mathbf{X} + \Delta\mathbf{x}/2, \mathbf{X} - \Delta\mathbf{x}/2; t) \tag{1.224}$$

satisfies, for a single particle of mass M in a potential $V(\mathbf{x})$, the Wigner-Liouville equation

$$\left(\partial_t + \mathbf{v}\cdot\boldsymbol{\nabla}_{\mathbf{X}}\right) W(\mathbf{X},\mathbf{p};t) = W_t(\mathbf{X},\mathbf{p};t), \qquad \mathbf{v}\equiv\frac{\mathbf{p}}{M}, \tag{1.225}$$

where

$$W_t(\mathbf{X},\mathbf{p};t) \equiv \frac{2}{\hbar}\int\frac{d^3q}{(2\pi\hbar)^3}\,W(\mathbf{X},\mathbf{p}-\mathbf{q};t)\int d^3\Delta\mathbf{x}\,V(\mathbf{X}-\Delta\mathbf{x}/2)e^{i\mathbf{q}\Delta\mathbf{x}/\hbar}. \tag{1.226}$$

In the limit $\hbar\to 0$, we may expand $W(\mathbf{X},\mathbf{p}-\mathbf{q};t)$ in powers of $\mathbf{q}$, and $V(\mathbf{X}-\Delta\mathbf{x}/2)$ in powers of $\Delta\mathbf{x}$, which we rewrite in front of the exponential $e^{i\mathbf{q}\Delta\mathbf{x}/\hbar}$ as powers of $-i\hbar\nabla_{\mathbf{q}}$. Then we perform the integral over $\Delta\mathbf{x}$ to obtain $(2\pi\hbar)^3\delta^{(3)}(\mathbf{q})$, and perform the integral over $\mathbf{q}$ to obtain the classical *Liouville equation* for the probability density of the particle in phase space

$$\left(\partial_t + \mathbf{v}\cdot\boldsymbol{\nabla}_{\mathbf{X}}\right) W(\mathbf{X},\mathbf{p};t) = -F(\mathbf{X})\boldsymbol{\nabla}_{\mathbf{p}}W(\mathbf{X},\mathbf{p};t), \qquad \mathbf{v}\equiv\frac{\mathbf{p}}{M}, \tag{1.227}$$

where $F(\mathbf{X})\equiv-\boldsymbol{\nabla}_{\mathbf{X}}V(\mathbf{X})$ is the force associated with the potential $V(\mathbf{X})$.

1.5.3 Generalization to Many Particles

All this development can be extended to systems of N distinguishable mass points with Cartesian coordinates $\mathbf{x}_1,\dots,\mathbf{x}_N$. If $H(\mathbf{p}_\nu,\mathbf{x}_\nu,t)$ is the Hamiltonian, the Schrödinger equation becomes

$$H(\hat{\mathbf{p}}_\nu,\hat{\mathbf{x}}_\nu,t)|\Psi(t)\rangle = i\hbar\partial_t|\Psi(t)\rangle. \tag{1.228}$$

We may introduce a complete local basis $|\mathbf{x}_1,\dots,\mathbf{x}_N\rangle$ with the properties

$$\begin{aligned}\langle\mathbf{x}_1,\dots,\mathbf{x}_N|\mathbf{x}'_1,\dots,\mathbf{x}'_N\rangle &= \delta^{(3)}(\mathbf{x}_1-\mathbf{x}'_1)\cdots\delta^{(3)}(\mathbf{x}_N-\mathbf{x}'_N),\\ \int d^3x_1\cdots d^3x_N|\mathbf{x}_1,\dots,\mathbf{x}_N\rangle\langle\mathbf{x}_1,\dots,\mathbf{x}_N| &= 1,\end{aligned} \tag{1.229}$$

and define

$$\begin{aligned}\langle\mathbf{x}_1,\dots,\mathbf{x}_N|\hat{\mathbf{p}}_\nu &= -i\hbar\partial_{\mathbf{x}_\nu}\langle\mathbf{x}_1,\dots,\mathbf{x}_N|,\\ \langle\mathbf{x}_1,\dots,\mathbf{x}_N|\hat{\mathbf{x}}_\nu &= \mathbf{x}_\nu\langle\mathbf{x}_1,\dots,\mathbf{x}_N|.\end{aligned} \tag{1.230}$$

The Schrödinger equation for N particles (1.107) follows from (1.228) by multiplying it from the left with the bra vectors $\langle\mathbf{x}_1,\dots,\mathbf{x}_N|$. In the same way, all other formulas given above can be generalized to N-body state vectors.

1.6 Time Evolution Operator

If the Hamiltonian operator possesses no explicit time dependence, the basis-independent Schrödinger equation (1.163) can be integrated to find the wave function $|\Psi(t)\rangle$ at any time t_b from the state at any other time t_a

$$|\Psi(t_b)\rangle = e^{-i(t_b-t_a)\hat{H}/\hbar}|\Psi(t_a)\rangle. \tag{1.231}$$

The operator

$$\hat{U}(t_b, t_a) = e^{-i(t_b - t_a)\hat{H}/\hbar} \tag{1.232}$$

is called the *time evolution operator*. It satisfies the differential equation

$$i\hbar\partial_{t_b}\hat{U}(t_b, t_a) = \hat{H}\,\hat{U}(t_b, t_a). \tag{1.233}$$

Its inverse is obtained by interchanging the order of t_b and t_a:

$$\hat{U}^{-1}(t_b, t_a) \equiv e^{i(t_b - t_a)\hat{H}/\hbar} = \hat{U}(t_a, t_b). \tag{1.234}$$

As an exponential of i times a Hermitian operator, $\hat{U}$ is a *unitary operator* satisfying

$$\hat{U}^\dagger = \hat{U}^{-1}. \tag{1.235}$$

Indeed,

$$\begin{aligned} \hat{U}^\dagger(t_b, t_a) &= e^{i(t_b - t_a)\hat{H}^\dagger/\hbar} \\ &= e^{i(t_b - t_a)\hat{H}/\hbar} = \hat{U}^{-1}(t_b, t_a). \end{aligned} \tag{1.236}$$

If $H(\hat{\mathbf{p}}, \hat{\mathbf{x}}, t)$ depends explicitly on time, the integration of the Schrödinger equation (1.163) is somewhat more involved. The solution may be found iteratively: For $t_b > t_a$, the time interval is sliced into a large number $N+1$ of small pieces of thickness ϵ with $\epsilon \equiv (t_b - t_a)/(N+1)$, slicing once at each time $t_n = t_a + n\epsilon$ for $n = 0, \ldots, N+1$. We then use the Schrödinger equation (1.163) to relate the wave function in each slice approximately to the previous one:

$$\begin{aligned} |\Psi(t_a+\epsilon)\rangle &\approx \left(1 - \frac{i}{\hbar}\int_{t_a}^{t_a+\epsilon} dt\,\hat{H}(t)\right)\Big|\Psi(t_a)\Big\rangle, \\ |\Psi(t_a+2\epsilon)\rangle &\approx \left(1 - \frac{i}{\hbar}\int_{t_a+\epsilon}^{t_a+2\epsilon} dt\,\hat{H}(t)\right)|\Psi(t_a+\epsilon)\rangle, \\ &\;\;\vdots \\ |\Psi(t_a+(N+1)\epsilon)\rangle &\approx \left(1 - \frac{i}{\hbar}\int_{t_a+N\epsilon}^{t_a+(N+1)\epsilon} dt\,\hat{H}(t)\right)|\Psi(t_a+N\epsilon)\rangle. \end{aligned} \tag{1.237}$$

From the combination of these equations we extract the evolution operator as a product

$$\hat{U}(t_b, t_a) \approx \left(1 - \frac{i}{\hbar}\int_{t_N}^{t_b} dt'_{N+1}\,\hat{H}(t'_{N+1})\right) \times \cdots \times \left(1 - \frac{i}{\hbar}\int_{t_a}^{t_1} dt'_1\,\hat{H}(t'_1)\right). \tag{1.238}$$

By multiplying out the product and going to the limit $N \to \infty$ we find the series

$$\begin{aligned} \hat{U}(t_b, t_a) &= 1 - \frac{i}{\hbar}\int_{t_a}^{t_b} dt'_1\,\hat{H}(t'_1) + \left(\frac{-i}{\hbar}\right)^2 \int_{t_a}^{t_b} dt'_2 \int_{t_a}^{t'_2} dt'_1\,\hat{H}(t'_2)\hat{H}(t'_1) \\ &\quad + \left(\frac{-i}{\hbar}\right)^3 \int_{t_a}^{t_b} dt'_3 \int_{t_a}^{t'_3} dt'_2 \int_{t_a}^{t'_2} dt'_1\,\hat{H}(t'_3)\hat{H}(t'_2)\hat{H}(t'_1) + \ldots\;, \end{aligned} \tag{1.239}$$

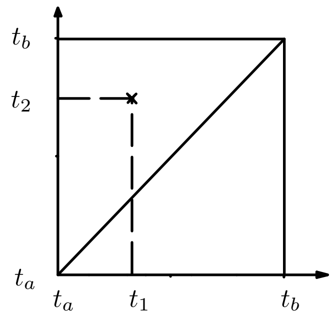

Figure 1.3 Illustration of time-ordering procedure in Eq. (1.243).

known as the *Neumann-Liouville expansion* or *Dyson series*. An interesting modification of this is the so-called *Magnus expansion* to be derived in Eq. (2A.25).

Note that each integral has the time arguments in the Hamilton operators ordered *causally*: Operators with later times stand to left of those with earlier times. It is useful to introduce a *time-ordering operator* which, when applied to an arbitrary product of operators,

$$\hat{O}_n(t_n)\cdots\hat{O}_1(t_1), \tag{1.240}$$

reorders the times successively. More explicitly we define

$$\hat{T}(\hat{O}_n(t_n)\cdots\hat{O}_1(t_1)) \equiv \hat{O}_{i_n}(t_{i_n})\cdots\hat{O}_{i_1}(t_{i_1}), \tag{1.241}$$

where $t_{i_n},\ldots,t_{i_1}$ are the times $t_n,\ldots,t_1$ relabeled in the causal order, so that

$$t_{i_n} > t_{i_{n-1}} > \ldots > t_{i_1}. \tag{1.242}$$

Any c-number factors in (1.241) can be pulled out in front of the $\hat{T}$ operator. With this formal operator, the Neumann-Liouville expansion can be rewritten in a more compact way. Take, for instance, the third term in (1.239)

$$\int_{t_a}^{t_b} dt_2 \int_{t_a}^{t_2} dt_1\, \hat{H}(t_2)\hat{H}(t_1). \tag{1.243}$$

The integration covers the triangle above the diagonal in the square $t_1, t_2 \in [t_a, t_b]$ in the (t_1, t_2) plane (see Fig. 1.2). By comparing this with the missing integral over the lower triangle

$$\int_{t_a}^{t_b} dt_2 \int_{t_2}^{t_b} dt_1\, \hat{H}(t_2)\hat{H}(t_1) \tag{1.244}$$

we see that the two expressions coincide except for the order of the operators. This can be corrected with the use of a time-ordering operator $\hat{T}$. The expression

$$\hat{T}\int_{t_a}^{t_b} dt_2 \int_{t_2}^{t_b} dt_1\, \hat{H}(t_2)\hat{H}(t_1) \tag{1.245}$$

is equal to (1.243) since it may be rewritten as

$$\int_{t_a}^{t_b} dt_2 \int_{t_2}^{t_b} dt_1\, \hat{H}(t_1)\hat{H}(t_2) \tag{1.246}$$

or, after interchanging the order of integration, as

$$\int_{t_a}^{t_b} dt_1 \int_{t_a}^{t_1} dt_2\, \hat{H}(t_1)\hat{H}(t_2). \tag{1.247}$$

Apart from the dummy integration variables $t_2 \leftrightarrow t_1$, this double integral coincides with (1.243). Since the time arguments are properly ordered, (1.243) can trivially be multiplied with the time-ordering operator. The conclusion of this discussion is that (1.243) can alternatively be written as

$$\frac{1}{2}\hat{T} \int_{t_a}^{t_b} dt_2 \int_{t_a}^{t_b} dt_1\, \hat{H}(t_2)\hat{H}(t_1). \tag{1.248}$$

On the right-hand side, the integrations now run over the full square in the t_1, t_2-plane so that the two integrals can be factorized into

$$\frac{1}{2}\hat{T} \left(\int_{t_a}^{t_b} dt\, \hat{H}(t) \right)^2. \tag{1.249}$$

Similarly, we may rewrite the nth-order term of (1.239) as

$$\begin{aligned} &\frac{1}{n!}\hat{T} \int_{t_a}^{t_b} dt_n \int_{t_a}^{t_b} dt_{n-1} \cdots \int_{t_a}^{t_b} dt_1\, \hat{H}(t_n)\hat{H}(t_{n-1}) \cdots \hat{H}(t_1) \\ &= \frac{1}{n!}\hat{T}\left[\int_{t_a}^{t_b} dt\, \hat{H}(t) \right]^n. \end{aligned} \tag{1.250}$$

The time evolution operator $\hat{U}(t_b, t_a)$ has therefore the series expansion

$$\begin{aligned} \hat{U}(t_b, t_a) &= 1 - \frac{i}{\hbar}\hat{T} \int_{t_a}^{t_b} dt\, \hat{H}(t) + \frac{1}{2!}\left(\frac{-i}{\hbar}\right)^2 \hat{T}\Big(\int_{t_a}^{t_b} dt\, \hat{H}(t)\Big)^2 \\ &\quad + \ldots + \frac{1}{n!}\left(\frac{-i}{\hbar}\right)^n \hat{T}\Big(\int_{t_a}^{t_b} dt\, \hat{H}(t)\Big)^n + \ldots\ . \end{aligned} \tag{1.251}$$

The right-hand side of $\hat{T}$ contains simply the power series expansion of the exponential so that we can write

$$\hat{U}(t_b, t_a) = \hat{T} \exp\left\{ -\frac{i}{\hbar} \int_{t_a}^{t_b} dt\, \hat{H}(t) \right\}. \tag{1.252}$$

If $\hat{H}$ does not depend on the time, the time-ordering operation is superfluous, the integral can be done trivially, and we recover the previous result (1.232).

Note that a small variation $\delta\hat{H}(t)$ of $\hat{H}(t)$ changes $\hat{U}(t_b, t_a)$ by

$$\begin{aligned} \delta\hat{U}(t_b, t_a) &= -\frac{i}{\hbar} \int_{t_a}^{t_b} dt'\ \hat{T}\exp\left\{ -\frac{i}{\hbar}\int_{t'}^{t_b} dt\, \hat{H}(t) \right\}\ \delta\hat{H}(t')\ \hat{T}\exp\left\{ -\frac{i}{\hbar}\int_{t_a}^{t'} dt\, \hat{H}(t) \right\} \\ &= -\frac{i}{\hbar} \int_{t_a}^{t_b} dt'\, \hat{U}(t_b, t')\, \delta\hat{H}(t')\, \hat{U}(t', t_a). \end{aligned} \tag{1.253}$$

A simple application for this relation is given in Appendix 1A.

1.7 Properties of the Time Evolution Operator

By construction, $\hat{U}(t_b, t_a)$ has some important properties:

a) Fundamental composition law
If two time translations are performed successively, the corresponding operators $\hat{U}$ are related by

$$\hat{U}(t_b, t_a) = \hat{U}(t_b, t')\hat{U}(t', t_a), \qquad t' \in (t_a, t_b). \tag{1.254}$$

This composition law makes the operators $\hat{U}$ a representation of the abelian group of time translations. For time-independent Hamiltonians with $\hat{U}(t_b, t_a)$ given by (1.232), the proof of (1.254) is trivial. In the general case (1.252), it follows from the simple manipulation valid for $t_b > t_a$:

$$\begin{aligned}
&\hat{T}\exp\left(-\frac{i}{\hbar}\int_{t'}^{t_b}\hat{H}(t)\,dt\right)\hat{T}\exp\left(-\frac{i}{\hbar}\int_{t_a}^{t'}\hat{H}(t)\,dt\right)\\
&= \hat{T}\left[\exp\left(-\frac{i}{\hbar}\int_{t'}^{t_b}\hat{H}(t)\,dt\right)\exp\left(-\frac{i}{\hbar}\int_{t_a}^{t'}\hat{H}(t)\,dt\right)\right]\\
&= \hat{T}\exp\left(-\frac{i}{\hbar}\int_{t_a}^{t_b}\hat{H}(t)\,dt\right).
\end{aligned} \tag{1.255}$$

b) Unitarity
The expression (1.252) for the time evolution operator $\hat{U}(t_b, t_a)$ was derived only for the *causal* (or *retarded*) time arguments, i.e., for t_b later than t_a. We may, however, define $\hat{U}(t_b, t_a)$ also for the *anticausal* (or *advanced*) case where t_b lies before t_a. To be consistent with the above composition law (1.254), we must have

$$\hat{U}(t_b, t_a) \equiv \hat{U}(t_a, t_b)^{-1}. \tag{1.256}$$

Indeed, when considering two states at successive times

$$|\Psi(t_a)\rangle = \hat{U}(t_a, t_b)|\Psi(t_b)\rangle, \tag{1.257}$$

the order of succession is inverted by multiplying both sides by $\hat{U}^{-1}(t_a, t_b)$:

$$|\Psi(t_b)\rangle = \hat{U}(t_a, t_b)^{-1}|\Psi(t_a)\rangle, \quad t_b < t_a. \tag{1.258}$$

The operator on the right-hand side is defined to be the time evolution operator $\hat{U}(t_b, t_a)$ from the later time t_a to the earlier time t_b.

If the Hamiltonian is independent of time, with the time evolution operator being

$$\hat{U}(t_a, t_b) = e^{-i(t_a - t_b)\hat{H}/\hbar}, \qquad t_a > t_b, \tag{1.259}$$

the unitarity of the operator $\hat{U}(t_b, t_a)$ is obvious:

$$\hat{U}^{\dagger}(t_b, t_a) = \hat{U}(t_b, t_a)^{-1}, \qquad t_b < t_a. \tag{1.260}$$

Let us verify this property for a general time-dependent Hamiltonian. There, a direct solution of the Schrödinger equation (1.163) for the state vector shows that the operator $\hat{U}(t_b, t_a)$ for $t_b < t_a$ has a representation just like (1.252), except for a reversed time order of its arguments. One writes this in the form [compare (1.252)]

$$\hat{U}(t_b, t_a) = \hat{\bar{T}} \exp\left\{\frac{i}{\hbar}\int_{t_a}^{t_b} \hat{H}(t)\, dt\right\}, \tag{1.261}$$

where $\hat{\bar{T}}$ denotes the time-antiordering operator, with an obvious definition analogous to (1.241), (1.242). This operator satisfies the relation

$$\left[\hat{T}\left(\hat{O}_1(t_1)\hat{O}_2(t_2)\right)\right]^\dagger = \hat{\bar{T}}\left(\hat{O}_2^\dagger(t_2)\hat{O}_1^\dagger(t_1)\right), \tag{1.262}$$

with an obvious generalization to the product of n operators. We can therefore conclude right away that

$$\hat{U}^\dagger(t_b, t_a) = \hat{U}(t_a, t_b), \quad t_b > t_a. \tag{1.263}$$

With $\hat{U}(t_a, t_b) \equiv \hat{U}(t_b, t_a)^{-1}$, this proves the unitarity relation (1.260), in general.

c) Schrödinger equation for $\hat{U}(t_b, t_a)$

Since the operator $\hat{U}(t_b, t_a)$ rules the relation between arbitrary wave functions at different times,

$$|\Psi(t_b)\rangle = \hat{U}(t_b, t_a)|\Psi(t_a)\rangle, \tag{1.264}$$

the Schrödinger equation (1.228) implies that the operator $\hat{U}(t_b, t_a)$ satisfies the corresponding equations

$$i\hbar\partial_t \hat{U}(t, t_a) = \hat{H}\hat{U}(t, t_a), \tag{1.265}$$

$$i\hbar\partial_t \hat{U}(t, t_a)^{-1} = -\hat{U}(t, t_a)^{-1}\hat{H}, \tag{1.266}$$

with the initial condition

$$\hat{U}(t_a, t_a) = 1. \tag{1.267}$$

1.8 Heisenberg Picture of Quantum Mechanics

The unitary time evolution operator $\hat{U}(t, t_a)$ may be used to give a different formulation of quantum mechanics bearing the closest resemblance to classical mechanics. This formulation, called the *Heisenberg picture* of quantum mechanics, is in a way more closely related to classical mechanics than the Schrödinger formulation. Many classical equations remain valid by simply replacing the canonical variables $p_i(t)$ and $q_i(t)$ in phase space by Heisenberg *operators*, to be denoted by $p_{Hi}(t)$, $q_{Hi}(t)$. Originally, Heisenberg postulated that they are matrices, but later it became clear that these matrices had to be functional matrix elements of operators, whose indices can be partly continuous. The classical equations hold for the Heisenberg operators and as long as the canonical commutation rules (1.93) are respected at any given time. In addition, $q_i(t)$ must be Cartesian coordinates. In this case we shall always use the notation x_i for the position variable, as in Section 1.4, rather than q_i. The corresponding Heisenberg operators are $\hat{x}_{Hi}(t)$. Suppressing the subscripts i, the canonical equal-time commutation rules are

$$\begin{aligned} [\hat{p}_H(t), \hat{x}_H(t)] &= -i\hbar, \\ [\hat{p}_H(t), \hat{p}_H(t)] &= 0, \\ [\hat{x}_H(t), \hat{x}_H(t)] &= 0. \end{aligned} \tag{1.268}$$

According to Heisenberg, classical equations involving Poisson brackets remain valid if the Poisson brackets are replaced by $i/\hbar$ times the matrix commutators at equal times. The canonical commutation relations (1.268) are a special case of this rule, recalling the fundamental Poisson brackets (1.24). The Hamilton equations of motion (1.23) turn into the Heisenberg equations

$$\begin{aligned} \frac{d}{dt}\hat{p}_H(t) &= \frac{i}{\hbar}\left[\hat{H}_H, \hat{p}_H(t)\right], \\ \frac{d}{dt}\hat{x}_H(t) &= \frac{i}{\hbar}\left[\hat{H}_H, \hat{x}_H(t)\right], \end{aligned} \tag{1.269}$$

where

$$\hat{H}_H \equiv H(\hat{p}_H(t), \hat{x}_H(t), t) \tag{1.270}$$

is the Hamiltonian in the Heisenberg picture. Similarly, the equation of motion for an arbitrary observable function $O(p_i(t), x_i(t), t)$ derived in (1.19) goes over into the matrix commutator equation for the Heisenberg operator

$$\hat{O}_H(t) \equiv O(\hat{p}_H(t), \hat{x}_H(t), t), \tag{1.271}$$

namely,

$$\frac{d}{dt}\hat{O}_H = \frac{i}{\hbar}[\hat{H}_H, \hat{O}_H] + \frac{\partial}{\partial t}\hat{O}_H. \tag{1.272}$$

These rules are referred to as *Heisenberg's correspondence principle*.

The relation between Schrödinger's and Heisenberg's picture is supplied by the time evolution operator. Let $\hat{O}$ be an arbitrary observable in the Schrödinger description

$$\hat{O}(t) \equiv O(\hat{p}, \hat{x}, t). \tag{1.273}$$

If the states $|\Psi_a(t)\rangle$ are an arbitrary complete set of solutions of the Schrödinger equation, where a runs through discrete and continuous indices, the operator $\hat{O}(t)$ can be specified in terms of its functional matrix elements

$$O_{ab}(t) \equiv \langle\Psi_a(t)|\hat{O}(t)|\Psi_b(t)\rangle. \tag{1.274}$$

We can now use the unitary operator $\hat{U}(t, 0)$ to go to a new time-independent basis $|\Psi_{Ha}\rangle$, defined by

$$|\Psi_a(t)\rangle \equiv \hat{U}(t, 0)|\Psi_{Ha}\rangle. \tag{1.275}$$

Simultaneously, we transform the Schrödinger operators of the canonical coordinates $\hat{p}$ and $\hat{x}$ into the time-dependent canonical *Heisenberg operators* $\hat{p}_H(t)$ and $\hat{x}_H(t)$ via

$$\hat{p}_H(t) \equiv \hat{U}(t,0)^{-1}\,\hat{p}\,\hat{U}(t,0), \tag{1.276}$$
$$\hat{x}_H(t) \equiv \hat{U}(t,0)^{-1}\,\hat{x}\,\hat{U}(t,0). \tag{1.277}$$

At the time $t=0$, the Heisenberg operators $\hat{p}_H(t)$ and $\hat{x}_H(t)$ coincide with the time-independent Schrödinger operators $\hat{p}$ and $\hat{x}$, respectively. An arbitrary observable $\hat{O}(t)$ is transformed into the associated Heisenberg operator as

$$\begin{aligned}\hat{O}_H(t) &\equiv \hat{U}(t,t_a)^{-1}O(\hat{p},\hat{x},t)\hat{U}(t,t_a)\\ &\equiv O\left(\hat{p}_H(t),\hat{x}_H(t),t\right).\end{aligned} \tag{1.278}$$

The Heisenberg matrices $O_H(t)_{ab}$ are then obtained from the Heisenberg operators $\hat{O}_H(t)$ by sandwiching $\hat{O}_H(t)$ between the time-independent basis vectors $|\Psi_{H\,a}\rangle$:

$$O_H(t)_{ab} \equiv \langle\Psi_{H\,a}|\hat{O}_H(t)|\Psi_{H\,b}\rangle. \tag{1.279}$$

Note that the time dependence of these matrix elements is now completely due to the time dependence of the operators,

$$\frac{d}{dt}O_H(t)_{ab} \equiv \langle\Psi_{H\,a}|\frac{d}{dt}\hat{O}_H(t)|\Psi_{H\,b}\rangle. \tag{1.280}$$

This is in contrast to the Schrödinger representation (1.274), where the right-hand side would have contained two more terms from the time dependence of the wave functions. Due to the absence of such terms in (1.280) it is possible to study the equation of motion of the Heisenberg matrices independently of the basis by considering directly the Heisenberg operators. It is straightforward to verify that they do indeed satisfy the rules of Heisenberg's correspondence principle. Consider the time derivative of an arbitrary observable $\hat{O}_H(t)$,

$$\begin{aligned}\frac{d}{dt}\hat{O}_H(t) &= \left(\frac{d}{dt}\hat{U}^{-1}(t,t_a)\right)\hat{O}(t)\hat{U}(t,t_a)\\ &+ \hat{U}^{-1}(t,t_a)\left(\frac{\partial}{\partial t}\hat{O}(t)\right)\hat{U}(t,t_a)+\hat{U}^{-1}(t,t_a)\hat{O}(t)\left(\frac{d}{dt}\hat{U}(t,t_a)\right),\end{aligned}$$

which can be rearranged as

$$\begin{aligned}&\left[\left(\frac{d}{dt}\hat{U}^{-1}(t,t_a)\right)\hat{U}(t,t_a)\right]\hat{U}^{-1}(t,t_a)\hat{O}(t)\hat{U}(t,t_a)\\ &+\left[\hat{U}^{-1}(t,t_a)\hat{O}(t)\hat{U}(t,t_a)\right]\hat{U}^{-1}(t,t_a)\frac{d}{dt}\hat{U}(t,t_a)+\hat{U}^{-1}(t,t_a)\left(\frac{\partial}{\partial t}\hat{O}(t)\right)\hat{U}(t,t_a).\end{aligned} \tag{1.281}$$

Using (1.265), we obtain

$$\frac{d}{dt}\hat{O}_H(t)=\frac{i}{\hbar}\left[\hat{U}^{-1}\hat{H}\hat{U},\hat{O}_H\right]+\hat{U}^{-1}\left(\frac{\partial}{\partial t}\hat{O}(t)\right)\hat{U}. \tag{1.282}$$

After inserting (1.278), we find the equation of motion for the Heisenberg operator:

$$\frac{d}{dt}\hat{O}_H(t) = \frac{i}{\hbar}\left[\hat{H}_H, \hat{O}_H(t)\right] + \left(\frac{\partial}{\partial t}\hat{O}\right)_H(t). \tag{1.283}$$

By sandwiching this equation between the complete time-independent basis states $|\Psi_a\rangle$ in the Hilbert space, it holds for the matrices and turns into the Heisenberg equation of motion. For the phase space variables $p_H(t)$, $x_H(t)$ themselves, these equations reduce, of course, to the Hamilton equations of motion (1.269).

Thus we have shown that Heisenberg's matrix quantum mechanics is completely equivalent to Schrödinger's quantum mechanics, and that the Heisenberg matrices obey the same Hamilton equations as the classical observables.

1.9 Interaction Picture and Perturbation Expansion

For some physical systems, the Hamiltonian operator can be split into two contributions

$$\hat{H} = \hat{H}_0 + \hat{V}, \tag{1.284}$$

where $\hat{H}_0$ is a so-called free Hamiltonian operator for which the Schrödinger equation $\hat{H}_0|\psi(t)\rangle = i\hbar\partial_t|\psi(t)\rangle$ can be solved, and $\hat{V}$ is an interaction potential which perturbs these solutions slightly. In this case it is useful to describe the system in Dirac's *interaction picture*. We remove the time evolution of the unperturbed Schrödinger solutions and define the states

$$|\psi_I(t)\rangle \equiv e^{i\hat{H}_0 t/\hbar}|\psi(t)\rangle. \tag{1.285}$$

Their time evolution comes entirely from the interaction potential $\hat{V}$. It is governed by the time evolution operator

$$\hat{U}_I(t_b, t_a) \equiv e^{iH_0 t_b/\hbar} e^{-iH(t_b - t_a)/\hbar} e^{-iH_0 t_a/\hbar}, \tag{1.286}$$

and reads

$$|\psi_I(t_b)\rangle = \hat{U}_I(t_b, t_a)|\psi_I(t_a)\rangle. \tag{1.287}$$

If $\hat{V} = 0$, the states $|\psi_I(t_b)\rangle$ are time-independent and coincide with the Heisenberg states (1.275) of the operator $\hat{H}_0$.

The operator $\hat{U}_I(t_b, t_a)$ satisfies the equation of motion

$$i\hbar\partial_{t_b}\hat{U}_I(t_b, t_a) = V_I(t_b)\hat{U}_I(t_b, t_a), \tag{1.288}$$

where

$$\hat{V}_I(t) \equiv e^{iH_0 t/\hbar}\hat{V}e^{-iH_0 t/\hbar} \tag{1.289}$$

is the potential in the interaction picture. This equation of motion can be turned into an integral equation

$$\hat{U}_I(t_b, t_a) = 1 - \frac{i}{\hbar}\int_{t_a}^{t_b} dt V_I(t)\hat{U}_I(t, t_a). \tag{1.290}$$

Inserting Eq. (1.289), this reads

$$\hat{U}_I(t_b,t_a) = 1 - \frac{i}{\hbar}\int_{t_a}^{t_b} dt\, e^{i\hat{H}_0 t/\hbar} V e^{-i\hat{H}_0 t/\hbar}\hat{U}_I(t,t_a). \tag{1.291}$$

This equation can be iterated to find a perturbation expansion for the operator $\hat{U}_I(t_b,t_a)$ in powers of the interaction potential:

$$\hat{U}_I(t_b,t_a) = 1 - \frac{i}{\hbar}\int_{t_a}^{t_b} dt\, e^{i\hat{H}_0 t/\hbar} V e^{-i\hat{H}_0 t/\hbar} + \left(-\frac{i}{\hbar}\right)^2 \int_{t_a}^{t_b} dt \int_{t_a}^{t} dt'\, e^{i\hat{H}_0 t/\hbar} V e^{-i\hat{H}_0 (t-t')/\hbar} V e^{-i\hat{H}_0 t'/\hbar} + \dots . \tag{1.292}$$

Inserting on the left-hand side the operator (1.286) and multiplying the equation from the left by $e^{-i\hat{H}_0 t_b/\hbar}$ and from the right by $e^{i\hat{H}_0 t_a/\hbar}$, this can also be rewritten as

$$e^{-iH(t_b-t_a)/\hbar} = e^{-iH_0(t_b-t_a)/\hbar} - \frac{i}{\hbar}\int_{t_a}^{t_b} dt\, e^{-i\hat{H}_0(t_b-t)/\hbar} V e^{-i\hat{H}_0(t-t_a)/\hbar} + \left(-\frac{i}{\hbar}\right)^2 \int_{t_a}^{t_b} dt \int_{t_a}^{t} dt'\, e^{-i\hat{H}_0(t_b-t)/\hbar} V e^{-i\hat{H}_0(t-t')/\hbar} V e^{-i\hat{H}_0(t'-t_a)/\hbar} + \dots . \tag{1.293}$$

This expansion is seen to be the recursive solution of the integral equation

$$e^{-iH(t_b-t_a)/\hbar} = e^{-iH_0(t_b-t_a)/\hbar} - \frac{i}{\hbar}\int_{t_a}^{t_b} dt\, e^{-i\hat{H}_0(t_b-t)/\hbar} V e^{-i\hat{H}(t-t_a)/\hbar}. \tag{1.294}$$

Note that the lowest-order correction agrees with the previous formula (1.253)

Another way of writing the expansion (1.293) is

$$e^{-iH(t_b-t_a)/\hbar} = e^{-iH_0 t_b/\hbar}\,\hat{T}\exp\left\{-\frac{i}{\hbar}\int_{t_a}^{t_b} dt\, e^{i\hat{H}_0 t/\hbar} V e^{-i\hat{H}_0 t/\hbar}\right\} e^{iHt_a/\hbar}. \tag{1.295}$$

This may be recorded as a mathematical operator formula

$$e^{T(\hat{A}+\hat{B})} = \hat{T} e^{\int_0^T dt e^{(T-t)\hat{A}}\hat{B}e^{t\hat{A}}} = e^{T\hat{A}}\hat{T}e^{\int_0^T dt e^{-t\hat{A}}\hat{B}e^{t\hat{A}}}. \tag{1.296}$$

Due to the time-ordering operator, the right-hand side cannot be evaluated with the help *Lie's expansion formula*, also known as *Hadamard's lemma*

$$e^{-t\hat{A}}\hat{B}e^{t\hat{A}} = \hat{B} - t[\hat{A},\hat{B}] + \frac{t^2}{2!}[\hat{A},[\hat{A},\hat{B}]] + \dots \ . \tag{1.297}$$

Due to $\hat{T}$, the evaluation is considerably more involved and relegated to Appendix 2A. The proper expression of the right-hand side is referred to as the Campbell-Baker-Hausdorff expansion

A simple consequence of Hadamard's lemma is a variation formula for a time-dependent operator $\hat{A}(t)$:

$$\delta e^{\hat{A}(t)} = \int_0^1 dt e^{(1-t)\hat{A}}\,\delta\hat{A}\,e^{t\hat{A}}, \tag{1.298}$$

which follows from (1.297) by setting $\hat{B} = \delta\hat{A}$ and expanding $e^{T(\hat{A}+\delta\hat{A})} - e^{T\hat{A}}$ to lowest order in $\delta\hat{A}$. This is, of course, just another way of expressing Eq. (1.253).

1.10 Time Evolution Amplitude

In the subsequent development, an important role will be played by the matrix elements of the time evolution operator in the localized basis states,

$$(\mathbf{x}_b t_b|\mathbf{x}_a t_a) \equiv \langle \mathbf{x}_b|\hat{U}(t_b, t_a)|\mathbf{x}_a\rangle. \tag{1.299}$$

They are referred to as *time evolution* amplitudes. The functional matrix $(\mathbf{x}_b t_b|\mathbf{x}_a t_a)$ is also called the *propagator* of the system. For a system with a time-independent Hamiltonian operator where $\hat{U}(t_b, t_a)$ is given by (1.259), the propagator is simply

$$(\mathbf{x}_b t_b|\mathbf{x}_a t_a) = \langle \mathbf{x}_b| \exp[-i\hat{H}(t_b - t_a)/\hbar]|\mathbf{x}_a\rangle. \tag{1.300}$$

Due to the operator equations (1.265), the propagator satisfies the Schrödinger equation

$$[H(-i\hbar\partial_{\mathbf{x}_b}, \mathbf{x}_b, t_b) - i\hbar\partial_{t_b}]\,(\mathbf{x}_b t_b|\mathbf{x}_a t_a) = 0. \tag{1.301}$$

In the quantum mechanics of nonrelativistic particles, only the propagators from earlier to later times will be relevant. It is therefore customary to introduce the so-called *causal time evolution operator* or *retarded time evolution operator*:[9]

$$\hat{U}^R(t_b, t_a) \equiv \begin{cases} \hat{U}(t_b, t_a), & t_b \geq t_a, \\ 0, & t_b < t_a, \end{cases} \tag{1.302}$$

and the associated *causal time evolution amplitude* or *retarded time evolution amplitude*

$$(\mathbf{x}_b t_b|\mathbf{x}_a t_a)^R \equiv \langle \mathbf{x}_b|\hat{U}^R(t_b, t_a)|\mathbf{x}_a\rangle. \tag{1.303}$$

Since this differs from (1.299) only for $t_b < t_a$, and since all formulas in the subsequent text will be used only for $t_b > t_a$, we shall often omit the superscript R. To abbreviate the case distinction in (1.302), it is convenient to use the *Heaviside function* defined by

$$\Theta(t) \equiv \begin{cases} 1 & \text{for} \quad t > 0, \\ 0 & \text{for} \quad t \leq 0, \end{cases} \tag{1.304}$$

and write

$$U^R(t_b, t_a) \equiv \Theta(t_b - t_a)\hat{U}(t_b, t_a), \qquad (\mathbf{x}_b t_b|\mathbf{x}_a t_a)^R \equiv \Theta(t_b - t_a)(\mathbf{x}_b t_b|\mathbf{x}_a t_a). \tag{1.305}$$

There exists also another Heaviside function which differs from (1.304) only by the value at $t_b = t_a$:

$$\Theta^R(t) \equiv \begin{cases} 1 & \text{for} \quad t \geq 0, \\ 0 & \text{for} \quad t < 0. \end{cases} \tag{1.306}$$

Both Heaviside functions have the property that their derivative yields Dirac's δ-function

$$\partial_t \Theta(t) = \delta(t). \tag{1.307}$$

[9]Compare this with the retarded *Green functions* to be introduced in Section 18.1

If it is not important which Θ-function is used we shall ignore the superscript.

The retarded propagator satisfies the Schrödinger equation

$$\left[H(-i\hbar\partial_{\mathbf{x}_b}, \mathbf{x}_b, t_b)^R - i\hbar\partial_{t_b}\right](\mathbf{x}_b t_b|\mathbf{x}_a t_a)^R = -i\hbar\delta(t_b - t_a)\delta^{(3)}(\mathbf{x}_b - \mathbf{x}_a). \tag{1.308}$$

The nonzero right-hand side arises from the extra term

$$-i\hbar\left[\partial_{t_b}\Theta(t_b - t_a)\right]\langle\mathbf{x}_b t_b|\mathbf{x}_a t_a\rangle = -i\hbar\delta(t_b - t_a)\langle\mathbf{x}_b t_b|\mathbf{x}_a t_a\rangle = -i\hbar\delta(t_b - t_a)\langle\mathbf{x}_b t_a|\mathbf{x}_a t_a\rangle \tag{1.309}$$

and the initial condition $\langle\mathbf{x}_b t_a|\mathbf{x}_a t_a\rangle = \langle\mathbf{x}_b|\mathbf{x}_a\rangle$, due to (1.267).

If the Hamiltonian does not depend on time, the propagator depends only on the time difference $t = t_b - t_a$. The retarded propagator vanishes for $t < 0$. Functions $f(t)$ with this property have a characteristic Fourier transform. The integral

$$\tilde{f}(E) \equiv \int_0^\infty dt\; f(t)e^{iEt/\hbar} \tag{1.310}$$

is an analytic function in the upper half of the complex energy plane. This analyticity property is necessary and sufficient to produce a factor $\Theta(t)$ when inverting the Fourier transform via the energy integral

$$f(t) \equiv \int_{-\infty}^{\infty} \frac{dE}{2\pi\hbar}\; \tilde{f}(E)e^{-iEt/\hbar}. \tag{1.311}$$

For $t < 0$, the contour of integration may be closed by an infinite semicircle in the upper half-plane at no extra cost. Since the contour encloses no singularities, it can be contracted to a point, yielding $f(t) = 0$.

The Heaviside function $\Theta(t)$ itself is the simplest retarded function, with a Fourier representation containing just a single pole just below the origin of the complex energy plane:

$$\Theta(t) = \int_{-\infty}^{\infty} \frac{dE}{2\pi}\frac{i}{E + i\eta}e^{-iEt}, \tag{1.312}$$

where η is an infinitesimally small positive number. The integral representation is undefined for $t = 0$ and there are, in fact, infinitely many possible definitions for the Heaviside function depending on the value assigned to the function at the origin. A special role is played by the average of the Heaviside functions (1.306) and (1.304), which is equal to 1/2 at the origin:

$$\bar{\Theta}(t) \equiv \begin{cases} 1 & \text{for} \quad t > 0, \\ \frac{1}{2} & \text{for} \quad t = 0, \\ 0 & \text{for} \quad t < 0. \end{cases} \tag{1.313}$$

Usually, the difference in the value at the origin does not matter since the Heaviside function appears only in integrals accompanied by some smooth function $f(t)$. This makes the Heaviside function a distribution with respect to smooth test functions $f(t)$ as defined in Eq. (1.162). All three distributions $\Theta_r(t)$, $\Theta^l(t)$, and $\bar{\Theta}(t)$ define the same linear functional of the test functions by the integral

$$\Theta[f] = \int dt\,\Theta(t - t')f(t'), \tag{1.314}$$

and this is an element in the linear space of all distributions.

As announced after Eq. (1.162), path integrals will specify, in addition, integrals over products of distribution and thus give rise to an important extension of the theory of distributions in Chapter 10. In this, the Heaviside function $\bar{\Theta}(t-t')$ plays the main role.

While discussing the concept of distributions let us introduce, for later use, the closely related distribution

$$\epsilon(t-t') \equiv \Theta(t-t') - \Theta(t'-t) = \bar{\Theta}(t-t') - \bar{\Theta}(t'-t), \tag{1.315}$$

which is a step function jumping at the origin from -1 to 1 as follows:

$$\epsilon(t-t') = \begin{cases} 1 & \text{for} \quad t > t', \\ 0 & \text{for} \quad t = t', \\ -1 & \text{for} \quad t < t'. \end{cases} \tag{1.316}$$

1.11 Fixed-Energy Amplitude

The Fourier-transform of the retarded time evolution amplitude (1.303)

$$(\mathbf{x}_b|\mathbf{x}_a)_E = \int_{-\infty}^{\infty} dt_b e^{iE(t_b-t_a)/\hbar} (\mathbf{x}_b t_b|\mathbf{x}_b t_a)^R = \int_{t_a}^{\infty} dt_b e^{iE(t_b-t_a)/\hbar} (\mathbf{x}_b t_b|\mathbf{x}_a t_a) \tag{1.317}$$

is called the *fixed-energy amplitude*.

If the Hamiltonian does not depend on time, we insert here Eq. (1.300) and find that the fixed-energy amplitudes are matrix elements

$$(\mathbf{x}_b|\mathbf{x}_a)_E = \langle \mathbf{x}_b|\hat{R}(E)|\mathbf{x}_a\rangle \tag{1.318}$$

of the so-called *resolvent operator*

$$\hat{R}(E) = \frac{i\hbar}{E - \hat{H} + i\eta}, \tag{1.319}$$

which is the Fourier transform of the retarded time evolution operator (1.302):

$$\hat{R}(E) = \int_{-\infty}^{\infty} dt_b \, e^{iE(t_b-t_a)/\hbar} \hat{U}^R(t_b, t_a) = \int_{t_a}^{\infty} dt_b \, e^{iE(t_b-t_a)/\hbar} \hat{U}(t_b, t_a). \tag{1.320}$$

Let us suppose that the time-independent Schrödinger equation is completely solved, i.e., that one knows all solutions $|\psi_n\rangle$ of the equation

$$\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle. \tag{1.321}$$

These satisfy the completeness relation

$$\sum_n |\psi_n\rangle\langle\psi_n| = 1, \tag{1.322}$$

which can be inserted on the right-hand side of (1.300) between the Dirac brackets leading to the *spectral representation*

$$(\mathbf{x}_b t_b|\mathbf{x}_a t_a) = \sum_n \psi_n(\mathbf{x}_b)\psi_n^*(\mathbf{x}_a) \exp\left[-iE_n(t_b - t_a)/\hbar\right], \tag{1.323}$$

with

$$\psi_n(\mathbf{x}) = \langle \mathbf{x}|\psi_n\rangle \tag{1.324}$$

being the wave functions associated with the eigenstates $|\psi_n\rangle$. Applying the Fourier transform (1.317), we obtain

$$(\mathbf{x}_b|\mathbf{x}_a)_E = \sum_n \psi_n(\mathbf{x}_b)\psi_n^*(\mathbf{x}_a) R_n(E) = \sum_n \psi_n(\mathbf{x}_b)\psi_n^*(\mathbf{x}_a)\frac{i\hbar}{E - E_n + i\eta}. \tag{1.325}$$

The fixed-energy amplitude (1.317) contains as much information on the system as the time evolution amplitude, which is recovered from it by the inverse Fourier transformation

$$(\mathbf{x}_b t_a|\mathbf{x}_a t_a) = \int_{-\infty}^{\infty} \frac{dE}{2\pi\hbar} e^{-iE(t_b - t_a)/\hbar} (\mathbf{x}_b|\mathbf{x}_a)_E. \tag{1.326}$$

The small $i\eta$-shift in the energy E in (1.325) may be thought of as being attached to each of the energies E_n, which are thus placed by an infinitesimal piece *below* the real energy axis. Then the exponential behavior of the wave functions is slightly damped, going to zero at infinite time:

$$e^{-i(E_n - i\eta)t/\hbar} \to 0. \tag{1.327}$$

This so-called $i\eta$-prescription ensures the causality of the Fourier representation (1.326). When doing the Fourier integral (1.326), the exponential $e^{iE(t_b - t_a)/\hbar}$ makes it always possible to close the integration contour along the energy axis by an infinite semicircle in the complex energy plane, which lies in the upper half-plane for $t_b < t_a$ and in the lower half-plane for $t_b > t_a$. The $i\eta$-prescription guarantees that for $t_b < t_a$, there is no pole inside the closed contour making the propagator vanish. For $t_b > t_a$, on the other hand, the poles in the lower half-plane give, via Cauchy's residue theorem, the spectral representation (1.323) of the propagator. An $i\eta$-prescription will appear in another context in Section 2.3.

If the eigenstates are nondegenerate, the residues at the poles of (1.325) render directly the products of eigenfunctions (barring degeneracies which must be discussed separately). For a system with a continuum of energy eigenvalues, there is a cut in the complex energy plane which may be thought of as a closely spaced sequence of poles. In general, the wave functions are recovered from the discontinuity of the amplitudes $(\mathbf{x}_b|\mathbf{x}_a)_E$ across the cut, using the formula

$$\mathrm{disc}\left(\frac{i\hbar}{E - E_n}\right) \equiv \frac{i\hbar}{E - E_n + i\eta} - \frac{i\hbar}{E - E_n - i\eta} = 2\pi\hbar\delta(E - E_n). \tag{1.328}$$

Here we have employed the relation[10], valid inside integrals over E:

$$\frac{1}{E-E_n\pm i\eta}=\frac{\mathcal{P}}{E-E_n}\mp i\pi\delta(E-E_n), \tag{1.329}$$

where $\mathcal{P}$ indicates that the principal value of the integral has to be taken.

The energy integral over the discontinuity of the fixed-energy amplitude (1.325) $(\mathbf{x}_b|\mathbf{x}_a)_E$ reproduces the completeness relation (1.322) taken between the local states $\langle\mathbf{x}_b|$ and $|\mathbf{x}_a\rangle$,

$$\int_{-\infty}^{\infty}\frac{dE}{2\pi\hbar}\text{disc}\,(\mathbf{x}_b|\mathbf{x}_a)_E=\sum_n\psi_n(\mathbf{x}_b)\psi_n^*(\mathbf{x}_a)=\langle\mathbf{x}_b|\mathbf{x}_a\rangle=\delta^{(D)}(\mathbf{x}_b-\mathbf{x}_a). \tag{1.330}$$

The completeness relation reflects the following property of the resolvent operator:

$$\int_{-\infty}^{\infty}\frac{dE}{2\pi\hbar}\text{disc}\,\hat{R}(E)=\hat{1}. \tag{1.331}$$

In general, the system possesses also a continuous spectrum, in which case the completeness relation contains a spectral integral and (1.322) has the form

$$\sum_n|\psi_n\rangle\langle\psi_n|+\int d\nu\,|\psi_\nu\rangle\langle\psi_\nu|=1. \tag{1.332}$$

The continuum causes a branch cut along in the complex energy plane, and (1.330) includes an integral over the discontinuity along the cut. The cut will mostly be omitted, for brevity.

1.12 Free-Particle Amplitudes

For a free particle with a Hamiltonian operator $\hat{H}=\hat{\mathbf{p}}^2/2M$, the spectrum is continuous. The eigenfunctions are (1.189) with energies $E(\mathbf{p})=\mathbf{p}^2/2M$. Inserting the completeness relation (1.187) into Eq. (1.300), we obtain for the time evolution amplitude of a free particle the Fourier representation

$$(\mathbf{x}_bt_b|\mathbf{x}_at_a)=\int\frac{d^Dp}{(2\pi\hbar)^D}\exp\left\{\frac{i}{\hbar}\left[\mathbf{p}(\mathbf{x}_b-\mathbf{x}_a)-\frac{\mathbf{p}^2}{2M}(t_b-t_a)\right]\right\}. \tag{1.333}$$

The momentum integrals can easily be done. First we perform a quadratic completion in the exponent and rewrite it as

$$\mathbf{p}(\mathbf{x}_b-\mathbf{x}_a)-\frac{1}{2M}\mathbf{p}^2(t_b-t_a)=\frac{1}{2M}\left(\mathbf{p}-\frac{1}{M}\frac{\mathbf{x}_b-\mathbf{x}_a}{t_b-t_a}\right)^2(t_b-t_a)-\frac{M}{2}\frac{(\mathbf{x}_b-\mathbf{x}_a)^2}{t_b-t_a}. \tag{1.334}$$

[10]This is often referred to as *Sochocki's formula*. It is the beginning of an expansion in powers of $\eta>0$: $1/(x\pm i\eta)=\mathcal{P}/x\mp i\pi\delta(x)+\eta\,[\pi\delta'(x)\pm id_x\mathcal{P}/x]+\mathcal{O}(\eta^2)$.

Then we replace the integration variables by the shifted momenta $\mathbf{p}' = \mathbf{p} - (\mathbf{x}_b - \mathbf{x}_a)/(t_b - t_a)M$, and the amplitude (1.333) becomes

$$(\mathbf{x}_b t_b|\mathbf{x}_a t_a) = F(t_b - t_a) \exp\left[\frac{i}{\hbar}\frac{M}{2}\frac{(\mathbf{x}_b - \mathbf{x}_a)^2}{t_b - t_a}\right], \tag{1.335}$$

where $F(t_b - t_a)$ is the integral over the shifted momenta

$$F(t_b - t_a) \equiv \int \frac{d^D p'}{(2\pi\hbar)^D} \exp\left\{-\frac{i}{\hbar}\frac{\mathbf{p}'^{\,2}}{2M}(t_b - t_a)\right\}. \tag{1.336}$$

This can be performed using the *Fresnel integral formula*

$$\int_{-\infty}^{\infty} \frac{dp}{\sqrt{2\pi}} \exp\left(i\,\frac{a}{2}p^2\right) = \frac{1}{\sqrt{|a|}} \begin{cases} \sqrt{i}, & a > 0, \\ 1/\sqrt{i}, & a < 0. \end{cases} \tag{1.337}$$

Here the square root $\sqrt{i}$ denotes the phase factor $e^{i\pi/4}$: This follows from the Gauss formula

$$\int_{-\infty}^{\infty} \frac{dp}{\sqrt{2\pi}} \exp\left(-\frac{\alpha}{2}p^2\right) = \frac{1}{\sqrt{\alpha}}, \qquad \mathrm{Re}\,\alpha > 0, \tag{1.338}$$

by continuing α analytically from positive values into the right complex half-plane. As long as $\mathrm{Re}\,\alpha > 0$, this is straightforward. On the boundaries, i.e., on the positive and negative imaginary axes, one has to be careful. At $\alpha = \pm ia + \eta$ with $a \gtrless 0$ and infinitesimal $\eta > 0$, the integral is certainly convergent yielding (1.337). But the integral also converges for $\eta = 0$, as can easily be seen by substituting $x^2 = z$. See Appendix 1B.

Note that differentiation of Eq. (1.338) with respect to α yields the more general Gaussian integral formula

$$\int_{-\infty}^{\infty} \frac{dp}{\sqrt{2\pi}}\, p^{2n} \exp\left(-\frac{\alpha}{2}p^2\right) = \frac{1}{\sqrt{\alpha}}\frac{(2n-1)!!}{\alpha^n} \qquad \mathrm{Re}\,\alpha > 0, \tag{1.339}$$

where $(2n-1)!!$ is defined as the product $(2n-1)\cdot(2n-3)\cdots 1$. For odd powers p^{2n+1}, the integral vanishes. In the Fresnel formula (1.337), an extra integrand p^{2n} produces a factor $(i/a)^n$.

Since the Fresnel formula is a special analytically continued case of the Gauss formula, we shall in the sequel always speak of Gaussian integrations and use Fresnel's name only if the imaginary nature of the quadratic exponent is to be emphasized.

Applying this formula to (1.336), we obtain

$$F(t_b - t_a) = \frac{1}{\sqrt{2\pi i\hbar(t_b - t_a)/M}^{\,D}}, \tag{1.340}$$

so that the full time evolution amplitude of a free massive point particle is

$$(\mathbf{x}_b t_b|\mathbf{x}_a t_a) = \frac{1}{\sqrt{2\pi i\hbar(t_b - t_a)/M}^{\,D}} \exp\left[\frac{i}{\hbar}\frac{M}{2}\frac{(\mathbf{x}_b - \mathbf{x}_a)^2}{t_b - t_a}\right]. \tag{1.341}$$

In the limit $t_b \to t_a$, the left-hand side becomes the scalar product $\langle \mathbf{x}_b|\mathbf{x}_a\rangle = \delta^{(D)}(\mathbf{x}_b - \mathbf{x}_a)$, implying the following limiting formula for the δ-function

$$\delta^{(D)}(\mathbf{x}_b - \mathbf{x}_a) = \lim_{t_b - t_a \to 0} \frac{1}{\sqrt{2\pi i\hbar(t_b - t_a)/M}^D} \exp\left[\frac{i}{\hbar}\frac{M}{2}\frac{(\mathbf{x}_b - \mathbf{x}_a)^2}{t_b - t_a}\right]. \tag{1.342}$$

Inserting Eq. (1.333) into (1.317), we have for the fixed-energy amplitude the integral representation

$$(\mathbf{x}_b|\mathbf{x}_a)_E = \int_0^\infty d(t_b - t_a) \int \frac{d^Dp}{(2\pi\hbar)^D} \exp\left\{\frac{i}{\hbar}\left[\mathbf{p}(\mathbf{x}_b - \mathbf{x}_a) + (t_b - t_a)\left(E - \frac{\mathbf{p}^2}{2M}\right)\right]\right\}. \tag{1.343}$$

Performing the time integration yields

$$(\mathbf{x}_b|\mathbf{x}_a)_E = \int \frac{d^Dp}{(2\pi\hbar)^D} \exp\left[i\mathbf{p}(\mathbf{x}_b - \mathbf{x}_a)\right] \frac{i\hbar}{E - \mathbf{p}^2/2M + i\eta}, \tag{1.344}$$

where we have inserted a damping factor $e^{-\eta(t_b - t_a)}$ into the integral to ensure convergence at large $t_b - t_a$. For a more explicit result it is more convenient to calculate the Fourier transform (1.341):

$$(\mathbf{x}_b|\mathbf{x}_a)_E = \int_0^\infty d(t_b - t_a) \frac{1}{\sqrt{2\pi i\hbar(t_b - t_a)/M}^D} \exp\left\{\frac{i}{\hbar}\left[E(t_b - t_a) + \frac{M}{2}\frac{(\mathbf{x}_b - \mathbf{x}_a)^2}{t_b - t_a}\right]\right\}. \tag{1.345}$$

For $E < 0$, we set

$$\kappa \equiv \sqrt{-2ME/\hbar^2}, \tag{1.346}$$

and perform the integral with the help of the formula[11]

$$\int_0^\infty dt\, t^{\nu-1} e^{-i\gamma t + i\beta/t} = 2\left(\frac{\beta}{\gamma}\right)^{\nu/2} e^{-i\nu\pi/2} K_{-\nu}(2\sqrt{\beta\gamma}), \tag{1.347}$$

where $K_\nu(z)$ is the modified Bessel function which satifies $K_\nu(z) = K_{-\nu}(z)$[12] The result is

$$(\mathbf{x}_b|\mathbf{x}_a)_E = -i\frac{2M}{\hbar}\frac{\kappa^{D-2}}{(2\pi)^{D/2}}\frac{K_{D/2-1}(\kappa R)}{(\kappa R)^{D/2-1}}, \tag{1.348}$$

where $R \equiv |\mathbf{x}_b - \mathbf{x}_a|$. The simplest modified Bessel function is[13]

$$K_{1/2}(z) = K_{-1/2}(z) = \sqrt{\frac{\pi}{2z}} e^{-z}, \tag{1.349}$$

[11] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, New York, 1980, Formulas 3.471.10, 3.471.11, and 8.432.6

[12] *ibid.*, Formula 8.486.16

[13] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1965, Formula 10.2.17.

so that we find for $D = 1, 2, 3$, the amplitudes

$$-i\frac{M}{\hbar}\frac{1}{\kappa}e^{-\kappa R}, \qquad -i\frac{M}{\hbar}\frac{1}{\pi}K_0(\kappa R), \qquad -i\frac{M}{\hbar}\frac{1}{2\pi R}e^{-\kappa R}. \tag{1.350}$$

At $R = 0$, the amplitude (1.348) is finite for all $D \leq 2$, where we can use the small-argument behavior of the associated Bessel function[14]

$$K_\nu(z) = K_{-\nu}(z) \approx \frac{1}{2}\Gamma(\nu)\left(\frac{z}{2}\right)^{-\nu} \quad \text{for} \quad \mathrm{Re}\,\nu > 0, \tag{1.351}$$

to obtain

$$(\mathbf{x}|\mathbf{x})_E = -i\frac{2M}{\hbar}\frac{\kappa^{D-2}}{(4\pi)^{D/2}}\Gamma(1 - D/2). \tag{1.352}$$

This result can be continued analytically to $D > 2$, which will be needed later (for example in Subsection 4.9.4).

For $E > 0$ we set

$$k \equiv \sqrt{2ME/\hbar^2} \tag{1.353}$$

and use the formula[15]

$$\int_0^\infty dt t^{\nu-1} e^{i\gamma t + i\beta/t} = i\pi\left(\frac{\beta}{\gamma}\right)^{\nu/2} e^{-i\nu\pi/2} H^{(1)}_{-\nu}(2\sqrt{\beta\gamma}), \tag{1.354}$$

where $H^{(1)}_\nu(z)$ is the Hankel function, to find

$$(\mathbf{x}_b|\mathbf{x}_a)_E = \frac{M\pi}{\hbar}\frac{k^{D-2}}{(2\pi)^{D/2}}\frac{H^{(1)}_{D/2-1}(kR)}{(kR)^{D/2-1}}. \tag{1.355}$$

The relation[16]

$$K_\nu(-iz) = \frac{\pi}{2} i e^{i\nu\pi/2} H^{(1)}_\nu(z) \tag{1.356}$$

connects the two formulas with each other when continuing the energy from negative to positive values, which replaces κ by $e^{-i\pi/2}k = -ik$.

For large distances, the asymptotic behavior[17]

$$K_\nu(z) \approx \sqrt{\frac{\pi}{2z}}e^{-z}, \qquad H^{(1)}_\nu(z) \approx \sqrt{\frac{2}{\pi z}}e^{i(z-\nu\pi/2-\pi/4)} \tag{1.357}$$

shows that the fixed-energy amplitude behaves for $E < 0$ like

$$(\mathbf{x}_b|\mathbf{x}_a)_E \approx -i\frac{M}{\hbar}\kappa^{D-2}\frac{1}{(2\pi)^{(D-1)/2}}\frac{1}{(\kappa R)^{(D-1)/2}}e^{-\kappa R/\hbar}, \tag{1.358}$$

and for $E > 0$ like

$$(\mathbf{x}_b|\mathbf{x}_a)_E \approx \frac{M}{\hbar}k^{D-2}\frac{1}{(2\pi i)^{(D-1)/2}}\frac{1}{(kR)^{(D-1)/2}}e^{ikR/\hbar}. \tag{1.359}$$

For $D = 1$ and 3, these asymptotic expressions hold for all R.

---

[14] *ibid.*, Formula 9.6.9.
[15] *ibid.*, Formulas 3.471.11 and 8.421.7.
[16] *ibid.*, Formula 8.407.1.
[17] *ibid.*, Formulas 8.451.6 and 8.451.3.

1.13 Quantum Mechanics of General Lagrangian Systems

An extension of the quantum-mechanical formalism to systems described by a set of completely general Lagrange coordinates $q_1, \dots, q_N$ is not straightforward. Only in the special case of q_i $(i = 1, \dots, N)$ being merely a curvilinear reparametrization of a D-dimensional Euclidean space are the above correspondence rules sufficient to quantize the system. Then $N = D$ and a variable change from x^i to q_j in the Schrödinger equation leads to the correct quantum mechanics. It will be useful to label the curvilinear coordinates by Greek superscripts and write q^μ instead of q_j. This will help when we write all ensuing equations in a form that is manifestly covariant under coordinate transformations. In the original definition of generalized coordinates in Eq. (1.1), this was unnecessary since transformation properties were ignored. For the Cartesian coordinates we shall use Latin indices alternatively as sub- or superscripts. The coordinate transformation $x^i = x^i(q^\mu)$ implies the relation between the derivatives $\partial_\mu \equiv \partial/\partial q^\mu$ and $\partial_i \equiv \partial/\partial x^i$:

$$\partial_\mu = e^i{}_\mu(q)\partial_i, \tag{1.360}$$

with the transformation matrix

$$e^i{}_\mu(q) \equiv \partial_\mu x^i(q) \tag{1.361}$$

called *basis D-ad* (in 3 dimensions triad, in 4 dimensions tetrad, etc.). Let $e_i{}^\mu(q) = \partial q^\mu/\partial x^i$ be the inverse matrix (assuming it exists) called the *reciprocal D-ad*, satisfying with $e^i{}_\mu$ the orthogonality and completeness relations

$$e^i{}_\mu \, e_i{}^\nu = \delta_\mu{}^\nu, \qquad e^i{}_\mu \, e_j{}^\mu = \delta^i{}_j. \tag{1.362}$$

Then, (1.360) is inverted to

$$\partial_i = e_i{}^\mu(q)\partial_\mu \tag{1.363}$$

and yields the curvilinear transform of the Cartesian quantum-mechanical momentum operators

$$\hat{p}_i = -i\hbar\partial_i = -i\hbar e_i{}^\mu(q)\partial_\mu. \tag{1.364}$$

The free-particle Hamiltonian operator

$$\hat{H}_0 = \hat{T} = \frac{1}{2M}\hat{\mathbf{p}}^2 = -\frac{\hbar^2}{2M}\boldsymbol{\nabla}^2 \tag{1.365}$$

goes over into

$$\hat{H}_0 = -\frac{\hbar^2}{2M}\Delta, \tag{1.366}$$

where Δ is the Laplacian expressed in curvilinear coordinates:

$$\begin{aligned} \Delta &= \partial_i^2 = e^{i\mu}\partial_\mu e_i{}^\nu\partial_\nu \\ &= e^{i\mu}e_i{}^\nu\partial_\mu\partial_\nu + (e^{i\mu}\partial_\mu e_i{}^\nu)\partial_\nu. \end{aligned} \tag{1.367}$$

At this point one introduces the *metric tensor*

$$g_{\mu\nu}(q) \equiv e_{i\mu}(q)e^{i}{}_{\nu}(q), \tag{1.368}$$

its inverse

$$g^{\mu\nu}(q) = e^{i\mu}(q)e_{i}{}^{\nu}(q), \tag{1.369}$$

defined by $g^{\mu\nu}g_{\nu\lambda} = \delta^{\mu}{}_{\lambda}$, and the so-called *affine connection*

$$\Gamma_{\mu\nu}{}^{\lambda}(q) = -e^{i}{}_{\nu}(q)\partial_{\mu}e_{i}{}^{\lambda}(q) = e_{i}{}^{\lambda}(q)\partial_{\mu}e^{i}{}_{\nu}(q). \tag{1.370}$$

Then the Laplacian takes the form

$$\Delta = g^{\mu\nu}(q)\partial_{\mu}\partial_{\nu} - \Gamma_{\mu}{}^{\mu\nu}(q)\partial_{\nu}, \tag{1.371}$$

with $\Gamma_{\mu}{}^{\lambda\nu}$ being defined as the *contraction*

$$\Gamma_{\mu}{}^{\lambda\nu} \equiv g^{\lambda\kappa}\Gamma_{\mu\kappa}{}^{\nu}. \tag{1.372}$$

The reason why (1.368) is called a metric tensor is obvious: An infinitesimal square distance between two points in the original Cartesian coordinates

$$ds^2 \equiv d\mathbf{x}^2 \tag{1.373}$$

becomes in curvilinear coordinates

$$ds^2 = \frac{\partial \mathbf{x}}{\partial q^{\mu}}\frac{\partial \mathbf{x}}{\partial q^{\nu}}dq^{\mu}dq^{\nu} = g_{\mu\nu}(q)dq^{\mu}dq^{\nu}. \tag{1.374}$$

The infinitesimal volume element $d^D x$ is given by

$$d^D x = \sqrt{g}\, d^D q, \tag{1.375}$$

where

$$g(q) \equiv \det\left(g_{\mu\nu}(q)\right) \tag{1.376}$$

is the determinant of the metric tensor. Using this determinant, we form the quantity

$$\Gamma_{\mu} \equiv g^{-1/2}(\partial_{\mu}g^{1/2}) = \frac{1}{2}g^{\lambda\kappa}(\partial_{\mu}g_{\lambda\kappa}) \tag{1.377}$$

and see that it is equal to the once-contracted connection

$$\Gamma_{\mu} = \Gamma_{\mu\lambda}{}^{\lambda}. \tag{1.378}$$

With the inverse metric (1.369) we have furthermore

$$\Gamma_{\mu}{}^{\mu\nu} = -\partial_{\mu}g^{\mu\nu} - \Gamma_{\mu}{}^{\nu\mu}. \tag{1.379}$$

We now take advantage of the fact that the derivatives $\partial_\mu, \partial_\nu$ applied to the coordinate transformation $x^i(q)$ commute causing $\Gamma_{\mu\nu}{}^{\lambda}$ to be symmetric in $\mu\nu$, i.e., $\Gamma_{\mu\nu}{}^{\lambda} = \Gamma_{\nu\mu}{}^{\lambda}$ and hence $\Gamma_\mu{}^{\nu\mu} = \Gamma^\nu$. Together with (1.377) we find the rotation

$$\Gamma_\mu{}^{\mu\nu} = -\frac{1}{\sqrt{g}}(\partial_\mu g^{\mu\nu}\sqrt{g}), \tag{1.380}$$

which allows the Laplace operator Δ to be rewritten in the more compact form

$$\Delta = \frac{1}{\sqrt{g}}\partial_\mu g^{\mu\nu}\sqrt{g}\partial_\nu. \tag{1.381}$$

This expression is called the *Laplace-Beltrami operator*.[18]

Thus we have shown that for a Hamiltonian in a Euclidean space

$$H(\hat{\mathbf{p}}, \mathbf{x}) = \frac{1}{2M}\hat{\mathbf{p}}^2 + V(\mathbf{x}), \tag{1.382}$$

the Schrödinger equation in curvilinear coordinates becomes

$$\hat{H}\psi(q,t) \equiv \left[-\frac{\hbar^2}{2M}\Delta + V(q)\right]\psi(q,t) = i\hbar\partial_t\psi(q,t), \tag{1.383}$$

where $V(q)$ is short for $V(\mathbf{x}(q))$. The scalar product of two wave functions $\int d^Dx\psi_2^*(\mathbf{x},t)\psi_1(\mathbf{x},t)$, which determines the transition amplitudes of the system, transforms into

$$\int d^Dq\sqrt{g}\,\psi_2^*(q,t)\psi_1(q,t). \tag{1.384}$$

It is important to realize that this Schrödinger equation would *not* be obtained by a straightforward application of the canonical formalism to the coordinate-transformed version of the Cartesian Lagrangian

$$L(\mathbf{x}, \dot{\mathbf{x}}) = \frac{M}{2}\dot{\mathbf{x}}^2 - V(\mathbf{x}). \tag{1.385}$$

With the velocities transforming as

$$\dot{x}^i = e^i{}_\mu(q)\dot{q}^\mu, \tag{1.386}$$

the Lagrangian becomes

$$L(q, \dot{q}) = \frac{M}{2}g_{\mu\nu}(q)\dot{q}^\mu\dot{q}^\nu - V(q). \tag{1.387}$$

Up to a factor M, the metric is equal to the Hessian metric of the system, which depends here only on q^μ [recall (1.12)]:

$$H_{\mu\nu}(q) = Mg_{\mu\nu}(q). \tag{1.388}$$

[18]More details will be given later in Eqs. (11.12)–(11.18).

The canonical momenta are

$$p_\mu \equiv \frac{\partial L}{\partial \dot{q}^\mu} = M g_{\mu\nu} \dot{q}^\nu. \tag{1.389}$$

The associated quantum-mechanical momentum operators $\hat{p}_\mu$ have to be Hermitian in the scalar product (1.384) and must satisfy the canonical commutation rules (1.268):

$$\begin{aligned} [\hat{p}_\mu, \hat{q}^\nu] &= -i\hbar \delta_\mu{}^\nu, \\ [\hat{q}^\mu, \hat{q}^\nu] &= 0, \\ [\hat{p}_\mu, \hat{p}_\nu] &= 0. \end{aligned} \tag{1.390}$$

An obvious solution is

$$\hat{p}_\mu = -i\hbar g^{-1/4} \partial_\mu g^{1/4}, \quad \hat{q}^\mu = q^\mu. \tag{1.391}$$

The commutation rules are true for $-i\hbar g^{-z}\partial_\mu g^z$ with any power z, but only $z = 1/4$ produces a Hermitian momentum operator:

$$\begin{aligned} \int d^3q \sqrt{g}\, \Psi_2^*(q,t)[-i\hbar g^{-1/4}\partial_\mu g^{1/4}\Psi_1(q,t)] &= \int d^3q\, g^{1/4}\, \Psi_2^*(q,t)[-i\hbar \partial_\mu g^{1/4}\Psi_1(q,t)] \\ &= \int d^3q \sqrt{g}\, [-i\hbar g^{-1/4}\partial_\mu g^{1/4}\Psi_2(q,t)]^* \Psi_1(q,t), \end{aligned} \tag{1.392}$$

as is easily verified by partial integration.

In terms of the quantity (1.377), this can also be rewritten as

$$\hat{p}_\mu = -i\hbar(\partial_\mu + \tfrac{1}{2}\Gamma_\mu). \tag{1.393}$$

Consider now the classical Hamiltonian associated with the Lagrangian (1.387), which by (1.389) is simply

$$H = p_\mu \dot{q}^\mu - L = \frac{1}{2M} g_{\mu\nu}(q) p^\mu p^\nu + V(q). \tag{1.394}$$

When trying to turn this expression into a Hamiltonian operator, we encounter the operator-ordering problem discussed in connection with Eq. (1.101). The correspondence principle requires replacing the momenta p_μ by the momentum operators $\hat{p}_\mu$, but it does not specify the position of these operators with respect to the coordinates q^μ contained in the inverse metric $g^{\mu\nu}(q)$. An important constraint is provided by the required Hermiticity of the Hamiltonian operator, but this is not sufficient for a unique specification. We may, for instance, define the canonical Hamiltonian operator as

$$\hat{H}_{\rm can} \equiv \frac{1}{2M} \hat{p}^\mu g_{\mu\nu}(q) \hat{p}^\nu + V(q), \tag{1.395}$$

in which the momentum operators have been arranged symmetrically around the inverse metric to achieve Hermiticity. This operator, however, is not equal to the correct Schrödinger operator in (1.383). The kinetic term contains what we may call the *canonical Laplacian*

$$\Delta_{\rm can} = (\partial_\mu + \tfrac{1}{2}\Gamma_\mu)\, g^{\mu\nu}(q)\, (\partial_\nu + \tfrac{1}{2}\Gamma_\nu). \tag{1.396}$$

It differs from the Laplace-Beltrami operator (1.381) in (1.383) by

$$\Delta - \Delta_{\rm can} \;=\; -\tfrac{1}{2}\partial_\mu(g^{\mu\nu}\Gamma_\nu) - \tfrac{1}{4}g^{\mu\nu}\Gamma_\nu\Gamma_\mu. \tag{1.397}$$

The correct Hamiltonian operator could be obtained by suitably distributing pairs of dummy factors of $g^{1/4}$ and $g^{-1/4}$ symmetrically between the canonical operators [5]:

$$\hat{H} = \frac{1}{2M} g^{-1/4}\hat{p}_\mu g^{1/4} g^{\mu\nu}(q) g^{1/4}\hat{p}_\nu g^{-1/4} + V(q). \tag{1.398}$$

This operator has the same classical limit (1.394) as (1.395). Unfortunately, the correspondence principle does not specify how the classical factors have to be ordered before being replaced by operators.

The simplest system exhibiting the breakdown of the canonical quantization rules is a free particle in a plane described by radial coordinates $q^1 = r, q^2 = \varphi$:

$$x^1 = r\,\cos\varphi, \quad x^2 = r\,\sin\varphi. \tag{1.399}$$

Since the infinitesimal square distance is $ds^2 = dr^2 + r^2 d\varphi^2$, the metric reads

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}_{\mu\nu}. \tag{1.400}$$

It has a determinant

$$g = r^2 \tag{1.401}$$

and an inverse

$$g^{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & r^{-2} \end{pmatrix}^{\mu\nu}. \tag{1.402}$$

The Laplace-Beltrami operator becomes

$$\Delta = \frac{1}{r}\partial_r r \partial_r + \frac{1}{r^2}\partial_\varphi{}^2. \tag{1.403}$$

The canonical Laplacian, on the other hand, reads

$$\begin{aligned} \Delta_{\rm can} &= (\partial_r + 1/2r)^2 + \frac{1}{r^2}\partial_\varphi{}^2 \\ &= \partial_r{}^2 + \frac{1}{r}\partial_r - \frac{1}{4r^2} + \frac{1}{r^2}\partial_\varphi{}^2. \end{aligned} \tag{1.404}$$

The discrepancy (1.397) is therefore

$$\Delta_{\rm can} - \Delta = -\frac{1}{4r^2}. \tag{1.405}$$

Note that this discrepancy arises even though there is no apparent ordering problem in the naively quantized canonical expression $\hat{p}^\mu g_{\mu\nu}(q)\,\hat{p}^\nu$ in (1.404). Only the need to introduce dummy $g^{1/4}$- and $g^{-1/4}$-factors creates such problems, and a specification of the order is required to obtain the correct result.

If the Lagrangian coordinates q_i do not merely reparametrize a Euclidean space but specify the points of a general geometry, we cannot proceed as above and derive the Laplace-Beltrami operator by a coordinate transformation of a Cartesian Laplacian. With the canonical quantization rules being unreliable in curvilinear coordinates there are, at first sight, severe difficulties in quantizing such a system. This is why the literature contains many proposals for handling this problem [6]. Fortunately, a large class of non-Cartesian systems allows for a unique quantum-mechanical description on completely different grounds. These systems have the common property that their Hamiltonian can be expressed in terms of the generators of a group of motion in the general coordinate frame. For symmetry reasons, the correspondence principle must then be imposed not on the Poisson brackets of the canonical variables p and q, but on those of the group generators and the coordinates. The brackets containing two group generators specify the structure of the group, those containing a generator and a coordinate specify the defining representation of the group in configuration space. The replacement of these brackets by commutation rules constitutes the proper generalization of the canonical quantization from Cartesian to non-Cartesian coordinates. It is called *group quantization*. The replacement rule will be referred to as the *group correspondence principle*. The canonical commutation rules in Euclidean space may be viewed as a special case of the commutation rules between group generators, i.e., of the *Lie algebra* of the group. In a Cartesian coordinate frame, the group of motion is the Euclidean group containing translations and rotations. The generators of translations and rotations are the momenta and the angular momenta, respectively. According to the group correspondence principle, the Poisson brackets between the generators and the coordinates are to be replaced by commutation rules. Thus, in a Euclidean space, the commutation rules between group generators and coordinates lead to the canonical quantization rules, and this appears to be the deeper reason why the canonical rules are correct. In systems whose energy depends on generators of the group of motion other than those of translations, for instance on the angular momenta, the commutators between the generators have to be used for quantization rather than the canonical commutators between positions and momenta.

The prime examples for such systems are a particle on the surface of a sphere or a spinning top whose quantization will now be discussed.

1.14 Particle on the Surface of a Sphere

For a particle moving on the surface of a sphere of radius r with coordinates

$$x^1 = r\sin\theta\cos\varphi, \quad x^2 = r\sin\theta\sin\varphi, \quad x^3 = r\cos\theta, \qquad (1.406)$$

the Lagrangian reads

$$L = \frac{Mr^2}{2}(\dot{\theta}^2 + \sin^2\theta\,\dot{\varphi}^2). \tag{1.407}$$

The canonical momenta are

$$p_\theta = Mr^2\dot{\theta}, \qquad p_\varphi = Mr^2 \sin^2\theta\,\dot{\varphi}, \tag{1.408}$$

and the classical Hamiltonian is given by

$$H = \frac{1}{2Mr^2}\left(p_\theta^2 + \frac{1}{\sin^2\theta}p_\varphi^2\right). \tag{1.409}$$

According to the canonical quantization rules, the momenta should become operators

$$\hat{p}_\theta = -i\hbar\frac{1}{\sin^{1/2}\theta}\partial_\theta \sin^{1/2}\theta, \qquad \hat{p}_\varphi = -i\hbar\partial_\varphi. \tag{1.410}$$

But as explained in the previous section, these momentum operators are not expected to give the correct Hamiltonian operator when inserted into the Hamiltonian (1.409). Moreover, there exists no proper coordinate transformation from the surface of the sphere to Cartesian coordinates[19] such that a particle on a sphere cannot be treated via the safe Cartesian quantization rules (1.268):

$$\begin{aligned} [\hat{p}_i, \hat{x}^j] &= -i\hbar\delta_i{}^j, \\ [\hat{x}^i, \hat{x}^j] &= 0, \\ [\hat{p}_i, \hat{p}_j] &= 0. \end{aligned} \tag{1.411}$$

The only help comes from the group properties of the motion on the surface of the sphere. The angular momentum

$$\mathbf{L} = \mathbf{x} \times \mathbf{p} \tag{1.412}$$

can be quantized uniquely in Cartesian coordinates and becomes an operator

$$\hat{\mathbf{L}} = \hat{\mathbf{x}} \times \hat{\mathbf{p}} \tag{1.413}$$

whose components satisfy the commutation rules of the Lie algebra of the rotation group

$$[\hat{L}_i, \hat{L}_j] \;=\; i\hbar\hat{L}_k \qquad (i,j,k \text{ cyclic}). \tag{1.414}$$

Note that there is no factor-ordering problem since the $\hat{x}^i$'s and the $\hat{p}_i$'s appear with different indices in each $\hat{L}_k$. An important property of the angular momentum

[19]There exist, however, certain infinitesimal nonholonomic coordinate transformations which are multivalued and can be used to transform infinitesimal distances in a curved space into those in a flat one. They are introduced and applied in Sections 10.2 and Appendix 10A, leading once more to the same quantum mechanics as the one described here.

operator is its homogeneity in $\mathbf{x}$. It has the consequence that when going from Cartesian to spherical coordinates

$$x^1 = r\sin\theta\cos\varphi, \quad x^2 = r\sin\theta\sin\varphi, \quad x^3 = r\cos\theta, \tag{1.415}$$

the radial coordinate cancels making the angular momentum a differential operator involving only the angles θ, φ:

$$\begin{aligned} \hat{L}_1 &= \quad i\hbar\left(\sin\varphi\,\partial_\theta + \cot\theta\cos\varphi\,\partial_\varphi\right), \\ \hat{L}_2 &= -i\hbar\left(\cos\varphi\,\partial_\theta - \cot\theta\sin\varphi\,\partial_\varphi\right), \\ \hat{L}_3 &= -i\hbar\partial_\varphi. \end{aligned} \tag{1.416}$$

There is then a natural way of quantizing the system which makes use of these operators $\hat{L}_i$. We re-express the classical Hamiltonian (1.409) in terms of the classical angular momenta

$$\begin{aligned} L_1 &= Mr^2\left(-\sin\varphi\,\dot{\theta} - \sin\theta\cos\theta\cos\varphi\,\dot{\varphi}\right), \\ L_2 &= Mr^2\left(\cos\varphi\,\dot{\theta} - \sin\theta\cos\theta\sin\varphi\,\dot{\varphi}\right), \\ L_3 &= Mr^2\sin^2\theta\,\dot{\varphi} \end{aligned} \tag{1.417}$$

as

$$H = \frac{1}{2Mr^2}\mathbf{L}^2, \tag{1.418}$$

and replace the angular momenta by the operators (1.416). The result is the Hamiltonian operator:

$$\hat{H} = \frac{1}{2Mr^2}\hat{\mathbf{L}}^2 = -\frac{\hbar^2}{2Mr^2}\left[\frac{1}{\sin\theta}\partial_\theta\left(\sin\theta\,\partial_\theta\right) + \frac{1}{\sin^2\theta}\partial_\varphi^2\right]. \tag{1.419}$$

The eigenfunctions diagonalizing the rotation-invariant operator $\hat{\mathbf{L}}^2$ are well known. They can be chosen to diagonalize simultaneously one component of $\hat{L}_i$, for instance the third one, $\hat{L}_3$, in which case they are equal to the spherical harmonics

$$Y_{lm}(\theta,\varphi) = (-1)^m\left[\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}\right]^{1/2} P_l^m(\cos\theta)e^{im\varphi}, \tag{1.420}$$

with $P_l^m(z)$ being the associated Legendre polynomials

$$P_l^m(z) = \frac{1}{2^l l!}(1-z^2)^{m/2}\frac{d^{l+m}}{dx^{l+m}}(z^2-1)^l. \tag{1.421}$$

The spherical harmonics are orthonormal with respect to the rotation-invariant scalar product

$$\int_0^\pi d\theta\sin\theta\int_0^{2\pi} d\varphi\; Y_{lm}^*(\theta,\varphi)Y_{l'm'}(\theta,\varphi) = \delta_{ll'}\delta_{mm'}. \tag{1.422}$$

Two important lessons can be learned from this group quantization. First, the correct Hamiltonian operator (1.419) does not agree with the canonically quantized one which would be obtained by inserting Eqs. (1.410) into (1.409). The correct result would, however, arise by distributing dummy factors

$$g^{-1/4} = r^{-1} \sin^{-1/2}\theta, \quad g^{1/4} = r \sin^{1/2}\theta \tag{1.423}$$

between the canonical momentum operators as observed earlier in Eq. (1.398). Second, just as in the case of polar coordinates, the correct Hamiltonian operator is equal to

$$\hat{H} = -\frac{\hbar^2}{2M}\Delta, \tag{1.424}$$

where Δ is the Laplace-Beltrami operator associated with the metric

$$g_{\mu\nu} = r^2 \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix}, \tag{1.425}$$

i.e.,

$$\Delta = \frac{1}{r^2}\left[\frac{1}{\sin\theta}\,\partial_\theta\left(\sin\theta\partial_\theta\right) + \frac{1}{\sin^2\theta}\partial_\varphi^2\right]. \tag{1.426}$$

1.15 Spinning Top

For a spinning top, the optimal starting point is again not the classical Lagrangian but the Hamiltonian expressed in terms of the classical angular momenta. In the symmetric case in which two moments of inertia coincide, it is written as

$$H = \frac{1}{2I_\xi}({L_\xi}^2 + {L_\eta}^2) + \frac{1}{2I_\zeta}{L_\zeta}^2, \tag{1.427}$$

where L_ξ, L_η, L_ζ are the components of the orbital angular momentum in the directions of the principal body axes with $I_\xi, I_\eta \equiv I_\xi, I_\zeta$ being the corresponding moments of inertia. The classical angular momentum of an aggregate of mass points is given by

$$\mathbf{L} = \sum_\nu \mathbf{x}_\nu \times \mathbf{p}_\nu, \tag{1.428}$$

where the sum over ν runs over all mass points. The angular momentum possesses a unique operator

$$\hat{\mathbf{L}} = \sum_\nu \hat{\mathbf{x}}_\nu \times \hat{\mathbf{p}}_\nu, \tag{1.429}$$

with the commutation rules (1.414) between the components $\hat{L}_i$. Since rotations do not change the distances between the mass points, they commute with the constraints of the rigid body. If the center of mass of the rigid body is placed at the origin, the only dynamical degrees of freedom are the orientations in space. They can uniquely be specified by the rotation matrix which brings the body from some standard orientation to the actual one. We may choose the standard orientation to have the principal body axes aligned with the x, y, z-directions, respectively. An arbitrary orientation is obtained by applying all finite rotations to each point of the body. They are specified by the 3×3 orthonormal matrices R_{ij}. The space of these matrices has three degrees of freedom. It can be decomposed, omitting the matrix indices as

$$R(\alpha, \beta, \gamma) = R_3(\alpha) R_2(\beta) R_3(\gamma), \tag{1.430}$$

where $R_3(\alpha)$, $R_3(\gamma)$ are rotations around the z-axis by angles α, γ, respectively, and $R_2(\beta)$ is a rotation around the y-axis by β. These rotation matrices can be expressed as exponentials

$$R_i(\delta) \equiv e^{-i\delta L_i/\hbar}, \tag{1.431}$$

where δ is the rotation angle and L_i are the 3×3 matrix generators of the rotations with the elements

$$(L_i)_{jk} = -i\hbar\epsilon_{ijk}. \tag{1.432}$$

It is easy to check that these generators satisfy the commutation rules (1.414) of angular momentum operators. The angles α, β, γ are referred to as *Euler angles*.

The 3×3 rotation matrices make it possible to express the infinitesimal rotations around the three coordinate axes as differential operators of the three Euler angles. Let $\psi(R)$ be the wave function of the spinning top describing the probability amplitude of the different orientations which arise from a standard orientation by the rotation matrix $R = R(\alpha, \beta, \gamma)$. Under a further rotation by $R(\alpha', \beta', \gamma')$, the wave function goes over into $\psi'(R) = \psi(R^{-1}(\alpha', \beta', \gamma')R)$. The transformation may be described by a unitary differential operator

$$\hat{U}(\alpha', \beta', \gamma') \equiv e^{-i\alpha'\hat{L}_3} e^{-i\beta'\hat{L}_2} e^{-i\gamma'\hat{L}_3}, \tag{1.433}$$

where $\hat{L}_i$ is the representation of the generators in terms of differential operators. To calculate these we note that the 3×3 -matrix $R^{-1}(\alpha, \beta, \gamma)$ has the following derivatives

$$\begin{aligned} -i\hbar\partial_\alpha R^{-1} &= R^{-1} L_3, \\ -i\hbar\partial_\beta R^{-1} &= R^{-1}(\cos\alpha\, L_2 - \sin\alpha\, L_1), \\ -i\hbar\partial_\gamma R^{-1} &= R^{-1}\left[\cos\beta\, L_3 + \sin\beta(\cos\alpha\, L_1 + \sin\alpha\, L_2)\right]. \end{aligned} \tag{1.434}$$

The first relation is trivial, the second follows from the rotation of the generator

$$e^{-i\alpha L_3/\hbar} L_2 e^{i\alpha L_3/\hbar} = \cos\alpha\, L_2 - \sin\alpha\, L_1, \tag{1.435}$$

which is a consequence of Lie's expansion formula (1.297) together with the commutation rules (1.432) of the 3×3 matrices L_i. The third requires, in addition, the rotation

$$e^{-i\beta L_2/\hbar} L_3 e^{i\beta L_2/\hbar} = \cos\beta L_3 + \sin\beta L_1. \tag{1.436}$$

Inverting the relations (1.434), we find the differential operators generating the rotations [7]:

$$\begin{aligned}\hat{L}_1 &= i\hbar\left(\cos\alpha\cot\beta\ \partial_\alpha + \sin\alpha\ \partial_\beta - \frac{\cos\alpha}{\sin\beta}\partial_\gamma\right),\\ \hat{L}_2 &= i\hbar\left(\sin\alpha\cot\beta\ \partial_\alpha - \cos\alpha\ \partial_\beta - \frac{\sin\alpha}{\sin\beta}\partial_\gamma\right),\\ \hat{L}_3 &= -i\hbar\partial_\alpha.\end{aligned} \tag{1.437}$$

After exponentiating these differential operators we derive

$$\begin{aligned}\hat{U}(\alpha',\beta',\gamma')R^{-1}\hat{U}^{-1}(\alpha',\beta',\gamma')(\alpha,\beta,\gamma) &= R^{-1}(\alpha,\beta,\gamma)R(\alpha',\beta',\gamma'),\\ \hat{U}(\alpha',\beta',\gamma')R(\alpha,\beta,\gamma)\hat{U}^{-1}(\alpha',\beta',\gamma') &= R^{-1}(\alpha',\beta',\gamma')R(\alpha,\beta,\gamma),\end{aligned} \tag{1.438}$$

so that $\hat{U}(\alpha',\beta',\gamma')\psi(R) = \psi'(R)$, as desired.

In the Hamiltonian (1.427), we need the components of $\hat{\mathbf{L}}$ along the body axes. They are obtained by rotating the 3×3 matrices L_i by $R(\alpha,\beta,\gamma)$ into

$$\begin{aligned}L_\xi &= RL_1R^{-1} = \cos\gamma\cos\beta(\cos\alpha\, L_1 + \sin\alpha\, L_2)\\ &\quad + \sin\gamma(\cos\alpha\, L_2 - \sin\alpha\, L_1) - \cos\gamma\sin\beta\, L_3,\\ L_\eta &= RL_2R^{-1} = -\sin\gamma\cos\beta(\cos\alpha\, L_1 + \sin\alpha\, L_2)\\ &\quad + \cos\gamma(\cos\alpha\, L_2 - \sin\alpha\, L_1) + \sin\gamma\sin\beta\, L_3,\\ L_\zeta &= RL_3R^{-1} = \cos\beta\, L_3 + \sin\beta(\cos\alpha\, L_1 + \sin\alpha\, L_2),\end{aligned} \tag{1.439}$$

and replacing $L_i \to \hat{L}_i$ in the final expressions. Inserting (1.437), we find the operators

$$\begin{aligned}\hat{L}_\xi &= i\hbar\left(-\cos\gamma\cot\beta\ \partial_\gamma - \sin\gamma\ \partial_\beta + \frac{\cos\gamma}{\sin\beta}\partial_\alpha\right),\\ \hat{L}_\eta &= i\hbar\left(\sin\gamma\cot\beta\ \partial_\gamma - \cos\gamma\ \partial_\beta - \frac{\sin\gamma}{\sin\beta}\partial_\alpha\right\},\\ \hat{L}_\zeta &= -i\hbar\partial_\gamma.\end{aligned} \tag{1.440}$$

Note that these commutation rules have an opposite sign with respect to those in Eqs. (1.414) of the operators $\hat{L}_i$:[20]

$$[\hat{L}_\xi, \hat{L}_\eta] = -i\hbar\hat{L}_\zeta, \quad \xi,\eta,\zeta = \text{cyclic}. \tag{1.441}$$

The sign is most simply understood by writing

$$\hat{L}_\xi = a_\xi^i\hat{L}_i, \quad \hat{L}_\eta = a_\eta^i\hat{L}_i, \quad \hat{L}_\zeta = a_\zeta^i\hat{L}_i, \tag{1.442}$$

[20]When applied to functions not depending on α, then, after replacing $\beta\to\theta$ and $\gamma\to\varphi$, the operators agree with those in (1.416), up to the sign of $\hat{L}_1$.

where a^i_ξ, a^i_η, a^i_ζ, are the components of the body axes. Under rotations these behave like $[\hat{L}_i, a^j_\xi] = i\hbar\epsilon_{ijk}a^k_\xi$, i.e., they are vector operators. It is easy to check that this property produces the sign reversal in (1.441) with respect to (1.414).

The correspondence principle is now applied to the Hamiltonian in Eq. (1.427) by placing operator hats on the L_a's. The energy spectrum and the wave functions can then be obtained by using only the group commutators between $\hat{L}_\xi, \hat{L}_\eta, \hat{L}_\zeta$. The spectrum is

$$E_{L\Lambda} = \hbar^2 \left[\frac{1}{2I_\xi} L(L+1) + \left(\frac{1}{2I_\zeta} - \frac{1}{2I_\xi} \right) \Lambda^2 \right], \tag{1.443}$$

where $L(L+1)$ with $L = 0, 1, 2, \dots$ are the eigenvalues of $\hat{\mathbf{L}}^2$, and $\Lambda = -L, \dots, L$ are the eigenvalues of $\hat{L}_\zeta$. The wave functions are the representation functions of the rotation group. If the Euler angles α, β, γ are used to specify the orientation of the body axes, the wave functions are

$$\psi_{L\Lambda m}(\alpha, \beta, \gamma) = D^L_{m\Lambda}(-\alpha, -\beta, -\gamma). \tag{1.444}$$

Here m' are the eigenvalues of $\hat{L}_3$, the magnetic quantum numbers, and $D^L_{m\Lambda}(\alpha, \beta, \gamma)$ are the representation matrices of angular momentum L. In accordance with (1.433), one may decompose

$$D^L_{mm'}(\alpha, \beta, \gamma) = e^{-i(m\alpha + m'\gamma)} d^L_{mm'}(\beta), \tag{1.445}$$

with the matrices

$$\begin{aligned} d^L_{mm'}(\beta) = & \left[\frac{(L+m')!(L-m')!}{(L+m)!(L-m)!} \right]^{1/2} \\ \times & \left(\cos\frac{\beta}{2} \right)^{m+m'} \left(-\sin\frac{\beta}{2} \right)^{m-m'} P^{(m'-m, m'+m)}_{L-m'}(\cos\beta). \end{aligned} \tag{1.446}$$

For $j = 1/2$, these form the spinor representation of the rotations around the y-axis

$$d^{1/2}_{m'm}(\beta) = \begin{pmatrix} \cos\beta/2 & -\sin\beta/2 \\ \sin\beta/2 & \cos\beta/2 \end{pmatrix}. \tag{1.447}$$

The indices have the order $+1/2, -1/2$. The full spinor representation function $D^{1/2}(\alpha, \beta, \gamma)$ in (1.445) is most easily obtained by inserting into the general expression (1.433) the representation matrices of spin 1/2 for the generators $\hat{L}_i$ with the commutation rules (1.414), the famous *Pauli spin matrices*:

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.448}$$

Thus we can write

$$D^{1/2}(\alpha, \beta, \gamma) = e^{-i\alpha\sigma_3/2} e^{-i\beta\sigma_2/2} e^{-i\gamma\sigma_3/2}. \tag{1.449}$$

The first and the third factor yield the pure phase factors in (1.445). The function $d^{1/2}_{m'm}(\beta)$ is obtained by a simple power series expansion of $e^{-i\beta\sigma^2/2}$, using the fact that $(\sigma^2)^{2n}=1$ and $(\sigma^2)^{2n+1}=\sigma^2$:

$$e^{-i\beta\sigma^2/2} = \cos\beta/2\, - i\sin\beta/2\,\sigma^2, \tag{1.450}$$

which is equal to (1.447).

For $j=1$, the representation functions (1.446) form the vector representation

$$d^1_{m'm}(\beta) = \begin{pmatrix} \frac{1}{2}(1+\cos\beta) & -\frac{1}{\sqrt{2}}\sin\beta & \frac{1}{2}(1-\cos\beta) \\ \frac{1}{\sqrt{2}}\sin\beta & \cos\beta & -\frac{1}{\sqrt{2}}\sin\beta \\ \frac{1}{2}(1-\cos\beta) & \frac{1}{\sqrt{2}}\sin\beta & \frac{1}{2}(1+\cos\beta) \end{pmatrix}. \tag{1.451}$$

where the indices have the order $+1/2, -1/2$. The vector representation goes over into the ordinary rotation matrices $R_{ij}(\beta)$ by mapping the states $|1m\rangle$ onto the spherical unit vectors $\boldsymbol{\epsilon}(0)=\hat{\mathbf{z}}$, $\boldsymbol{\epsilon}(\pm1)=\mp(\hat{\mathbf{x}}\pm i\hat{\mathbf{y}})/2$ using the matrix elements $\langle i|1m\rangle = \epsilon^i(m)$. Hence $R(\beta)\boldsymbol{\epsilon}(m)=\sum_{m'=-1}^{1}\boldsymbol{\epsilon}(m')d^1_{m'm}(\beta)$.

The representation functions $D^1(\alpha,\beta,\gamma)$ can also be obtained by inserting into the general exponential (1.433) the representation matrices of spin 1 for the generators $\hat{L}_i$ with the commutation rules (1.414). In Cartesian coordinates, these are simply $(\hat{L}_i)_{jk} = -i\epsilon_{ijk}$, where ϵ_{ijk} is the completely antisymmetric tensor with $\epsilon_{123}=1$. In the spherical basis, these become $(\hat{L}_i)_{mm'} = \langle m|i\rangle(\hat{L}_i)_{ij}\langle j|m'\rangle = \epsilon_i^*(m)(\hat{L}_i)_{ij}\boldsymbol{\epsilon}_j(m')$. The exponential $(e^{-i\beta\hat{L}_2})_{mm'}$ is equal to (1.451).

The functions $P_l^{(\alpha,\beta)}(z)$ are the Jacobi polynomials [8], which can be expressed in terms of hypergeometric functions as

$$P_l^{(\alpha,\beta)} \equiv \frac{(-1)^l}{l!}\,\frac{\Gamma(l+\beta+1)}{\Gamma(\beta+1)}F(-l,l+1+\alpha+\beta;1+\beta;(1+z)/2), \tag{1.452}$$

where

$$F(a,b;c;z) \equiv 1+\frac{ab}{c}z+\frac{a(a+1)\,b(b+1)}{c(c+1)}\frac{z^2}{2!}+\dots\ . \tag{1.453}$$

The rotation functions $d^L_{mm'}(\beta)$ satisfy the differential equation

$$\left(-\frac{d^2}{d\beta^2}-\cot\beta\frac{d}{d\beta}+\frac{m^2+m'^2-2mm'\cos\beta}{\sin^2\beta}\right)d^L_{mm'}(\beta) = L(L+1)d^L_{mm'}(\beta). \tag{1.454}$$

The scalar products of two wave functions have to be calculated with a measure of integration that is invariant under rotations:

$$\langle\psi_2|\psi_1\rangle \equiv \int_0^{2\pi}\int_0^{\pi}\int_0^{2\pi} d\alpha d\beta\sin\beta d\gamma\ \psi_2^*(\alpha,\beta,\gamma)\psi_1(\alpha,\beta,\gamma). \tag{1.455}$$

The above eigenstates (1.445) satisfy the orthogonality relation

$$\int_0^{2\pi}\int_0^{\pi}\int_0^{2\pi} d\alpha d\beta\sin\beta d\gamma\ D^{L_1\,*}_{m_1'm_1}(\alpha,\beta,\gamma)D^{L_2}_{m_2'm_2}(\alpha,\beta,\gamma)$$
$$= \delta_{m_1'm_2'}\,\delta_{m_1m_2}\delta_{L_1L_2}\,\frac{8\pi^2}{2L_1+1}. \tag{1.456}$$

Let us also contrast in this example the correct quantization via the commutation rules between group generators with the canonical approach which would start out with the classical Lagrangian. In terms of Euler angles, the Lagrangian reads

$$L = \frac{1}{2}[I_\xi({\omega_\xi}^2 + {\omega_\eta}^2) + I_\zeta {\omega_\zeta}^2], \tag{1.457}$$

where $\omega_\xi, \omega_\eta, \omega_\zeta$ are the angular velocities measured along the principal axes of the top. To find these we note that the components in the rest system $\omega_1, \omega_2, \omega_3$ are obtained from the relation

$$\omega_k L_k = i\dot{R}R^{-1} \tag{1.458}$$

as

$$\begin{aligned} \omega_1 &= -\dot{\beta}\sin\alpha + \dot{\gamma}\sin\beta\cos\alpha, \\ \omega_2 &= \dot{\beta}\cos\alpha + \dot{\gamma}\sin\beta\sin\alpha, \\ \omega_3 &= \dot{\gamma}\cos\beta + \dot{\alpha}. \end{aligned} \tag{1.459}$$

After the rotation (1.439) into the body-fixed system, these become

$$\begin{aligned} \omega_\xi &= \dot{\beta}\sin\gamma - \dot{\alpha}\sin\beta\cos\gamma, \\ \omega_\eta &= \dot{\beta}\cos\gamma + \dot{\alpha}\sin\beta\sin\gamma, \\ \omega_\zeta &= \dot{\alpha}\cos\beta + \dot{\gamma}. \end{aligned} \tag{1.460}$$

Explicitly, the Lagrangian is

$$L = \frac{1}{2}[I_\xi(\dot{\beta}^2 + \dot{\alpha}^2\sin^2\beta) + I_\zeta(\dot{\alpha}\cos\beta + \dot{\gamma})^2]. \tag{1.461}$$

Considering α, β, γ as Lagrange coordinates q^μ with $\mu = 1, 2, 3$, this can be written in the form (1.387) with the Hessian metric [recall (1.12) and (1.388)]:

$$g_{\mu\nu} = \begin{pmatrix} I_\xi \sin^2\beta + I_\zeta\cos^2\beta & 0 & I_\zeta\cos\beta \\ 0 & I_\xi & 0 \\ I_\zeta\cos\beta & 0 & I_\zeta \end{pmatrix}, \tag{1.462}$$

whose determinant is

$$g = I_\xi^2 I_\zeta \sin^2\beta. \tag{1.463}$$

Hence the measure $\int d^3q\sqrt{g}$ in the scalar product (1.384) agrees with the rotation-invariant measure (1.455) up to a trivial constant factor. Incidentally, this is also true for the asymmetric top with $I_\xi \neq I_\eta \neq I_\zeta$, where $g = I_\xi^2 I_\zeta \sin^2\beta$, although the metric $g_{\mu\nu}$ is then much more complicated (see Appendix 1C).

The canonical momenta associated with the Lagrangian (1.457) are, according to (1.387),

$$\begin{aligned} p_\alpha &= \partial L/\partial\dot{\alpha} = I_\xi\,\dot{\alpha}\sin^2\beta + I_\zeta\cos\beta(\dot{\alpha}\cos\beta + \dot{\gamma}), \\ p_\beta &= \partial L/\partial\dot{\beta} = I_\xi\,\dot{\beta}, \\ p_\gamma &= \partial L/\partial\dot{\gamma} = I_\zeta\,(\dot{\alpha}\cos\beta + \dot{\gamma}). \end{aligned} \tag{1.464}$$

After inverting the metric to

$$g^{\mu\nu} = \frac{1}{I_\xi \sin^2\beta} \begin{pmatrix} 1 & 0 & -\cos\beta \\ 0 & \sin^2\beta & 0 \\ -\cos\beta & 0 & \cos^2\beta + I_\xi \sin^2\beta / I_\zeta \end{pmatrix}^{\mu\nu}, \tag{1.465}$$

we find the classical Hamiltonian

$$H = \frac{1}{2}\left[\frac{1}{I_\xi}{p_\beta}^2 + \left(\frac{\cos^2\beta}{I_\xi \sin^2\beta} + \frac{1}{I_\zeta}\right){p_\gamma}^2 + \frac{1}{I_\xi \sin^2\beta}{p_\alpha}^2 - \frac{2\cos\beta}{I_\xi \sin^2\beta} p_\alpha p_\gamma\right]. \tag{1.466}$$

This Hamiltonian has no apparent ordering problem. One is therefore tempted to replace the momenta simply by the corresponding Hermitian operators which are, according to (1.391),

$$\begin{aligned} \hat p_\alpha &= -i\hbar\partial_\alpha, \\ \hat p_\beta &= -i\hbar(\sin\beta)^{-1/2}\partial_\beta(\sin\beta)^{1/2} = -i\hbar(\partial_\beta + \frac{1}{2}\cot\beta), \\ \hat p_\gamma &= -i\hbar\partial_\gamma. \end{aligned} \tag{1.467}$$

Inserting these into (1.466) gives the canonical Hamiltonian operator

$$\hat H_{\rm can} = \hat H + \hat H_{\rm discr}, \tag{1.468}$$

with

$$\begin{aligned} \hat H \equiv -\frac{\hbar^2}{2I_\xi}\Bigg[{\partial_\beta}^2 + \cot\beta\partial_\beta + \left(\frac{I_\xi}{I_\zeta} + \cot^2\beta\right){\partial_\gamma}^2 \\ + \frac{1}{\sin^2\beta}{\partial_\alpha}^2 - \frac{2\cos\beta}{\sin^2\beta}\partial_\alpha\partial_\gamma\Bigg] \end{aligned} \tag{1.469}$$

and

$$\hat H_{\rm discr} \equiv \frac{1}{2}(\partial_\beta \cot\beta) + \frac{1}{4}\cot^2\beta = \frac{1}{4\sin^2\beta} - \frac{3}{4}. \tag{1.470}$$

The first term $\hat H$ agrees with the correct quantum-mechanical operator derived above. Indeed, inserting the differential operators for the body-fixed angular momenta (1.440) into the Hamiltonian (1.427), we find $\hat H$. The term $\hat H_{\rm discr}$ is the discrepancy between the canonical and the correct Hamiltonian operator. It exists even though there is no apparent ordering problem, just as in the radial coordinate expression (1.404). The correct Hamiltonian could be obtained by replacing the classical ${p_\beta}^2$ term in H by the operator $g^{-1/4}\hat p_\beta g^{1/2}\hat p_\beta g^{-1/4}$, by analogy with the treatment of the radial coordinates in $\hat H$ of Eq. (1.398).

As another similarity with the two-dimensional system in radial coordinates and the particle on the surface of the sphere, we observe that while the canonical quantization fails, the Hamiltonian operator of the symmetric spinning top is correctly given by the Laplace-Beltrami operator (1.381) after inserting the metric (1.462) and the inverse (1.465). It is straightforward although tedious to verify that this is also true for the completely asymmetric top [which has quite a complicated metric given in Appendix 1C, see Eqs. (1C.2), and (1C.4)]. This is an important nontrivial result, since for a spinning top, the Lagrangian cannot be obtained by reparametrizing a particle in a Euclidean space with curvilinear coordinates. The result suggests that a replacement

$$g_{\mu\nu}(q)p^\mu p^\nu \to -\hbar^2\Delta \tag{1.471}$$

produces the correct Hamiltonian operator in any non-Euclidean space.[21]

What is the characteristic non-Euclidean property of the α, β, γ space? As we shall see in detail in Chapter 10, the relevant quantity is the curvature scalar R. The exact definition will be found in Eq. (10.42). For the asymmetric spinning top we find (see Appendix 1C)

$$R = \frac{(I_\xi + I_\eta + I_\zeta)^2 - 2(I_\xi^2 + I_\eta^2 + I_\zeta^2)}{2I_\xi I_\eta I_\zeta}. \tag{1.472}$$

Thus, just like a particle on the surface of a sphere, the spinning top corresponds to a particle moving in a space with constant curvature. In this space, the correct correspondence principle can also be deduced from symmetry arguments. The geometry is most easily understood by observing that the α, β, γ space may be considered as the surface of a sphere in four dimensions, as we shall see in more detail in Chapter 8.

An important non-Euclidean space of physical interest is encountered in the context of general relativity. Originally, gravitating matter was assumed to move in a spacetime with an arbitrary local curvature. In newer developments of the theory one also allows for the presence of a nonvanishing torsion. In such a general situation, where the group quantization rule is inapplicable, the correspondence principle has always been a matter of controversy [see the references after (1.405)] to be resolved in this text. In Chapters 10 and 8 we shall present a new quantum equivalence principle which is based on an application of simple geometrical principles to path integrals and which will specify a natural and unique passage from classical to quantum mechanics in any coordinate frame.[22] The configuration space may carry curvature and a certain class of torsions (gradient torsion). Several arguments suggest that our principle is correct. For the above systems with a Hamiltonian which can be expressed entirely in terms of generators of a group of motion in the underlying space, the new quantum equivalence principle will give the same results as the group quantization rule.

---

[21] If the space has curvature and no torsion, this is the correct answer. If torsion is present, the correct answer will be given in Chapters 10 and 8.

[22] H. Kleinert, Mod. Phys. Lett. A *4*, 2329 (1989) (`http://www.physik.fu-berlin.de/~kleinert/199`); Phys. Lett. B *236*, 315 (1990) (*ibid.*`http/202`).

1.16 Scattering

Most observations of quantum phenomena are obtained from scattering processes of fundamental particles.

1.16.1 Scattering Matrix

Consider a particle impinging with a momentum $\mathbf{p}_a$ and energy $E = E_a = \mathbf{p}_a^2/2M$ upon a nonzero potential concentrated around the origin. After a long time, it will be found far from the potential with some momentum $\mathbf{p}_b$. The energy will be unchanged: $E = E_b = \mathbf{p}_b^2/2M$. The probability amplitude for such a process is given by the time evolution amplitude in the momentum representation

$$(\mathbf{p}_b t_b|\mathbf{p}_a t_a) \equiv \langle \mathbf{p}_b|e^{-i\hat{H}(t_b-t_a)/\hbar}|\mathbf{p}_a\rangle, \tag{1.473}$$

where the limit $t_b \to \infty$ and $t_a \to -\infty$ has to be taken. Long before and after the collision, this amplitude oscillates with a frequency $\omega = E/\hbar$ characteristic for free particles of energy E. In order to have a time-independent limit, we remove these oscillations, from (1.473), and define the scattering matrix (S-matrix) by the limit

$$\langle \mathbf{p}_b|\hat{S}|\mathbf{p}_a\rangle \;\; \equiv \lim_{t_b-t_a\to\infty} e^{i(E_bt_b-E_at_a)/\hbar}\langle \mathbf{p}_b|e^{-i\hat{H}(t_b-t_a)/\hbar}|\mathbf{p}_a\rangle. \tag{1.474}$$

Most of the impinging particles will not scatter at all, so that this amplitude must contain a leading term, which is separated as follows:

$$\langle \mathbf{p}_b|\hat{S}|\mathbf{p}_a\rangle = \langle \mathbf{p}_b|\mathbf{p}_a\rangle + \langle \mathbf{p}_b|\hat{S}|\mathbf{p}_a\rangle', \tag{1.475}$$

where

$$\langle \mathbf{p}_b|\mathbf{p}_a\rangle = \langle \mathbf{p}_b|e^{-i\hat{H}(t_b-t_a)/\hbar}|\mathbf{p}_a\rangle = (2\pi\hbar)^3\delta^{(3)}(\mathbf{p}_b-\mathbf{p}_a) \tag{1.476}$$

shows the normalization of the states [recall (1.186)]. This leading term is commonly subtracted from (1.474) to find the true scattering amplitude. Moreover, since potential scattering conserves energy, the remaining amplitude contains a δ-function ensuring energy conservation, and it is useful to divide this out, defining the so-called T-matrix by the decomposition

$$\langle \mathbf{p}_b|\hat{S}|\mathbf{p}_a\rangle \equiv (2\pi\hbar)^3\delta^{(3)}(\mathbf{p}_a-\mathbf{p}_a) - 2\pi\hbar i\delta(E_b-E_a)\langle \mathbf{p}_b|\hat{T}|\mathbf{p}_a\rangle. \tag{1.477}$$

From the definition (1.474) and the hermiticity of $\hat{H}$ it follows that the scattering matrix is a unitary matrix. This expresses the physical fact that the total probability of an incident particle to re-emerge at some time is unity (in quantum field theory the situation is more complicated due to emission and absorption processes).

In the basis states $|\mathbf{p}^{\mathbf{m}}\rangle$ introduced in Eq. (1.180) which satisfy the completeness relation (1.182) and are normalized to unity in a finite volume V, the unitarity is expressed as

$$\sum_{\mathbf{m}'}\langle \mathbf{p}^{\mathbf{m}}|\hat{S}^\dagger|\mathbf{p}^{\mathbf{m}'}\rangle\langle \mathbf{p}^{\mathbf{m}'}|\hat{S}|\mathbf{p}^{\mathbf{m}''}\rangle = \sum_{\mathbf{m}'}\langle \mathbf{p}^{\mathbf{m}}|\hat{S}|\mathbf{p}^{\mathbf{m}'}\rangle\langle \mathbf{p}^{\mathbf{m}'}|\hat{S}^\dagger|\mathbf{p}^{\mathbf{m}''}\rangle = 1. \tag{1.478}$$

Remembering the relation (1.185) between the discrete states $|\mathbf{p}^{\mathbf{m}}\rangle$ and their continuous limits $|\mathbf{p}\rangle$, we see that

$$\langle \mathbf{p}_b{}^{\mathbf{m}'}|\hat{S}|\mathbf{p}_a{}^{\mathbf{m}}\rangle \approx \frac{1}{L^3}\langle \mathbf{p}_b|\hat{S}|\mathbf{p}_a\rangle, \tag{1.479}$$

where L^3 is the spatial volume, and $\mathbf{p}_b^{\mathbf{m}}$ and $\mathbf{p}_a^{\mathbf{m}}$ are the discrete momenta closest to $\mathbf{p}_b$ and $\mathbf{p}_a$. In the continuous basis $|\mathbf{p}\rangle$, the unitarity relation reads

$$\int \frac{d^3p}{(2\pi\hbar)^3}\langle \mathbf{p}_b|\hat{S}^\dagger|\mathbf{p}\rangle\langle \mathbf{p}|\hat{S}|\mathbf{p}_a\rangle = \int \frac{d^3p}{(2\pi\hbar)^3}\langle \mathbf{p}_b|\hat{S}|\mathbf{p}\rangle\langle \mathbf{p}|\hat{S}^\dagger|\mathbf{p}_a\rangle = 1. \tag{1.480}$$

1.16.2 Cross Section

The absolute square of $\langle \mathbf{p}_b|\hat{S}|\mathbf{p}_a\rangle$ gives the probability $P_{\mathbf{p}_b\leftarrow\mathbf{p}_a}$ for the scattering from the initial momentum state $\mathbf{p}_a$ to the final momentum state $\mathbf{p}_b$. Omitting the unscattered particles, we have

$$P_{\mathbf{p}_b\leftarrow\mathbf{p}_a} = \frac{1}{L^6}\, 2\pi\hbar\delta(0)\, 2\pi\hbar\delta(E_b - E_a)|\langle \mathbf{p}_b|\hat{T}|\mathbf{p}_a\rangle|^2. \tag{1.481}$$

The factor $\delta(0)$ at zero energy is made finite by imagining the scattering process to take place with an incident time-independent plane wave over a finite total time T. Then $2\pi\hbar\delta(0) = \int dt\, e^{iEt/\hbar}|_{E=0} = T$, and the probability is proportional to the time T:

$$P_{\mathbf{p}_b\leftarrow\mathbf{p}_a} = \frac{1}{L^6}\, T\, 2\pi\hbar\delta(E_b - E_a)|\langle \mathbf{p}_b|\hat{T}|\mathbf{p}_a\rangle|^2. \tag{1.482}$$

By summing this over all discrete final momenta, or equivalently, by integrating this over the phase space of the final momenta [recall (1.184)], we find the total probability per unit time for the scattering to take place

$$\frac{dP}{dt} = \frac{1}{L^6}\int \frac{d^3p_b L^3}{(2\pi\hbar)^3}\, 2\pi\hbar\delta(E_b - E_a)|\langle \mathbf{p}_b|\hat{T}|\mathbf{p}_a\rangle|^2. \tag{1.483}$$

The momentum integral can be split into an integral over the final energy and the final solid angle. For non-relativistic particles, this goes as follows

$$\int \frac{d^3p_b}{(2\pi\hbar)^3} \quad = \quad \frac{1}{(2\pi\hbar)^3}\frac{M}{(2\pi\hbar)^3}\int d\Omega \int_0^\infty dE_b\, p_b, \tag{1.484}$$

where $d\Omega = \mathrm{d}\phi_b \mathrm{d}\cos\theta_b$ is the element of solid angle into which the particle is scattered. The energy integral removes the δ-function in (1.483), and makes p_b equal to p_a.

The differential scattering cross section $d\sigma/d\Omega$ is defined as the probability that a single impinging particle ends up in a solid angle $d\Omega$ per unit time and unit current density. From (1.483) we identify

$$\frac{d\sigma}{d\Omega} = \frac{d\dot{P}}{d\Omega}\frac{1}{j} = \frac{1}{L^3}\frac{Mp}{(2\pi\hbar)^3}2\pi\hbar|T_{\mathbf{p}_b\mathbf{p}_a}|^2\frac{1}{j}, \tag{1.485}$$

where we have set

$$\langle \mathbf{p}_b|\hat{T}|\mathbf{p}_a\rangle \equiv T_{\mathbf{p}_b\mathbf{p}_a}, \tag{1.486}$$

for brevity. In a volume L^3, the current density of a single impinging particle is given by the velocity $v = p/M$ as

$$j = \frac{1}{L^3}\frac{p}{M}, \tag{1.487}$$

so that the differential cross section becomes

$$\frac{d\sigma}{d\Omega} = \frac{M^2}{(2\pi\hbar)^2}|T_{\mathbf{p}_b\mathbf{p}_a}|^2. \tag{1.488}$$

If the scattered particle moves relativistically, we have to replace the constant mass M in (1.484) by $E = \sqrt{p^2+M^2}$ inside the momentum integral, where $p = |\mathbf{p}|$, so that

$$\begin{aligned}\int \frac{d^3p}{(2\pi\hbar)^3} &= \frac{1}{(2\pi\hbar)^3}\int d\Omega \int_0^\infty dp\, p^2 \\ &= \frac{1}{(2\pi\hbar)^3}\int d\Omega \int_0^\infty dE E\, p.\end{aligned} \tag{1.489}$$

In the relativistic case, the initial current density is not proportional to p/M but to the relativistic velocity $v = p/E$ so that

$$j = \frac{1}{L^3}\frac{p}{E}. \tag{1.490}$$

Hence the cross section becomes

$$\frac{d\sigma}{d\Omega} = \frac{E^2}{(2\pi\hbar)^2}|T_{\mathbf{p}_b\mathbf{p}_a}|^2. \tag{1.491}$$

1.16.3 Born Approximation

To lowest order in the interaction strength, the operator $\hat{S}$ in (1.474) is

$$\hat{S} \approx 1 - i\hat{V}/\hbar. \tag{1.492}$$

For a time-independent scattering potential, this implies

$$T_{\mathbf{p}_b\mathbf{p}_a} \approx V_{\mathbf{p}_b\mathbf{p}_a}/\hbar, \tag{1.493}$$

where

$$V_{\mathbf{p}_b\mathbf{p}_a} \equiv \langle \mathbf{p}_b|\hat{V}|\mathbf{p}_a\rangle = \int d^3x\, e^{i(\mathbf{p}_b-\mathbf{p}_a)\mathbf{x}/\hbar}V(\mathbf{x}) = \tilde{V}(\mathbf{p}_b-\mathbf{p}_a) \tag{1.494}$$

is a function of the *momentum transfer* $\mathbf{q} \equiv \mathbf{p}_b - \mathbf{p}_a$ only. Then (1.491) reduces to the so called *Born approximation* (Born 1926)

$$\frac{d\sigma}{d\Omega} \approx \frac{E^2}{(2\pi\hbar)^2\hbar^2}|V_{\mathbf{p}_b\mathbf{p}_a}|^2. \tag{1.495}$$

The amplitude whose square is equal to the differential cross section is usually denoted by $f_{\mathbf{p}_b\mathbf{p}_a}$, i.e., one writes

$$\frac{d\sigma}{d\Omega} = |f_{\mathbf{p}_b\mathbf{p}_a}|^2. \tag{1.496}$$

By comparison with (1.495) we identify

$$f_{\mathbf{p}_b\mathbf{p}_a} \equiv -\frac{M}{2\pi\hbar}R_{\mathbf{p}_b\mathbf{p}_a}, \tag{1.497}$$

where we have chosen the sign to agree with the convention in the textbook by Landau and Lifshitz [9].

1.16.4 Partial Wave Expansion and Eikonal Approximation

The scattering amplitude is usually expanded in partial waves with the help of Legendre polynomials $P_l(z) \equiv P_l^0(z)$ [see (1.421)] as

$$f_{\mathbf{p}_b\mathbf{p}_a} = \frac{\hbar}{2ip}\sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta)\left(e^{2i\delta_l(p)}-1\right) \tag{1.498}$$

where $p \equiv |\mathbf{p}| = |\mathbf{p}_b| = |\mathbf{p}_a|$ and θ is the scattering defined by $\cos\theta \equiv \mathbf{p}_b\mathbf{p}_b/|\mathbf{p}_b||\mathbf{p}_a|$. In terms of θ, the momentum transfer $\mathbf{q} = \mathbf{p}_b - \mathbf{p}_a$ has the size $|\mathbf{q}| = 2p\sin(\theta/2)$.

For small θ, we can use the asymptotic form of the Legendre polynomials[23]

$$P_l^{-m}(\cos\theta) \approx \frac{1}{l^m}J_m(l\theta), \tag{1.499}$$

to rewrite (1.498) approximately as an integral

$$f^{\rm ei}_{\mathbf{p}_b\mathbf{p}_a} = \frac{p}{i\hbar}\int db\, b\, J_0(qb)\left\{\exp\left[2i\delta_{pb/\hbar}(p)\right]-1\right\}, \tag{1.500}$$

where $b \equiv l\hbar/p$ is the so called *impact parameter* of the scattering process. This is the *eikonal approximation* to the scattering amplitude. As an example, consider Coulomb scattering where $V(r) = Ze^2/r$ and (2.751) yields

$$\chi^{\rm ei}_{\mathbf{b},\mathbf{P}}[\mathbf{v}] = -\frac{Ze^2M}{|\mathbf{P}|}\frac{1}{\hbar}\int_{-\infty}^{\infty}dz\,\frac{1}{\sqrt{b^2+z^2}}. \tag{1.501}$$

[23] M. Abramowitz and I. Stegun, op. cit., Formula 9.1.71.

The integral diverges logarithmically, but in a physical sample, the potential is screened at some distance R by opposite charges. Performing the integral up to R yields

$$\chi^{\rm ei}_{\mathbf{b},\mathbf{P}}[\mathbf{v}] = -\frac{Ze^2M}{|\mathbf{P}|}\frac{1}{\hbar}\int_b^R dr\,\frac{1}{\sqrt{r^2-b^2}} = -\frac{Ze^2M}{|\mathbf{P}|}\frac{1}{\hbar}\log\frac{R+\sqrt{R^2-b^2}}{b}$$
$$\approx -2\frac{Ze^2M}{|\mathbf{P}|}\frac{1}{\hbar}\log\frac{2R}{b}. \tag{1.502}$$

This implies

$$\exp\left(\chi^{\rm ei}_{\mathbf{b},\mathbf{P}}\right) \approx \left(\frac{b}{2R}\right)^{2i\gamma}, \tag{1.503}$$

where

$$\gamma \equiv \frac{Ze^2M}{|\mathbf{P}|}\frac{1}{\hbar} \tag{1.504}$$

is a dimensionless quantity since $e^2=\hbar c\alpha$ where α is the dimensionless *fine-structure constant*[24]

$$\alpha = \frac{e^2}{\hbar c} = 1/137.035\,997\,9\dots\ . \tag{1.505}$$

The integral over the impact parameter in (1.500) can now be performed and yields

$$f^{\rm ei}_{\mathbf{p}_b\mathbf{p}_a} \approx \frac{\hbar}{2ip}\frac{1}{\sin^{2+2i\gamma}(\theta/2)}\frac{\Gamma(1+i\gamma)}{\Gamma(-i\gamma)}e^{-2i\gamma\log(2pR/\hbar)}. \tag{1.506}$$

Remarkably, this is the exact quantum mechanical amplitude of Coulomb scattering, except for the last phase factor which accounts for a finite screening length. This amplitude contains poles at momentum variables $p=p_n$ whenever

$$i\gamma_n \equiv \frac{Ze^2M\hbar}{p_n} = -n, \quad n=1,2,3,\dots\ . \tag{1.507}$$

This corresponds to energies

$$E^{(n)} = -\frac{p_n^2}{2M} = -\frac{MZ^2e^4}{\hbar^2}\frac{1}{2n^2}, \tag{1.508}$$

which are the well-known energy values of hydrogen-like atoms with nuclear charge Ze. The prefactor $E_H \equiv e^2/a_H = Me^4/\hbar^2 = 4.359\times10^{-11}$ erg $=27.210$ eV, is equal to twice the *Rydberg energy* (see also p. 964).

[24]Throughout this book we use electromagnetic units where the electric field $\mathbf{E}=-\boldsymbol{\nabla}\phi$ has the energy density $\mathcal{H}=\mathbf{E^2}/8\pi+\rho\phi$, where ρ is the charge density, so that $\boldsymbol{\nabla}\cdot\mathbf{E}=4\pi\rho$ and $e^2=\hbar c\,\alpha$. The fine-structure constant is measured most precisely via the *quantum Hall effect*, see M.E. Cage et al., IEEE Trans. Instrum. Meas. *38*, 284 (1989). The magnetic field satisfies Ampère's law $\boldsymbol{\nabla}\times\mathbf{B}=4\pi\mathbf{j}$, where $\mathbf{j}$ is the current density.

1.16.5 Scattering Amplitude from Time Evolution Amplitude

There exists a heuristic formula expressing the scattering amplitude as a limit of the time evolution amplitude. For this we express the δ-function in the energy as a large-time limit

$$\delta(E_b - E_a) = \frac{M}{p_b}\delta(p_b - p_a) = \frac{M}{p_b}\lim_{t_b\to\infty}\left(\frac{t_b}{2\pi\hbar M/i}\right)^{1/2}\exp\left[-\frac{i}{\hbar}\frac{t_b}{2M}(p_b - p_a)^2\right], \tag{1.509}$$

where $p_b = |\mathbf{p}_b|$. Inserting this into Eq. (1.477) and setting sloppily $p_b = p_a$ for elastic scattering, the δ-function is removed and we obtain the following expression for the scattering amplitude

$$f_{\mathbf{p}_b\mathbf{p}_a} = \frac{p_b}{M}\frac{\sqrt{2\pi\hbar M/i}^3}{(2\pi\hbar)^3}\lim_{t_b\to\infty}\frac{1}{t_b^{1/2}}e^{iE_b(t_b-t_a)/\hbar}\left[(\mathbf{p}_b t_b|\mathbf{p}_a t_a) - \langle\mathbf{p}_b|\mathbf{p}_a\rangle\right]. \tag{1.510}$$

This treatment of a δ-function is certainly unsatisfactory. A satisfactory treatment will be given in the path integral formulation in Section 2.22. At the present stage, we may proceed with more care with the following operator calculation. We rewrite the limit (1.474) with the help of the time evolution operator (2.5) as follows:

$$\begin{aligned}\langle\mathbf{p}_b|\hat{S}|\mathbf{p}_a\rangle &\equiv \lim_{t_b-t_a\to\infty} e^{i(E_b t_b - E_a t_a)/\hbar}(\mathbf{p}_b t_b|\mathbf{p}_a t_a)\\ &= \lim_{t_b,-t_a\to\infty}\langle\mathbf{p}_b|\hat{U}_I(t_b,t_a)|\mathbf{p}_a\rangle,\end{aligned} \tag{1.511}$$

where $\hat{U}_I(t_b,t_a)$ is the time evolution operator in Dirac's interaction picture (1.286).

1.16.6 Lippmann-Schwinger Equation

From the definition (1.286) it follows that the operator $\hat{U}_I(t_b,t_a)$ satisfies the same composition law (1.254) as the ordinary time evolution operator $\hat{U}(t,t_a)$:

$$\hat{U}_I(t,t_a) = \hat{U}_I(t,t_b)\hat{U}_I(t_b,t_a). \tag{1.512}$$

Now we observe that

$$e^{-iH_0t/\hbar}\hat{U}_I(t,t_a) = e^{-iHt/\hbar}\hat{U}_I(0,t_a) = \hat{U}_I(0,t_a - t)e^{-iH_0t/\hbar}, \tag{1.513}$$

so that in the limit $t_a \to -\infty$

$$e^{-iH_0t/\hbar}\hat{U}_I(t,t_a) = e^{-iHt/\hbar}\hat{U}_I(0,t_a) \longrightarrow \hat{U}_I(0,t_a)e^{-iH_0t/\hbar}, \tag{1.514}$$

and therefore

$$\lim_{t_a\to-\infty}\hat{U}_I(t_b,t_a) = \lim_{t_a\to-\infty}e^{iH_0t_b/\hbar}e^{-iHt_b/\hbar}\hat{U}_I(0,t_a) = \lim_{t_a\to-\infty}e^{iH_0t_b/\hbar}\hat{U}_I(0,t_a)e^{-iH_0t_b/\hbar}; \tag{1.515}$$

which allows us to rewrite the scattering matrix (1.511) as

$$\langle \mathbf{p}_b|\hat{S}|\mathbf{p}_a\rangle \equiv \lim_{t_b,-t_a\to\infty} e^{i(E_b-E_a)t_b/\hbar}\langle \mathbf{p}_b|\hat{U}_I(0,t_a)|\mathbf{p}_a\rangle. \tag{1.516}$$

Note that in contrast to (1.474), the time evolution of the initial state goes now only over the negative time axis rather than the full one.

Taking the matrix elements of Eq. (1.291) between free-particle states $\langle \mathbf{p}_b|$ and $|\mathbf{p}_b\rangle$, and using Eqs. (1.291) and (1.514), we obtain at $t_b = 0$

$$\langle \mathbf{p}_b|\hat{U}_I(0,t_a)|\mathbf{p}_b\rangle = \langle \mathbf{p}_b|\mathbf{p}_b\rangle - \frac{i}{\hbar}\int_{-\infty}^{0} dt\, e^{i(E_b-E_a-i\eta)t/\hbar}\langle \mathbf{p}_b|\hat{V}\hat{U}_I(0,t_a)|\mathbf{p}_b\rangle. \tag{1.517}$$

A small damping factor $e^{\eta t/\hbar}$ is inserted to ensure convergence at $t = -\infty$. For a time-independent potential, the integral can be done and yields

$$\langle \mathbf{p}_b|\hat{U}_I(0,t_a)|\mathbf{p}_b\rangle = \langle \mathbf{p}_b|\mathbf{p}_b\rangle - \frac{1}{E_b-E_a-i\eta}\langle \mathbf{p}_b|\hat{V}\hat{U}_I(0,t_a)|\mathbf{p}_b\rangle. \tag{1.518}$$

This is the famous *Lippmann-Schwinger equation*. Inserting this into (1.516), we obtain the equation for the scattering matrix

$$\langle \mathbf{p}_b|\hat{S}|\mathbf{p}_a\rangle = \lim_{t_b,-t_a\to\infty} e^{i(E_b-E_a)t_b}\left[\langle \mathbf{p}_b|\mathbf{p}_a\rangle - \frac{1}{E_b-E_a-i\eta}\langle \mathbf{p}_b|\hat{V}\hat{U}_I(0,t_a)|\mathbf{p}_b\rangle\right]. \tag{1.519}$$

The first term in brackets is nonzero only if the momenta $\mathbf{p}_a$ and $\mathbf{p}_b$ are equal, in which case also the energies are equal, $E_b = E_a$, so that the prefactor can be set equal to one. In front of the second term, the prefactor oscillates rapidly as the time t_b grows large, making any finite function of E_b vanish, as a consequence of the Riemann-Lebesgue lemma. The second term contains, however, a pole at $E_b = E_a$ for which the limit has to be done more carefully. The prefactor has the property

$$\lim_{t_b\to\infty}\frac{e^{i(E_b-E_a)t_b/\hbar}}{E_b-E_a-i\eta} = \begin{cases} 0, & E_b \neq E_a,\\ i/\eta, & E_b = E_a. \end{cases} \tag{1.520}$$

It is easy to see that this property defines a δ-function in the energy:

$$\lim_{t_b\to\infty}\frac{e^{i(E_b-E_a)t_b/\hbar}}{E_b-E_a-i\eta} = 2\pi i\delta(E_b-E_a). \tag{1.521}$$

Indeed, let us integrate the left-hand side together with a smooth function $f(E_b)$, and set

$$E_b \equiv E_a + \xi/t_b. \tag{1.522}$$

Then the E_b-integral is rewritten as

$$\int_{-\infty}^{\infty} d\xi \frac{e^{i\xi}}{\xi + i\eta} f\left(E_a + \xi/t_a\right). \tag{1.523}$$

In the limit of large t_a, the function $f(E_a)$ can be taken out of the integral and the contour of integration can then be closed in the upper half of the complex energy plane, yielding $2\pi i$. Thus we obtain from (1.519) the formula (1.477), with the T-matrix

$$\langle \mathbf{p}_b|\hat{T}|\mathbf{p}_a\rangle = \frac{1}{\hbar}\langle \mathbf{p}_b|\hat{V}\hat{U}_I(0,t_a)|\mathbf{p}_b\rangle. \tag{1.524}$$

For a small potential $\hat{V}$, we approximate $\hat{U}_I(0,t_a)\approx 1$, and find the Born approximation (1.493).

The Lippmann-Schwinger equation can be recast as an integral equation for the T-matrix. Multiplying the original equation (1.518) by the matrix $\langle \mathbf{p}_b|\hat{V}|\mathbf{p}_a\rangle = V_{\mathbf{p}_b\mathbf{p}_c}$ from the left, we obtain

$$T_{\mathbf{p}_b\mathbf{p}_a} = V_{\mathbf{p}_b\mathbf{p}_a} - \int \frac{d^3p_c}{(2\pi\hbar)^3} V_{\mathbf{p}_b\mathbf{p}_c}\frac{1}{E_c - E_a - i\eta}T_{\mathbf{p}_c\mathbf{p}_a}. \tag{1.525}$$

To extract physical information from the T-matrix (1.524) it is useful to analyze the behavior of the interacting state $\hat{U}_I(0,t_a)|\mathbf{p}_a\rangle$ in $\mathbf{x}$-space. From Eq. (1.514), we see that it is an eigenstate of the full Hamiltonian operator $\hat{H}$ with the initial energy E_a. Multiplying this state by $\langle\mathbf{x}|$ from the left, and inserting a complete set of momentum eigenstates, we calculate

$$\langle\mathbf{x}|\hat{U}_I(0,t_a)|\mathbf{p}_a\rangle = \int\frac{d^3p}{(2\pi\hbar)^3}\langle\mathbf{x}|\mathbf{p}\rangle\langle\mathbf{p}|\hat{U}_I(0,t_a)|\mathbf{p}_a\rangle = \int\frac{d^3p}{(2\pi\hbar)^3}\langle\mathbf{x}|\mathbf{p}\rangle\langle\mathbf{p}|\hat{U}_I(0,t_a)|\mathbf{p}_a\rangle.$$

Using Eq. (1.518), this becomes

$$\langle\mathbf{x}|\hat{U}_I(0,t_a)|\mathbf{p}_a\rangle = \langle\mathbf{x}|\mathbf{p}_a\rangle + \int d^3x'\int\frac{d^3p_b}{(2\pi\hbar)^3}\frac{e^{i\mathbf{p}_b(\mathbf{x}-\mathbf{x}')/\hbar}}{E_a - \mathbf{p}_b^2/2M + i\eta}V(\mathbf{x}')\langle\mathbf{x}'|\hat{U}_I(0,t_a)|\mathbf{p}_a\rangle. \tag{1.526}$$

The function

$$(\mathbf{x}|\mathbf{x}')_{E_a} = \int\frac{d^3p_b}{(2\pi\hbar)^3}e^{i\mathbf{p}_b(\mathbf{x}-\mathbf{x}')/\hbar}\frac{i\hbar}{E_a - \mathbf{p}^2/2M + i\eta} \tag{1.527}$$

is recognized as the fixed-energy amplitude (1.344) of the free particle. In three dimensions it reads [see (1.359)]

$$(\mathbf{x}|\mathbf{x}')_{E_a} = -\frac{2Mi}{\hbar}\frac{e^{ip_a|\mathbf{x}-\mathbf{x}'|/\hbar}}{4\pi|\mathbf{x}-\mathbf{x}'|}, \qquad p_a = \sqrt{2ME_a}. \tag{1.528}$$

In order to find the scattering amplitude, we consider the wave function (1.526) far away from the scattering center, i.e., at large $|\mathbf{x}|$. Under the assumption that $V(\mathbf{x}')$ is nonzero only for small $\mathbf{x}'$, we approximate $|\mathbf{x}-\mathbf{x}'|\approx r - \hat{\mathbf{x}}\mathbf{x}'$, where $\hat{\mathbf{x}}$ is the unit vector in the direction of $\mathbf{x}$, and (1.526) becomes

$$\langle\mathbf{x}|\hat{U}_I(0,t_a)|\mathbf{p}_a\rangle \approx e^{i\mathbf{p}_a\mathbf{x}/\hbar} - \frac{e^{ip_ar}}{4\pi r}\int d^4x' e^{-ip_a\hat{\mathbf{x}}\mathbf{x}'}\frac{2M}{\hbar^2}V(\mathbf{x}')\langle\mathbf{x}'|\hat{U}_I(0,t_a)|\mathbf{p}_a\rangle. \tag{1.529}$$

In the limit $t_a \to -\infty$, the factor multiplying the spherical wave factor $e^{ip_a r/\hbar}/r$ is the scattering amplitude $f(\hat{\mathbf{x}})_{\mathbf{p}_a}$, whose absolute square gives the cross section. For scattering to a final momentum $\mathbf{p}_b$, the outgoing particles are detected far away from the scattering center in the direction $\hat{\mathbf{x}} = \hat{\mathbf{p}_b}$. Because of energy conservation, we may set $p_a\hat{\mathbf{x}} = \mathbf{p}_b$ and obtain the formula

$$f_{\mathbf{p}_b\mathbf{p}_a} = \lim_{t_a\to-\infty} -\frac{M}{2\pi\hbar^2}\int d^4x_b e^{-i\mathbf{p}_b\mathbf{x}_b}V(\mathbf{x}_b)\langle\mathbf{x}_b\,|\hat{U}_I(0,t_a)|\mathbf{p}_a\rangle. \tag{1.530}$$

By studying the interacting state $\hat{U}_I(0,t_a)|\mathbf{p}_a\rangle$ in $\mathbf{x}$-space, we have avoided the singular δ-function of energy conservation.

We are now prepared to derive formula (1.510) for the scattering amplitude. We observe that in the limit $t_a \to -\infty$, the amplitude $\langle\mathbf{x}_b\,|\hat{U}_I(0,t_a)|\mathbf{p}_a\rangle$ can be obtained from the time evolution amplitude $(\mathbf{x}_b t_b|\mathbf{x}_a t_a)$ as follows:

$$\begin{aligned}\langle\mathbf{x}_b\,|\hat{U}_I(0,t_a)|\mathbf{p}_a\rangle &= \langle\mathbf{x}_b\,|\hat{U}(0,t_a)|\mathbf{p}_a\rangle e^{-iE_at_a/\hbar} \\ &= \lim_{t_a\to-\infty}\left(\frac{-2\pi i\hbar t_a}{M}\right)^{3/2}(\mathbf{x}_bt_b|\mathbf{x}_at_a)e^{i(\mathbf{p}_a\mathbf{x}_a-p_a^2t_a/2M)/\hbar}\Big|_{\mathbf{x}_a=\mathbf{p}_at_a/M}.\end{aligned} \tag{1.531}$$

This follows directly from the Fourier transformation

$$\langle\mathbf{x}_b\,|\hat{U}(0,t_a)|\mathbf{p}_a\rangle e^{-iE_at_a/\hbar} = \int d^3x_a(\mathbf{x}_bt_b|\mathbf{x}_at_a)e^{i(\mathbf{p}_a\mathbf{x}_a-p_a^2t_a/2M)/\hbar}, \tag{1.532}$$

by substituting the dummy integration variable $\mathbf{x}_a$ by $\mathbf{p}t_a/M$. Then the right-hand side becomes

$$\left(\frac{-t_a}{M}\right)^3\int d^3p\,(\mathbf{x}_b\,0|\mathbf{p}t_a\;t_a)e^{i(\mathbf{p}_a\mathbf{p}-p_a^2)t_a/2M\hbar}. \tag{1.533}$$

Now, for large $-t_a$, the momentum integration is squeezed to $\mathbf{p} = \mathbf{p}_a$, and we obtain (1.531). The appropriate limiting formula for the δ-function

$$\delta^{(D)}(\mathbf{p}_b-\mathbf{p}_a) = \lim_{t_a\to-\infty}\frac{(-t_a)^{D/2}}{\sqrt{2\pi i\hbar M}^D}\exp\left\{-\frac{i}{\hbar}\frac{t_a}{2M}(\mathbf{p}_b-\mathbf{p}_a)^2\right\} \tag{1.534}$$

is easily obtained from Eq. (1.342) by an obvious substitution of variables. Its complex conjugate for $D = 1$ was written down before in Eq. (1.509) with t_a replaced by $-t_b$. The exponential on the right-hand side can just as well be multiplied by a factor $e^{i(p_b^2-p_a^2)^2/2M\hbar}$ which is unity when both sides are nonzero, so that it becomes $e^{-i(\mathbf{p}_a\mathbf{p}-p_a^2)t_a/2M\hbar}$. In this way we obtain a representation of the δ-function by which the Fourier integral (1.533) goes over into (1.531). The phase factor $e^{i(\mathbf{p}_a\mathbf{x}_a-p_a^2t_a/2M)/\hbar}$ on the right-hand side of Eq. (1.531), which is unity in the limit performed in that equation, is kept in Eq. (4.580) for later convenience.

Formula (1.531) is a reliable starting point for extracting the scattering amplitude $f_{\mathbf{p}_b\mathbf{p}_a}$ from the time evolution amplitude in $\mathbf{x}$-space $(\mathbf{x}_b\,0|\mathbf{x}_at_a)$ at $\mathbf{x}_a = \mathbf{p}_at_a/M$ by extracting the coefficient of the outcoming spherical wave $e^{ip_ar/\hbar}/r$.

As a cross check we insert the free-particle amplitude (1.341) into (1.531) and obtain the free undisturbed wave function $e^{i\mathbf{p}_a\mathbf{x}}$, which is the correct first term in Eq. (1.526) associated with unscattered particles.

1.17 Classical and Quantum Statistics

Consider a physical system with a constant number of particles N whose Hamiltonian has no explicit time dependence. If it is brought into contact with a thermal reservoir at a temperature T and has reached equilibrium, its thermodynamic properties can be obtained through the following rules: At the level of classical mechanics, each volume element in phase space

$$\frac{dp\,dq}{h} = \frac{dp\,dq}{2\pi\hbar} \tag{1.535}$$

is occupied with a probability proportional to the *Boltzmann factor*

$$e^{-H(p,q)/k_BT}, \tag{1.536}$$

where k_B is the *Boltzmann constant*,

$$k_B = 1.3806221(59) \times 10^{-16}\,\text{erg/Kelvin}. \tag{1.537}$$

The number in parentheses indicates the experimental uncertainty of the two digits in front of it. The quantity $1/k_BT$ has the dimension of an inverse energy and is commonly denoted by β. It will be called the *inverse temperature*, forgetting about the factor k_B. In fact, we shall sometimes take T to be measured in energy units k_B times Kelvin rather than in Kelvin. Then we may drop k_B in all formulas.

The integral over the Boltzmann factors of all phase space elements,[25]

$$Z_{\rm cl}(T) \equiv \int \frac{dp\,dq}{2\pi\hbar}\, e^{-H(p,q)/k_BT}, \tag{1.538}$$

is called the *classical partition function*. It contains all classical thermodynamic information of the system. Of course, for a general Hamiltonian system with many degrees of freedom, the phase space integral is $\prod_n \int dp_n\, dq_n/2\pi\hbar$. The reader may wonder why an expression containing Planck's quantum $\hbar$ is called *classical*. The reason is that $\hbar$ can really be omitted in calculating any thermodynamic average. In classical statistics it merely supplies us with an irrelevant normalization factor which makes Z dimensionless.

1.17.1 Canonical Ensemble

In quantum statistics, the Hamiltonian is replaced by the operator $\hat{H}$ and the integral over phase space by the trace in the Hilbert space. This leads to the *quantum-statistical partition function*

$$Z(T) \equiv \text{Tr}\left(e^{-\hat{H}/k_BT}\right) \equiv \text{Tr}\left(e^{-H(\hat{p},\hat{x})/k_BT}\right), \tag{1.539}$$

[25]In the sequel we shall always work at a fixed volume V and therefore suppress the argument V everywhere.

where $\mathrm{Tr}\,\hat{O}$ denotes the trace of the operator $\hat{O}$. If $\hat{H}$ is an N-particle Schrödinger Hamiltonian, the quantum-statistical system is referred to as a *canonical ensemble*. The right-hand side of (1.539) contains the position operator $\hat{x}$ in Cartesian coordinates rather than $\hat{q}$ to ensure that the system can be quantized canonically. In cases such as the spinning top, the trace formula is also valid but the Hilbert space is spanned by the representation states of the angular momentum operators. In more general Lagrangian systems, the quantization has to be performed differently in the way to be described in Chapters 10 and 8.

At this point we make an important observation: The quantum partition function is related in a very simple way to the quantum-mechanical time evolution operator. To emphasize this relation we shall define the trace of this operator for time-independent Hamiltonians as the *quantum-mechanical partition function*:

$$Z_{\mathrm{QM}}(t_b - t_a) \equiv \mathrm{Tr}\left(\hat{U}(t_b, t_a)\right) = \mathrm{Tr}\left(e^{-i(t_b - t_a)\hat{H}/\hbar}\right). \tag{1.540}$$

Obviously the quantum-statistical partition function $Z(T)$ may be obtained from the quantum-mechanical one by continuing the time interval $t_b - t_a$ to the negative imaginary value

$$t_b - t_a = -\frac{i\hbar}{k_B T} \equiv -i\hbar\beta. \tag{1.541}$$

This simple formal relation shows that the trace of the time evolution operator contains all information on the thermodynamic equilibrium properties of a quantum system.

1.17.2 Grand-Canonical Ensemble

For systems containing many bodies it is often convenient to study their equilibrium properties in contact with a particle reservoir characterized by a chemical potential μ. For this one defines what is called the *grand-canonical quantum-statistical partition function*

$$Z_G(T, \mu) = \mathrm{Tr}\left(e^{-(\hat{H} - \mu\hat{N})/k_B T}\right). \tag{1.542}$$

Here $\hat{N}$ is the operator counting the number of particles in each state of the ensemble. The combination of operators in the exponent,

$$\hat{H}_G = \hat{H} - \mu\hat{N}, \tag{1.543}$$

is called the *grand-canonical Hamiltonian*.

Given a partition function $Z(T)$ at a fixed particle number N, the *free energy* is defined by

$$F(T) = -k_B T \log Z(T). \tag{1.544}$$

Its grand-canonical version at a fixed chemical potential is

$$F_G(T, \mu) = -k_B T \log Z_G(T, \mu). \tag{1.545}$$

The *average energy* or *internal energy* is defined by

$$E = \mathrm{Tr}\left(\hat{H}e^{-\hat{H}/k_BT}\right)\Big/\mathrm{Tr}\left(e^{-\hat{H}/k_BT}\right). \tag{1.546}$$

It may be obtained from the partition function $Z(T)$ by forming the temperature derivative

$$E = Z^{-1}k_BT^2\frac{\partial}{\partial T}Z(T) = k_BT^2\frac{\partial}{\partial T}\log Z(T). \tag{1.547}$$

In terms of the free energy (1.544), this becomes

$$E = T^2\frac{\partial}{\partial T}\left(-F(T)/T\right) = \left(1 - T\frac{\partial}{\partial T}\right)F(T). \tag{1.548}$$

For a grand-canonical ensemble we may introduce an *average particle number* defined by

$$N = \mathrm{Tr}\left(\hat{N}e^{-(\hat{H}-\mu\hat{N})/k_BT}\right)\Big/\mathrm{Tr}\left(e^{-(\hat{H}-\mu\hat{N})/k_BT}\right). \tag{1.549}$$

This can be derived from the grand-canonical partition function as

$$N = {Z_G}^{-1}(T,\mu)k_BT\frac{\partial}{\partial\mu}Z_G(T,\mu) = k_BT\frac{\partial}{\partial\mu}\log Z_G(T,\mu), \tag{1.550}$$

or, using the grand-canonical free energy, as

$$N = -\frac{\partial}{\partial\mu}F_G(T,\mu). \tag{1.551}$$

The average energy in a grand-canonical system,

$$E = \mathrm{Tr}\left(\hat{H}e^{-(\hat{H}-\mu\hat{N})/k_BT}\right)\Big/\mathrm{Tr}\left(e^{-(\hat{H}-\mu\hat{N})/k_BT}\right), \tag{1.552}$$

can be obtained by forming, by analogy with (1.547) and (1.548), the derivative

$$\begin{aligned} E - \mu N &= {Z_G}^{-1}(T,\mu)k_BT^2\frac{\partial}{\partial T}Z_G(T,\mu) \\ &= \left(1 - T\frac{\partial}{\partial T}\right)F_G(T,\mu). \end{aligned} \tag{1.553}$$

For a large number of particles, the density is a rapidly growing function of energy. For a system of N free particles, for example, the number of states up to energy E is given by

$$N(E) = \sum_{\mathbf{p}_i}\Theta(E - \sum_{i=1}^{N}\mathbf{p}_i^2/2M), \tag{1.554}$$

where each of the particle momenta $\mathbf{p}_i$ is summed over all discrete momenta $\mathbf{p}^{\mathbf{m}}$ in (1.179) available to a single particle in a finite box of volume $V = L^3$. For a large V, the sum can be converted into an integral[26]

$$N(E) = V^N\prod_{i=1}^{N}\left[\int\frac{d^3p_i}{(2\pi\hbar)^3}\right]\Theta(E - \sum_{i=1}^{N}\mathbf{p}_i^2/2M), \tag{1.555}$$

[26] Remember, however, the exception noted in the footnote to Eq. (1.184) for systems possessing a condensate.

which is simply $[V/(2\pi\hbar)^3]^N$ times the volume Ω_{3N} of a $3N$-dimensional sphere of radius $\sqrt{2ME}$:

$$\begin{aligned} N(E) &= \left[\frac{V}{(2\pi\hbar)^3}\right]^N \Omega_{3N} \\ &\equiv \left[\frac{V}{(2\pi\hbar)^3}\right]^N \frac{(2\pi ME)^{3N/2}}{\Gamma\left(\frac{3}{2}N+1\right)}. \end{aligned} \tag{1.556}$$

Recall the well-known formula for the volume of a unit sphere in D dimensions:

$$\Omega_D = \pi^{D/2}/\Gamma(D/2+1). \tag{1.557}$$

The surface is [see Eqs. (8.117) and (8.118) for a derivation]

$$S_D = 2\pi^{D/2}/\Gamma(D/2). \tag{1.558}$$

This follows directly from the integral[27]

$$S_D = \int d^Dp\,\delta(p-1) = \int d^Dp\,2\delta(p^2-1) = \int d^Dp \int_{-\infty}^{\infty} \frac{d\lambda}{\pi} e^{i\lambda(p^2-1)} \tag{1.559}$$

$$= \int_{-\infty}^{\infty} \frac{d\lambda}{\pi} \left(\frac{\pi}{-i\lambda}\right)^{D/2} e^{-i\lambda} = \frac{2\pi^{D/2}}{\Gamma(D/2)} \tag{1.560}$$

Therefore, the density per energy $\rho = \partial N/\partial E$ is given by

$$\rho(E) = \left[\frac{V}{(2\pi\hbar)^3}\right]^N 2\pi M \frac{(2\pi ME)^{3N/2-1}}{\Gamma(\frac{3}{2}N)}. \tag{1.561}$$

It grows with the very large power $E^{3N/2}$ in the energy. Nevertheless, the integral for the partition function (1.582) is convergent, due to the overwhelming exponential falloff of the Boltzmann factor, e^{-E/k_BT}. As the two functions $\rho(e)$ and e^{-e/k_BT} are multiplied with each other, the result is a function which peaks very sharply at the average energy E of the system. The position of the peak depends on the temperature T. For the free N particle system, for example,

$$\rho(E)e^{-E/k_BT} \sim e^{(3N/2-1)\log E - E/k_BT}. \tag{1.562}$$

This function has a sharp peak at

$$E(T) = k_BT\left(\frac{3N}{2}-1\right) \approx k_BT\frac{3N}{2}. \tag{1.563}$$

The width of the peak is found by expanding (1.562) in $\delta E = E - E(T)$:

$$\exp\left\{\frac{3N}{2}\log E(T) - \frac{E(T)}{k_BT} - \frac{1}{2E^2(T)}\frac{3N}{2}(\delta E)^2 + \ldots\right\}. \tag{1.564}$$

[27] I. S. Gradshteyn and I. M. Ryzhik, op. cit., Formula 3.382.7.

Thus, as soon as δE gets to be of the order of $E(T)/\sqrt{N}$, the exponential is reduced by a factor of two with respect to $E(T) \approx k_B T\, 3N/2$. The deviation is of a relative order $1/\sqrt{N}$, i.e., the peak is very sharp. With N being very large, the peak at $E(T)$ of width $E(T)/\sqrt{N}$ can be idealized by a δ-function, and we may write

$$\rho(E)e^{-E/k_B T} \approx \delta(E - E(T))N(T)e^{-E(T)/k_B T}. \tag{1.565}$$

The quantity $N(T)$ measures the total number of states over which the system is distributed at the temperature T.

The entropy $S(T)$ is now defined in terms of $N(T)$ by

$$N(T) = e^{S(T)/k_B}. \tag{1.566}$$

Inserting this with (1.565) into (1.582), we see that in the limit of a large number of particles N:

$$Z(T) = e^{-[E(T)-TS(T)]/k_B T}. \tag{1.567}$$

Using (1.544), the free energy can thus be expressed in the form

$$F(T) = E(T) - TS(T). \tag{1.568}$$

By comparison with (1.548) we see that the entropy may be obtained from the free energy directly as

$$S(T) = -\frac{\partial}{\partial T}F(T). \tag{1.569}$$

For grand-canonical ensembles we may similarly consider

$$Z_G(T,\mu) = \int dE\, dn\, \rho(E,n)e^{-(E-\mu n)/k_B T}, \tag{1.570}$$

where

$$\rho(E,n)e^{-(E-\mu n)/k_B T} \tag{1.571}$$

is now strongly peaked at $E = E(T,\mu)$, $n = N(T,\mu)$ and can be written approximately as

$$\begin{aligned}\rho(E,n)e^{-(E-\mu n)/k_B T} &\approx \delta\left(E - E(T,\mu)\right)\delta\left(n - N(T,\mu)\right) \\ &\times\; e^{S(T,\mu)/k_B}e^{-[E(T,\mu)-\mu N(T,\mu)]/k_B T}.\end{aligned} \tag{1.572}$$

Inserting this back into (1.570) we find for large N

$$Z_G(T,\mu) = e^{-[E(T,\mu)-\mu N(T,\mu)-TS(T,\mu)]/k_B T}. \tag{1.573}$$

For the grand-canonical free energy (1.545), this implies the relation

$$F_G(T,\mu) = E(T,\mu) - \mu N(T,\mu) - TS(T,\mu). \tag{1.574}$$

By comparison with (1.553) we see that the entropy can be calculated directly from the derivative of the grand-canonical free energy

$$S(T,\mu) = -\frac{\partial}{\partial T}F_G(T,\mu). \tag{1.575}$$

The particle number is, of course, found from the derivative (1.551) with respect to the chemical potential, as follows directly from the definition (1.570).

The canonical free energy and the entropy appearing in the above equations depend on the particle number N and the volume V of the system, i.e., they are more explicitly written as $F(T, N, V)$ and $S(T, N, V)$, respectively.

In the arguments of the grand-canonical quantities, the particle number N is replaced by the chemical potential μ.

Among the arguments of the grand-canonical free energy $F_G(T, \mu, V)$, the volume V is the only one which grows with the system. Thus $F_G(T, \mu, V)$ must be directly proportional to V. The proportionality constant defines the *pressure* p of the system:

$$F_G(T, \mu, V) \equiv -p(T, \mu, V)V. \tag{1.576}$$

Under infinitesimal changes of the three variables, $F_G(T, \mu, V)$ changes as follows:

$$dF_G(T, \mu, V) = -SdT - Nd\mu - pdV. \tag{1.577}$$

The first two terms on the right-hand side follow from varying Eq. (1.574) at a fixed volume. When varying the volume, the definition (1.576) renders the last term.

Inserting (1.576) into (1.574), we find *Euler's relation*:

$$E = TS - Nd\mu - pV. \tag{1.578}$$

The energy has S, N, V as natural variables. Equivalently, we may write

$$F = -\mu N - pV, \tag{1.579}$$

where T, N, V are the natural variables.

1.18 Density of States and Tracelog

In many thermodynamic calculations, a quantity of fundamental interest is the density of states. To define it, we express the canonical partition function

$$Z(T) = \mathrm{Tr}\left(e^{-\hat{H}/k_BT}\right) \tag{1.580}$$

as a sum over the Boltzmann factors of all eigenstates $|n\rangle$ of the Hamiltonian:, i.e.

$$Z(T) = \sum_n e^{-E_n/k_BT}. \tag{1.581}$$

This can be rewritten as an integral:

$$Z(T) = \int dE\, \rho(E) e^{-E/k_BT}. \tag{1.582}$$

The quantity

$$\rho(E) = \sum_n \delta(E - E_n) \tag{1.583}$$

specifies the *density of states* of the system in the energy interval $(E, E + dE)$. It may also be written formally as a trace of the density of states operator $\hat{\rho}(E)$:

$$\rho(E) = \mathrm{Tr}\, \hat{\rho}(E) \equiv \mathrm{Tr}\, \delta(E - \hat{H}). \tag{1.584}$$

The density of states is obviously the Fourier transform of the canonical partition function (1.580):

$$\rho(E) = \int_{-i\infty}^{\infty} \frac{d\beta}{2\pi i}\, e^{\beta E}\, \mathrm{Tr}\, \left(e^{-\beta \hat{H}}\right) = \int_{-i\infty}^{\infty} \frac{d\beta}{2\pi i}\, e^{\beta E}\, Z(1/k_B\beta). \tag{1.585}$$

The integral

$$N(E) = \int^{E} dE'\, \rho(E') \tag{1.586}$$

is the number of states up to energy E. The integration may start anywhere below the ground state energy. The function $N(E)$ is a sum of Heaviside step functions (1.313):

$$N(E) = \sum_n \Theta(E - E_n). \tag{1.587}$$

This equation is correct only with the Heaviside function which is equal to 1/2 at the origin, not with the one-sided version (1.306), as we shall see later. Indeed, if integrated to the energy of a certain level E_n, the result is

$$N(E_n) = (n + 1/2). \tag{1.588}$$

This formula will serve to determine the energies of bound states from approximations to $\omega(E)$ in Section 4.7, for instance from the Bohr-Sommerfeld condition (4.190) via the relation (4.210). In order to apply this relation one must be sure that all levels have different energies. Otherwise $N(E)$ jumps at E_n by half the degeneracy of this level. In Eq. (4A.9) we shall exhibit an example for this situation.

An important quantity related to $\rho(E)$ which will appear frequently in this text is the trace of the logarithm, short *tracelog*, of the operator $\hat{H} - E$.

$$\mathrm{Tr}\log(\hat{H} - E) = \sum_n \log(E_n - E). \tag{1.589}$$

It may be expressed in terms of the density of states (1.584) as

$$\mathrm{Tr}\log(\hat{H} - E) = \mathrm{Tr} \int_{-\infty}^{\infty} dE'\, \delta(E' - \hat{H}) \log(E' - E) = \int_{-\infty}^{\infty} dE'\, \rho(E') \log(E' - E). \tag{1.590}$$

The tracelog of the Hamiltonian operator itself can be viewed as a limit of an *operator zeta function* associated with $\hat{H}$:

$$\hat{\zeta}_{\hat{H}}(\nu) = \mathrm{Tr}\, \hat{H}^{-\nu}, \tag{1.591}$$

whose trace is the generalized zeta-function

$$\zeta_{\hat{H}}(\nu) \equiv \mathrm{Tr}\, \left[\hat{\zeta}_{\hat{H}}(\nu)\right] = \mathrm{Tr}\, (\hat{H}^{-\nu}) = \sum_n E_n^{-\nu}. \tag{1.592}$$

For a linearly spaced spectrum $E_n = n$ with $n = 1, 2, 3 \ldots$, this reduces to Riemann's zeta function (2.521).

From the generalized zeta function we can obtain the tracelog by forming the derivative

$$\operatorname{Tr}\log \hat{H} = -\partial_\nu \, \zeta_{\hat{H}}(\nu)\big|_{\nu=0}. \tag{1.593}$$

By differentiating the tracelog (1.589) with respect to E, we find the trace of the resolvent (1.319):

$$\partial_E \operatorname{Tr}\log(\hat{H} - E) = \operatorname{Tr}\frac{1}{E - \hat{H}} = \sum_n \frac{1}{E - E_n} = \frac{1}{i\hbar}\sum_n R_n(E) = \frac{1}{i\hbar}\operatorname{Tr}\hat{R}(E). \tag{1.594}$$

Recalling Eq. (1.329) we see that the imaginary part of this quantity slightly above the real E-axis yields the density of states

$$-\frac{1}{\pi}\operatorname{Im}\partial_E \operatorname{Tr}\log(\hat{H} - E - i\eta) = \sum_n \delta(E - E_n) = \rho(E). \tag{1.595}$$

By integrating this over the energy we obtain the number of states function $N(E)$ of Eq. (1.586):

$$-\frac{1}{\pi}\operatorname{Im}\operatorname{Tr}\log(E - \hat{H}) = \sum_n \Theta(E - E_n) = N(E). \tag{1.596}$$

Appendix 1A Simple Time Evolution Operator

Consider the simplest nontrivial time evolution operator of a spin-1/2 particle in a magnetic field $\mathbf{B}$. The reduced Hamiltonian operator is $\hat{H}_0 = -\mathbf{B}\cdot\boldsymbol{\sigma}/2$, so that the time evolution operator reads, in natural units with $\hbar = 1$,

$$e^{-i\hat{H}_0(t_b - t_a)} = e^{i(t_b - t_a)\mathbf{B}\cdot\boldsymbol{\sigma}/2}. \tag{1A.1}$$

Expanding this as in (1.293) and using the fact that $(\mathbf{B}\cdot\boldsymbol{\sigma})^{2n} = B^{2n}$ and $(\mathbf{B}\cdot\boldsymbol{\sigma})^{2n+1} = B^{2n}(\mathbf{B}\cdot\boldsymbol{\sigma})$, we obtain

$$e^{-i\hat{H}_0(t_b - t_a)} = \cos[B(t_b - t_a)/2] + i\hat{\mathbf{B}}\cdot\boldsymbol{\sigma}\,\sin[B(t_b - t_a)/2]\,, \tag{1A.2}$$

where $\hat{\mathbf{B}} \equiv \mathbf{B}/|\mathbf{B}|$. Suppose now that the magnetic field is not constant but has a small time-dependent variation $\delta\mathbf{B}(t)$. Then we obtain from (1.253) [or the lowest expansion term in (1.293)]

$$\delta e^{-i\hat{H}_0(t_b - t_a)} = \int_{t_a}^{t_b} dt\, e^{-i\hat{H}_0(t_b - t)}\delta\mathbf{B}(t)\cdot\boldsymbol{\sigma} e^{-i\hat{H}_0(t - t_a)}. \tag{1A.3}$$

Using (1A.2), the integrand on the right-hand side becomes

$$\Big\{\cos[B(t_b - t)/2] + i\hat{\mathbf{B}}\cdot\boldsymbol{\sigma}\sin[B(t_b - t)/2]\Big\}\delta\mathbf{B}(t)\cdot\boldsymbol{\sigma}\Big\{\cos[B(t - t_a)/2] + i\hat{\mathbf{B}}\cdot\boldsymbol{\sigma}\sin[B(t - t_a)/2]\Big\}. \tag{1A.4}$$

We simplify this with the help of the formula [recall (1.448)]

$$\sigma^i\sigma^j = \delta_{ij} + i\epsilon_{ijk}\sigma^k \tag{1A.5}$$

so that

$$\hat{\mathbf{B}}\cdot\boldsymbol{\sigma}\;\delta\mathbf{B}(t)\cdot\boldsymbol{\sigma} = \hat{\mathbf{B}}\cdot\delta\mathbf{B}(t) + i[\hat{\mathbf{B}}\times\delta\mathbf{B}(t)]\cdot\boldsymbol{\sigma}, \quad \delta\mathbf{B}(t)\cdot\boldsymbol{\sigma}\;\hat{\mathbf{B}}\cdot\boldsymbol{\sigma} = \hat{\mathbf{B}}\cdot\delta\mathbf{B}(t) - i[\hat{\mathbf{B}}\times\delta\mathbf{B}(t)]\cdot\boldsymbol{\sigma}, \tag{1A.6}$$

and

$$\hat{\mathbf{B}}\cdot\boldsymbol{\sigma}\;\delta\mathbf{B}(t)\cdot\boldsymbol{\sigma}\;\hat{\mathbf{B}}\cdot\boldsymbol{\sigma} = \left[\hat{\mathbf{B}}\cdot\delta\mathbf{B}(t)\right]\hat{\mathbf{B}}\cdot\boldsymbol{\sigma} + i[\hat{\mathbf{B}}\times\delta\mathbf{B}(t)]\cdot\boldsymbol{\sigma}\;\hat{\mathbf{B}}\cdot\boldsymbol{\sigma}$$
$$= i[\hat{\mathbf{B}}\times\delta\mathbf{B}(t)]\cdot\hat{\mathbf{B}} + \left\{[\hat{\mathbf{B}}\cdot\delta\mathbf{B}(t)]\hat{\mathbf{B}} - [\hat{\mathbf{B}}\times\delta\mathbf{B}(t)]\times\hat{\mathbf{B}}\right\}\cdot\boldsymbol{\sigma}. \quad (1A.7)$$

The first term on the right-hand side vanishes, the second term is equal to $\delta\mathbf{B}$, since $\hat{\mathbf{B}}^2 = 1$. Thus we find for the integrand in (1A.4):

$$\begin{aligned}
&\cos B(t_b-t)/2\cos B(t-t_a)/2\,\delta\mathbf{B}(t)\cdot\boldsymbol{\sigma}\\
&+i\sin B(t_b-t)/2\cos B(t-t_a)/2\{\hat{\mathbf{B}}\cdot\delta\mathbf{B}(t)+i[\hat{\mathbf{B}}\times\delta\mathbf{B}(t)]\cdot\boldsymbol{\sigma}\}\\
&+i\cos B(t_b-t)/2\sin B(t-t_a)/2\{\hat{\mathbf{B}}\cdot\delta\mathbf{B}(t)-i[\hat{\mathbf{B}}\times\delta\mathbf{B}(t)]\cdot\boldsymbol{\sigma}\}\\
&+\sin B(t_b-t)/2\sin B(t-t_a)/2\;\delta\mathbf{B}\cdot\boldsymbol{\sigma}
\end{aligned} \quad (1A.8)$$

which can be combined to give

$$\left\{\cos B[(t_b+t_a)/2-t]\,\delta\mathbf{B}(t)-\sin B[(t_b+t_a)/2-t]\;[\hat{\mathbf{B}}\times\delta\mathbf{B}(t)]\right\}\cdot\boldsymbol{\sigma}+i\sin[B(t_b-t_a)/2]\,\hat{\mathbf{B}}\cdot\delta\mathbf{B}(t). \quad (1A.9)$$

Integrating this from t_a to t_b we obtain the variation (1A.3).

Appendix 1B Convergence of the Fresnel Integral

Here we prove the convergence of the Fresnel integral (1.337) by relating it to the Gauss integral. According to Cauchy's integral theorem, the sum of the integrals along the three pieces of the

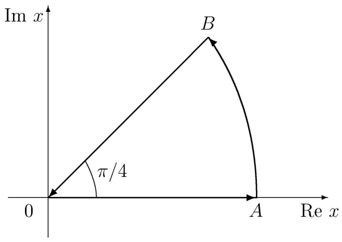

Figure 1.4 Triangular closed contour for Cauchy integral

closed contour shown in Fig. 1.4 vanishes since the integrand e^{-z^2} is analytic in the triangular domain:

$$\oint dz e^{-z^2} = \int_0^A dz e^{-z^2} + \int_A^B dz e^{-z^2} + \int_B^O dz e^{-z^2} = 0. \quad (1B.1)$$

Let R be the radius of the arc. Then we substitute in the three integrals the variable z as follows:

$$\begin{array}{llll}
0\,A: & z = p, & dz = dp, & z^2 = p^2\\
B\,0: & z = pe^{i\pi/4}, & dz = dp\,e^{i\pi/4}, & z^2 = ip^2\\
AB: & z = R\,e^{i\varphi}, & dz = i\,Rdp, & z^2 = p^2,
\end{array}$$

and obtain the equation

$$\int_0^R dp\,e^{-p^2} + e^{i\pi/4}\int_R^0 dp\,e^{-ip^2} + \int_0^{\pi/4} d\varphi\,iR\,e^{-R^2(\cos 2\varphi + i\sin 2\varphi)+i\varphi} = 0. \quad (1B.2)$$

The first integral converges rapidly to $\sqrt{\pi}/2$ for $R\to\infty$. The last term goes to zero in this limit. To see this we estimate its absolute value as follows:

$$\left|\int_0^{\pi/4} d\varphi\,iR\,e^{-R^2(\cos 2\varphi + i\sin 2\varphi)+i\varphi}\right| < R\int_0^{\pi/4} d\varphi\;e^{-R^2\cos 2\varphi}. \quad (1B.3)$$

The right-hand side goes to zero exponentially fast, except for angles φ close to $\pi/4$ where the cosine in the exponent vanishes. In the dangerous regime $\alpha \in (\pi/4-\epsilon, \pi/4)$ with small $\epsilon > 0$, one certainly has $\sin 2\varphi > \sin 2\alpha$, so that

$$R\int_{\alpha}^{\pi/4} d\varphi\, e^{-R^2\cos 2\varphi} < R\int_{\alpha}^{\pi/4} d\varphi\, \frac{\sin 2\varphi}{\sin 2\alpha}\, e^{-R^2\cos 2\varphi}. \tag{1B.4}$$

The right-hand integral can be performed by parts and yields

$$\alpha R\, e^{-R^2\cos 2\alpha} + \frac{1}{R\sin 2\alpha}\left[e^{-R^2\cos 2\varphi}\right]_{\varphi=\alpha}^{\varphi=\pi/4}, \tag{1B.5}$$

which goes to zero like $1/R$ for large R. Thus we find from (1B.2) the limiting formula $\int_\infty^0 dp\, e^{-ip^2} = -e^{-i\pi/4}\sqrt{\pi}/2$, or

$$\int_\infty^\infty dp\, e^{-ip^2} = e^{-i\pi/4}\sqrt{\pi}, \tag{1B.6}$$

which goes into Fresnel's integral formula (1.337) by substituting $p \to p\sqrt{a/2}$.

Appendix 1C The Asymmetric Top

The Lagrangian of the asymmetric top with three different moments of inertia reads

$$L = \frac{1}{2}[I_\xi {\omega_\xi}^2 + I_\eta {\omega_\eta}^2 + I_\zeta {\omega_\zeta}^2]. \tag{1C.1}$$

It has the Hessian metric [recall (1.12) and (1.388)]

$$\begin{aligned}
g_{11} &= I_\xi \sin^2\beta + I_\zeta\cos^2\beta - (I_\xi - I_\eta)\sin^2\beta\sin^2\gamma,\\
g_{21} &= -(I_\xi - I_\eta)\sin\beta\sin\gamma\cos\gamma,\\
g_{31} &= I_\zeta\cos\beta,\\
g_{22} &= I_\eta + (I_\xi - I_\eta)\sin^2\gamma,\\
g_{32} &= 0,\\
g_{33} &= I_\zeta,
\end{aligned} \tag{1C.2}$$

rather than (1.462). The determinant is

$$g = I_\xi I_\eta I_\zeta \sin^2\beta, \tag{1C.3}$$

and the inverse metric has the components

$$\begin{aligned}
g^{11} &= \frac{1}{g}\{I_\eta + (I_\xi - I_\eta)\sin^2\gamma\}I_\zeta,\\
g^{21} &= \frac{1}{g}\sin\beta\sin\gamma\cos\gamma(I_\xi - I_\eta)I_\zeta,\\
g^{31} &= \frac{1}{g}\{\cos\beta[-\sin^2\gamma(I_\xi - I_\eta) - I_\eta]\}I_\zeta,\\
g^{22} &= \frac{1}{g}\{\sin^2\beta[I_\xi - \sin^2\gamma(I_\xi - I_\eta)]\}I_\zeta,\\
g^{32} &= \frac{1}{g}\{\sin\beta\cos\beta\sin\gamma\cos\gamma(I_\eta - I_\xi)\}I_\zeta,\\
g^{33} &= \frac{1}{g}\{\sin^2\beta I_\xi I_\eta + \cos^2\beta I_\eta I_\zeta + \cos^2\beta\sin^2\gamma\,(I_\xi - I_\eta)I_\zeta\}.
\end{aligned} \tag{1C.4}$$

From this we find the components of the Riemann connection, the Christoffel symbol defined in Eq. (1.70):

$$
\begin{aligned}
\bar{\Gamma}_{11}{}^{1} &= [\cos\beta\cos\gamma\sin\gamma(I_\eta^2 - I_\eta I_\zeta - I_\xi^2 + I_\xi I_\zeta)]/I_\xi I_\eta,\\
\bar{\Gamma}_{21}{}^{1} &= \{\cos\beta[\sin^2\gamma(I_\xi^2 - I_\eta^2 - (I_\xi - I_\eta)I_\zeta)\\
&\qquad + I_\eta(I_\xi + I_\eta - I_\zeta)]\}/2\sin\beta I_\xi I_\eta,\\
\bar{\Gamma}_{31}{}^{1} &= \{\cos\gamma\sin\gamma[I_\eta^2 - I_\xi^2 + (I_\xi - I_\eta)I_\zeta]\}/2I_\xi I_\eta,\\
\bar{\Gamma}_{22}{}^{1} &= 0,\\
\bar{\Gamma}_{32}{}^{1} &= [\sin^2\gamma(I_\xi^2 - I_\eta^2 - (I_\xi - I_\eta)I_\zeta) - I_\eta(I_\xi - I_\eta + I_\zeta)]/2\sin\beta I_\xi I_\eta,\\
\bar{\Gamma}_{33}{}^{1} &= 0,\\
\bar{\Gamma}_{11}{}^{2} &= \{\cos\beta\sin\beta[\sin^2\gamma(I_\xi^2 - I_\eta^2 - I_\zeta(I_\xi - I_\eta)) - I_\xi(I_\xi - I_\zeta)]\}/I_\xi I_\eta,\\
\bar{\Gamma}_{21}{}^{2} &= \{\cos\beta\cos\gamma\sin\gamma[I_\xi^2 - I_\eta^2 - I_\zeta(I_\xi - I_\eta)]\}/2I_\xi I_\eta,\\
\bar{\Gamma}_{31}{}^{2} &= \{\sin\beta[\sin^2\gamma(I_\xi^2 - I_\eta^2 - I_\zeta(I_\xi - I_\eta)) - I_\xi(I_\xi - I_\eta - I_\zeta)]\}/2I_\xi I_\eta,\\
\bar{\Gamma}_{22}{}^{2} &= 0,\\
\bar{\Gamma}_{32}{}^{2} &= [\cos\gamma\sin\gamma(I_\xi^2 - I_\eta^2 - I_\zeta(I_\xi - I_\eta))]/2I_\xi I_\eta,\\
\bar{\Gamma}_{33}{}^{2} &= 0,\\
\bar{\Gamma}_{11}{}^{3} &= \{\cos\gamma\sin\gamma[\sin^2\beta(I_\xi I_\eta(I_\xi - I_\eta) - I_\zeta(I_\xi^2 - I_\eta^2) + I_\zeta^2(I_\xi - I_\eta))\\
&\qquad + (I_\xi^2 - I_\eta^2)I_\zeta - I_\zeta^2(I_\xi - I_\eta)]\}/I_\xi I_\eta I_\zeta,\\
\bar{\Gamma}_{21}{}^{3} &= \{\sin^2\beta[\sin^2\gamma(2I_\xi I_\eta(I_\eta - I_\xi) + I_\zeta(I_\xi^2 - I_\eta^2) - I_\zeta^2(I_\xi - I_\eta))\\
&\qquad + I_\xi I_\eta(I_\xi - I_\eta) + I_\eta I_\zeta(I_\eta - I_\zeta)] - \sin^2\gamma((I_\xi^2 - I_\eta^2)I_\zeta - I_\zeta^2(I_\xi - I_\eta))\\
&\qquad - I_\eta I_\zeta(I_\xi + I_\eta - I_\zeta)\}/2\sin\beta I_\xi I_\eta I_\zeta,\\
\bar{\Gamma}_{31}{}^{3} &= [\cos\beta\cos\gamma\sin\gamma(I_\xi^2 - I_\eta^2 - I_\zeta(I_\xi - I_\eta))]/2I_\xi I_\eta,\\
\bar{\Gamma}_{22}{}^{3} &= \cos\gamma\sin\gamma(I_\eta - I_\xi)/I_\zeta,\\
\bar{\Gamma}_{32}{}^{3} &= \{\cos\beta[\sin^2\gamma(I_\eta^2 - I_\xi^2 + (I_\xi - I_\eta)I_\zeta) + I_\eta(I_\xi - I_\eta + I_\zeta)]\}/2\sin\beta I_\eta I_\xi,\\
\bar{\Gamma}_{33}{}^{3} &= 0.
\end{aligned}
\tag{1C.5}
$$

The other components follow from the symmetry in the first two indices $\bar{\Gamma}_{\mu\nu}{}^{\lambda} = \bar{\Gamma}^{\lambda}_{\nu\mu}$. From this Christoffel symbol we calculate the Ricci tensor, to be defined in Eq. (10.41),

$$
\begin{aligned}
\bar{R}_{11} &= \{\sin^2\beta[\sin^2\gamma(I_\eta^3 - I_\xi^3 - (I_\xi I_\eta - I_\zeta^2)(I_\xi - I_\eta))\\
&\qquad + ((I_\xi + I_\zeta)^2 - I_\eta^2)(I_\xi - I_\zeta)] + I_\zeta^3 - I_\zeta(I_\xi - I_\eta)^2\}/2I_\xi I_\eta I_\zeta,\\
\bar{R}_{21} &= \{\sin\beta\sin\gamma\cos\gamma[I_\eta^3 - I_\xi^3 + (I_\xi I_\eta - I_\zeta^2)(I_\eta - I_\xi)]\}/2I_\xi I_\eta I_\zeta,\\
\bar{R}_{31} &= -\{\cos\beta[(I_\xi - I_\eta)^2 - I_\zeta^2]\}/2I_\xi I_\eta,\\
\bar{R}_{22} &= \{\sin^2\gamma[I_\xi^3 - I_\eta^3 + (I_\xi I_\eta - I_\zeta^2)(I_\xi - I_\eta)] + I_\eta^3 - (I_\xi - I_\zeta)^2 I_\eta\}/2I_\xi I_\eta I_\zeta,\\
\bar{R}_{32} &= 0,\\
\bar{R}_{33} &= -[(I_\xi - I_\eta)^2 - I_\zeta^2]/2I_\xi I_\eta.
\end{aligned}
\tag{1C.6}
$$

Contraction with $g^{\mu\nu}$ gives the curvature scalar

$$\bar{R} = [2(I_\xi I_\eta + I_\eta I_\zeta + I_\zeta I_\xi) - I_\xi^2 - I_\eta^2 - I_\zeta^2]/2I_\xi I_\eta I_\zeta. \tag{1C.7}$$

Since the space under consideration is free of torsion, the Christoffel symbol $\bar{\Gamma}_{\mu\nu}{}^{\lambda}$ is equal to the full affine connection $\Gamma_{\mu\nu}{}^{\lambda}$. The same thing is true for the curvature scalars $\bar{R}$ and R calculated from $\bar{\Gamma}_{\mu\nu}{}^{\lambda}$ and $\Gamma_{\mu\nu}{}^{\lambda}$, respectively.

Notes and References

For more details see some standard textbooks:
I. Newton, *Mathematische Prinzipien der Naturlehre*, Wiss. Buchgesellschaft, Darmstadt, 1963;
J.L. Lagrange, *Analytische Mechanik*, Springer, Berlin, 1887;
G. Hamel, *Theoretische Mechanik*, Springer, Berlin, 1949;
A. Sommerfeld, *Mechanik*, Harri Deutsch, Frankfurt, 1977;
W. Weizel, *Lehrbuch der Theoretischen Physik*, Springer, Berlin, 1963;
H. Goldstein, *Classical Mechanics*, Addison-Wesley, Reading, 1950;
L.D. Landau and E.M. Lifshitz, *Mechanics*, Pergamon, London, 1965;
R. Abraham and J.E. Marsden, *Foundations of Mechanics*, Benjamin, New York, 1967;
C.L. Siegel and J.K. Moser, *Lectures on Celestial Mechanics*, Springer, Berlin, 1971;
P.A.M. Dirac, *The Principles of Quantum Mechanics*, Clarendon, Oxford, 1958;
L.D. Landau and E.M. Lifshitz, *Quantum Mechanics*, Pergamon, London, 1965;
A. Messiah, *Quantum Mechanics*, Vols. I and II, North-Holland , Amsterdam, 1961;
L.I. Schiff, *Quantum Mechanics*, 3rd ed., McGraw-Hill, New York, 1968;
E. Merzbacher, *Quantum Mechanics*, 2nd ed, Wiley, New York, 1970;
L.D. Landau and E.M. Lifshitz, *Statistical Physics*, Pergamon, London, 1958;
E.M. Lifshitz and L.P. Pitaevski, *Statistical Physics*, Vol. 2, Pergamon, London, 1987.

The particular citations in this chapter refer to the publications

[1] For an elementary introduction see the book
H.B. Callen, *Classical Thermodynamics*, John Wiley and Sons, New York, 1960. More details are also found later in Eqs. (4.56) and (4.57).

[2] The integrability conditions are named after the mathematician of complex analysis H.A. Schwarz, a student of K. Weierstrass, who taught at the Humboldt-University of Berlin from 1892–1921.

[3] L. Schwartz, *Théorie des distributions*, Vols.I-II, Hermann & Cie, Paris, 1950-51;
I.M. Gelfand and G.E. Shilov, *Generalized functions*, Vols.I-II, Academic Press, New York-London, 1964-68.

[4] An exception occurs in the theory of Bose-Einstein condensation where the single state $\mathbf{p} = 0$ requires a separate treatment since it collects a large number of particles in what is called a *Bose-Einstein condensate*. See p. 169 in the above-cited textbook by L.D. Landau and E.M. Lifshitz on *Statistical Mechanics*. Bose-Einstein condensation will be discussed in Sections 7.2.1 and 7.2.4.

[5] This was first observed by
B. Podolsky, Phys. Rev. **32**, 812 (1928).

[6] B.S. DeWitt, Rev. Mod. Phys. **29**, 377 (1957);
K.S. Cheng, J.Math. Phys. **13**, 1723 (1972);
H. Kamo and T. Kawai, Prog. Theor. Phys. **50**, 680, (1973);
T. Kawai, Found. Phys. **5**, 143 (1975);
H. Dekker, Physica A **103**, 586 (1980);
G.M. Gavazzi, Nuovo Cimento **101**A, 241 (1981).
See also the alternative approach by
N.M.J. Woodhouse, *Geometric Quantization*, Oxford University Press, Oxford, 1992.

[7] C. van Winter, Physica **20**, 274 (1954).

[8] For detailed properties of the representation matrices of the rotation group, see
A.R. Edmonds, *Angular Momentum in Quantum Mechanics*, Princeton University Press, Princeton, 1960.

[9] L.D. Landau and E.M. Lifshitz, *Quantum Mechanics*, Pergamon, London, 1965.