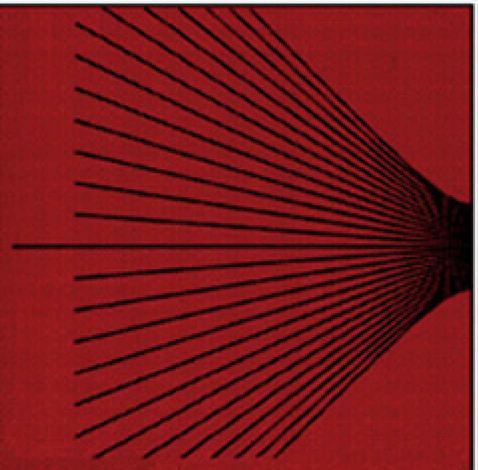
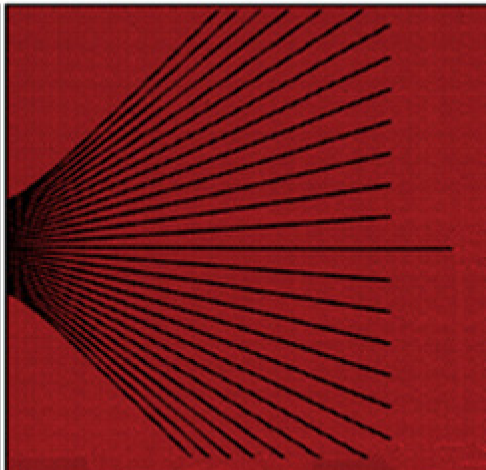
THE QUANTUM THEORY OF MOTION

An Account of the de Broglie–Bohm Causal Interpretation of Quantum Mechanics

PETER R. HOLLAND





CAMBRIDGE

more information – www.cambridge.org/9780521354042

This page intentionally left blank

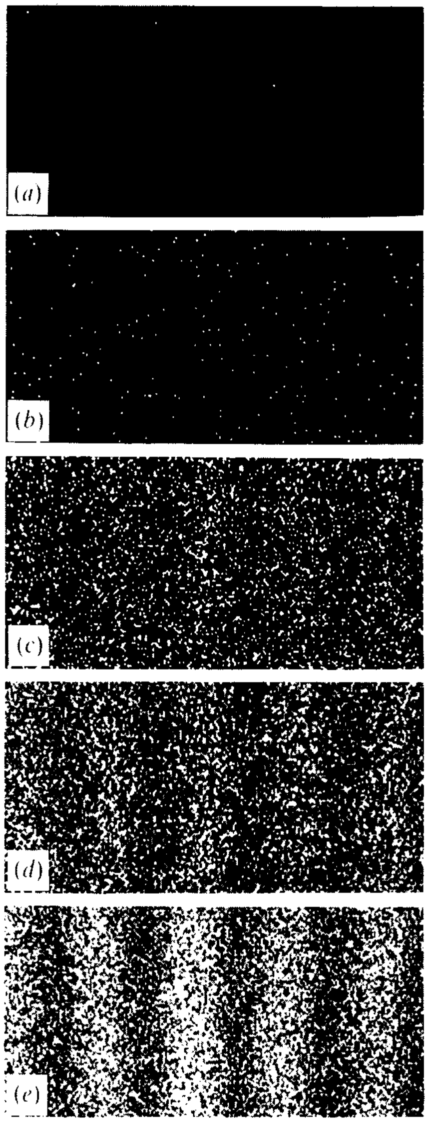
This book presents the first comprehensive exposition of the interpretation of quantum mechanics pioneered by Louis de Broglie and David Bohm. The purpose is to explain how quantum processes may be visualized without ambiguity or confusion in terms of a simple physical model.

Developing the theme that a material system such as an electron is a particle guided by a surrounding quantum wave, a detailed examination of the classic phenomena of quantum theory is presented to show how the spacetime orbits of an ensemble of particles can reproduce the statistical quantum predictions. The mathematical and conceptual aspects of the theory are developed carefully from first principles and topics covered include self-interference, tunnelling, the stability of matter, spin $\frac{1}{2}$, and nonlocality in many-body systems. The theory provides a novel and satisfactory framework for analysing the classical limit of quantum mechanics and Heisenberg's relations, and implies a theory of measurement without wavefunction collapse. It also suggests a strikingly novel view of relativistic quantum theory, including the Dirac equation, quantum field theory and the wavefunction of the universe.

This book provides the first comprehensive technical overview of an approach which brings clarity to a subject notorious for its conceptual difficulties. The book will therefore appeal to all physicists with an interest in the foundations of their subject, and will stimulate all students and research workers in physics who seek an intuitive understanding of the quantum world.

THE QUANTUM THEORY OF MOTION

An account of the de Broglie–Bohm causal interpretation of quantum mechanics



Frontispiece. The cumulative pattern generated by the 'self-interference' of electrons sent one by one through a two-slit interferometer. Number of electrons: (*a*) 10, (*b*) 100, (*c*) 3000, (*d*) 20 000, (*e*) 70 000 (from Tonomura *et al.* (1989)).

THE QUANTUM THEORY OF MOTION

An account of the de Broglie–Bohm causal interpretation of quantum mechanics

PETER R. HOLLAND
Green College, University of Oxford

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, United Kingdom
40 West 20th Street, New York, NY 10011-4211, USA
10 Stamford Road, Oakleigh, VIC 3166, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa
http://www.cambridge.org

First published 1993
First paperback edition 1995
Reprinted 1997, 2000

A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publiation Data

Holland, P. R.
The quantum theory of motion : an account of the de Broglie–Bohm
casual interpretation of quantum mechanics / P. R. Holland
p. cm.
ISBN 0 521 35404 8 (hbk) ISBN 0 521 48543 6 (pbk)
1. Motion 2. Quantum theory 3. Relativistic quantum theory
I. Title
QC133.H65 1993
530.1′2-dc20 92-20205 CIP

ISBN 0 521 35404 8 hardback
ISBN 0 521 48543 6 paperback

Transferred to digital printing 2004

KW

To Jackie, Lucette & Jack
who gave far more
than they know

Contents

Preface

The *frontispiece* portrays a sequence of pictures in which a pattern reminiscent of the interference of waves is progressively built up by a series of individual events in which one electron at a time is sent through a two-slit interferometer and arrives at a detecting screen. Each event is unpredictable, yet over time a definite and reproducible pattern is formed. It is not just arbitrary. What causes the electrons to aggregate in this way? Is there some force acting on each individual electron as it passes through the device which impels it, on the average, to land in certain regions of the screen rather than others?

The historically dominant view in quantum mechanics regards this seemingly natural question as meaningless. For the query rests on the supposition that matter can be at least conceptually analysed, and that when we speak of an 'electron' we really mean an autonomous entity that can move in spacetime and be subject to forces. In contrast, it is generally asserted that our theoretical account of physics in the microdomain must stop short at predicting the likely outcomes when observations at the classical macroscopic scale are performed (the density of dots on the screen). That two electrons in the same quantum state appear at different points does not permit one to infer that they are, in fact, physically distinguished in some way. The uncertainty is postulated to be intrinsic to the system.

This implies the following difficulty: our most basic physical theory contains no account of the constitution and structure of matter, corresponding to the interacting particles and fields of classical physics. It is a means to compute the statistical results of macroobservations carried out on systems that are unspecified and, indeed, unspecifiable. The word 'electron' does not actually *mean* anything at all – it is simply shorthand for a mathematical function. Quantum mechanics is the subject where we never know what we are talking about.

Yet there is a way to know what we are talking about. It was suggested by Louis de Broglie in 1927 and developed into a physical theory by David Bohm in 1952, and has been almost universally ignored since. According to de Broglie and Bohm, the novelty of quantum mechanics is not that we have to revise our customary notions regarding the reality of physical systems, but that their conception must be extended. In describing the electron the classical corpuscle may be retained but must be supplemented by a new type of physical field mathematically described by Schrödinger's wavefunction. The wave guides the particle so that it performs quantum motions rather than classical ones, which gives us the title of this book. The probability distribution of quantum mechanics, such as that displayed in the *frontispiece*, is reproduced because the particle orbits tend to congregate where the wave is most intense. Complete specification of the state of an individual system then requires both aspects: not wave *or* particle but wave *and* particle. The wave-particle composite continuously evolves according to a set of deterministic laws.

This theory of motion is not at all foreign to physicists' practice, for it is tacitly invoked whenever one desires to discuss things to which quantum mechanics potentially applies that no one doubts have *positions* (e.g. meters). Perhaps part of the reason why this approach has been spurned is that its motivation has been misunderstood. The aim of the de Broglie–Bohm theory is *not* to attempt a return to classical physics, or even particularly to invent a deterministic theory. *Its goal is a complete description of an individual real situation as it exists independently of acts of observation.* According to Einstein, that is the programmatic aim of physics. *Determinism is a means to this end.* What emerges is a highly nonclassical theory in which the parts of a larger system are subject to organization by the whole, something not anticipated in the classical paradigm.

There are unsatisfactory aspects of the de Broglie–Bohm model but it has been discarded by the wider scientific community in the absence of a sustained technical examination. This book is intended as a contribution to filling that vacuum. We do not wish to appear too evangelical in our prosecution of the theory, but close study does reveal that the gains in conceptual clarity over the current rivals far outweigh the drawbacks. In the first instance what is at issue is the task of *imagining* the quantum world and that is our principal concern. There now exists a core of established results in this field, some of which are reported here for the first time, and as a research area there are rich prospects. A recurring theme in the book is how the classical and quantum paradigms are connected, a problem that seems to be insoluble in the conventional approach since the usual quantum notion of state does not contain the classical conception as a special case. Another problem where there is great scope for a rational input from the causal model is the interpretation of relativistic quantum theory. The usual presentation of this is extraordinarily vague and provides a classic example of how the rules regarding the meaning of the quantum formalism as pertaining to measurement outcomes are quietly put aside and the physical disposition of matter discussed as if it were objectively real. Most importantly, there is the possibility of bringing the de Broglie–Bohm theory into the experimental arena, a subject that is currently under investigation.

It is a pleasure to thank the following friends and colleagues for their support and encouragement during the writing of this book, and for their courtesy in providing valuable information and/or constructive criticism of the final manuscript: Michael Berry, Piotr Garbaczewski, Dipankar Home, Phil Jacobs, Tassos Kyprianidis, Mioara Mugur-Schächter, Chris Philippidis, Helmut Rauch, Euan Squires and Jean-Pierre Vigier. I have tried to take into account their suggestions but the responsibility for the outcome is mine alone. I am indebted also to D. & J. Canavan for help with the typing at a critical moment, to L. Holland and M. Anderson for preparing many of the figures, and to A. Tonomura for providing the *frontispiece.*

The volume was written while I was a visiting fellow at the Laboratoire de Physique Théorique, Institut Henri Poincaré, Paris, to whom thanks are due for hospitality and generous provision of facilities. Grants from the Royal Society, the SERC and the French government provided partial support in the early stages and the final chapter was written during the tenure of a Leverhulme Trust fellowship. All these agencies are cordially acknowledged. During the bulk of the writing I worked as an editor for Elsevier Science Publishers BV. Thanks are due to Elsevier for its generosity, and to Joost Kircz and Jean-Pierre Vigier without whose assistance and understanding the book would not have been finished.

Paris *PRH*
May, 1992

Note added to Preface for paperback edition

According to the theory of evolution the path of natural phenomena may be regarded as explicable yet unpredictable. Darwin's theory comprises an observer-independent, causal reconstruction of events whose random character precludes evolutionary forecasting. Of course, the historically important issue at the heart of the theory of evolution is the quality of its *explanation* rather than its immediate predictive power. The achievement of the de Broglie–Bohm theory in the domain of quantum physical phenomena is the production of a science similar in its scope and conceptual structure to the elements of Darwinism just cited. Yet what is perceived as an intellectually satisfying and fertile mode of explanation in evolutionary biology has been fiercely resisted in quantum physics. Darwin would not fare well if judged by the criteria of prediction and control currently dominant in physics. But how can we conceive of a testable prediction of an explanatory theory unless we contemplate the explanation it offers?

It is instructive how impractical, inhibiting ideas came to dominate and distort the entire development of a fundamental field of physics. The early quantum physicists attributed to nature a limitation we can now see was simply a deficiency of contemporary thought. The biological comparison shows the arbitrariness in the historical development of physical ideas, but it is not merely a source of analogy. Quantum mechanicists often claim their theory is universal and applicable, in principle, to the entire universe. But biological processes, which are surely part of the same universe, lie beyond the reach of physical explanation as we currently understand it. As far as we can tell, they are not instances of quantum mechanics. Actually, the failure of quantum mechanics to reproduce the valid results of other sciences already occurs at a more basic level of macro-experience. The difficulty arises in attempts to derive classical particle and field theory from the quantum theory in cases where the former are known to be valid. The de Broglie–Bohm theory suggests that this programme is generally unrealisable; generic classical processes are inaccessible starting from quantum ones. Even in those cases where the Correspondence Principle of the de Broglie–Bohm model is obeyed (as gauged by the relative effectiveness of the quantum potential) only a subclass of admissible classical behaviour may be recovered in general. Taking into account the entire spectrum of physical processes to which they have been applied, we cannot therefore assert that the quantum theory of matter and motion is *better* than the corresponding classical theory. Rather, they emerge as *different* theories whose domain of overlap is smaller than is customarily believed. Some examples illustrating this point, hitherto unnoticed even by partisans of the de Broglie–Bohm model but which may prove to be its most significant new insight, are given in Chap. 6.

It is pleasing to report that since this book was completed many new workers have been drawn to the de Broglie–Bohm theory. Work is still in progress so I have restricted myself in this reprint to the correction of minor errors. The comments of many correspondents are much appreciated.

Paris,
July 1994

PRH

1

Quantum mechanics and its interpretation

1.1 The nature of the problem

The quantum world is inexplicable in classical terms. The predictions pertaining to the interaction of matter and light embodied in Newton's laws for the motion of particles and Maxwell's equations governing the propagation of electromagnetic fields are in flat contradiction with the experimental facts at the microscopic scale. A key feature of quantum effects is their apparent indeterminism, that individual atomic events are unpredictable, uncontrollable and literally seem to have no cause. Regularities emerge only when one considers a large ensemble of such events. This indeed is generally considered to constitute the heart of the conceptual problem posed by quantum phenomena, necessitating a fundamental revision of the deterministic classical world view.

Some of the principal phenomena, among those discussed in this book, are as follows:

(*a*) *Self-interference:* a beam of electrons sent one at a time through a barrier containing two apertures builds up through a series of localized detection events on a screen an interference pattern characteristic of waves (cf. the *frontispiece* and §5.1).
(*b*) *Tunnelling:* an α-particle trapped in a nucleus can pass through a potential barrier in a manner forbidden to a classical particle (§5.3).
(*c*) *The stability of matter:* atoms and molecules are found to exist only in certain discrete, or 'stationary', energy states. For this reason they do not 'collapse', the result predicted by classical electrodynamics. During transitions between stationary states (quantum jumps) an atom exchanges a discrete quantity of energy with the electromagnetic field (§§4.5, 7.6).
(*d*) *Spin* $\frac{1}{2}$*:* a beam of atoms sent through an inhomogeneous magnetic field is split into a discrete set of subbeams (the Stern–Gerlach experiment). This reveals a novel type of nonclassical internal angular momentum (§9.5).

(*e*) *Nonlocal correlations:* the properties of one system can depend on those of another arbitrarily distant system with which it has interacted in the past (Chap. 11).

The discrete, statistical and nonlocal character of these phenomena is clearly in conflict with the continuous, determinist and generally local structure of the world according to classical particle and field physics.

It is remarkable that a single coherent mathematical theory, quantum mechanics, could be devised to correlate the heterogeneous empirical data just cited (and many more). Through the Schrödinger equation, quantum mechanics describes the laws of evolution of statistical ensembles of similarly prepared systems. To date, the results obtained are in accord with all experimental evidence. But, insofar as it only predicts the outcomes of measurements performed on statistical aggregates of physical systems, quantum mechanics does not in itself provide an *explanation* of the experimental facts. What is missing is *a description of the actual individual events of experience*, of which the statistical phenomena would be functions.

The challenge is to develop a theory of individual material systems, each obeying its own law of motion, whose mean behaviour over an ensemble reproduces the statistical predictions of quantum mechanics. The empirical record would then be explained as the outcome of a sequence of well-defined processes undergone by systems possessing properties that exist independently of acts of observation.

A way to do this was found by de Broglie and Bohm. It turns out that a causal representation of quantum phenomena may be constructed which leaves intact the basic mathematical formalism of the theory, if this is reinterpreted and extended in a certain way. The detailed working out of the de Broglie–Bohm idea is the subject of this book.

1.2 The wavefunction and the Schrödinger equation

In classical physics the state of an individual material or field (e.g. electromagnetic) system is uniquely defined by the position $\mathbf{x}(t)$ of the object in the first case and the real or complex amplitude $\phi(\mathbf{x}, t)$ in the second, as functions of the time t. Here $\mathbf{x} = (x, y, z)$ represents the Cartesian coordinates of a point in space. The equations of motion of classical physics, Newton's laws and Maxwell's equations, specify these state variables for all time if they and the corresponding momenta are precisely determined at one instant.

The quantum theory developed in the 1920s is connected with its classical predecessor by the mathematical procedure of 'quantization', in which classical dynamical variables are replaced by operators. In the process, a new entity appears which the operators act on, the *wavefunction*. For a single-body matter system this is a complex function, $\psi(\mathbf{x}, t)$, and for a field it is a complex functional, $\psi[\phi, t]$. The wavefunction constitutes a new notion of the state of a physical system.

The historical problem of interpreting quantum theory may be formulated as follows: In prosecuting their quantization procedure, the Founding Fathers introduced the new notion of state not in *addition* to the classical state variables but *instead* of them (Fig. 1.1). They could not see, and finally did not want to see even when presented with a consistent example, how to retain in some form the assumption at the heart of the classical paradigm that matter has substance and form independently of whether or not it is observed. The ψ-function alone was adopted as characterizing the state of a system. Since there is no way to describe individual processes using just the wavefunction, it seemed natural to claim that these are indeterminate and unanalysable in principle.

How the state function is to be interpreted, at least partially, and how it evolves in time are questions addressed by the axioms of quantum mechanics, which we now summarize in a fairly informal way. These are generally agreed upon and will be accepted here. It is not possible to state the rules without some interpretative elements intruding, particularly Born's probabilistic interpretation of ψ. But the further significance of ψ, if any, is open to debate and will be discussed later. Although the formalism is interwoven with the interpretation, the latter is not unique and one should not equate either of

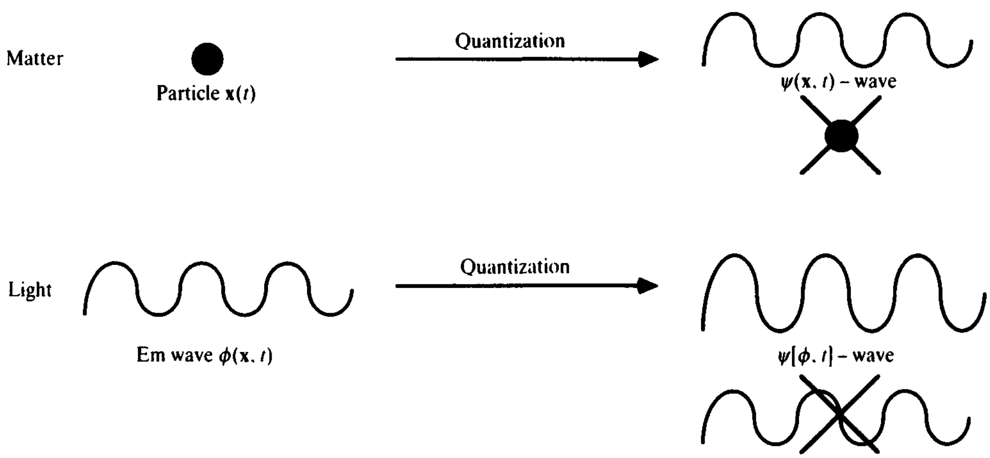


Fig. 1.1 When the wavefunction was introduced, the classical particle and field variables characterizing the states of individual physical systems were discarded.

them with the Copenhagen interpretation (§1.3). More detailed and rigorous presentations will be found in the standard sources such as von Neumann (1955), Messiah (1961), Schiff (1968) and Dirac (1974). Summaries are given by Bjorken & Drell (1964), Jammer (1974, Chap. 1) and d'Espagnat (1976).

We shall consider just the one-body problem. The wavefunction introduced above is referred to as 'the state in the position representation'. This means the following. We denote the state of the system by the ket $|\psi\rangle$, an abstract vector in a linear vector space. In this space we introduce a set of axes $|\mathbf{x}\rangle$, one axis for each value of $\mathbf{x}$. Then the wavefunction is the set of components (one for each $\mathbf{x}$) of the state vector with respect to this basis: $\psi(\mathbf{x}, t) = \langle\mathbf{x}|\psi(t)\rangle$.

The set of complex numbers, i.e., the wavefunction, has the following interpretation. Suppose that the wavefunction is square-integrable, so that it lies in a Hilbert space $\mathscr{H}$, and is normalized:

$$\langle\psi|\psi\rangle = \int_{-\infty}^{\infty} |\psi|^2\, \mathrm{d}^3x = 1. \tag{1.2.1}$$

Then if at time t a measurement is performed to determine the position of the system described by the function $\psi(\mathbf{x}, t)$, the probability that the result lies in the element of space d^3x around the point $\mathbf{x}$ is given by

$$P(\mathbf{x}, t)\, \mathrm{d}^3x = |\psi(\mathbf{x}, t)|^2 \mathrm{d}^3x. \tag{1.2.2}$$

This probability interpretation may be generalized and applied to the measurements of other physical quantities. We first note that all observables in quantum mechanics are represented by linear Hermitian operators acting in the Hilbert space. For example, the classical canonical momentum $\mathbf{p}$ is replaced, in the position representation, as follows:

$$\mathbf{p} \to \hat{\mathbf{p}} = -\mathrm{i}\hbar\nabla \tag{1.2.3}$$

where $\mathrm{i} = \sqrt{-1}$, $\hbar = h/2\pi$, h is Planck's constant and ∇ is the gradient operator. This is the procedure of quantization. In the following we shall often denote the operator corresponding to a classical variable A by $\hat{A}$, but where it is clear from the context that an operator is intended we sometimes omit the caret.

The outcome of the measurement of an observable $\hat{A}$ is one of its eigenvalues, a, defined by the equation

$$\hat{A}|a\rangle = a|a\rangle \tag{1.2.4}$$

where a is real and $|a\rangle$ is the corresponding eigenstate (we assume the spectrum is discrete and ignore degeneracy where several independent eigenstates may correspond to the same eigenvalue). In the position representation we shall write $\langle \mathbf{x}|a\rangle = \psi_a(\mathbf{x})$. The eigenfunctions corresponding to distinct eigenvalues a, a' are orthonormal,

$$\int_{-\infty}^{\infty} \psi_{a'}^*(\mathbf{x})\psi_a(\mathbf{x})\,\mathrm{d}^3x = \delta_{aa'}, \tag{1.2.5}$$

and form a complete set so that an arbitrary wavefunction can be expanded in terms of them:

$$\psi(\mathbf{x}) = \sum_a c_a \psi_a(\mathbf{x}), \tag{1.2.6}$$

where c_a are complex numbers. We have $c_a = \langle a|\psi\rangle$ so that these numbers are the components of the vector $|\psi\rangle$ with respect to the basis $|a\rangle$ in $\mathscr{H}$.

If now a measurement of $\hat{A}$ is performed on a system in the general normalized state (1.2.6), the probability of the outcome a is given by $|c_a|^2$. As a result of the measurement the system has 'collapsed' into the state $|a\rangle$. Notice that $|c_a|^2$ does not refer to the probability that the system is *in* the state $|a\rangle$, independently of the performance of a measurement.

The expectation value of the operator $\hat{A}$ in the state $|\psi\rangle$ is given by

$$\langle \hat{A}\rangle = \langle\psi|\hat{A}|\psi\rangle = \sum_a a|c_a|^2. \tag{1.2.7}$$

By a suitable choice of operators, all the testable predictions of quantum mechanics can be expressed in terms of expectation values in the state $|\psi\rangle$. To do this, we introduce the 'projection operator' $\hat{P}_a = |a\rangle\langle a|$. Then the probabilities are given by

$$|c_a|^2 = \langle\psi|\hat{P}_a|\psi\rangle. \tag{1.2.8}$$

Similarly, we find for the eigenvalues the expressions

$$\left.\begin{aligned} a &= \langle a|\hat{A}|a\rangle \\ &= \langle\psi|\hat{P}_a\hat{A}\hat{P}_a|\psi\rangle/\langle\psi|\hat{P}_a|\psi\rangle. \end{aligned}\right\} \tag{1.2.9}$$

In quantum mechanics the classical Hamiltonian function, $H = \mathbf{p}^2/2m + V(\mathbf{x})$ where V is the external potential energy and m is the mass of the system, is replaced by the Hermitian operator $\hat{H} = \hat{\mathbf{p}}^2/2m + \hat{V}(\hat{\mathbf{x}})$. This determines the evolution of the quantum state via the Schrödinger equation $\mathrm{i}\hbar\,\partial|\psi\rangle/\partial t = \hat{H}|\psi\rangle$. The eigenvalues of $\hat{H}$ are the possible energies of the system corresponding to stationary states. In the position representation the Schrödinger equation becomes the partial differential equation

$$\mathrm{i}\hbar\frac{\partial\psi(\mathbf{x},t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{x},t) + V(\mathbf{x})\psi(\mathbf{x},t). \tag{1.2.10}$$

This law of motion has two notable features: (1) It admits a linear superposition principle. If ψ_1 and ψ_2 are two solutions, then the function

$$\psi(\mathbf{x}, t) = c_1\psi_1(\mathbf{x}, t) + c_2\psi_2(\mathbf{x}, t), \tag{1.2.11}$$

where c_1, c_2 are complex constants, is also a solution. (2) The hermiticity of $\hat{H}$ implies a conservation law for the flow of probability:

$$(\mathrm{d}/\mathrm{d}t)\int_{-\infty}^{\infty} |\psi|^2 \,\mathrm{d}^3x = 0. \tag{1.2.12}$$

To solve (1.2.10) for all time we must specify the initial state function $\psi_0(\mathbf{x}) = \psi(\mathbf{x}, 0)$. In addition, ψ and its derivatives must obey certain boundary and subsidiary conditions. Thus, it is required in regions where V is finite that ψ and $\nabla\psi$ be bounded, continuous and single-valued functions of $\mathbf{x}$. If the potential V has a discontinuity along a surface with normal unit vector $\mathbf{n}$, ψ and $\nabla\psi\cdot\mathbf{n}$ must be continuous across it. If, however, the potential step is infinite, $\psi = 0$ along the surface and $\mathbf{n}\cdot\nabla\psi$ is indeterminate.

Further details of the formalism, such as the extension to many-body systems, systems with spin, density matrices and the relativistic formulation, will be introduced as we need them.

As presented so far, quantum mechanics appears essentially as a set of working rules for computing the likely outcomes of certain as yet undefined processes called 'measurements'. One might well ask what happened to the original programme embodied in the old quantum theory of explaining the stability of atoms as objective structures in spacetime. In fact, quantum mechanics leaves the primitive notion of 'system' undefined; it contains no statement regarding the objective constitution of matter corresponding to the conception of particles and fields employed in classical physics. There are no 'electrons' or 'atoms' in the sense of distinct localized entities beyond the act of observation. These are simply names attributed to the mathematical symbol ψ to distinguish one functional form from another. The original quest to comprehend atomic structure culminated in a set of rules governing laboratory practice.

It is not clear that one can apply the formalism to concrete problems without at least some mental image of the systems studied. For example, in applying quantum mechanics to the formation of molecules (§7.6), we have in mind some informal picture of the physical distribution of matter in space so that certain interaction energies may be deemed 'small' and treated as perturbations. And many orthodox accounts often slip into describing the motion of a wave packet through space as if it were something 'real'. Thus, in applying the theory physicists tacitly consider, and perhaps have to consider, quantum mechanics to be something more than just a means of correlating experimental results, and attribute to it the ability to describe some kind of reality beyond the phenomena.

The quantum formalism is not an entirely well-defined, closed structure. For example, there is no general rule for defining the order of the operators when products of classical variables are quantized, and the treatment of time is an open question (cf. §5.5). But even for that part which is uncontroversial, including the rules just stated, the 'interpretations' that may be put on the formalism are open-ended constructions about which there is no global consensus. Their limits and consistency are not established. Thus the common assumption that experiment cannot decide between the various interpretations if these reproduce all the usual quantal predictions should be viewed critically.

1.3 The completeness assumption

It was observed in §1.2 that the rules of quantum calculus cannot be stated without some reference to the physical interpretation of the mathematical symbols. In this connection it may be noted that the notion of the 'quantum state' used in the sense of some attribute *of* a physical system, although widely employed in presentations of the subject, is not a concept that appears in Bohr's analysis of the interpretative problems of quantum mechanics (e.g. Bohr, 1934, 1949, 1958). Bohr emphasizes the necessity of a complete specification of the experimental conditions for an unambiguous application of the formalism. Moreover, he requires that these conditions must be expressed in terms drawn from classical physics, for it is only in this way that he considers unambiguous communication in physics is possible. The predictions embodied in the ψ-function then pertain to the total classically-described experimental phenomenon and not to the independent properties of an 'object'. Bohr (1948) summarized his view thus:

> The entire formalism is to be considered as a tool for deriving predictions, of definite or statistical character, as regards information obtainable under experimental conditions described in classical terms and specified by means of parameters entering into the algebraic or differential equations of which the matrices or the wave-functions, respectively, are solutions. These symbols themselves, as is indicated already by the use of imaginary numbers, are not susceptible to pictorial interpretation; and even derived real functions like densities and currents are only to be regarded as expressing the probabilities for the occurrence of individual events observable under well-defined experimental conditions.

Words like 'position' and 'spin' are not meant to be understood in the way we might think, or indeed how we really do continue to think, but as convenient shorthands for macroscopic and technically complex instruments designed to 'measure' these 'properties'. And the following words are attributed to Bohr (quoted by Jammer (1974, p. 204)): 'There is no quantum world, there is only an abstract quantum physical description. It is wrong to think the task of physics is to find out how nature *is*. Physics concerns what we can say about nature.'

Much of Bohr's argument is concerned with demonstrating that any attempt to construct a physical explanation of the formalism inevitably leads to ambiguity and confusion. He saw in quantum mechanics ground for an application of his philosophy of complementarity, but we emphasize that this view cannot be *derived* from quantum mechanics. In this approach the wave and particle aspects of matter are supposed to be revealed by complementary but mutually exclusive experimental arrangements. The purpose of complementarity is not to go beyond classical concepts but to preserve them by limiting their range of application. In this sense quantum mechanics is claimed to introduce no new physical concepts. Nevertheless, in the process of arguing for this programme Bohr brought out, perhaps in a rather indirect manner, what may be the really novel feature of the quantum description, namely its holistic character, that the observer cannot be excluded from an account of the physics she observes.

Bohr assumes that everyday language, and its refinement in classical physics (particles, fields, position, momentum, energy), is a natural and necessary mode of discourse for human beings to communicate their experiences unambiguously, and that it is unproblematic. This seems to ignore an essential component of language that it is a function of, and contributor to, a changing social context. Classical physics took millennia to develop and takes years of schooling to learn. The range and content of 'everyday experience' is constantly being enriched and altered and it is hard to see what is natural about its state in our particular epoch. Why should everyday language not eventually come to include concepts that do enable quantum phenomena to be conceptually analysed? Actually, this insistence of Bohr on the necessity of a classical mode of discourse is only of historical interest, for we now know that we can indeed develop concepts to analyse the individuality of quantum experience (for example, through the theory described in this book), *if we countenance the introduction of new concepts that go beyond the classical paradigm.* In this connection we might note that the appearance of complex numbers in quantum mechanics, adduced by Bohr in the passage cited above as indicative of the impossibility of a pictorial representation of the phenomena, is actually irrelevant to this question (cf. the presentation of classical mechanics in terms of complex fields in Chap. 2). Further difficulties with Bohr's approach are discussed in Chap. 8.

Although the claim is widely disseminated that Bohr's 'Copenhagen interpretation' is the most consistent and satisfactory interpretation of quantum mechanics available, it is remarkable that there is no source one can turn to for an unambiguous rendering of Bohr's position about which there is general agreement. Scholars give varying accounts of his work (cf., e.g., Folse (1985) and Murdoch (1989)). Textbooks do not apply Bohr literally. When one investigates the properties of the hydrogen atom, for example, one solves the Schrödinger equation for two charged particles interacting via a Coulomb potential, i.e., for a system in its own right. One hardly bothers to specify the experimental arrangements necessary to verify the predictions contained within the wavefunction that are calculated in this way. Indeed, the natural inclination of scientists is to attempt to visualize the inner workings of an atom, for example by treating $|\psi|^2$ as a kind of objective 'charge density' (cf. §4.5). In doing this they move beyond the interpretation of ψ that pertains simply to the relative frequencies of measurement results to attempt a description of the essence behind the phenomena. This is the unspoken contradiction at the heart of quantum physics: physicists do want to find out 'how nature is' and feel they are doing this with quantum mechanics, yet the official view which most workers claim to follow rules out the attempt as meaningless!

The views of the principal exponents of the conventional interpretation of quantum theory (Bohr, Heisenberg and von Neumann) will be examined in the course of this book, particularly in Chap. 8. Although their analyses differ in certain important respects, they share a common theme which is at the core of the conventional interpretation presented in most textbooks. This may be stated as the following assumption of 'completeness':

The wavefunction is associated with an individual physical system. It provides the most complete description of the system that is, in principle, possible. The nature of the description is statistical, and concerns the probabilities of the outcomes of all conceivable measurements that may be performed on the system.

In practice one checks these statistical predictions through the relative frequencies of the results of measurements carried out on an ensemble of identically prepared (i.e., same ψ) systems. What this assumption boils down to is the proposal that if ψ is held as descriptive of a single system, the rules set out in §1.2 constitute *in themselves* the entire physical theory. Notice that it is asserted that not only is this postulate sufficient for interpreting quantum theory, but that it is *necessary*.

We emphasize the speculative character of the completeness assumption. It is not forced upon us by the experimental facts and the more detailed theory of individual processes that it forbids is not excluded by the generally agreed formalism. A difficulty with it is that it takes for granted that the formalism is essentially closed and unambiguous. But the latter is not completely well defined and in an area such as the problem of time (cf. §5.5) the completeness assumption becomes vulnerable and open to question.

Because it is postulated that predicting the results of experiments exhausts the possibilities of description of an individual system that are even conceivable, the indeterminism implied in the statistical interpretation of ψ is not of the kind encountered in classical statistical mechanics. There the evolution of a system is unpredictable due to our ignorance of an, in fact, well-defined physical state of an individual system. Here no meaning is to be attributed to the notion of a state beyond that encoded in ψ and hence the indeterminism is in some sense supposed to be intrinsic to the very nature of the system. Of course, if one relinquishes the assumption that material systems of the same kind (e.g., electrons) are distinguished from one another by virtue of having a well-defined location in spacetime, physical processes involving them will appear to be inherently indeterminist since systems that are prepared identically in the quantum mechanical sense (same initial wavefunction) and subjected to the same forces (same Hamiltonian) behave differently (points appear apparently at random on a detecting screen).

Following Einstein (§1.4) we shall locate the origin of the difficulties and paradoxes encountered in quantum theory in the attempt to squeeze the entire physical content of the theory of an individual system into the straitjacket of the statistically interpreted ψ-function. It has allowed quantum mechanics to acquire an aura of 'magic' in which 'smoky dragons' are invoked in discussions of two-slit interference effects and mind is held actively to influence generally occurring physical processes. The various difficulties of interpretation largely evaporate if the completeness assumption is relinquished in favour of less restrictive assumptions where one admits the possibility that other physical elements may enter the theory and that ψ has a significance beyond the mere specification of probabilities. Then one can indeed use quantum mechanics to find out 'how nature is'.

The historical triumph of the Copenhagen interpretation seems to be a rather fortuitous affair and we do not consider that the exclusion of other valid points of view has had a beneficial influence on the development of theoretical science. The legacy of the orthodox view has been a lopsided presentation of the subject where great emphasis is laid on what we *cannot* know about nature. The only research it appears to have stimulated is the attempt to understand and/or refute it, and the latter has hardly been encouraged. Because there is nothing in the theory to discuss but the results of experiments, it has also contributed to the notion that science is essentially concerned with the prediction and control of physical phenomena, and that progress in physics is most likely to come about through the manipulation of formalisms rather than the sharpening of our conception of reality. This is somewhat odd in a field that prides itself on the clarity and exactness of its thought and expression. The elegance of modern theoretical physics is largely to be found in its formal languages, not in the images with which it seeks to comprehend the world. Too often a concept is judged on what it 'predicts' and on whether a test can be proposed at its moment of inception to demonstrate its 'truth'. But analysed in these terms the transformation between Ptolemaic and Copernican cosmology is incomprehensible, for what was at issue there was a novel perception of how nature is, and this could not be immediately 'proved' by a new 'prediction'. To find a test, ideas must first be nurtured; in the meantime, they should be assessed according to different criteria, such as their explanatory power.

In the following the important issue is not so much the denial of causality in the processes governed by quantum mechanics, but the claim that no model at all can be constructed of an individual system. This latter point is a fundamental component in Einstein's critical analysis, as we now see.

1.4 Einstein's point of view

There is a popular view of Einstein in relation to the quantum theory which holds that he was unable to assimilate the revolutionary changes in world view apparently required by the new theory, that what bothered him most was the elimination of determinism from fundamental physics ('God does not play dice'), and that he 'wasted' the last 30 years of his scientific life in a fruitless quest to reestablish old-fashioned classical determinism as the ground of physical theory, having been 'beaten' by Bohr in their famous dialogue. Contrary to this view, consideration of what Einstein actually said in his public and private writings in the quantum period from the late 1920s to his death reveals a rather different picture. He, in fact, offered a carefully crafted critique that has never been satisfactorily answered, or in some ways even addressed, by the quantum establishment, and which, moreover, endures today. Pertinent references are the articles of Einstein (1936, 1940, 1948, 1949), the letters in Prizbram (1967) and Born (1971), and the analyses of Ballentine (1972) and Fine (1986).

A consistent theme in Einstein's commentary is the incompleteness of the theoretical description provided by the ψ-function, and his advocacy of a definite counter-interpretation.

To illustrate what he means by 'incomplete', Einstein (1949, p. 667) considers the case of radioactive decay in which an α-particle is emitted from an atom localized practically at a point. This system may be modelled by a closed potential barrier which at $t = 0$ encloses the α-particle. As time passes, the wavefunction, initially finite only inside the barrier, leaks into the surrounding space. According to the usual prescription this function yields the probability that at some instant the particle is found in a certain portion of the external space. Yet the wave may take many centuries to expand into the outer space whereas the particle may be found there after only a relatively short time. The wavefunction does not therefore imply any assertion concerning the time instant of the disintegration of the radioactive atom. That is, it does not describe the actual individual event revealed by the detector, including its cause.

If it is reasonable to suppose that the individual atom really does have a definite moment of decay, one may conclude, according to Einstein, that the ψ-function does not provide a complete description of the individual; it must be considered incomplete.

An objection may be raised that one is concerned here with a microscopic system of which we can have no direct knowledge, and that one can only claim there is a definite moment of disintegration if this can be determined empirically. To counter this, Einstein amplifies his example of incompleteness to the macroscopic scale (his version of the cat problem of Schrödinger (1935b)) by including the detector (Geiger counter) and a registration-strip, upon which a mark is made when the detector fires, in the entire system to which quantum mechanics is to be applied. After a suitable time has elapsed we expect to find a single mark on the strip. Yet the theory only offers the relative probabilities for the location of the mark *if this is observed.* It does not describe the objective definiteness of the mark as a property of the total system *per se.*

In this example the lack of description of the actual moment of decay is translated into a failure to describe the location of the mark on the strip. Einstein admits as a logical possibility that the mark becomes definite only when the strip is observed but states that, because we are now entirely within the sphere of macroscopic concepts (Einstein, 1949, p. 671),

> ... there is hardly likely to be anyone who would be inclined to consider it seriously ... One arrives at very implausible theoretical conceptions, if one attempts to maintain the thesis that the statistical quantum theory is in principle capable of producing a complete description of an individual physical system.

Einstein gave further arguments that quantum mechanics works with an inadequate description of individual systems in connection with the classical limit of the theory (a similar example to the macroscopic example above – see §6.5) and correlations in many-body systems (Chap. 11). His method was to show that those who adhere to the completeness assumption are compelled to adopt 'unnatural theoretical interpretations', in which according to him no one seriously believes.

It will be noted that in the example just described Einstein's concern is the *indeterminateness* of the world according to the usual quantum description of physical events, i.e., its failure to describe a reality comprising independent substantial objects beyond the phenomena. The complex function ψ simply does not exhibit any feature that could be put into correspondence with the (presumed) real state of affairs *viz.* that matter has substance and form independently of whether or not it is observed. Einstein's argument does not signify a general distaste for statistical theories in physics (he did after all develop the theory of Brownian motion). In his view, the indeterministic aspect of quantum mechanics follows from the failure to provide a complete description and not because it is an intrinsic characteristic of matter. In a letter written to Schrödinger in 1950 he says (Prizbram, 1967, p. 40) '... it seems certain to me that the fundamentally statistical character of the theory is simply a consequence of the incompleteness of the description.' In Einstein's programme, resolving the difficulty of describing a determinate reality entails constructing a causal (determinist) description, because he felt that this is a basic requirement of a complete physical theory (cf. Fine (1986, p. 103)). That is, in the process of making microphysics determinate, it would cease to be intrinsically statistical.

Of course, although linked in Einstein's critique, the requirements of determinism and determinateness are logically distinct. It is useful to separate the problem of completing quantum mechanics as a broad research programme from the specific brand of completion that invokes a deterministic description. As an analogy we may think of the classical theory of Brownian motion (Ballentine, 1972). This provides an objective description of a particle where individual events have definite antecedents, but the process is innately indeterminist and furnishes only statistical predictions.

As a way of resolving the interpretative dilemma within the terms of reference employed in quantum mechanics, Einstein proposed his own interpretation in which he *advocates* Born's statistical postulate, but interpreted in the sense that ψ pertains not to a single physical system but rather to an *ensemble* of systems. In this view ψ is admitted to be an incomplete representation of actual physical states and plays a role roughly analogous to the distribution function in classical statistical mechanics. Unfortunately, Einstein did not develop this idea sufficiently far for it to be clear exactly what it entails. He did not present it as an independent point of view but rather invoked it in the context of his examples designed to illustrate the untenability of the completeness assumption. He consistently claimed that the 'ensemble interpretation' dissolves the various difficulties and paradoxes flowing from the postulate of completeness, but he never explained precisely how.

Einstein's interpretation of the ψ-function has subsequently been developed into the 'statistical interpretation of quantum mechanics' (Ballentine, 1970; for a review see Home & Whitaker, 1992). Here it is asserted that ψ refers only to ensembles and that these are composed of particles pursuing definite (but unknown) spacetime trajectories. The only meaning attributed to $|\psi|^2$ is a statistical one, that it determines the relative frequency with which positions are realized in an ensemble of similarly prepared systems. Although it adopts a more modest and temperate line than the Copenhagen interpretation, a difficulty with this proposal is that one is not told what the laws are that govern the motions of the particles which are now admitted to exist objectively. Clearly, they cannot be the classical Newtonian laws because the ensemble of individual motions must reproduce the predictions of quantum mechanics, which in general contradict those of classical mechanics. In remaining agnostic on the issue of the laws obeyed by the particles, the statistical interpretation does not seem to offer any greater insight into the nature of quantum phenomena than the usual view. It is akin to treating the evolution of systems in classical statistical mechanics as defined just by the Liouville equation and not enquiring into the latter's origin in Hamilton's equations governing the underlying ensemble elements. This gap in the statistical interpretation is, of course, filled by the causal interpretation. But in the process it is found that the wavefunction cannot have the significance simply of encoding statistical information, but actually acquires the highly nonclassical property of being a physical component of each ensemble element, something not envisaged in the statistical interpretation. In this sense, the causal interpretation may be viewed as conceptually lying somewhere between the Copenhagen and statistical interpretations; it posits that ψ is in itself an incomplete description of an individual system, but in completing the theory, ψ is associated with the individual.

Einstein's critique, particularly his insistence that a deeper explanation of the phenomena correctly correlated by quantum mechanics is possible and necessary, provides the context for the theory presented in this book. Many statements made subsequently evoke his sentiments, without always directly referring to one of his utterances (for his reaction to the causal interpretation see §1.5.2).

1.5 The causal interpretation

1.5.1 De Broglie and Bohm

There was a general movement in theoretical physics in the 1920s against the idea that individual atomic events could be visualized as parts of causally connected sequences of spacetime processes. In the paper where he proposed the probability interpretation of the wavefunction, Born (1926) wrote: 'I myself am inclined to give up determinism in the world of atoms. But that is a philosophical question for which physical arguments alone are not decisive.'

It seems to have been regarded as almost axiomatic that the trajectory concept of classical mechanics is incompatible with wave mechanics. An argument in this direction was offered by Schrödinger (1926a,c, 1928), based on Hamilton's analogy between the rays of geometrical optics and mechanical paths. Schrödinger first asserted that in optics the conception of rays is well defined only in the geometrical limit and that in domains where the finiteness of the wavelength becomes relevant it loses nearly all significance because, even in homogeneous media, the 'rays' would be curved and appear mutually to influence one another. Then by analogy he claimed that the notion of the path of a mechanical system in ordinary mechanics likewise becomes inapplicable in cases where the de Broglie wavelength is comparable to characteristic lengths associated with the orbit, and indeed that it entails a contradiction. Although he was not a supporter of the emerging Copenhagen interpretation, Schrödinger proposed that the path should be replaced by the wave and have only an approximate significance in what would be the analogue of the geometrical optics limit. An obvious reply to this is as follows. Whatever the merits of the argument for excluding the ray concept from undulatory optics may be (and this is not a closed subject, see §12.6), the analogy drawn by Schrödinger between wave mechanics and wave optics is not an exact one (the mathematical theories are not in one-to-one correspondence) and the case against the meaningfulness of material paths in wave mechanics cannot therefore be regarded as proved, at least not on these grounds.

The idea that one should introduce the new matter wave not instead of the material point but as coexisting with it was advanced in this period by Louis de Broglie (1926a,b, 1927a,b,c). In the nonrelativistic Schrödinger case de Broglie suggested that the wavefunction is associated with an ensemble of identical particles differing in their positions and distributed in space according to the usual quantum formula, $|\psi|^2$. But he recognized a dual role for the ψ-function; not only does it determine the likely location of a particle, it also *influences* the location by exerting a force on the orbit. It thus acts as a 'pilot-wave' that guides the particles (only one of which actually accompanies each wave) into regions where ψ is most intense. For a scalar external potential the law of motion of the system point is that of the classical Hamilton–Jacobi theory according to which the possible paths are orthogonal to the surfaces of constant phase (de Broglie generally presented his theory in terms of the relativistic Klein–Gordon equation which was subsequently found to present severe interpretative problems – see §12.1). In fact, the pilot-wave idea is a truncated version of de Broglie's complete proposal in which he envisaged the particle as being represented by a singularity in a second field introduced in addition to the ψ-wave, rather than just treating it as a classical-style point. We shall not discuss this 'double-solution' interpretation further (see e.g. de Broglie (1956); Jammer (1974, p. 44)).

One may say that the contradiction perceived by Schrödinger, that the ensemble of paths seem mutually to influence one another, had been removed in de Broglie's approach by treating the wave itself as an agent that causes the paths to curve (in addition to classical forces).

De Broglie presented his proposal at the 1927 Solvay conference (*Electrons et Photons*, 1928, pp. 105–32). In particular, he applied his guidance formula to compute the orbits for the hydrogen atom stationary states (see §4.5). The approach met with a general lack of enthusiasm. Although it was discussed, only Einstein said de Broglie was right to search in the direction of a particle interpretation, although he did not endorse the specific model described (*Electrons et Photons*, 1928, p. 256). Kramers, while not questioning the possibility of tracing precise orbits, remarked that he could see no *advantage* in doing so (*Electrons et Photons*, 1928, p. 266). Pauli presented a detailed objection (*Electrons et Photons*, 1928, pp. 280–2; see §7.5.2) which de Broglie attempted to answer. However, the unfavourable climate, presumably compounded by Heisenberg's discovery of the 'uncertainty' relations, eventually led him to abandon his programme and indeed he soon began to propagate arguments against it (de Broglie, 1930). De Broglie returned to research in this area only 25 years later when in 1952 David Bohm rediscovered the approach and developed it to the level of a fully fledged physical theory (the story is told in de Broglie (1953a, 1956)).

In the intervening period Rosen (1945) had noted the possibility of retaining the particle picture in the quantum domain but was led to give up his proposed interpretation because he felt it was inconsistent with the existence of interference phenomena.

Bohm's classic pair of papers (Bohm, 1952a,b) remain the starting point for anyone wishing to find out about the de Broglie–Bohm theory. Bohm showed conclusively by developing a consistent counterexample that the assumption of completeness described in §1.3, a notion that pervaded practically all contemporary quantal discourse, was not logically necessary. One *could* analyse the causes of individual atomic events in terms of an intuitively clear and precisely definable conceptual model which ascribed reality to processes independently of acts of observation, *and* reproduce all the empirical predictions of quantum mechanics. Bohm's model is essentially de Broglie's pilot-wave theory carried to its logical conclusion.

But its significance goes beyond a mere existence proof, a kind of theoretical game of no direct value to the practising quantum physicist. A phrase such as 'an electron moves along the x-axis' is no longer simply an aid to calculating the wavefunction but refers to an objective process engaged in by a material system possessing its own properties through which the appearances (the results of successive measurements) are continuously and causally connected. It is thus very much a 'physicist's theory' and indeed puts on a consistent footing the way in which many scientists instinctively think about the world anyway. This comes about not merely through an *extension* of classical notions but requires the development of a new physical intuition. Bohm locates the novelty of quantum mechanics not in its statistical or discrete aspects but in a new physical conception of the state of a system (mathematically described by Schrödinger's wavefunction) that manifests itself in the motion of particles through a new type of potential, the quantum potential. The resulting theory stands in a clear and obvious relation to its classical counterpart (Chap. 6). The principal feature it shares with the classical paradigm is that the individuality of experience is comprehensible, but it diverges in its new notion of state.

Bohm applied his interpretation to a range of examples drawn from nonrelativistic quantum mechanics and speculated on possible alterations in the particle and field laws of motion such that the predictions of the (modified) theory continue to agree with those of quantum mechanics where this is tested but disagree in as yet unexplored domains. In Bohm's theory Born's statistical postulate is not dropped but is incorporated as a special case of the more general conception that $|\psi|^2$ represents the likely *current* location of a particle. Born's assumption follows when the theory is applied to measurement processes, the analysis of which is a significant feature of Bohm's contribution. He lays great emphasis on the creation of the outcome of the measurement through the interaction between the system and apparatus; in general, this does not passively reveal the premeasurement value of a physical quantity. Bohm extended his approach to include the second quantized electromagnetic field and also answered the earlier objections of Pauli, de Broglie and Rosen cited above. It should be noted that Bohm took issue with de Broglie's conception of light in which 'photons' are conceived as massive corpuscles moving within the electromagnetic guiding field, and proposed instead that the only 'real' parameters are the field coordinates and their conjugate momenta (see Chap. 12).

Because the *causes* of microevents could be analysed in their individuality and were no longer treated as irreducibly indeterminist and inexplicable, the de Broglie–Bohm proposal came to be known as the 'causal interpretation' of quantum mechanics. But while it indeed sets up a correspondence between all the mathematical symbols appearing in the formalism and physical properties of the wave-particle composite, the theory actually goes beyond the mere interpretative debate concerning the assignation of meanings to the symbols because it adds not just to the concepts but to the formalism itself, through the particle law of motion (cf. §5.5). Its description as a causal 'interpretation' therefore appears to be inadequate. Since it is essentially a novel theory of material motion (the first in quantum mechanics, in fact), a more appropriate title is 'the quantum theory of motion'. We shall use the various terms 'pilot-wave', 'causal interpretation' and 'quantum theory of motion' interchangeably in this book. But however we think of it, this theory of motion is not presented as a conceptually closed edifice offering the final word on quantum mechanics, and its originators never intended it to be this. Rather, it is a view worth developing for the insight it provides, and as a clue for possible future avenues of enquiry.

A few papers appeared in the wake of Bohm's articles raising technical questions connected with his approach, notably by Takabayasi (1952, 1953; replied to by Bohm, 1953b). The issues raised in these various articles will be discussed in later chapters. In this period de Broglie developed a small school of collaborators, including Jean-Pierre Vigier (see de Broglie (1956), Vigier (1956)). A good review of developments in the theory up to the mid-1950s is given by Freistadt (1957) and a qualitative account by Bohm (1957). A point to note is the attempt of Bohm (1953a) to demonstrate within the framework of the causal axioms why the probability distribution should be $|\psi|^2$, and also the attempts to include in the theory nonrelativistic systems with spin (Chap. 9) and systems governed by the Dirac equation (Bohm, 1953b; de Broglie, 1956; Takabayasi, 1957; see §12.2).

In the main the commentary of other physicists, particularly the Founding

Fathers (see below), was, where it existed, negative, and the theory did not enter the mainstream of physics either as a research topic or in textbooks. In fact, in the following 25 years only occasional and sporadic references were made to the de Broglie–Bohm theory. Books were written, courses taught and research conducted as if what Bohm had demonstrated to be possible was still, in fact, impossible. Although de Broglie continued to advertise the idea in his books, and Bohm worked on hidden-variable theories, no development or application of the pilot-wave theory was made.

A solitary and notable exception is the work of Bell (1966) who made reference to the nonlocality inherent in the de Broglie–Bohm picture (Chaps. 7, 11). Raising the question of whether this is a generic feature of all 'hidden-variable' completions of quantum mechanics, Bell (1964) was led to his inequality distinguishing a class of local theories from quantum mechanics, a step which brought the issue within the realm of experimental physics.

Belinfante (1973) and Jammer (1974) respectively published technical and nontechnical abridged accounts of the theory but it was not until the late 1970s that serious interest was rekindled when the trajectories corresponding to the two-slit experiment were explicitly displayed in computer graphics (Philippidis, Dewdney & Hiley, 1979; §5.1). During the 1980s developments took place in several branches of the theory, as will be discussed in detail later. This was motivated in part by continued dissatisfaction with the conventional solution to the problems of interpretation, and a certain thawing in the attitude that these matters were all settled long ago. Also, what had for a generation been *gedanken* experiments, such as single-photon interferometry and Einstein–Podolsky–Rosen correlations, could now actually be performed. In this recent period the de Broglie–Bohm idea has featured in some popular accounts of quantum mechanics (e.g., Squires (1986)) and Bohm has returned to actively develop the theory, the principal new element in his work being the proposal that the quantum potential may be interpreted as a kind of 'information potential' (Bohm, 1987). The approach has been eloquently defended in several articles by Bell (1987). Some criticisms of the approach have appeared (e.g., Tipler (1984, 1987) (for a reply see Dewdney, Holland, Kyprianidis & Vigier (1986)), de Muynck (1987)) and comparisons have been drawn between the pilot-wave theory and interpretations other than the 'conventional' one, such as the 'many-worlds' picture (e.g., Bell (1987), Bohm & Hiley (1987), Zeh (1988)), a subject we shall not discuss.

Nevertheless, at the time of writing it is fair to say that the de Broglie–Bohm theory of motion is still marginalized and, when it is referred to, often misrepresented. The sustained technical examination of its novel features, a prerequisite before any decision can be taken regarding its value, has not been undertaken. There has been some discussion of why this should be so (e.g., Bohm & Peat (1989) and the following two subsections) but there clearly remains considerable scope for analysing the social relations of science in the context of this physical theory.

The causal interpretation may be viewed in the context of Einstein's critique as a concrete proposal for how a complete description of individual events may be obtained, but the deterministic model employed is by no means the only possible completion that is conceivable. As observed in the Preface, reintroducing determinism into microphysics was a means to the end defined as accounting for the individuality of physical systems, but other avenues are open through which one may achieve the same goal. An attempt at providing a complete description along indeterministic lines was proposed by Bohm & Vigier (1954). The idea is to suppose that the particle is constantly subjected to random perturbations coming from some background source, such as random fluctuations in the ψ-field, so that its motion is akin to a kind of Brownian movement and hence deviates from the deterministic law of the basic pilot-wave theory. The latter now only describes the mean motion and the particle may jump between the mean flow lines. By means of this further postulate, it can be proved that an arbitrary initial distribution of positions will decay in the course of time to the $|\psi|^2$-distribution of quantum mechanics (for the proof see also Belinfante (1973, p. 186); see also Valentini (1991)). This approach is connected with the subsequently developed stochastic interpretation of quantum mechanics (Nelson, 1966, 1985; Jammer, 1974, Chap. 9; Vigier, 1982) in which it is demonstrated how the Schrödinger equation is implied if particles in a stochastic process are subjected to a particular kind of force law. It is to be emphasized that the Bohm–Vigier and Nelson programme of deriving the laws of quantum mechanics from the laws obeyed by particles at some deeper or subquantum-mechanical level, whatever its merits, is logically unconnected with the basic de Broglie–Bohm pilot-wave idea which constitutes in itself a self-contained and consistent theory of motion that does not *require* the assumption of a further as yet unrevealed layer of physical reality. In this book we shall be concerned solely with working out the original deterministic de Broglie–Bohm proposal and do not discuss possible stochastic extensions.

1.5.2 What the great men said

The reaction to Bohm's work by the Copenhagen establishment was generally unfavourable, unrestrained and at times vitriolic (e.g. Rosenfeld (1958)).

We first consider the response of Heisenberg (1955; 1962, Chap. 8) to

Bohm's contribution, and see in outline how the points he raises may be answered (see also Bohm (1962) for a reply). Heisenberg first questions what it means to say that a wave propagating in configuration space is 'real'. His objection to this notion is based on his assertion that only 'things' in three-dimensional space are 'real'. He offers no logical or scientific argument to show that examining the possiblity of the physical reality of multi-dimensional spaces is a fruitless enterprise. It is useful to recall here the Kaluza–Klein programme in general relativity where physicists contemplate spacetimes of dimension greater than four as a valuable aid to comprehending and unifying the basic physical interactions.

Heisenberg goes on to bemoan what he considers an asymmetrical treatment of position and momentum in the causal interpretation, and the apparent breaking of a fundamental symmetry of the quantum theory. This criticism seems to confound the quite reasonable asymmetry in the physical interpretation, which assumes that the preferred arena for physical processes is position space, with the symmetry exhibited by the mathematical theory (cf. §3.12). The asymmetry is also connected with the nature of measuring processes which generally entail active transformations and do not passively reveal preexisting states (Chap. 8). Classical physics exhibits in the canonical formalism an analogous feature of mathematical symmetry but physical asymmetry.

The 'hidden parameters', i.e., the particle orbits, are denounced by Heisenberg as a "superfluous 'ideological superstructure'" having little to do with immediate physical reality because the causal formulation generates the same empirical results as the Copenhagen view. But in Bohm's theory it is precisely the positions of particles that are recorded in experiments; they are the immediately sensed 'reality'.

Finally, Heisenberg mentions Bohm's tentative proposals for modifying the quantum laws of motion so as to permit an experimental test of the trajectory interpretation in a domain where the quantum theory may conceivably break down. This possibility he dismisses as akin to the 'strange hope' that someday it will turn out that sometimes $2 \times 2 = 5$. That is, Heisenberg believes that any alternative or even modification to the current quantum theory is logically impossible. He considers that Bohm's purpose was to return to classical physics and thus misses the key point about the de Broglie–Bohm proposal: it refutes the view that the actual individual facts of experience are in principle unimaginable.

The objection pertaining to the asymmetrical role of the position and momentum variables in the causal interpretation had been raised in an earlier paper by Pauli (1953). He considers that this feature renders the theory 'artificial metaphysics' and that any proposal to modify the formalism so as to demonstrate empirically the existence of particle tracks will inevitably conflict with established experimental facts. A point of interest raised by Pauli is why, in a theory which treats the laws of the ensemble as functions of individual laws, the probability distribution should be given by the quantum formula $|\psi|^2$ and not arbitrarily specified (a similar observation is made by Keller (1953)). Although a physical justification of this assumption is not a logical requirement of a causal theory (one may just take the probability to be $|\psi|^2$ with no further discussion, as indeed the orthodox view does), Bohm & Vigier (1954) did, in fact, supply an argument justifying the $|\psi|^2$-distribution through a development of the original model, as noted above. Commenting on Pauli's article, Born (1971, p. 207) wrote to Einstein that '... Pauli has come up with an idea ... which slays Bohm not only philosophically but physically as well'!

At a 1957 conference in Bristol, Rosenfeld repeated the charge that Bohm was engaged in 'metaphysics'. Bopp summarized the discussion thus: '... we say that Bohm's theory cannot be refuted, adding ... that we don't believe in it' (Körner, 1957, p. 51; Jammer, 1974, p. 296).

Viewed in retrospect, what is striking about the reaction to Bohm is not that the proponents of the orthodoxy should attempt to defend themselves vigorously, but the inadequate character of the arguments they adduced in their defence. Personal distaste regarding specific features of the theory, such as its asymmetry, or the *ad hoc* manner in which the quantum potential is introduced, or its intrinsic nonlocality, seems to us to be beside the point (they are among the criticisms that have been levelled by advocates of the approach). The point is that it demonstrates that quantum phenomena need not be sealed in black boxes and forever hidden from our conceptual gaze, as claimed by the Copenhagen lobby. Such a demonstration is not invalidated by dubbing it 'metaphysical' (this term carries with it the implied rebuke that Bohm was doing 'mere philosophy', the physicist's ultimate censure). In a climate more disposed to a spirit of free enquiry, Bohm's work would have been acclaimed rather than treated as an inconvenience and then ignored. As an example of how new physics may be generated by a consideration of alternative theories, we have already cited the case of Bell's theorem which followed directly from contemplation of the de Broglie–Bohm model and is widely regarded as one of the seminal discoveries of twentieth century physics.

The fact is, of course, that the assumptions made by the Founding Fathers in regard to the possibilities of visualizing the origin of quantum phenomena were equally 'metaphysical' and devoid of empirical support. One can indeed argue that Bohm had provided at least theoretical evidence against the orthodox view, whereas there was no counterevidence forthcoming, either theoretical or experimental, to exclude the pilot-wave. In fact, the establishment did have in its possession theoretical evidence which indeed purported to demonstrate the impossibility of theories, of the type propounded by Bohm, that add supplementary variables to the quantum formalism while reproducing all its empirical predictions: the 1932 theorem of von Neumann (1955). Given the existence of Bohm's counterexample, one might have expected that the causal interpretation and von Neumann's theorem would both have been subjected to a sustained theoretical analysis to discover in which the error lay. While the quantum nobility expressed (private) doubts that Bohm had been able to circumvent von Neumann's theorem (Hanson, 1969, p. 174), this analysis simply was not carried out (the story is told by Pinch (1977, 1979)). More than 10 years elapsed before von Neumann's assumptions were critically analysed and found to be wanting in that he supposes that the probability distribution of the supplementary variables has the same properties as the quantum mechanical distribution (Mugur-Schächter, 1964). This supposition entails, for example, that the mean values of the new variables are linearly superposable in the same way that the means of quantum mechanical observables are, and hence that the outcome of the measurement of a linear sum of operators is the linear sum of the outcomes of measurements of the operators individually (Bell, 1966). The causal interpretation does not fall within the scope of this assumption because for it measurements entail transformations of the system under investigation. These transformations are very specific to the observable 'measured' and, as we have said, such processes do not passively reveal premeasurement values. Bell (1966, 1982) identified similar problems with other 'impossibility proofs'. Eventually, it proved possible to show that quantum mechanics can always be supplemented by 'hidden variables' (Gudder, 1970).

In letters written in the early 1950s, Einstein (1989, p. 60) expressed solidarity with de Broglie in his search for a complete representation of physical reality. But he does not seem to have been overly impressed by the specific solution advanced in the pilot-wave theory. To Born (1971, p. 192) he wrote in 1952 'Have you noticed that Bohm believes (as de Broglie did, by the way, 25 years ago) that he is able to interpret the quantum theory in deterministic terms? That way seems too cheap to me.' While the de Broglie–Bohm proposal might be considered as a deterministic completion carried out in accordance with Einstein's critique, Einstein did not think that fundamental progress towards the discovery of a deterministic substructure could be made by a completion from within, i.e., simply by appending supplementary physical variables to an essentially unmodified formalism.

He expressed this in 1954 as follows (quoted by Fine (1986, p. 57)): 'I think it is not possible to get rid of the statistical character of the present quantum theory by merely adding something to the latter, without changing the fundamental concepts about the whole structure.'

Einstein was after a more radical completion which countenanced going beyond the classical concepts that Bohr had retained in his interpretation and which de Broglie and Bohm were using to describe the individual process (material points and forces). Generally, he felt that the quantum theory did not serve as a useful point of departure. De Broglie's notion of the 'double-solution' was closer in spirit to Einstein's field-theoretic approach than the basic pilot-wave model but was at that stage (and is still today) largely an unfulfilled programme. Still, despite his reservations, Einstein (1953) took Bohm's work seriously enough to offer an objection which raised an interesting issue, as we discuss in §6.5.

It might be argued that Einstein's negative attitude was a tactical mistake and that, whatever he perceived as its drawbacks, some model such as that of de Broglie and Bohm was better than none at all in countering the prevailing vagueness in interpretation, at least as a makeshift before a more satisfactory foundation could be found (the view taken by de Broglie and Bohm towards their theory). But it seems unlikely that Einstein's endorsement would have made much historical difference given that for 25 years he had been branded as a quantum dissident and that the precise nature of his own critique was not widely understood. After all, his solution to the problem of interpretation within the quantum scheme, the view that ψ refers only to ensembles, was widely advertised in his public writings and was itself ignored (although admittedly he never developed the idea very far).

1.5.3 Some objections

There is a sense in which physicists do not understand how it is even *possible* to have a causal theory of the de Broglie–Bohm type, and feel uneasy that there must be some inconsistency in it if it really does what it claims and covers all the empirical ground accounted for by quantum mechanics. But beyond aesthetic displeasure, which has always loomed large in discussions of it, to our knowledge no serious technical objections have ever been raised against the de Broglie–Bohm world view. Anticipating our subsequent detailed analysis, we here summarize and offer preliminary answers to some of the typical objections that have been advanced against this theory.

(1) *You cannot prove the trajectories are there*

Insofar as the quantum theory of motion reproduces the assertion of quantum mechanics that one cannot perform a precise measurement of position simultaneously with a precise measurement of momentum, this statement is true (§8.4). But this cannot be adduced as evidence against the tenability of the trajectory concept. Science would not exist if ideas were only admitted when evidence for them already exists. One cannot after all empirically prove the completeness postulate. The argument in favour of the trajectory lies elsewhere, in its capacity to make intelligible a large swathe of empirical facts.

(2) *It predicts nothing new*

First we note that the postulate of completeness predicts nothing about the details of a process beyond the distribution of measurement results. In contrast, the quantum theory of motion permits more detailed predictions to be made pertaining to the individual process (cf. $5.1.3). Whether these may be subjected to an experimental test is an open question (cf. §§5.5, 8.8).

(3) *It attempts to return to classical physics*

The deterministic model of wave and particle causally evolving from the past into the future is a particular solution to the problem of describing a *determinate* reality. Although de Broglie and Bohm have often been chastised for reintroducing the classical paradigm, this misses the key point that they are invoking a concept not anticipated in classical physics, that of a 'state' of a mechanical system that lies beyond the material points. The role of the trajectory is to bring out this new concept so sharply it cannot be ignored. This essentially nonclassical programme should be contrasted with that of Bohr who strove to leave intact as far as possible the classical concepts by restricting their applicability.

(4) *The price to be paid is nonlocality*

Nonlocality is an intrinsic feature of the de Broglie–Bohm theory (Chaps. 7, 11). This property does not contradict the statistical predictions of relativistic quantum mechanics (Chap. 12) but it is considered to be in some way a defect, the implication being that entirely local theories are preferable. Yet nonlocality seems to be a small price to pay if the alternative is to forego any account of objective processes at all (including local ones). Also, it is inconsistent to deny the logical possibility of a pictorial representation of the phenomena, and then lay down conditions for what such a picture should consist of when one is produced.

(5) *It is more complicated than quantum mechanics*

Mathematically, the quantum theory of motion requires a reformulation of the quantum formalism (not an alteration). The reason is that the usual presentation of the theory is not the one most appropriate to the physical interpretation. But, mathematically, the theory remains quantum mechanics. In particular, the quantum potential is implicit in the Schrödinger equation.

(6) *It is counterintuitive*

It certainly runs counter to classical intuition. The concept of 'intuition' is like that of 'human nature': it is a function of history and not eternally frozen. The notion that a body persists in a state of uniform motion unless acted upon by a resultant force would be counterintuitive to Aristotle but natural for Galileo. Quantum phenomena require the creation of quantum intuition.

(7) *There is no reciprocal action of the particle on the wave*

In classical physics there is a dialectical interplay between particle and field, each generating the dynamics of the other. In the pilot-wave model the dynamical connection is one way. Among the many nonclassical properties exhibited by this theory (cf. §3.3), one is that the particle does not react dynamically on the wave it is guided by. But while it may be reasonable to require reciprocity of actions in classical theory, this cannot be regarded as a logical requirement of all theories that employ the particle and field concepts, especially one involving a nonclassical field.

It will be noted how, having been chided for its classical pretensions (point (3)), the causal interpretation is admonished in points (4), (6) and (7) because it is not classical enough!

2

Hamilton–Jacobi theory

2.1 The need for a common language

If we wish to compare the methods, content, claims and experimental predictions of two physical theories we have to find some common ground between them. What is needed ideally is a *language* which embraces the essential elements of each theory as parts of a broader structure which transcends them both. There are two components to such a language, which might be called the formal and the informal. Briefly, by 'formal' we mean a precisely defined set of concepts and their relationships from which one can deduce unambiguous conclusions by a series of logical steps (mathematics); by 'informal' we mean the intuitive concepts and pictures that a theory employs in order to render intelligible the 'reality' for which it seeks to account. These two aspects are naturally closely connected. The possibility of constructing a language of this type is not given *a priori*. It may turn out that it is possible to develop only one of the two components, and the theories we want to compare may or may not be commensurable, or only partly so.

We shall examine this question in connection with the relation between two physical theories of matter: classical mechanics and quantum mechanics. This relation is subtle and operates on several levels and we shall return to it throughout the book. For the present we confine ourselves to some general remarks. Roughly speaking, we may say that classical mechanics as a distinct discipline constitutes a language possessing both formal and informal aspects. Newton's laws (or their refinements in the Lagrangian, Hamiltonian and Hamilton–Jacobi formalisms) allow us not only to predict the results of experiments on fields and particles, but also provide an *explanation* of these results in terms of a definite world view – that of mass points pursuing well-defined trajectories in space and time and interacting via preassigned potentials. Formally this is a theory of continuous functions in configuration or phase space and this aspect derives its physical meaning from a definite theory of matter and motion with which it is inextricably linked.

Quantum mechanics on the other hand possesses a sophisticated and highly developed formal language of linear operators in Hilbert space but, in the orthodox interpretation, only provides hints of what an informal language which would explain the results the formalism predicts might be like. Indeed, the absence of a clear physical picture has tended to lead to an identification of 'quantum reality' with Hilbert space, i.e., with the formal language.

At the present stage in the development of physics the possibility of an intuitive account is closely connected with the possibility of a theory of matter and motion as a process in space and time. In this book we shall therefore be mainly concerned with developing a language which allows us to comprehend microphysical phenomena in this way. Naturally, on its own the spacetime theory of matter and motion of classical mechanics is not broad or deep enough to include quantum mechanics. We have to develop a new language.

Several avenues of enquiry are open to us. We might try to express classical mechanics in a Hilbert space language with the hope that eventually this may lead to new insights into the quantum theory. Formally this can indeed be done; for example, canonical transformations may be shown to be equivalent to unitary transformations in a Hilbert space of square integrable functions on phase space (Koopman, 1931). While this implies the possibility of a formal comparison between classical and quantum mechanics, it does not seem to lead to a clear physical conception of the latter. In particular it does not lead to a spacetime picture of quantum processes.

Alternatively, one may try to introduce a kind of 'phase space' into quantum mechanics. Again, this can indeed be done and leads to a formulation of quantum mechanics in terms of, for example, Wigner functions (§8.4.3). This is a more promising approach but again it does not result in a theory of material processes in space and time.

We pursue here a different approach, one that associates a well-defined phase space (i.e., simultaneously real position and momentum variables) with a quantum mechanical system, but which is immediately tied to a picture of physical events as processes in spacetime. To locate what is new in quantum mechanics in relation to classical mechanics we formulate both as particular instances of (a suitably generalized) Hamilton–Jacobi theory. It turns out that the Hamilton–Jacobi theory admits a natural generalization which provides a language broad enough to embrace both theories, formally and informally. Formally there is a clear mathematical procedure as to how one passes from the quantum to the classical domain (see Chap. 6), and informally the language provides an unambiguous physical picture of waves, rays (trajectories) and their interrelationships. Although historically connected with just classical mechanics and field theory and the semiclassical approximation to quantum mechanics, the Hamilton–Jacobi method transcends its origins.

The aim of this chapter is to provide the relevant classical background so as to put the generalization required by quantum mechanics into context. It is not an intrinsic requirement of a causal spacetime theory of quantum processes that we employ the Hamilton–Jacobi language of waves and rays. Indeed a 'minimalist' version of the causal interpretation of quantum mechanics can be formulated without reference to it (Chap. 3). However, the Hamilton–Jacobi language affords considerable insight into the formal and conceptual structure of quantum mechanics, in particular its relation with classical mechanics, and we shall make extensive use of it in this book.

2.2 The Hamilton–Jacobi method in classical mechanics

2.2.1 Hamilton's principal function

We begin by reviewing some basic results in classical mechanics, in particular how a certain function, a solution of the Hamilton–Jacobi partial differential equation, facilitates the solution of the equations of motion. This method is due to Jacobi. For definiteness we shall talk in terms of the motions of particles, i.e., material points (in three-dimensional space or configuration space), but all our remarks and results apply to any physical system as usually treated by the methods of classical mechanics (e.g., rigid bodies and fields). Our treatment is nonrelativistic and we use absolute time as an evolution parameter. Further details may be found in the texts of Synge (1954), Landau & Lifschitz (1960), Lanczos (1970), Arnold (1978) and Goldstein (1980).

The following notation is used. The generalized coordinates are denoted by q_i, $i = 1, \ldots, n$, where n is the number of degrees of freedom of the system, $\mathbf{x} = (x, y, z)$ are the coordinates of a body in a Cartesian system, p_i represents the momenta in both general and Cartesian coordinates, and t is the time. Often we shall neglect the indices and write just q and p. These then stand for the points whose coordinates are q_i and p_i. It is assumed that a metric is given on configuration space so that we can form invariants. Under coordinate transformations q_i transforms as the components of a contravariant vector and p_i as the components of a covariant vector, but we shall not distinguish this in the notation (all indices are subscripts). No attempt is made at a rigorous presentation.

Starting from the Lagrangian $L = L(q, \dot{q}, t)$, where along a trajectory

$q = q(t)$ and $\dot{q} = \mathrm{d}q/\mathrm{d}t$, we define the momentum canonical to the coordinate q_i to be

$$p_i = \partial L/\partial \dot{q}_i. \tag{2.2.1}$$

The Hamiltonian is defined as a function on the phase space (q, p) by a Legendre transformation:

$$H(q, p, t) = \sum_i p_i \dot{q}_i - L(q, \dot{q}, t). \tag{2.2.2}$$

The equations of motion may be derived as follows. Consider two points in state space, (q_0, t_0) and (q, t), and the paths that may join them. Then the actual path traversed by the physical system is that which extremizes the action integral

$$I(q, t; q_0, t_0) = \int_{q_0,t_0}^{q,t} L(q, \dot{q}, t)\,\mathrm{d}t. \tag{2.2.3}$$

This is Hamilton's Principle which we write as

$$\delta I = 0 \tag{2.2.4}$$

and it implies the n second-order Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \qquad i = 1, \ldots, n. \tag{2.2.5}$$

Alternatively, we can substitute for L in (2.2.3) from (2.2.2) and write the action function as

$$I = \int \sum_i p_i\,\mathrm{d}q_i - H\,\mathrm{d}t. \tag{2.2.6}$$

Varying this with respect to the $2n$ independent variables q_i and p_i yields Hamilton's set of $2n$ first-order differential equations:

$$\dot{q}_i = \partial H/\partial p_i|_{q_j = q_j(t), p_j = p_j(t)}, \quad \dot{p}_i = -\partial H/\partial q_i|_{q_j = q_j(t), p_j = p_j(t)}. \tag{2.2.7}$$

for all i, j. Solution of eqs. (2.2.5) or (2.2.7) requires that we specify the initial coordinates and the initial velocities or canonical momenta.

In the Hamiltonian formulation of the theory we may replace the independent variables q_i, p_i by a new set of $2n$ independent variables Q_i, P_i:

$$Q = Q(q, p, t), \qquad P = P(q, p, t). \tag{2.2.8}$$

This is a coordinate transformation in phase space. The new set of coordinates is canonical if Hamilton's equations (2.2.7) retain their form under the transformation. If $K = K(Q, P, t)$ is the new Hamiltonian, we require

$$\dot{Q}_i = \partial K/\partial P_i|_{Q=Q(t),P=P(t)}, \qquad \dot{P}_i = -\partial K/\partial Q_i|_{Q=Q(t),P=P(t)}. \tag{2.2.9}$$

These equations may be derived from a variational principle applied to an action function of the form (2.2.6):

$$I = \int \sum_i P_i \, \mathrm{d}Q_i - K \, \mathrm{d}t. \tag{2.2.10}$$

The integrands in (2.2.6) and (2.2.10) then differ by a total time derivative:

$$\sum_i P_i \dot{Q}_i - K = \sum_i p_i \dot{q}_i - H - \mathrm{d}F/\mathrm{d}t, \tag{2.2.11}$$

where F is an arbitrary differentiable function of q, p, Q, P and t. By (2.2.8) only $2n$ of the total of $4n$ coordinates q, p, Q and P are independent and so F depends on just $2n$ of these coordinates, together with t. For example, it may have the form $F(q, Q, t)$, $F(q, P, t)$, $F(p, Q, t)$ or $F(p, P, t)$ or some other mixture of q, p, Q and P. Note that the new coordinates may not have the dimensions of position and momentum – they may represent some conjugate pair such as energy and time or action-angle variables.

Consider the case where $F = F(q, Q, t)$. Then from (2.2.11) we deduce that

$$p_i = \partial F/\partial q_i, \tag{2.2.12a}$$

$$P_i = -\partial F/\partial Q_i, \tag{2.2.12b}$$

$$K = H + \partial F/\partial t. \tag{2.2.12c}$$

Given F we can reconstruct from these relations the canonical transformation (2.2.8). To do this we solve (2.2.12a) for Q in terms of p, q and t and then (2.2.12b) for P. The new and old Hamiltonians are related by (2.2.12c). In view of this, F is called the *generating function* of the canonical transformation. Notice that F is not a function on phase space but relates two sets of coordinate systems on that space.

In order that we may use the relations (2.2.12) to solve the dynamical problem we need a further result: the motion of the system in time is equivalent to *the continuous unfolding of a canonical transformation.* To see this, consider the infinitesimal canonical transformation $Q_i = q_i + \delta q_i$, $P_i = p_i + \delta p_i$, where δq_i, δp_i are infinitesimal changes in the coordinates and momenta. To the first order in small quantities, the generating function can in this case be taken to be a function on the phase space labelled by q, p and it may be shown that

$$\delta q_i = \varepsilon \, \partial G/\partial p_i, \qquad \delta p_i = -\varepsilon \, \partial G/\partial q_i$$

where ε is an infinitesimal parameter and $G = G(q, p, t)$. Consider the case where $\varepsilon = \mathrm{d}t$, a small time interval, and $G = H$, the Hamiltonian. Then

$$\delta q_i = \mathrm{d}t\, \partial H/\partial p_i = \mathrm{d}t\, \dot{q}_i = \mathrm{d}q_i,$$

$$\delta p_i = -\mathrm{d}t\, \partial H/\partial q_i = \mathrm{d}t\, \dot{p}_i = \mathrm{d}p_i.$$

In other words, the infinitesimal canonical transformation generated by the Hamiltonian is precisely the physical change undergone by the generalized canonical coordinates of the system during the time interval $\mathrm{d}t$. Since an arbitrary canonical transformation can be built from a succession of infinitesimal transformations we conclude that the actual motion during a finite time interval of any system governed by Hamilton's equations is a continuous canonical transformation:

$$q = q(q_0, p_0, t), \qquad p = p(q_0, p_0, t) \tag{2.2.13}$$

where q_0, p_0 are the initial canonical coordinates. We shall find the generating function of this *finite* transformation below.

Inverting (2.2.13) to give q_0, p_0 in terms of q, p and t we see that the problem of motion is solved if we can find a canonical transformation from coordinates q, p at time t to a set of constant (in time) coordinates q_0, p_0 at some initial time t_0. Returning to (2.2.9), we see that this may be achieved if $K = 0$, for this ensures that the new coordinates are constant in time:

$$\dot{Q} = \dot{P} = 0. \tag{2.2.14}$$

From (2.2.12c) K will be zero if F satisfies the equation $\partial F/\partial t + H(q, p, t) = 0$. It is usual to denote F by S in this case. If we suppose for definiteness that F is a function of q, Q and t, and substitute for p in H from (2.2.12a), we obtain the *Hamilton–Jacobi equation*:

$$\partial S(q, Q, t)/\partial t + H(q, \partial S(q, Q, t)/\partial q, t) = 0. \tag{2.2.15}$$

The function S is called *Hamilton's principal function.*

Eq. (2.2.15) is a first-order partial differential equation in the $(n + 1)$ variables $q_1, \ldots, q_n, t$. Since S itself does not appear in the equation, a complete solution involves n nonadditive constants $\alpha_1, \ldots, \alpha_n$: $S = S(q, \alpha, t)$. We can take the αs to be the new coordinates as in (2.2.15): $\alpha_i = Q_i$. But more generally the αs may be any function of the Qs or, if some other form is assumed for the generating function, of some combination of the Qs and Ps.

In order to solve the equations of motion by the Hamilton–Jacobi method we proceed as follows. Given a Hamiltonian $H = H(q, p, t)$ we substitute $p = \partial S/\partial q$ and write down the Hamilton–Jacobi equation

$$\partial S(q, \alpha, t)/\partial t + H(q, \partial S(q, \alpha, t)/\partial q, t) = 0. \tag{2.2.16}$$

We seek a complete solution for S in terms of n nontrivial integration constants $\alpha_1, \ldots, \alpha_n$. The new constant coordinates Q_i may be chosen as any n independent functions of α_i: $Q_i = \gamma_i(\alpha_1, \ldots, \alpha_n)$, so that we may write $S = S(q, \gamma, t)$. The transformation equation (2.2.12b) introduces the new constant momenta

$$P_i = -\partial S/\partial \gamma_i = \beta_i. \tag{2.2.17}$$

This relation is the Jacobi law of motion: it can, in principle, be solved algebraically for q_i in terms of t and the $2n$ constants β_i, γ_i. This is possible if $\det(\partial^2 S/\partial q_i\, \partial \gamma_j) \neq 0$; p_i is given by (2.2.12a). The solution is completed by expressing β_i, γ_i in terms of the actual initial conditions q_{0i}, p_{0i} of the problem. Evaluating the transformation equation (2.2.12a) at time t_0 gives

$$p_0 = \partial S/\partial q|_{q=q_0, t=t_0} \tag{2.2.18}$$

which implies a relation between q_0, p_0 and γ, and evaluating (2.2.17) at t_0 gives a relation between q_0, γ and β. Finally then we obtain

$$q = q(q_0, p_0, t), \qquad p = p(q_0, p_0, t). \tag{2.2.19}$$

By an identical procedure we can derive (2.2.19) if we use a generating function in which the nonadditive constants are expressed in terms of the new momenta: $P_i = \gamma_i(\alpha_1, \ldots, \alpha_n)$. Then the transformation equation (2.2.12a) is unchanged and (2.2.12b) is replaced by an expression for the new coordinates, $Q_i = \partial S/\partial \gamma_i = \beta_i$, which is now the Jacobi law of motion.

We have thus constructed a function, Hamilton's principal function, which generates a canonical transformation to coordinates and momenta that are constant along a trajectory. This establishes an equivalence between the $2n$ first-order Hamilton equations and the single first-order Hamilton–Jacobi partial differential equation. This type of relation is well known in the theory of differential equations where the mechanical paths are the characteristics of the Cauchy problem associated with the Hamilton–Jacobi equation.

2.2.2 The action function

An alternative perspective on the meaning of the Hamilton–Jacobi function S may be gained as follows. In formulating Hamilton's Principle (2.2.4) we considered the motion of a system between two given instants t_0 and t at which its coordinates are q_0 and q respectively. The values of I for neighbouring paths linking these fixed limits are compared, and the true path is the one for which I is an extremum. We shall now regard the action function I as a quantity associated with just the actual path traversed by the system and consider the change in I induced by variations in the final coordinate q and time t and the initial coordinate q_0, keeping the initial time t_0 fixed. That is, define

$$I(q, t; q_0, t_0) = \int_\gamma L(q, \dot{q}, t)\, \mathrm{d}t, \tag{2.2.20}$$

where the integral is evaluated along the extremal γ joining the point (q_0, t_0) to (q, t).

Taking the differential of I as a function of the variable coordinates it is possible to show that

$$\mathrm{d}I = p\, \mathrm{d}q - H\, \mathrm{d}t - p_0\, \mathrm{d}q_0, \tag{2.2.21}$$

where $p = \partial L/\partial \dot{q}$, $H = p\dot{q} - L$, $\dot{q}$ is the terminal velocity of the trajectory γ, and similar relations hold for q_0 and p_0. Writing

$$\mathrm{d}I = (\partial I/\partial q)\, \mathrm{d}q + (\partial I/\partial t)\, \mathrm{d}t + (\partial I/\partial q_0)\, \mathrm{d}q_0$$

and comparing with the right hand side of (2.2.21) we deduce that

$$p = \partial I/\partial q, \qquad H = -\partial I/\partial t, \qquad p_0 = -\partial I/\partial q_0$$

and hence that I satisfies the Hamilton–Jacobi equation:

$$\partial I/\partial t + H(q, \partial I/\partial q, t) = 0. \tag{2.2.22}$$

Indeed, (2.2.21) has the form of a canonical transformation (2.2.11) which trivializes the motion ($K = 0$), where I is the generating function and $q_0 \equiv Q$, $p_0 \equiv P$. The action function I is therefore a particular complete integral of the Hamilton–Jacobi equation, the n nonadditive constants being the initial positions q_{0i}, $i = 1, \dots, n$ and the new momenta being the actual initial momentum coordinates. In view of the identity of the action function with a form of Hamilton's principal function we shall henceforth denote both by the symbol S.

We started by showing how Hamilton's equations could be solved by integrating the Hamilton–Jacobi equation. The result just proved enables us to do the converse – to solve the Cauchy problem for the Hamilton–Jacobi equation (i.e., solve (2.2.16) subject to the initial condition $S(q, t_0) = S_0(q)$, ignoring the constants) by writing down the equivalent set of Hamilton equations and solving these by some other method. The resulting trajectory is a characteristic of the partial differential equation and $S(q, t)$ may be constructed by integrating the Lagrangian (obtained from the Hamiltonian by an inverse Legendre transformation) along the characteristic. The result is

$$S(q, t) = S_0(q_0(q, t), t_0) + \int_\gamma L(q, \dot{q}, t)\, \mathrm{d}t. \tag{2.2.23}$$

Notice that the function S obtained in this way depends on S_0 and may be quite different in form from a complete integral for the same problem found by a separation of variables, even though the paths associated with the two functions are the same. This point is considered in more detail in §2.3.

2.2.3 A single particle

In the following we shall often refer to the case of a single body of mass m in a scalar external potential V and described by a Cartesian coordinate system. The Lagrangian is given by

$$L(\mathbf{x}, \dot{\mathbf{x}}, t) = \tfrac{1}{2}m\dot{\mathbf{x}}^2 - V(\mathbf{x}, t) \tag{2.2.24}$$

from which the canonical momentum (2.2.1) is found to be

$$\mathbf{p} = m\dot{\mathbf{x}} \tag{2.2.25}$$

and the Hamiltonian (2.2.2) is

$$H(\mathbf{x}, \mathbf{p}, t) = \mathbf{p}^2/2m + V(\mathbf{x}, t). \tag{2.2.26}$$

Eq. (2.2.26) is evidently the total energy. The Euler–Lagrange equations (2.2.5) imply the equation of motion in Newton's form

$$m\ddot{\mathbf{x}} = -\nabla V|_{\mathbf{x}=\mathbf{x}(t)}. \tag{2.2.27}$$

From (2.2.26) we obtain the Hamilton–Jacobi equation

$$\partial S/\partial t + (\nabla S)^2/2m + V = 0. \tag{2.2.28}$$

When the potential is time-dependent the energy of the particle, $(-\partial S/\partial t)$ evaluated along the trajectory $\mathbf{x} = \mathbf{x}(t)$, is not conserved in general.

We shall also study the case where the particle carries a charge e and is acted upon by an external electromagnetic field with scalar potential $A_0(\mathbf{x}, t)$ and vector potential $\mathbf{A}(\mathbf{x}, t)$. If c is the speed of light, the Lagrangian, canonical momentum and Hamiltonian are respectively

$$\left.\begin{aligned} L &= \tfrac{1}{2}m\dot{\mathbf{x}}^2 + (e/c)\mathbf{A}\cdot\dot{\mathbf{x}} - eA_0 - V \\ \mathbf{p} &= m\dot{\mathbf{x}} + (e/c)\mathbf{A} \\ H &= (1/2m)[\mathbf{p} - (e/c)\mathbf{A}]^2 + eA_0 + V. \end{aligned}\right\} \tag{2.2.29}$$

The equation of motion includes the Lorentz force:

$$m\ddot{\mathbf{x}} = -\nabla V + e(\mathbf{E} + c^{-1}\dot{\mathbf{x}} \times \mathbf{B}) \tag{2.2.30}$$

where $\mathbf{E} = -\nabla A_0 - (1/c)\,\partial\mathbf{A}/\partial t$, $\mathbf{B} = \nabla \times \mathbf{A}$ and the Hamilton–Jacobi equation becomes

$$\partial S/\partial t + (1/2m)[\nabla S - (e/c)\mathbf{A}]^2 + eA_0 + V = 0. \tag{2.2.31}$$

The case of a rotator, for which the generalized coordinates are the Euler angles, is treated in Chap. 10.

2.3 Properties of the Hamilton–Jacobi function

2.3.1 The nonuniqueness of S for a given mechanical problem

The Hamilton–Jacobi theory is often taught as a technique for solving the equations of motion of mechanics, and this is indeed how we have introduced it in the last section. Yet it presents certain features that point the way to a development of the conceptual framework of Newtonian mechanics, and in this and the next sections we call attention to some of these.

To solve a dynamical problem in classical mechanics using Jacobi's transformation theory as set out in §2.2, we require a complete integral of the Hamilton–Jacobi equation depending on as many nonadditive constants as there are degrees of freedom in the system. Such a solution may be found for example by a separation of variables. In order to understand clearly the difference with the analogous case of solutions to the quantum Hamilton–Jacobi equation studied later, it is important to emphasize that *any* complete solution will do in determining the particle motion. A given Hamilton–Jacobi equation may have many different complete integrals, for fixed particle initial conditions, in the sense that the functional dependence of S on q, α and t may vary. But, for the given initial coordinates and momenta, the particle motion associated with all these functions will be the same.

While the S-function corresponding to a given mechanical problem (i.e., prescribed Hamiltonian, q_0 and p_0) is not unique, the various solutions are nevertheless distinguished by virtue of the fact that S is evidently a field function in configuration space. Thus, while the aim of Jacobi's method is the computation of a single orbit, S is actually connected with an infinite set of potential trajectories pursued by an ensemble of identical particles. This set is obtained by varying the constants β for fixed α. The relation (2.2.12a), $p = \partial S/\partial q$, gives at each point q at each instant the canonical momentum of a system that may potentially pass through that point. The physical momentum follows from (2.2.1). The ensemble of motions is characterized by the form of S as a function of q and t and the values of the constants α.

The different functional forms of S connected with the same Hamiltonian imply different types of ensemble. However, at one point at one moment the various S-functions will have the same gradient and they will all imply the same subsequent (and preceding) motion for a system passing through that point. For other space points where the gradients of the S-function do not coincide, the motions of the other ensemble elements are different.

These points are easily illustrated by a simple example. It suffices to treat a free particle in Cartesian coordinates, for which (2.2.28) becomes

$$\partial S/\partial t + (\nabla S)^2/2m = 0. \qquad (2.3.1)$$

First of all we solve this equation by separating the variables, i.e., we treat S as a sum of four functions depending on x, y, z and t respectively. The result is

$$S(x, y, z, P_1, P_2, P_3, t) = -(1/2m)(P_1^2 + P_2^2 + P_3^2)t + P_1 x + P_2 y + P_3 z, \quad (2.3.2)$$

where the nonadditive constants P_1, P_2, P_3 are the components of a momentum vector. Eq. (2.2.12a), $p_i = \partial S/\partial x_i$, $i = 1, 2, 3$, yields $p_i = P_i$. The trajectory is found by writing $\partial S/\partial P_i = Q_i$, where Q_i are constant coordinates. We obtain $-(\mathbf{P}/m)t + \mathbf{x} = \mathbf{Q}$, showing that $\mathbf{Q}$ is the initial (we choose $t_0 = 0$) coordinate vector. Writing $\mathbf{x}_0 = \mathbf{Q}$ and $\mathbf{v} = \mathbf{P}/m$ we therefore find

$$\mathbf{x}(t) = \mathbf{x}_0 + \mathbf{v}t. \qquad (2.3.3)$$

The motion is uniform and rectilinear, starting from the point $\mathbf{x}_0$ with velocity $\mathbf{v}$. The ensemble described by (2.3.2) pursues a set of parallel straight line motions of momentum $\mathbf{P}$ generated by varying $\mathbf{x}_0$ (the βs of this example).

To obtain a different functional form for S we construct the latter from (2.2.20) by integrating along the trajectory (2.3.3). The Lagrangian is $L = \frac{1}{2}m\dot{\mathbf{x}}^2$ and so the free particle action function is given by the expression

$$S(\mathbf{x}, t; \mathbf{x}_0, 0) = (m/2t)(\mathbf{x} - \mathbf{x}_0)^2. \qquad (2.3.4)$$

Here the nonadditive constants are the initial coordinates $\mathbf{x}_0$. Applying the Jacobi theory we can reconstruct the path from (2.3.4): $\partial S/\partial x_{0i} = -P_i$ implies $-m(\mathbf{x} - \mathbf{x}_0)/t = -\mathbf{P}$. Writing again $\mathbf{P}/m = \mathbf{v}$ we recover (2.3.3). In this case, the function (2.3.4) describes an ensemble of particles which all emanate from the point $\mathbf{x}_0$ with a range of momenta $\mathbf{P}$ (i.e. β). The motions generated by (2.3.2) and (2.3.4) coincide when $\mathbf{x}_0$ and $\mathbf{P}$ are chosen to be the same in both cases.

We shall see that in quantum theory two S-functions are distinguished not only globally through the ensembles they generate but in a stronger sense: two particles that start with the same $\mathbf{x}_0$, $\mathbf{p}_0$ do *not* in general pursue the same trajectory in the given potential V.

2.3.2 The basic law of motion

It is by no means guaranteed that we will have available a complete integral of the Hamilton–Jacobi equation. There is no rule for obtaining such a solution; the possibility of solving by a separation of variables depends on finding a suitable coordinate system and for many problems of physical interest (e.g. the three-body problem) this is impossible. It may be possible to find a solution depending on less than the required number of nonadditive constants, but even this may not be feasible. It follows that if, for whatever reason, we do not possess a complete solution of the Hamilton–Jacobi equation, the Jacobi law of motion $\partial S/\partial \gamma = \beta$ cannot be used to solve the dynamical problem completely, and perhaps cannot be used at all.

Suppose though that we have available a *general* solution to the Hamilton–Jacobi equation, that is a function of $S(q, t)$ not depending explicitly on any constants. Then, although we cannot use $\partial S/\partial \gamma = \beta$ to solve for the motion of the system point algebraically, we may employ the other set of transformation equations (2.2.12a). The covariant momentum is

$$p_i = \partial S/\partial q_i \tag{2.3.5}$$

and expressing p in terms of q, $\dot{q}$ and t using (2.2.1) we may solve for $q(t)$ by directly integrating (2.3.5) and specifying $q_0 = q(0)$.

Of course, there should be consistency between the two laws of motion ($\beta = \partial S/\partial \gamma$ and $p = \partial S/\partial q$) in the case where we have a complete integral and both are applicable. This is ensured by (2.2.18), which is simply (2.3.5) evaluated at a definite point q_0, and indeed we must always use this in conjunction with (2.2.17) to solve for the motion completely. Thus, however the trajectory is calculated, we always make use of the relation (2.3.5), either to fix the initial conditions in the case of a complete integral or to integrate directly when we have no complete integral. It follows that, since $p = \partial S/\partial q$ applies in all cases, it is natural to regard this as the basic law of particle motion in the Hamilton–Jacobi theory.

We now give illustrations of this technique (albeit applied to complete integrals) for the case of the two S-functions given in §2.3.1. There $\mathbf{p} = \nabla S$ and from (2.2.25), $\mathbf{p} = m\dot{\mathbf{x}}$.

For S of the form (2.3.2) we have

$$m\dot{\mathbf{x}} = \nabla S = \mathbf{P}, \tag{2.3.6}$$

which is a constant. If the initial condition is $\mathbf{x}(0) = \mathbf{x}_0$ this integrates immediately to give (2.3.3). On the other hand, the form (2.3.4) gives

$$m\dot{\mathbf{x}} = \nabla S = m(\mathbf{x} - \mathbf{x}_0)/t. \tag{2.3.7}$$

This is again readily integrated to yield (2.3.3) if $\mathbf{x}(0) = \mathbf{x}_0$. Notice the quite different functional dependence of ∇S in (2.3.6) and (2.3.7). In one case it is constant in space and time and in the other it is variable. Nevertheless, both forms imply the same physical trajectory if the initial canonical coordinates are the same, and coincide along that trajectory.

Indeed, all functions of S representing particle properties will take the same value when evaluated along the same trajectory. Apart from the momentum ∇S, these include the energy $(-\partial S/\partial t)$ and the angular momentum $\mathbf{x} \times \nabla S$. The two functions differ in their global properties.

What lies behind the more general method of solution is the formulation of dynamical evolution in terms of a Cauchy problem for the Hamilton–Jacobi equation:

$$\partial S/\partial t + H(q, \partial S/\partial q, t) = 0. \tag{2.3.8}$$

Specification of the initial S-function S_0 for all q implies a unique solution for $S(q, t)$. It also has the effect of fixing the initial canonical momentum everywhere in configuration space:

$$p_0 = \partial S_0/\partial q. \tag{2.3.9}$$

Obviously, the specification of an ensemble of p_0s considerably overdetermines the mechanical problem since all that is required is p_0 at one point q_0. But, of course, one is tacitly making a choice of $S_0(q)$ for all q when one seeks a complete integral of the Hamilton–Jacobi equation. For example, for a conservative one-body system we may seek a solution by separating the time variable:

$$S(\mathbf{x}, E, t) = W(\mathbf{x}, E) - Et. \tag{2.3.10}$$

The function $W(\mathbf{x}, E)$ is called *Hamilton's characteristic function* and satisfies the equation

$$E = (\nabla W)^2/2m + V. \tag{2.3.11}$$

Clearly, W is nothing more than the initial S-function, $W(\mathbf{x}) = S_0(\mathbf{x})$. The function (2.3.10) describes an ensemble of particles with the same energy E and variable momentum $\mathbf{p} = \nabla W$. An example is the function (2.3.2) which corresponds to choosing $S_0 = \mathbf{P}\cdot\mathbf{x}$. Any solution to the Hamilton–Jacobi equation for the given time-independent potential V will have the property that $(-\partial S/\partial t) = E$ when evaluated along the trajectory.

To summarize so far, we have shown the following. The problem of dynamics as defined by Hamilton's canonical equations may be formulated in terms of a partial differential equation (2.3.8) determining the evolution of a field $S(q, t)$. This function determines at each point and at each instant the momentum of a system that may be potentially placed there through the relation (2.3.5). For one body the basic law of motion is $\dot{\mathbf{x}} = \nabla S/m$. The function S is thus connected with an ensemble of identical systems rather than a single orbit. It is in this way that the S-functions may be physically distinguished. For fixed q_0, p_0 all S-functions imply the same time development $q(t)$. This reflects the fact that the state of a material system is completely exhausted by specifying its position and momentum – the S-function plays no role in either defining the state or in determining the dynamics.

2.3.3 Multivalued trajectory fields

The set of classical orbits moving in a given potential forms a single-valued congruence when represented in phase space, i.e., only one trajectory may pass through each phase space point. It is a common property of classical force fields that when mapped into configuration space the trajectory field is multivalued: at an instant t more than one orbit may pass through a point q. The degree to which this happens depends on the nature of the force and the particular ensemble chosen (i.e., on $S_0(q)$) and is reflected in the value of $S(q, t)$ (which may, for example, include square roots and hence possess different branches). Most interesting ensembles (bound or scattering) in most interesting force fields are of this sort.

An example is an ensemble of one-dimensional harmonic oscillators of frequency ω having potential energy $V = \frac{1}{2}m\omega^2 x^2$ and energy E. Each orbit is distinguished by the initial position x_0:

$$x(t) = x_0 \cos \omega t + (a^2 - x_0^2)^{1/2} \sin \omega t \tag{2.3.12}$$

where $a = (2E/m\omega^2)^{1/2}$ is the amplitude of the motion. Two trajectories cross each point x with equal and opposite velocities (see Fig. 2.1 which shows

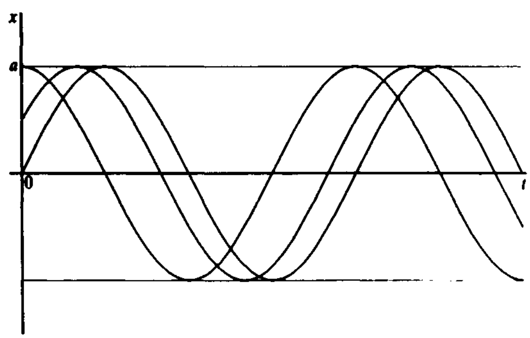


Fig. 2.1 Three orbits of a harmonic oscillator (corresponding to $x_0 = 0, a/2, a$). Each orbit has two segments, each belonging to a single-valued trajectory field.