14

Unified kinematics and dynamics

This chapter sews together some of the threads from the earlier chapters to show the relationships between apparently disparate dynamical descriptions of physics.

14.1 Classical Hamiltonian particle dynamics

The traditional formulation Schrödinger quantum mechanics builds on a Hamiltonian formulation of dynamical systems, in which the dynamics describe not only particle coordinates q but also their momenta p. The interesting feature of the Hamiltonian formulation, in classical mechanics, is that one deals only with quantities which have a direct physical interpretation. The disadvantage with the Hamiltonian approach in field theory is its lack of manifest covariance under Lorentz transformations: time is singled out explicitly in the formulation.[1] Some important features of the Hamiltonian formulation are summarized here in order to provide an alternative view to dynamics with some different insights.

The Hamiltonian formulation begins with the definition of the momentum p_i conjugate to the particle coordinate q_i. This quantity is introduced since it is expected to have a particular physical importance. Ironically, one begins with the Lagrangian, which is unphysical and is to be eliminated from the discussion. The Lagrangian is generally considered to be a function of the particle coordinates q_i and their time derivatives or velocities $\dot{q}^i$. The momentum is then conveniently defined from the Lagrangian,

$$p_i = \frac{\partial L}{\partial \dot{q}_i}. \tag{14.1}$$

---

[1] Actually, time is singled out in a special way even in the fully covariant Lagrangian formulation, since time plays a fundamentally different role from space as far as the dynamics are concerned. The main objection which is often raised against the Hamiltonian formulation is the fact that the derivation of covariant results is somewhat clumsy.

This is not the only way in which one could define a momentum, but it is convenient to use a definition which refers only to the abstract quantities L and $\dot{q}_i$ in cases where the Lagrangian and its basic variables are known, but other physical quantities are harder to identify. This extends the use of the formalism to encompass objects which one would not normally think of as positions and momenta. The total time derivative of the Lagrangian is

$$\frac{\mathrm{d}L(q, \dot{q}, t)}{\mathrm{d}t} = \frac{\partial L}{\partial q_i}\dot{q}_i + \frac{\partial L}{\partial \dot{q}_i}\ddot{q} + \frac{\partial L}{\partial t}, \tag{14.2}$$

which may be written

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\{\dot{q}_i\frac{\partial L}{\partial \dot{q}_i} - L\right\}. \tag{14.3}$$

Now, if the Lagrangian is not explicitly time-dependent, $\frac{\partial L}{\partial t} = 0$, then the quantity in the curly braces must be constant with respect to time, so, using eqn. (14.1), we may define the Hamiltonian H by

$$H = \text{const.} = p\dot{q} - L. \tag{14.4}$$

Notice that this definition involves time derivatives. When we consider the relativistic case, timelike quantities are often accompanied by a sign from the metric tensor, so the form of the Hamiltonian above should not be treated as sacred.

14.1.1 Hamilton's equations of motion

The equations of motion in terms of the new variables may be obtained in the usual way from the action principle, but now treating q_i and p_i as independent variables. Using the Lagrangian directly to obtain the action gives us

$$S = \int \mathrm{d}t \,\{p\dot{q} - L\}. \tag{14.5}$$

However, from earlier discussions about symmetrical derivatives, we know that the correct action is symmetrized about the derivatives. Thus, the action is given by

$$S = \int \mathrm{d}t \,\left\{\frac{1}{2}(p\dot{q} - q\dot{p}) - L\right\}. \tag{14.6}$$

Varying this action with fixed end-points, one obtains (integrating the $p\dot{q}$ term by parts)

$$\frac{\delta S}{\delta q(t)} = -\dot{p} - \frac{\partial H}{\partial q} = 0$$
$$\frac{\delta S}{\delta p(t)} = \dot{q} - \frac{\partial H}{\partial p}. \tag{14.7}$$

Hence, Hamilton's two equations of motion result:

$$\dot{p} = -\frac{\partial H}{\partial q} \tag{14.8}$$

$$\dot{q} = \frac{\partial H}{\partial p}. \tag{14.9}$$

Notice that this is a pair of equations. This is a result of our insistence on introducing an extra variable (the momentum) into the formulation.

14.1.2 Symmetry and conservation

One nice feature of the Hamiltonian formulation is that invariances of the equations of motion are all identifiable as a generalized translational invariance. If the action is independent of a given coordinate

$$\frac{\partial L}{\partial q_n} = 0, \tag{14.10}$$

then

$$\frac{\partial L}{\partial \dot{q}_n} = p_n = \text{const.}; \tag{14.11}$$

i.e. the momentum associated with that coordinate is constant, or is *conserved*. The coordinate q_n is then called an *ignorable* coordinate.

14.1.3 Symplectic transformations

We started originally with an action principle, which treated only $q(t)$ as a dynamical variable, and later introduced (artificially) the independent momentum variable p.[2] The fact that we now have twice the number of dynamical variables seems unnecessary. This intuition is further borne out by the observation that, if we make the substitution

$$q \to -p \tag{14.12}$$

$$p \to q \tag{14.13}$$

in eqn. (14.9), then we end up with an identical set of equations, with only the roles of the two equations switched. This transformation represents an

---

[2] In many textbooks, the Lagrangian formulation is presented as a function of coordinates q and velocities $\dot{q}$. Here we have bypassed this discussion by working directly with variations of the action, where it is possible to integrate by parts and perform functional variations. This makes the usual classical Lagrangian formalism redundant.

invariance of the Hamilton equations of motion, and hints that the positions and momenta are really just two sides of the same coin.

Based on the above, one is motivated to look for a more general linear transformation in which p and q are interchanged. In doing so, one must be a little cautious, since positions and momenta clearly have different engineering dimensions. Let us therefore introduce the quantities $\hat{p}$ and $\hat{q}$, which are re-scaled by a constant Ω with dimensions of mass per unit time in order that they have the same dimensions:

$$\hat{p} = p/\sqrt{\Omega}$$
$$\hat{q} = q\sqrt{\Omega}. \tag{14.14}$$

The product of $\hat{q}$ and $\hat{p}$ is independent of this scale, and this implies that the form of the equations of motion is unchanged:

$$\dot{\hat{p}} = -\frac{\partial H}{\partial \hat{q}} \tag{14.15a}$$

$$\dot{\hat{q}} = \frac{\partial H}{\partial \hat{p}}. \tag{14.15b}$$

Let us consider, then, general linear combinations of q and p and look for all those combinations which leave the equations of motion invariant. In matrix form, we may write such a transformation as

$$\begin{pmatrix} \hat{q}' \\ \hat{p}' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \hat{q} \\ \hat{p} \end{pmatrix}. \tag{14.16}$$

The derivatives associated with the new coordinates are

$$\frac{\partial}{\partial \hat{q}'} = \frac{1}{2}\left(\frac{1}{a}\frac{\partial}{\partial \hat{q}} + \frac{1}{b}\frac{\partial}{\partial \hat{p}}\right)$$
$$\frac{\partial}{\partial \hat{p}'} = \frac{1}{2}\left(\frac{1}{c}\frac{\partial}{\partial \hat{q}} + \frac{1}{d}\frac{\partial}{\partial \hat{p}}\right). \tag{14.17}$$

We may now substitute these transformed coordinates into the Hamilton equations of motion (14.15) and determine the values of a, b, c, d for which the equations of motion are preserved. From eqn. (14.15b), one obtains

$$a\dot{\hat{q}} + b\dot{\hat{p}} = \frac{1}{2}\left(\frac{1}{c}\frac{\partial H}{\partial q} + \frac{1}{d}\frac{\partial H}{\partial p}\right). \tag{14.18}$$

This equation is a linear combination of the original equation in (14.15) provided that we identify

$$2ad = 1$$
$$2bc = -1. \tag{14.19}$$

Substitution into eqn. (14.15a) confirms this. The determinant of the transformation matrix is therefore

$$ad - bc = 1 \tag{14.20}$$

and we may write, in full generality:

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathrm{e}^{\mathrm{i}\theta} & \mathrm{i}\mathrm{e}^{\mathrm{i}\phi} \\ \mathrm{i}\mathrm{e}^{-\mathrm{i}\phi} & \mathrm{e}^{-\mathrm{i}\theta.} \end{pmatrix} \tag{14.21}$$

This is the most general transformation of $\hat{p}$, $\hat{q}$ pairs which leaves the equations of motion invariant. The set of transformations, which leaves the Poisson bracket invariant, forms a group known as the symplectic group $sp(2, C)$. If we generalize the above discussion by adding indices to the coordinates and momenta $i = 1, \ldots, n$, then the group becomes $sp(2n, C)$.

Since we have now shown that p and q play virtually identical roles in the dynamical equations, there is no longer any real need to distinguish them with separate symbols. In *symplectic notation*, many authors write both coordinates and momenta as Q_i, where $i = 1, \ldots, 2n$ grouping both together as *generalized coordinates*.

14.1.4 Poisson brackets

Having identified a symmetry group of the equations of motion which is general (i.e. which follows entirely from the definition of the conjugate momentum in terms of the Lagrangian), the next step is to ask which quantities are invariant under this symmetry group. A quantity of particular interest is the so-called *Poisson bracket*.

If we apply the group transformation to the derivative operators,

$$\begin{pmatrix} \hat{D}_+ \\ \hat{D}_- \end{pmatrix} \equiv U(\theta, \phi) \begin{pmatrix} \frac{\partial}{\partial \hat{q}} \\ \frac{\partial}{\partial \hat{p}} \end{pmatrix}, \tag{14.22}$$

then it is a straightforward algebraic matter to show that, for any two functions of the dynamical variables A, B, the Poisson bracket, defined by

$$(D_+ X)(D_- Y) - (D_- X)(D_+ Y) \equiv [X, Y]_{\hat{p}\hat{q}}, \tag{14.23}$$

is independent of θ and ϕ and is given in all bases by

$$[X, Y]_{pq} = \frac{\partial X}{\partial q} \frac{\partial Y}{\partial p} - \frac{\partial Y}{\partial q} \frac{\partial X}{\partial p}. \tag{14.24}$$

Notice, in particular that, owing to the product of pq in the denominators, this bracket is even independent of the re-scaling by Ω in eqn. (14.14).

We shall return to the Poisson bracket to demonstrate its importance to the variational formalism and dynamics after a closer look at symmetry transformations.

14.1.5 General canonical transformations

The linear combinations of p, q described in the previous section form a symmetry which has its origins in the linear formulation of the Hamiltonian method. Symplectic symmetry is not the only symmetry which might leave the equations of motion invariant, however. More generally, we might expect the coordinates and momenta to be changed into quite different functions of the dynamical variables:

$$q \to q'(p, q, t)$$
$$p \to p'(p, q, t). \qquad (14.25)$$

Changes of variable fit this general description, as does the time development of p and q. We might, for example, wish to change from a Cartesian description of physics to a polar coordinate basis, which better reflects the symmetries of the problem. Any such change which preserves the form of the field equations is called a canonical transformation.

It turns out that one can effect general infinitesimal transformations of coordinates by simply adding total derivatives to the Lagrangian. This is closely related to the discussion of continuity in section 4.1.4. Consider the following addition

$$L \to L + \frac{\mathrm{d}F}{\mathrm{d}t}, \qquad (14.26)$$

for some arbitrary function $F(q, p, t)$. Normally, one ignores total derivatives in the Lagrangian, for the reasons mentioned in section 4.4.2. This is because the action is varied, with the end-points of the variation fixed. However, if one relaxes this requirement and allows the end-points to vary about *dynamical variables which obey the equations of motion*, then these total derivatives (often referred to as surface terms in field theory), have a special and profound significance. Our programme and its notation are the following.

- We add the total time derivative of a function $F(q, p, t)$ to the Lagrangian so that

$$S \to S + \int \mathrm{d}t \, \frac{\mathrm{d}F}{\mathrm{d}t}$$
$$= S + F \Big|_{t_1}^{t_2}. \qquad (14.27)$$

- We vary the action and the additional term and define the quantity G_ξ, which will play a central role in transformation theory, by

$$\delta_\xi F \equiv G_\xi, \qquad (14.28)$$

  so that

$$\delta S \to \delta S + G. \qquad (14.29)$$

- We may optionally absorb the change in the variation of the action G into the generalized coordinates by making a transformation, which we write formally as

$$q \to q + \delta q = q + R_\xi \, \delta \xi. \qquad (14.30)$$

R_ξ is called the auxiliary function of the transformation, and is related to G_ξ, which is called the generator of the transformation. This transforms the coordinates and momenta by an infinitesimal amount.

Let us now illustrate the procedure in practice. The variation of the action may be written

$$\delta S = \delta \int \mathrm{d}t \, (p\dot{q} - H) + \int \mathrm{d}t \, \delta \dot{F},$$

$$= \int \mathrm{d}t \left(\left(-\dot{p} - \frac{\partial H}{\partial q} \right) \delta q + \frac{\partial \dot{G}}{\partial q} \delta q + \frac{\partial \dot{G}}{\partial q} \delta q \right). \qquad (14.31)$$

In the last line we have expanded the infinitesimal change $\delta F = G$ in terms of its components along q and p. This can always be done, regardless of what general change G represents. We can now invoke the modified action principle and obtain the equations of motion:

$$\frac{\delta S}{\delta q} = 0 = -\dot{p} - \frac{\delta H}{\delta q} + \frac{\delta \dot{G}}{\delta q} = -(\dot{p} + \delta\dot{p}) - \frac{\delta H}{\delta q}$$

$$\frac{\delta S}{\delta p} = 0 = \dot{q} - \frac{\delta H}{\delta p} + \frac{\delta \dot{G}}{\delta p} = (\dot{q} + \delta\dot{q}) - \frac{\delta H}{\delta p}, \qquad (14.32)$$

where we have identified

$$\delta\dot{p} = -\frac{\partial \dot{G}}{\partial q}$$

$$\delta\dot{q} = \frac{\partial \dot{G}}{\partial p}, \qquad (14.33)$$

or, on integrating,

$$\delta p = -\frac{\partial G}{\partial q} = R_a^p \delta\xi^a$$

$$\delta q = \frac{\partial G}{\partial p} = R_a^q \delta\xi^a. \qquad (14.34)$$

Notice that G is infinitesimal, by definition, so we may always write it in terms of a set of infinitesimal parameters $\delta\xi$, but ξ need not include q, p now since the q, p dependence was already removed to infinitesimal order in eqn. (14.31).[3] It is now possible to see why G is referred to as the generator of infinitesimal canonical transformations.

---

[3] The expressions are not incorrect for p, q variations, but they become trivial cases.

14.1.6 Variations of dynamical variables and Poisson brackets

One of the most important observations about variational dynamics, as far as the extension to quantum field theory is concerned, is that variational changes in any dynamical variable can be expressed in terms of the invariant Poisson bracket between that variable and the generator of the variation:

$$\delta X(p, q, t) = [X, G_\xi]_{pq} \tag{14.35}$$

To see this, it is sufficient to use eqns. (14.34) in the differential expansion of the function:

$$\delta X = \frac{\partial X}{\partial q_i} \delta q_i + \frac{\partial X}{\partial p_i} \delta p_i. \tag{14.36}$$

Substituting for δq_i and δp_i gives eqn. (14.35). These relations are exemplified as follows.

- Generator of time translations: $G_t = -H \delta t$;

$$\delta X = [X, H] \delta t. \tag{14.37}$$

Noting that the change in X is the dynamical evolution of the function, but that the numerical value of t is unaltered owing to linearity and the infinitesimal nature of the change, we have that

$$\delta X = -\left(\frac{\mathrm{d}X}{\mathrm{d}t} - \frac{\partial X}{\partial t}\right) \mathrm{d}t = [X, H] \delta t. \tag{14.38}$$

Thus, we arrive at the equation of motion for the dynamical variable X:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = [X, H] + \frac{\partial X}{\partial t}. \tag{14.39}$$

This result has the following corollaries:

$$\begin{aligned} \dot{q} &= [q, H] \\ \dot{p} &= [p, H] \\ 1 &= [H, t]. \end{aligned} \tag{14.40}$$

The first two equations are simply a thinly concealed version of the Hamilton equations (14.9). The third, which is most easily obtained from eqn. (14.37), is an expression of the time independence of the Hamiltonian. An object which commutes with the Hamiltonian is conserved.

- Generator of coordinate translations: $G_q = p\delta q$.

  It is interesting to note that, if we consider the variation in the coordinate q with respect to the generator for q itself, the result is an identity which summarizes the completeness of our dynamical system:

$$\delta q = [q, G_q]_{pq}$$
$$\delta q = [q, p]_{pq}\delta q$$
$$\Rightarrow 1 = [q, p]_{pq}. \tag{14.41}$$

  In Lorentz-covariant notation, one may write

$$[x^\mu, p_\nu] = \delta^\mu_\nu, \tag{14.42}$$

  where $p_\mu = (-H/c, \mathbf{p})$. This result pervades almost all of dynamics arising from Lagrangian/Hamiltonian systems. In the quantum theory it is supplanted by commutation relations, which have the same significance as the Poisson bracket, though they are not directly related.

14.1.7 Derivation of generators from the action

Starting from the correctly symmetrized action in eqn. (14.6), the generator of infinitesimal canonical transformations for a variable ξ is obtained from the surface contribution to the variation, with respect to ξ.[4] For example,

$$\delta_q S = \int dt \left(-\dot{p} - \frac{\partial H}{\partial q} \right) \delta q + \frac{1}{2} p\delta q$$
$$= 0 + \frac{1}{2} G_q, \tag{14.43}$$

where we have used the field equation to set the value of the integral in the first line to zero, and we identify

$$G_q = p\delta q. \tag{14.44}$$

Similarly,

$$\delta_p S = \int dt \left(\dot{q} - \frac{\partial H}{\partial p} \right) \delta p - \frac{1}{2} q\delta p$$
$$= \frac{1}{2} G_p, \tag{14.45}$$

---

[4] The constants of proportionality are rather inconsistent in this Hamiltonian formulation. If one begins with the action defined in terms of the Lagrangian, the general rule is: for actions which are linear in the time derivatives, the surface contribution is one-half the corresponding generator; for actions which are quadratic in the time derivatives, the generator is all of the surface contribution.

hence

$$G_p = -q\delta p. \tag{14.46}$$

For time variations,

$$\delta_t S = -H\delta t$$
$$= G_t. \tag{14.47}$$

The generators are identified, with numerical values determined by convention. The factors of one-half are introduced here in order to eliminate irrelevant constants from the Poisson bracket. This is of no consequence, as mentioned in the next section; it is mainly for aesthetic purposes.

Suppose we write the action in the form

$$S = \int \{p\mathrm{d}q - H\mathrm{d}t\}, \tag{14.48}$$

where we have cancelled an infinitesimal time differential in the first term. It is now straightforward to see that

$$\frac{\partial S}{\partial t} + H = 0. \tag{14.49}$$

This is the Hamilton–Jacobi equation of classical mechanics. From the action principle, one may see that this results from boundary activity, by applying a general boundary disturbance F:

$$S \to S + \int (\mathrm{d}x)\, \partial_\mu F. \tag{14.50}$$

$\delta F = G$ is the generator of infinitesimal canonical transformations, and

$$\frac{\partial G}{\partial q} = \int \mathrm{d}\sigma^\mu\, R^a_\mu \delta\xi_a. \tag{14.51}$$

Notice from eqn. (11.43) that

$$\int \mathrm{d}\sigma^\mu G_\mu = \int \mathrm{d}\sigma^\mu (\Pi_\mu \delta q - \theta_{\mu\nu}\delta x^\nu), \tag{14.52}$$

which is to be compared with

$$\delta S = p\delta q - H\mathrm{d}t. \tag{14.53}$$

Moreover, from this we have the Hamilton–Jacobi equation (see eqn. (11.78)),

$$\frac{\delta S}{\delta x^0} = -\frac{1}{c} \int \mathrm{d}\sigma^\mu\, \theta_{\mu 0} = -\frac{H}{c} \tag{14.54}$$

or

$$\frac{\delta S}{\delta t} + H = 0. \tag{14.55}$$

14.1.8 Conjugate variables and dynamical completeness

The commutator functions we have evaluated above are summarized by

$$[q_A, p_B]_{pq} = \delta_{AB}$$
$$[t, H]_{pq} = 1. \tag{14.56}$$

These equations are a formal expression of the completeness of the set of variables we have chosen to parametrize the dynamical equations. Not every variational equation necessarily has coordinates and momenta, but every set of conservative dynamical equations has pairs of conjugate variables which play the roles of p and q. If one is in possession of a complete set of such variables (i.e. a set which spans all of phase space), then an arbitrary state of the dynamical system can be represented in terms of those variables, and it can be characterized as being canonical.

Ignorable coordinates imply that the dimension of phase space is effectively reduced, so there is no contradiction in the presence of symmetries.

Given the definition of the Poisson bracket in eqn. (14.24), the value of $[q, p]_{pq} = 1$ is unique. But we could easily have defined the derivative differently up to a constant, so that we had obtained

$$[q_A, p_B]'_{pq} = \alpha \, \delta_{AB}. \tag{14.57}$$

What is important is not the value of the right hand side of this expression, but the fact that it is constant for all conjugate pairs. In any closed, conservative system, the Hamiltonian time relation is also constant, but again the normalization could easily be altered by an arbitrary constant. These are features which are basic to the geometry of phase space, and they carry over to the quantum theory for commutators. There it is natural to choose a different value for the constant and a different but equivalent definition of completeness.

14.1.9 The Jacobi identity and group algebra

The anti-symmetrical property of the Poisson bracket alone is responsible for the canonical group structure which it generates and the completeness argument above. This may be seen from an algebraic identity known as the Jacobi identity. Suppose that we use the bracket $[A, B]$ to represent *any* object which has the property

$$[A, B] = -[B, A]. \tag{14.58}$$

The Poisson bracket and the commutator both have this property. It may be seen, by writing out the combinations explicitly, that

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0. \tag{14.59}$$

This result does not depend on whether A, B, C commute. This equation is known as the Jacobi identity. It is closely related to the Bianchi identity in eqn. (2.27).

Any objects which satisfy this identity also satisfy a Lie algebra. This is easily seen if we identify a symbol

$$T_A(B) \equiv [A, B]. \tag{14.60}$$

Then, re-writing eqn. (14.59) so that all the C elements are to the right,

$$[A, [B, C]] - [B, [A, C]] - [[A, B], C] = 0, \tag{14.61}$$

we have

$$T_A T_B(C) - T_B T_A(C) = T_{[A,B]}(C), \tag{14.62}$$

or

$$[T_A, T_B] = T_{[A,B]}(C). \tag{14.63}$$

14.2 Classical Lagrangian field dynamics

14.2.1 Spacetime continuum

In the traditional classical mechanics, one parametrizes a system by the coordinates and momenta of pointlike particles. In order to discuss continuous matter or elementary systems, we move to field theory and allow a smooth dependence on the **x** coordinate.

In field theory, one no longer speaks of discrete objects with positions or trajectories (world-lines) $q(t)$ or $x(\tau)$. Rather x, t take on the roles of a ruler or measuring rod, which is positioned and oriented by the elements of the Galilean or Lorentz symmetry groups. Schwinger expresses this by saying that space and time play the role of an abstract measurement apparatus [119], which means that **x** is no longer $q(t)$, the position of an existing particle. It is simply a point along some ruler, or coordinate system, which is used to measure position. The position might be occupied by a particle or by something else; then again, it might not be.

14.2.2 Poisson brackets of fields

The Poisson bracket is not really usable in field theory, but it is instructive to examine its definition as an invariant object. We begin with the relativistic scalar field as the prototype.

The Poisson bracket of two functions X and Y may be written in one of two ways. Since the dynamical variables in continuum field theory are now $\phi_A(x)$

and $\Pi_A(x)$, one obvious definition is the direct transcription of the classical particle result for the canonical field variables, with only an additional integral over all space owing to the functional dependence on **x**. By analogy with Poisson brackets in particle mechanics, the bracketed quantities are evaluated at equal times.

$$[X, Y]_{\phi\Pi} = \int d\sigma_x \left(\frac{\partial X}{\partial \phi_A(x)} \frac{\partial Y}{\partial \Pi_A(x)} - \frac{\partial Y}{\partial \phi_A(x)} \frac{\partial X}{\partial \Pi_A(x)} \right). \quad (14.64)$$

With this definition we have

$$[\phi(x), \Pi(x')]_{\phi\Pi}\Big|_{t=t'} = \delta(\mathbf{x}, \mathbf{x}')$$

$$\int d\sigma \, [\phi(x), \Pi(x')]_{\phi\Pi}\Big|_{t=t'} = 1; \quad (14.65)$$

thus, the familiar structure is reproduced. It should be noted, however, that the interpretation of these results is totally different to that for classical particle mechanics. Classically, $q_A(t)$ is the position of the Ath particle as a function of time. $\phi_A(x)$ on the other hand refers to the Ath species of scalar field (representing some unknown particle symmetry, or different discrete states, but there is no inference about localized particles at a definite position and particular time). To think of $\phi(x), \Pi(x)$ as an infinite-dimensional phase space (independent variables at every new value of **x**) is not a directly useful concept. The above form conceals a number of additional subtleties, which are best resolved by abandoning the Hamiltonian variables in favour of a pure description in terms of the field and its Green functions.

It is now possible to define the Poisson bracket using the fields and Green functions, ignoring the Hamiltonian idea of conjugate momentum. In this language, we may write the invariant Poisson bracket in terms of a directional functional derivative, for any two functions X and Y.

$$[X, Y]_\phi \equiv \mathcal{D}_X Y - \mathcal{D}_Y X, \quad (14.66)$$

where

$$\mathcal{D}_X Y \equiv \int (dx) \frac{\delta Y}{\delta A(x)} \lim_{\xi \to 0} \delta_X \phi^A(x), \quad (14.67)$$

and

$$\delta_X \phi_A(x) = \int (dx') G^r_{AB}(x, x') \delta X_B(x'). \quad (14.68)$$

Since we are looking at Lorentz-invariant quantities, there are several possible choices of causal boundary conditions, and we must define the causal nature of the variations. The natural approach is to use a retarded variation by introducing

the retarded Green function explicitly in order to connect the source δX_B to the response $\delta_X \phi$. In terms of the small parameter ξ, we may write $\delta X_B = X_{,B} \xi$, or

$$\delta_X \phi_A(x) = \int (dx') G^r_{AB}(x, x') X_{,B} \xi. \tag{14.69}$$

Using this in eqn. (14.66), we obtain, in condensed notation,

$$[X, Y]_\phi = X_{,A} \, G^{AB}_r \, Y_{,B} - Y_{,A} \, G^{AB}_r \, X_{,B}, \tag{14.70}$$

or, in uncondensed notation,

$$[X, Y]_\phi = \int (dx)(dx') \left(\frac{\delta X}{\delta A(x)} \, G^{AB}_r(x, x') \, \frac{\delta Y}{\delta B(x')} \right.$$
$$\left. - \frac{\delta Y}{\delta A(x)} \, G^{AB}_r(x', x) \, \frac{\delta X}{\delta B(x')} \right). \tag{14.71}$$

Now, using eqns. (5.74) and (5.71), we note that

$$2 \tilde{G}_{AB}(x, x') = G^{AB}_r(x, x') - G^{BA}_r(x', x), \tag{14.72}$$

so, re-labelling indices in the second term of eqn. (14.71), we have (condensed)

$$[X, Y]_\phi = 2 Y_{,A} \tilde{G}^{AB} X_{,B} \tag{14.73}$$

or (uncondensed)

$$[X, Y]_\phi = 2 \int (dx)(dx') \frac{\delta Y}{\delta \phi_A(x)}(x) \tilde{G}^{AB}(x, x') \frac{\delta X}{\delta \phi_B(x')}. \tag{14.74}$$

The connection between this expression and the operational definition in terms of Hamiltonian variables in eqn. (14.64) is not entirely obvious from this expression, but we hand-wave the reasonableness of the new expression by stretching formalism. From eqn. (5.73), one can write formally

$$\tilde{G}_{AB}(x, x') \bigg|_{t=t'} = \delta_{AB} \delta(\mathbf{x}, \mathbf{x}') \frac{1}{\partial_0} \tag{14.75}$$

and thus, hand-wavingly, at equal times,

$$\frac{\delta}{\delta \phi_A} \tilde{G}_{AB} \frac{\delta}{\delta \phi_A} \sim \frac{\delta}{\delta \phi_A} \frac{\delta}{\delta (\partial_0 \phi_A)} \sim \frac{\delta}{\delta \phi_A} \frac{\delta}{\delta \Pi_A}. \tag{14.76}$$

Although we have diverged from a covariant expression in eqn. (14.74) by singling out a spacelike hyper-surface in eqn. (14.76), this occurs in a natural way as a result of the retarded boundary conditions implicit in the causal

variations. Manifest covariance of notation cannot alter the fact that time plays a special role in dynamical systems. Clearly, one has

$$
[\phi(x), \Pi(x')]_\phi \Big|_{t=t'} = \int (\mathrm{d}y)(\mathrm{d}y') \, \frac{\delta \phi(x)}{\delta \phi_A(y)} \tilde{G}_{AB}(y, y') \frac{\delta (\partial_0 \phi(x'))}{\delta \phi_B(y')}
$$
$$
= \int (\mathrm{d}y)(\mathrm{d}y') \, \frac{\delta \phi(x)}{\delta \phi_A(y)} - \tilde{\partial}_0 G_{AB}(y, y') \frac{\delta \phi(x')}{\delta \phi_B(y')}
$$
$$
= \delta(\mathbf{x}, \mathbf{x}). \tag{14.77}
$$

The Poisson bracket is only unique if the variables are observable, i.e. if they are invariant quantities.

14.3 Classical statistical mechanics

Statistical mechanics provides a natural point of departure from particle mechanics. Although tethered to classical particle notions in the form of canonical Hamiltonian relations, it seeks to take the limit $N \to \infty$ of infinite numbers of discrete particles. It thereby moves towards a continuum representation of matter, which is a step towards field theory. To understand field theory fully, it is necessary to acknowledge a few of its roots in statistical mechanics. By definition, statistical mechanics is about many-particle systems.

14.3.1 Ensembles and ergodicity

An ensemble is formally a collection of 'identical' systems. The systems in an ensemble are identical in the sense that they contain the same dynamical variables and properties, not in the sense that each system is an exact image of every other with identical values for all its variables (that would be a useless concept). The concept of ensembles is useful for discussing the random or (more correctly) unpredictable development of systems under sufficiently similar conditions. An ensemble is a model for the possible ways in which one system might develop, taking into account a random or unpredictable element. If one takes a snapshot of systems in an ensemble at any time, the outcome could have happened in any of the systems, and may indeed happen in the future in any or all of them if they were allowed to run for a sufficient period of time. Ensembles are used to discuss the process of averaging over possible outcomes.

The ergodic hypothesis claims that the time average of a system is the same as an ensemble average in the limit of large times and large numbers of ensembles. In the limit of infinite time and ensembles, this hypothesis can be proven. The implication is that it does not matter how we choose to define the average properties of a complex (statistical) system, the same results will always be obtained. The ergodic hypothesis is therefore compatible with the continuum hypothesis, but can be expected to fail when one deals with measurably finite

times or countably finite ensembles. Much of statistical mechanics and much of quantum theory assumes the truth of this hypothesis.

14.3.2 Expectation values and correlations

Macroscopic observables are the expectation values of variables, averaged over time, or over many similar particle systems (an ensemble). The expectation value of a dynamical variable $X(q, p)$ is defined by the ensemble average. For N particles in a fixed volume V, one has

$$\langle X \rangle_{pq} = \overline{X}(t) = \frac{\int d^N q \, d^N p \; \rho(q, p, t) \, X(q, p, t)}{\int d^N q \, d^N p \; \rho(q, p, t)}, \qquad (14.78)$$

where ρ is the density of states in phase space in the fixed volume V. This is sometimes written

$$\langle X \rangle_{pq} = \mathrm{Tr}(\rho \, X) \qquad (14.79)$$

The integral in eqn. (14.78) is interpreted as an ensemble average because it integrates over every possible position and momentum for the particles. All possible outcomes are taken into account because the integral averages over all possible outcomes for the system, which is like averaging over a number of systems in which one (by the rules of chance) expects every possibility to occur randomly.

Suppose one defines the generating or partition functional $Z_{pq}[J(t)]$ by

$$Z_{pq}[J(t)] = \int d^N q \, d^N p \; \rho(q, p, t) \, e^{-\int J_X \, X \, dt'}, \qquad (14.80)$$

and the 'transformation function' by

$$W_{pq}[J(t)] = -\ln Z_{pq}[X(t)], \qquad (14.81)$$

then the average value of X can be expressed as a functional derivative in the following way:

$$\langle X(t) \rangle = -\frac{\delta W[J(t)]}{\delta J(t)}. \qquad (14.82)$$

Similarly, the correlation function is

$$\langle X(t) X(t') \rangle = \frac{\delta^2 W[J(t)]}{\delta J(t) \delta J(t')}. \qquad (14.83)$$

Notice how this is essentially the Feynman Green function, providing a link between statistical physics and mechanics through this symmetrical Green function.

14.3.3 Liouville's theorem

An important theorem in statistical mechanics clarifies the definition of time evolution in many-particle systems, where it is impractical to follow the trajectory of every particle. This theorem applies to closed, conservative systems.

A given point in phase space represents a specific state of a many-particle system. The density ρ of points in phase space can itself be thought of as a dynamical variable which deforms with time in such a way that the number of points in an infinitesimal volume element is constant. The overall density of points is constant in time since the number of particles is constant and the volume is a constant, by assumption:

$$\frac{d\rho}{dt} = 0, \tag{14.84}$$

or, equivalently,

$$\frac{\partial \rho}{\partial t} + [\rho, H]_{pq} = 0. \tag{14.85}$$

This last form is an expression of how the local density at a fixed point (q, p) in phase space (a fixed state) varies in time. When a dynamical system is in static equilibrium, the density of states at any point must be a constant, thus

$$[\rho, H]_{pq} = 0. \tag{14.86}$$

In a classical Hamiltonian time development, regions of phase space tend to spread out, distributing themselves over the whole of phase space (this is the essence of ergodicity); Liouville's theorem tells us that they do so in such a way as to occupy the same total volume when the system is in statistical equilibrium.

Another way of looking at this is in terms of the distribution function for the field. If the number of states does not change, as is the case for a free field, then

$$\frac{d}{dt} f(p, x) = 0. \tag{14.87}$$

By the chain-rule we may write

$$\left[\frac{\partial}{\partial t} + \left(\frac{\partial x^i}{\partial t} \right) \partial_i + \left(\frac{\partial p^i}{\partial t} \right) \frac{\partial}{\partial p^i} \right] f(p, x) = 0. \tag{14.88}$$

The rate of change of momentum is just the force. In a charged particle field (plasma) this is the Lorentz force $F_i = q E_i + \epsilon_{ijk} v^j B^k$.

14.3.4 Averaged dynamical variations

Since the expectation value is a simple product-weighted average, Liouville's theorem tells us that the time variation of expectation values is simply the

expectation value of the time variation, i.e. these two operations commute because the time derivative of ρ is zero:

$$
\begin{aligned}
\frac{d}{dt}\langle X\rangle_{pq} &= \frac{d}{dt}\text{Tr}(\rho\, X) \\
&= \text{Tr}\frac{d\rho}{dt}X + \text{Tr}\rho\frac{dX}{dt} \\
&= \text{Tr}\rho\frac{dX}{dt} \\
&= \left\langle\frac{dX}{dt}\right\rangle_{pq}.
\end{aligned}
\tag{14.89}
$$

This may also be written as

$$
\frac{d}{dt}\langle X\rangle_{pq} = \left\langle\frac{\partial X}{\partial t} + [X, H]\right\rangle_{pq},
\tag{14.90}
$$

or, more generally for variations, as

$$
\langle\delta_\xi X\rangle_{pq} = \langle[X, G_\xi]\rangle_{pq}.
\tag{14.91}
$$

Again, the similarity to the mechanical theory is striking.

14.4 Quantum mechanics

The discovery of de Broglie waves in electron diffraction experiments by Davisson and Germer (1927) and Thomson (1928) effectively undermined the status of particle coordinates as a fundamental dynamical variable in the quantum theory. The wavelike nature of light and matter cannot be reconciled with discrete labels $q_A(t)$ at the microscopic level. A probabilistic element was necessary to explain quantum mechanics. This is true even of single particles; it is not merely a continuum feature in the limit of large numbers of particles, such as one encounters in statistical mechanics. Instead it was necessary to find a new, more fundamental, description of matter in which both the wavelike properties and impulsive particle properties could be unified. Such a description is only possible by a more careful study of the role played by invariance groups.

Because of the cumbersome nature of the Poisson bracket for continuum theory, continuum theories are not generally described with Poisson algebras. Instead, an equivalent algebra arises naturally from the de Broglie relation $p_\mu = \hbar k_\mu$: namely commutator algebras. The important properties one wishes to preserve are the anti-symmetry of the conjugate pair algebra, which leads to the canonical invariances.

In classical mechanics, $q(t)$ does not transform like a vector under the action of symmetry groups, dynamical or otherwise. A more direct route to the

development of the system is obtained by introducing an eigenvector basis in group space which does transform like a vector and which employs operators to extract the dynamical information.

14.4.1 Non-relativistic quantum mechanics in terms of groups and operators

Schrödinger's formulation of quantum mechanics is postulated by starting with the Galilean energy conservation relation

$$E = \frac{\mathbf{p}^2}{2m} + V \qquad (14.92)$$

and making the operator replacements $E \rightarrow i\hbar\partial_t$ and $\mathbf{p} \rightarrow -i\hbar\nabla$. The solution of this equation, together with the interpretation of the wavefunction and a specification of boundary conditions, is quantum mechanics. It is interesting nonetheless to examine quantum mechanics as a dynamical system in order to identify its relationship to classical mechanics. The main physical assumptions of quantum mechanics are the de Broglie relation $p_\mu = \hbar k_\mu$ and the interpretation of the wavefunction as a probability amplitude for a given measurement. The remainder is a purely group theoretical framework for exploiting the underlying symmetries and structure.

From a group theoretical viewpoint, quantum mechanics is simpler than classical mechanics, and has a more natural formulation. The use of Poisson brackets to describe field theory is not practical, however. Such a formulation would require intimate knowledge of Green functions and boundary conditions, and would involve complicated functional equations. To some degree, this is the territory of quantum field theory, which is beyond the scope of this work. In the usual approach, canonical invariances are made possible by the introduction of a vector space description of the dynamics. It is based upon the algebra of observables and the method of eigenvalues.

The wavefunction or field Since a particle position cannot be a wave (a particle is by definition a localized object), and neither can its momentum, it is postulated that the wavelike nature of quantum propagation is embodied in a *function of state* for the particle system called the *wavefunction* and that all physical properties (called *observables*) can be extracted from this function by Hermitian operators. The wavefunction $\psi(\mathbf{x}, t)$ is postulated to be a vector in an abstract multi-dimensional Hilbert space, whose magnitude and direction contains all the information about the particle, in much the same way that phase space plays the same role for classical particle trajectories.

The fact that the wavefunction is a vector is very convenient from the point of view of the dynamics (see section 8.1.3), since it means that the generators of invariance groups can operate directly on them by multiplication. This leads to a

Table 14.1. Dynamical formulations.

Classical	Schrödinger	Heisenberg
$q(t)$	$\hat{x}\psi(\mathbf{x}, t)$	$\hat{x}(t)\psi(\mathbf{x})$
$p(t)$	$\hat{p}\psi(\mathbf{x}, t)$	$\hat{p}(t)\psi(\mathbf{x})$

closer connection with the group theory that explains so much of the dynamics. It means that any change in the system, characterized by a group transformation U, can be expressed in the operational form

$$\psi' = U\,\psi. \tag{14.93}$$

This is much simpler than the pair of equations in (14.34). It is, in fact, more closely related to eqn. (14.35) because of the group structure in eqn. (14.63), as we shall see below.

Operator-valued position and momentum $q(t)$ and $p(t)$ may be effectively supplanted as the dynamical variables by the wavefunction. To represent the position and momentum of particles, one makes a correspondence with operators according to one of two equivalent prescriptions (table 14.1). The choice depends on whether one wishes to place the time development of the system in the definition of the operators, or whether it should be placed in the wavefunction, along with all the other dynamical parameters. These two descriptions are equivalent to one another in virtue of the group combination law. We shall mainly use the Schrödinger representation here since this is more in tune with the group theoretical ideology of symmetries and generators.

As explained in section 11.1, it is the operators themselves, for dimensional reasons, which are the positions and momenta, not the operators multiplying the fields. The observable values which correspond to the classical quantities are extracted from this function by considering the eigenvalues of the operators. Since the wavefunction $\psi(x)$ can always be written as a linear combination of the complete set of eigenvectors $E(x)$ belonging to any operator on Hilbert space, with constants λ_a,

$$\psi(x) = \sum_a \lambda_a E_a(x), \tag{14.94}$$

there is always a well defined eigenvalue problem which can convert a Hermitian operator into a real eigenvalue.

Commutation relations Since the field Poisson bracket is unhelpful, we look for
a representation of position and momentum which distills the important property
in eqn. (14.57) from the classical canonical theory and injects it into the quantum
theory. One sees that, on choosing the following algebraic representation of the
underlying Galilean symmetry group for the wavefunction[5]

$$\psi(x) = \sum_k a_k e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}, \tag{14.95}$$

a simple representation of the operators $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$ may be constructed from

$$\hat{\mathbf{x}} = \mathbf{x}$$
$$\hat{\mathbf{p}} = -i\hbar\nabla. \tag{14.96}$$

These operators live on the vector space of the Galilean group (i.e. real space),
so it is natural to use their operator property directly in forming a canonical
algebra. They are complete, as may be verified by computing the straightforward
commutator

$$[\hat{\mathbf{x}}, \hat{\mathbf{p}}] = \hat{\mathbf{x}}\hat{\mathbf{p}} - \hat{\mathbf{p}}\hat{\mathbf{x}} = i\hbar. \tag{14.97}$$

This clearly satisfies eqn. (14.57). Thus, with this representation of position
and momentum, based directly on the underlying symmetry of spacetime, there
is no need to introduce an abstract phase space in order to construct a set
of vectors spanning the dynamics. The representations of the Galilean group
suffice. The only contribution from empirical quantum theory is the expression
of the wavenumber $\mathbf{k}$ and the frequency ω in terms of the de Broglie relation. In
fact, this cancels from eqn. (14.97).

Dirac notation: bases In Dirac notation, Hilbert space vectors are usually
written using angle brackets ($|x\rangle$, $|\psi\rangle$). To avoid confusing this notation with
that for expectation values, used earlier, we shall use another fairly common
notation, $|x)$, $|\psi)$, here. The components of such a vector are defined with
respect to a particular basis. Adjoint vectors are written $(x|$, $(\psi|$, and so on.

The scalar product of two such vectors is independent of the basis, and is
written

$$(\psi_1|\psi_2) = \int d\sigma\, \psi_1^\dagger(x)\psi_2(x)$$
$$= \int (dp)\psi_1^\dagger(p)\psi_2(p). \tag{14.98}$$

In Dirac notation one considers the functional dependence of wavefunctions to
be the basis in which they are defined. Thus, $\psi(x)$ is likened to the components

---

[5] Note that the representations of the Galilean group are simply the Fourier expansion.

Table 14.2.   Matrix elements and operator bases.

$\hat{O}$	$(x'|\hat{O}|x)$	$(p'|\hat{O}|p)$
$\hat{\mathbf{p}}$	$-i\hbar\nabla\,\delta\,(\mathbf{x}, \mathbf{x}')$	$\mathbf{p}\,\delta(\mathbf{p}, \mathbf{p}')$
$\hat{\mathbf{x}}$	$\mathbf{x}\delta(\mathbf{x}, \mathbf{x}')$	$-i\hbar\frac{\partial}{\partial\mathbf{p}}\delta(\mathbf{p}, \mathbf{p}')$

of the general function ψ in an x basis. Similarly, the Fourier transform $\psi(p)$ is thought of as the components of ψ in a p basis. As in regular geometry, the components of a vector are obtained by taking the scalar product of the vector with a basis vector. In Dirac notation, the wavefunction and its Fourier transform are therefore written as

$$\psi(x) = (x|\psi)$$
$$\psi(p) = (p|\psi), \qquad (14.99)$$

as a projection of the vector onto the basis vectors. The basis vectors $|x)$ and $|p)$ form a complete set of eigenstates of their respective operators, satisfying the completeness relation

$$(\mathbf{x}, \mathbf{x}') = \delta(\mathbf{x}, \mathbf{x}'). \qquad (14.100)$$

Similarly, a matrix, or operator is also defined by an outer product according to what basis, or type of variable, it operates on. The identity operator in a basis x is

$$\hat{I} = \int d\sigma_x |x)(x|, \qquad (14.101)$$

and similarly, in an arbitrary basis ξ, one has

$$\hat{I} = \int d\sigma_\xi |\xi)(\xi|. \qquad (14.102)$$

See table 14.2. This makes the scalar product basis-independent:

$$(\psi_1|\psi_2) = \int d\sigma_x\,(\psi_1|x)(x|\psi_2), \qquad (14.103)$$

as well as the expectation value of $\hat{O}$ with respect to the state $|\psi)$:

$$(\psi|\hat{O}|\psi) = \int d\sigma_x d\sigma_{x'}\,(\psi|x')(x'|\hat{O}|x)(x|\psi), \qquad (14.104)$$

Transformation function The scalar product $(\psi_1|\psi_2)$ represents an overlap of one state of the system with another; thus, the implication is that transitions or transformations from one state to another can take place between these two states. $(\psi_2(x_2)|\psi_1(x_1))$ is often called the transformation function. It represents the probability amplitude of a transition from $\psi_1(x_1)$ to $\psi_2(x_2)$. The quantity

$$A = (\psi'|\hat{O}|\psi) \tag{14.105}$$

is not an expectation value, since it refers to two separate states; rather, it is to be interpreted as another transition amplitude, perturbed by the operation $\hat{O}$, since

$$\hat{O}|\psi) = |\psi''). \tag{14.106}$$

Thus

$$A = (\psi'|\psi''), \tag{14.107}$$

which is just another transition function. The transformation function plays a central role in Schwinger's action principle for quantum mechanics, and is closely related to the path integral formulation.

Operator variations and unitary transformations In order to define a variational theory of quantum mechanics, meaning must be assigned to the variation of an operator. An operator has no meaning without a set of vectors on which to operate, so the notion of an operator variation must be tied to changes in the states on which it operates. States change when they are multiplied by the elements of a transformation group U:

$$|\psi) \to U|\psi). \tag{14.108}$$

Similarly, the adjoint transforms by

$$(\psi| \to (\psi|U^\dagger. \tag{14.109}$$

The invariance of the scalar product $(\psi|\psi)$ implies that U must be a unitary transformation, satisfying

$$U^\dagger = U^{-1}. \tag{14.110}$$

Consider an infinitesimal unitary transformation with generator G such that $U = \exp(-iG/\hbar)$.

$$|\psi) \to e^{-iG/\hbar}|\psi) = (1 - iG/\hbar)|\psi). \tag{14.111}$$

The change in an expectation value due to an operator variation $\hat{X} \to \hat{X} + \delta\hat{X}$,

$$(\psi|\hat{X} + \delta\hat{X}|\psi) = (\psi|e^{iG/\hbar}\hat{X}e^{-i/\hbar G}|\psi)$$
$$= (\psi|(1 + iG/\hbar)\hat{X}(1 - iG/\hbar)|\psi), \tag{14.112}$$

or, equating $\delta\hat{X}$ to the first infinitesimal order on the right hand side,

$$\delta\hat{X} = \frac{1}{\mathrm{i}\hbar}[\hat{X}, G]. \tag{14.113}$$

Eqn. (14.113) may be taken as the definition of operator variations, affected by unitary transformations. It can be compared with eqn. (14.35) for the canonical variations. It is this definition which permits an action principle to be constructed for quantum mechanics. From eqn. (14.113), one can define the expectation value

$$\mathrm{i}\hbar\,\langle\delta X\rangle = \langle\xi|[X, G]|\xi\rangle, \tag{14.114}$$

and, by a now familiar argument for the time variation $G_t = -H\delta t$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle X\rangle = \left\langle \frac{\partial X}{\partial t} + \frac{1}{\mathrm{i}\hbar}[X, H]\right\rangle, \tag{14.115}$$

where the expectation value is interpreted with respect to a basis ξ in Hilbert space:

$$\langle\ldots\rangle = \int \mathrm{d}\xi\,(\xi|\ldots|\xi). \tag{14.116}$$

This relation can be compared with eqn. (14.90) from classical statistical mechanics.

It is straightforward to verify Hamilton's equations for the operators by taking

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V, \tag{14.117}$$

so that

$$\dot{\hat{\mathbf{p}}} = \frac{1}{\mathrm{i}\hbar}[\hat{\mathbf{p}}, \hat{H}] = -\nabla\hat{H}$$

$$\dot{\hat{\mathbf{q}}} = \frac{1}{\mathrm{i}\hbar}[\hat{\mathbf{q}}, \hat{H}] = \mathrm{i}\hbar\left[-\frac{\nabla^2}{2m}\hat{x}\right] = -\mathrm{i}\hbar\frac{\nabla}{m} = \frac{\hat{\mathbf{p}}}{m} = \frac{\partial\hat{H}}{\partial\hat{\mathbf{p}}}. \tag{14.118}$$

The last step here is formal, since the derivation with respect to an operator is really a distribution or Green function, which includes a specification of boundary conditions. In this case, the only possibility is for causal, retarded, boundary conditions, and thus the expression is unambiguous.

Statistical interpretation By comparing quantum expectation values, or scalar products, with statistical mechanics, it is possible to see that the states referred to in quantum mechanics must have a probabilistic interpretation. This follows

directly from the canonical structure and from the analogy between the quantum state function and the density operator in eqn. (14.78).

If it were not already clear from phenomenology, it should now be clear from eqn. (14.37) that the quantum theory has the form of a statistical theory. Thus, the wavefunction can be regarded as a probabilistic object, and the waves which give rise to quantum interference are 'probability waves'.

The basis-independence of the quantum expectation value is analogous to the ergodicity property of classical mechanics: it implies that it is not important what variable one chooses to average over. A 'dynamically complete' average will lead to the same result.

The formalism of quantum theory makes no statements about wave–particle duality and no confusion arises with regard to this concept. Quantum mechanics must be regarded as a straightforward generalization of classical canonical mechanics, which admits a greater freedom of expression in terms of group theory and invariances.

Classical correspondence Although sometimes stated misleadingly, the correspondence between the Poisson bracket in classical mechanics and the commutator in quantum mechanics is not such that one recovers the Poisson bracket formulation from the classical limit of the commutator. They are completely independent, since they refer to different spaces. While the commutator function exists in the classical limit $\hbar \to 0$, the wavefunction does not, since $\mathbf{k} \to \infty$ and $\omega \to \infty$. Thus, the basis vectors cease to exist.

The true correspondence with classical physics is through expectation values of quantum operators, since these are independent of the operator basis. The classical theory is through the equations

$$\left(\psi \left| -\frac{\hbar^2}{2m}\nabla^2 + V = i\hbar\frac{\partial}{\partial t} \right| \psi \right) \to \frac{p^2}{2m} + V = E, \qquad (14.119)$$

and

$$
\begin{aligned}
\left\langle \frac{d\hat{\mathbf{p}}}{dt} \right\rangle &= \frac{d\langle\hat{\mathbf{p}}\rangle}{dt} \\
&= \frac{-i}{\hbar}\langle[\hat{\mathbf{p}}, H]\rangle \\
&= \frac{-i}{\hbar}(\hat{\mathbf{p}}V(x) - V(x)\hat{\mathbf{p}}) \\
&= \frac{-i}{\hbar}(-i\hbar\nabla V(x)) \\
&= -\nabla V(x)), \qquad (14.120)
\end{aligned}
$$

which is Newton's law.

14.4.2 Quantum mechanical action principle

Schwinger has shown that the complete unitary, dynamical structure of quantum mechanics can be derived from a quantum action principle, based on operator variations. The quantum action principle is directly analogous to its classical counterpart. We shall return to this quantum action principle in chapter 15 since it plays a central role in modern quantum field theory. For now, the action principle will not be proven; instead we summarize the main results. The algebraic similarities to the classical action principle are quite remarkable.

The central object in the quantum theory is the transformation function or transition amplitude $(\psi|\psi')$. The quantum action principle states that the action is a generating functional, which induces changes in the transformation function,

$$\delta(\psi(t_2)|\psi(t_1)) = \frac{1}{i\hbar}(\psi(t_2)|\delta\hat{S}_{12}|\psi(t_1)), \qquad (14.121)$$

where $\hat{S}$ is the action operator, which is constructed from the classical action by replacing each dynamical object by its operator counterpart:

$$\hat{S}_{12} = \int_{t_1}^{t_2} dt \left(\frac{1}{2}(\hat{\mathbf{p}}\dot{\hat{\mathbf{q}}} - \hat{\mathbf{q}}\dot{\hat{\mathbf{p}}}) - \hat{H} \right). \qquad (14.122)$$

In this simple case, the ordering of the operators is unambiguous. The variation in the action contains contributions only from within the time values corresponding to the end-points of the transformation function for causality.

If one now introduces the identity $I = |x) \times (x|$ into the transformation function and substitutes the real-space representations of the operators, eqn. (14.121) becomes

$$\delta(\psi(t_2)|\psi(t_1)) =$$
$$\frac{1}{i\hbar}\delta \int_{t_1}^{t_2} (dx)\psi^\dagger(x)\left[\frac{1}{2}(-i\hbar\nabla\dot{\mathbf{x}} + i\hbar\dot{\mathbf{x}}\nabla) + \frac{\hbar^2}{2m}\nabla^2 - V \right]\psi(x),$$
$$(14.123)$$

which is equal to

$$\delta(\psi(t_2)|\psi(t_1)) =$$

$$\frac{1}{i\hbar}\delta \int_{t_1}^{t_2} (dx)\ \psi^\dagger(x)\left[-\frac{i\hbar}{2}\left(\overrightarrow{\partial_t} - \overleftarrow{\partial_t}\right) + \frac{\hbar^2}{2m}\nabla^2 - V \right]\psi(x). \quad (14.124)$$

δ refers only to the contents of the square brackets. This expression may be compared with the action for the Schrödinger field in eqn. (17.4). For the purposes of computing the variation, the form in eqn. (14.122) is most

convenient. In fact, it is easy to see that – given that the variation of an operator is well defined – the results are exactly analogous to the classical case.

If the end-points of the variation satisfy fixed boundary conditions, then

$$\delta(\psi(t_2)|\psi(t_1)) = 0, \tag{14.125}$$

since the field is constrained to admit no unitary transformations there, thus the right hand side of eqn. (14.121) is also zero. This, in turn, implies that the variation of the action operator vanishes, and this leads to the *operator equations of motion*, analogous to Hamilton's equations:

$$\delta_{\hat{\mathbf{x}}}\hat{S} = \int \mathrm{d}t \left(-\dot{\hat{\mathbf{p}}} - \frac{\partial \hat{H}}{\partial \mathbf{x}} \right) \delta\mathbf{x}, \tag{14.126}$$

whence

$$\dot{\hat{\mathbf{p}}} = -\frac{\partial \hat{H}}{\partial \mathbf{x}}. \tag{14.127}$$

Similarly, the variation with respect to the momentum operator leads to

$$\dot{\hat{\mathbf{x}}} = \frac{\partial \hat{H}}{\partial \mathbf{p}}, \tag{14.128}$$

whose consistency was verified in eqns. (14.118). This tells us that quantum mechanics, with commutators in place of Poisson brackets and differential operators acting on a Hilbert space, forms a well defined Hamiltonian system. Eqn. (14.124) shows that this is compatible with Schrödinger field theory. The final piece of the puzzle is to generalize the variations of the action to include non-fixed end-points, in a way analogous to that in section 14.1.7. Then, using the equations of motion to set the bulk terms to zero, one has

$$\delta(\psi(t_2)|\psi(t_1)) = \frac{1}{\mathrm{i}\hbar}(\psi(t_2)|G_2 - G_1|\psi(t_1)), \tag{14.129}$$

which shows that the extended variation merely induces an infinitesimal unitary transformation at the end-points of the variation. This variation is in accord with eqn. (14.113), and one may verify that

$$\delta\hat{\mathbf{x}} = \frac{1}{\mathrm{i}\hbar}[\hat{\mathbf{x}}, G_{\mathbf{x}}]$$
$$= \frac{1}{\mathrm{i}\hbar}[\hat{\mathbf{x}}, \hat{\mathbf{p}}\delta\hat{\mathbf{x}}], \tag{14.130}$$

which immediately gives the fundamental commutation relations

$$[\hat{\mathbf{x}}, \hat{\mathbf{p}}] = \mathrm{i}\hbar. \tag{14.131}$$

This final piece of the puzzle verifies that the operator variational principle is self-consistent for quantum mechanics. In fact, it can be generalized to other operators too, as we shall see in chapter 15, when we consider the fully quantized theory of fields.

14.4.3 Relativistic quantum mechanics

A Lorentz-invariant theory of quantum mechanics may be obtained by repeating the previous construction for the non-relativistic theory, replacing the non-relativistic energy relation in eqn. (14.92) with

$$E^2 = \mathbf{p}^2 c^2 + m^2 c^4. \tag{14.132}$$

One writes

$$(-\hat{E}^2 + \hat{\mathbf{p}}^2 c^2 + m^2 c^4)\phi(x) = 0, \tag{14.133}$$

where $\hat{E} = i\hbar \partial_t$ and $\hat{\mathbf{p}} = -i\hbar \nabla$, and we call the field $\phi(x)$ to distinguish it from the non-relativistic field. This leads us directly to the Klein–Gordon equation

$$(-\hbar^2 c^2 \square + m^2 c^4)\phi = 0. \tag{14.134}$$

However, all is not straightforward. The interpretation of this equation is full of subtleties, which leads inexorably to a full quantum field theory. To begin with its quadratic nature implies that it has solutions of both arbitrarily large positive and negative energy (see section 5.1.3). This further implies that the conserved quantities normally used to define probability measures can also be negative; this is difficult to interpret. Ultimately, the assumptions of quantum field theory save the relativistic formulation. Leaning on these, relativistic quantum mechanics survives as an approximation to the more complete quantum field theory under conditions of 'sufficient stability'.[6]

State vectors and wavefunctions In non-relativistic quantum mechanics it was easy to choose state vectors satisfying the Schrödinger equation because of the simple form of the conserved quantities arising from the linear time derivative (see eqn. (12.39)). The structural symmetry of the natural inner product:

$$(\psi_1, \psi_2) = \int d\sigma \ \psi_1^\dagger \psi_2, \tag{14.135}$$

means that the state vectors $|\psi_1\rangle$ and the adjoint $\langle\psi_1|$ were simply Hermitian conjugates of one another. In the case of the Klein–Gordon equation, matters are more complicated. The corresponding invariant inner product is

$$(\phi_1, \phi_2) = -i\hbar c^2 \int d\sigma \ \phi_1^* \overset{\leftrightarrow}{\partial_0} \phi_2 \tag{14.136}$$

---

[6] To make this woolly statement precise, one needs to address issues in the language of quantum field theory or renormalization group, which we shall only touch on in chapter 15.

the symmetry of which is made less obvious by the time derivative, and one is now faced with both positive and negative energy solutions. These two sets of solutions decouple, however. If one splits the field into its positive and negative energy parts,

$$\phi(x) = \phi^{(+)}(x) + \phi^{(-)}(x), \tag{14.137}$$

then one has, for a real scalar field,

$$
\begin{aligned}
(\phi(x), \phi(x)) &= (\phi^{(+)}(x), \phi^{(+)}(x)) + (\phi^{(-)}(x), \phi^{(-)}(x)) \\
&= 0;
\end{aligned} \tag{14.138}
$$

i.e.

$$
\begin{aligned}
(\phi^{(+)}(x), \phi^{(+)}(x)) &= -(\phi^{(-)}(x), \phi^{(-)}(x)) \\
(\phi^{(+)}(x), \phi^{(-)}(x)) &= 0.
\end{aligned} \tag{14.139}
$$

or, more generally,

$$(\phi_A, \phi_B) = -(\phi_B, \phi_A)^*. \tag{14.140}$$

By analogy with the non-relativistic case, we wish to view this scalar product as the definition of a vector space with vectors $|\phi)$ and adjoint vectors $(\phi|$, such that

$$(\phi_1 | \phi_2) = (\phi_1, \phi_2), \tag{14.141}$$

i.e. the inner product on the vector space is identified with the conserved quantity for the field. The ϕ_A satisfy the Klein–Gordon equation:

$$
\begin{aligned}
\phi(x) &= \int \frac{\mathrm{d}^{n+1}k}{(2\pi)^{n+1}} \mathrm{e}^{ikx} \phi(k) \delta(p^2 c^2 + m^2 c^4) \\
&= \int \frac{\mathrm{d}^n k}{(2\pi)^n} \frac{1}{2k_0 c^2 \hbar^2} \mathrm{e}^{ikx} (\phi(p_0, \mathbf{p}) + \phi(-p_0, \mathbf{p})) \\
&= \int \mathrm{d}V_k \mathrm{e}^{ikx} (\phi^{(+)}(\mathbf{p}) + \phi^{(-)}(\mathbf{p})).
\end{aligned} \tag{14.142}
$$

What makes the relativistic situation different is the fact that the energy constraint surface is quadratic in k_0. The volume measure on momentum space is constrained by this energy relation. This is the so-called mass shell. On the manifold of only positive energy solutions, the volume measure is

$$
\begin{aligned}
V_k &= \int \frac{\mathrm{d}^{n+1}k}{(2\pi)^{n+1}} \delta(p^2 c^2 + m^2 c^4) \theta(k_0) \\
&= \int \frac{\mathrm{d}^n k}{(2\pi)^n} \frac{1}{2k_0 c^2 \hbar^2} \\
\mathrm{d}V_k &= \frac{\mathrm{d}^n k}{(2\pi)^n} \frac{1}{2k_0 c^2 \hbar^2}.
\end{aligned} \tag{14.143}
$$

Thus, if we examine complete sets of position and momentum eigenfunctions on this constraint manifold, we find that the normalization of momentum eigenfunctions is dictated by this constraint:

$$
\begin{aligned}
(\mathbf{x}, \mathbf{x}') &\equiv \delta(\mathbf{x}, \mathbf{x}') \\
&= \int \frac{\mathrm{d}^n k}{(2\pi)^n} \mathrm{e}^{\mathrm{i}\hat{\mathbf{k}}\cdot(\mathbf{x}-\mathbf{x}')} \\
&= 2k_0\hbar^2 c^2 \int \mathrm{d}V_k \mathrm{e}^{\mathrm{i}\hat{\mathbf{k}}\cdot(\mathbf{x}-\mathbf{x}')}.
\end{aligned}
\tag{14.144}
$$

From this expression, it follows that

$$
(\hat{\mathbf{k}}|\mathbf{x}) = \sqrt{2k_0\hbar^2 c^2}\ \mathrm{e}^{\mathrm{i}\hat{\mathbf{k}}\cdot(\mathbf{x})}
\tag{14.145}
$$

$$
(\hat{\mathbf{k}}|\hat{\mathbf{k}}') = \int \mathrm{d}\sigma\,(\hat{\mathbf{k}}|\mathbf{x})(\mathbf{x}|\hat{\mathbf{k}}') = 2k_0\hbar^2 c^2\ \delta(\hat{\mathbf{k}} - \hat{\mathbf{k}}').
\tag{14.146}
$$

Thus the one-particle positive energy wavefunction is

$$
\begin{aligned}
\psi \equiv \phi_1(x) = (\mathbf{x}, \phi) &= \int \mathrm{d}V_k (\mathbf{x}|\hat{\mathbf{k}})(\hat{\mathbf{k}}|\phi) \\
&= N \int \mathrm{d}V_k \sqrt{2k_0\hbar^2 c^2}\ \mathrm{e}^{\mathrm{i}\hat{\mathbf{k}}\cdot\mathbf{x}}.
\end{aligned}
\tag{14.147}
$$

Compare this with the re-scaling in eqn. (13.7). It is normalized such that

$$
\begin{aligned}
(\psi, \psi) = (\phi_1|\phi_1) &= 1 \\
&= N^2 \int \mathrm{d}V_k \mathrm{d}V_{k'} \phi_1^*(k)\,(2k_0\hbar^2 c^2)\,\phi_1(k)\delta(\hat{\mathbf{k}} - \hat{\mathbf{k}}') \\
&= N^2 \int \frac{\mathrm{d}^{n+1}k}{(2\pi)^{n+1}} |\phi_1(k)|^2.
\end{aligned}
\tag{14.148}
$$

The normalization factor, N, is fixed, at least in principle, by this relation, perhaps through box normalization. This inner product is unambiguously positive, owing to the restriction to only positive energies. An example is the one-particle wavefunction in $3 + 1$ dimensions:

$$
\begin{aligned}
\psi = \phi_1(x) &= N \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \frac{\mathrm{e}^{\mathrm{i}\hat{\mathbf{k}}\cdot\mathbf{x}}}{\sqrt{2k_0\hbar^2 c^2}} \\
&= \text{const.} \left(\frac{m}{\mathbf{x}}\right)^{\frac{5}{4}} H^{(1)}_{\frac{5}{4}}(\mathrm{i}m\mathbf{x}),
\end{aligned}
\tag{14.149}
$$

where $H^{(1)}_{\frac{5}{4}}(x)$ is a Hankel (Bessel) function. What is significant here is that the one-particle wavefunction is not localized like a delta function. Indeed, it would be impossible to construct a delta function from purely positive energy functions. Rather, it extends in space, decaying over a distance of order $\hbar/mc^2$. See also section 11.2.

14.5 Canonically complete theories

The operational view of classical, statistical and quantum mechanics, which has been lain out above, could seem sterile from a physical perspective. In presenting it as a formal system of canonical equations, one eschews phenomenology entirely and uses only elementary notions based on symmetry. That such an approach is possible is surely an important insight. Mechanics should be regarded for what it is: a description of dynamics in terms of algebraic rules determined from necessary symmetries. Given the mathematical structure, more physical or philosophical discussions can follow of their own accord.

The Hamiltonian dynamical formulation can, for the most part, be circumvented completely by direct use of the action formalism in chapter 4. Again, we use a version of the action principle in which we allow infinitesimal canonical changes at the end-points of dynamical variations.

The quantum theory, being linear, is essentially a theory of small disturbances. The imprint left on the action by variation with respect to some variable is that variable's conjugate quantity. The conjugate quantity is said to be the generator of the variation of disturbance. If one varies the action with respect to a set of parameters ξ^i, and the action is invariant under changes of these parameters, the variation must be zero. Manipulating the symbols in the action and separating out the variation $\delta\xi$ to first order, one can write the infinitesimal variation in the form

$$\delta_\xi S = \int \mathrm{d}\sigma^\mu \; G_i \delta\xi^i = 0, \tag{14.150}$$

where $\mathrm{d}\sigma^\mu$ represents a spacelike hyper-surface. The quantity G_i is the *generator* of the symmetry in ξ^i. It is also called the variable conjugate to ξ^i. Notice that an external source J_{ext}, such that

$$S \to S + \int (\mathrm{d}x) \; J_{\text{ext}}\phi \tag{14.151}$$

acts as a generator for the field, throughout the spacetime volume

$$\delta S \to 0 + \int (\mathrm{d}x) \; J_{\text{ext}}\delta\phi, \tag{14.152}$$

since the dynamical variation of the regular action vanishes. This observation has prompted Schwinger to develop the quantum theory of fields almost entirely with the aid of 'sources' or generators for the different dynamical and symmetrical entities [119] (see table 14.3).

In this chapter, we have compared the way in which classical and quantum mechanics are derivable from action principles and obey canonical completeness relations, in the form of Poisson brackets or commutators. This is no accident. Since the action principle always generates a conjugate momentum, as exhibited

Table 14.3. Some canonical transformations.

T_i	$\delta\xi^i$	Symmetry
$\int \mathrm{d}\sigma^\mu \theta_{\mu\nu}$	δx^ν	Lorentz invariance
$\int \mathrm{d}\sigma^\mu T_{\mu\nu}$	δx^ν	conformal invariance
$\mathbf{p}$	$\delta\mathbf{x}$	translational invariance
H	δt	time translation invariance
Π	$\delta q, \delta\phi, \delta\psi$	spacetime/canonical
J_{ext}	$\delta\phi, \delta\psi$	field canonical/unitary

by eqns. (4.62) and (4.66), one could *define* a canonical theory to be one which follows from an action principle. Thus, the action principle is always a good starting point for constructing a dynamical model of nature. To complete our picture of dynamics, it is necessary to resolve some of the problems which haunt the fringes of the classical relativistic theory. To do this, one must extend the dynamical content of the theory even more to allow the fields themselves to form a dynamical basis. As we saw in section 14.2.2, this was impractical using the Poisson bracket, so instead one looks for a commutator relation to generate infinitesimal variations. This completes the consistency of the formalism for relativistic fields.