P J SHEPHERD

A COURSE IN

Theoretical Physics

Central concepts, laws and results, with full mathematical derivations, in

Quantum Mechanics
Thermal and Statistical Physics
Many-Body Theory
Classical Theory of Fields
Special and General Relativity
Relativistic Quantum Mechanics
Gauge Theories in Particle Physics

WILEY

A Course in Theoretical Physics

A Course in Theoretical Physics

P. J. SHEPHERD

Central concepts, laws and results, with full mathematical derivations, in:

Quantum Mechanics
Thermal and Statistical Physics
Many-Body Theory
Classical Theory of Fields
Special and General Relativity
Relativistic Quantum Mechanics
Gauge Theories in Particle Physics

A John Wiley & Sons, Ltd., Publication

This edition first published 2013

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

Library of Congress Cataloging-in-Publication Data applied for.

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

HB ISBN: 9781118481349
PB ISBN: 9781118481424

Typeset in 10/12pt Times by Aptara Inc., New Delhi, India

To N.

Contents

Notation

The following types of alphabetic symbol are used in the mathematical expressions in this book:

Nonbold italic letters (both Latin-script and Greek-script) are used for quantities taking values (or for functions taking a range of values) in the form of single numbers (in some system of units). Examples are energy (E), wave function (ψ), magnitude of magnetic induction (B), polarizability (α), temperature (T), and the speed of light (c). Nonbold italic letters are also used for vector and tensor indices.

Bold italic letters (mostly Latin script) are used to denote three-component vectors and pseudovectors. Examples are the radius vector ($\boldsymbol{r}$), three-momentum ($\boldsymbol{p}$), magnetic induction ($\boldsymbol{B}$), and angular momentum ($\boldsymbol{L}$).

Regular (that is, nonbold and upright) Latin-script letters are used to denote standard functions (sin, cos, exp, etc.) and physical units (kg, m, s, J, W, etc.). They are also used (as are regular Greek-script letters) to denote matrices, including row matrices and column matrices. Examples are (3×3) rotation matrices $\mathrm{\lambda}$, (2×2) Pauli matrices σ_x, σ_y, σ_z, (4×4) Dirac matrices γ^0, γ^1, γ^2, γ^3, and (4×1) Dirac column-matrix wave functions Ψ. Unit and null matrices are denoted by $\mathrm{1}$ and $\mathrm{0}$, respectively.

Bold upright letters (both Latin-script and Greek-script) are used to denote three-component vector matrices. An example is the vector Pauli matrix $\boldsymbol{\sigma}$ with components σ_x, σ_y, and σ_z.

Preface

The twentieth century saw the birth and development of two great theories of physics – quantum mechanics and relativity, leading to an extraordinarily accurate description of the world, on all scales from subatomic to cosmological.

Indeed, the transformation of our understanding of the physical world during the past century has been so profound and so wide in its scope that future histories of science will surely acknowledge that the twentieth century was the century of physics.

Should not the great intellectual achievements of twentieth-century physics by now be part of our culture and of our way of viewing the world? Perhaps they should, but this is still very far from being the case. To take an example, Einstein's general theory of relativity is one of the most beautiful products of the human mind, in any field of endeavor, providing us with a tool to understand our universe and its evolution. Yet, far from having become part of our culture, general relativity is barely studied, even in university physics degree courses, and the majority of graduate physicists cannot even write down Einstein's field equations, let alone understand them.

The reason for this is clear. There is a consensus amongst university physics professors (reinforced by the fact that as students they were mostly victims of the same consensus) that the mathematics of general relativity is "too difficult". The result is that the general theory of relativity rarely finds its way into the undergraduate physics syllabus.

Of course, any mathematical methods that have not been carefully explained first will always be "too difficult". This applies in all fields of physics – not just in general relativity. Careful, and appropriately placed, explanations of mathematical concepts are essential. Serious students are unwilling to take results on trust, but at the same time may not wish to spend valuable time on detailed and often routine mathematical derivations. It is better that the latter be provided, so that a student has at least the choice as to whether to read them or omit them. Claims that "it can be shown that . . ." can lead to relatively unproductive expenditure of time on the filling of gaps – time that students could better spend on gaining a deeper understanding of physics, or on improving their social life.

Several excellent first-year physics degree texts have recognized the likely mathematical limitations of their readers, and, while covering essentially the entire first-year syllabus, ensure that their readers are not left struggling to fill in the gaps in the derivations.

This book is the result of extending this approach to an account of theoretical topics covered in the second, third and fourth years of a degree course in physics in most universities. These include quantum mechanics, thermal and statistical physics, special relativity, and the classical theory of fields, including electromagnetism. Other topics covered here, such as general relativity, many-body theory, relativistic quantum mechanics and gauge theories, are not always all covered in such courses, even though they are amongst the central achievements of twentieth-century theoretical physics.

The book covers all these topics in a single volume. It is self-contained in the sense that a reader who has completed a higher-school mathematics course (for example, A-level mathematics in the United Kingdom) should find within the book everything he or she might need in order to understand every part of it. Details of all derivations are supplied.

The theoretical concepts and methods described in this book provide the basis for commencing doctoral research in a number of areas of theoretical physics, and indeed some are often covered only in first-year postgraduate modules designed for induction into particular research areas. To illustrate their power, the book also includes, with full derivations of all expressions, accounts of Nobel-prize-winning work such as the Bardeen–Cooper–Schrieffer theory of superconductivity and the Weinberg–Salam theory of the electroweak interaction.

The book is divided into five parts, each of which provides a mathematically detailed account of the material typically covered in a forty-lecture module in the corresponding subject area.

Each of the modules is, necessarily, shorter than a typical textbook specializing in the subject of that module. This is not only because many detailed applications of the theory are not included here (nor would they be included in most university lecture modules), but also because, in each of the modules beyond the first, the necessary background can all be found in preceding modules of the book.

The book is the product of over thirty years of discussing topics in theoretical physics with students on degree courses at Exeter University, and of identifying and overcoming the (often subtle) false preconceptions that can impede our understanding of even basic concepts. My deepest thanks go to these students.

Like most course textbooks in such a highly developed (some would say, almost completed) subject area as theoretical physics, this book lays no claim to originality of the basic ideas expounded. Its distinctness lies more in its scope and in the detail of its exposition. The ideas themselves have been described in very many more-specialized textbooks, and my inclusion of some of these books in the lists of recommended supplementary reading is also to be regarded as my grateful acknowledgment of their authors' influence on my chosen approaches to the relevant subject matter.

John Shepherd
Exeter, July 2012

Module I

Nonrelativistic Quantum Mechanics

1

Basic Concepts of Quantum Mechanics

1.1 Probability interpretation of the wave function

In quantum mechanics the **state** of some particle moving in time and space is described by a (complex) **wave function** $\psi(\boldsymbol{r}, t)$ of the space coordinate $\boldsymbol{r}$ $(= \boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z)$ and time coordinate t. The state of the particle is "described" by $\psi(\boldsymbol{r}, t)$ in the sense that knowledge of the precise form of this function enables us to make precise predictions of the probabilities of the various possible outcomes of any given type of physical measurement on the particle. In particular, in a **position** measurement the probability of finding the particle to be in an infinitesimal box of volume $\mathrm{d}x\mathrm{d}y\mathrm{d}z \equiv \mathrm{d}^3\boldsymbol{r}$ with centre at $\boldsymbol{r}$ at time t is

$$P(\boldsymbol{r}, t)\mathrm{d}x\mathrm{d}y\mathrm{d}z = \psi^*(\boldsymbol{r}, t)\psi(\boldsymbol{r}, t)\mathrm{d}x\mathrm{d}y\mathrm{d}z, \tag{1.1}$$

where $\psi^*(\boldsymbol{r}, t)$ is the complex conjugate of the function $\psi(\boldsymbol{r}, t)$. Equation (1.1) may be written alternatively as

$$P(\boldsymbol{r}, t)\mathrm{d}^3\boldsymbol{r} = |\psi(\boldsymbol{r}, t)|^2\,\mathrm{d}^3\boldsymbol{r}. \tag{1.2}$$

Clearly, since any probability must be a dimensionless number, and $\mathrm{d}^3\boldsymbol{r}$ has the dimensions of volume $[L]^3$, the quantity $P(\boldsymbol{r}, t)$ is a **probability density**, with dimensions $[L]^{-3}$.

The wave function $\psi(\boldsymbol{r}, t)$ is also known as the **probability amplitude**, and has (in three-dimensional space) dimensions $[L]^{-3/2}$.

Since the particle must be found somewhere in space, the continuous sum (i.e., integral) of the probabilities (1.2) over infinitesimal boxes filling all space must be equal to unity:

$$\int_{\text{all space}} |\psi|^2\,\mathrm{d}^3\boldsymbol{r} = 1. \tag{1.3}$$

A wave function ψ satisfying (1.3) is said to be **normalized**.

A Course in Theoretical Physics, First Edition. P. J. Shepherd.

1.2 States of definite energy and states of definite momentum

If a particle is in a state of **definite energy** E, the corresponding wave function $\psi(\boldsymbol{r}, t)$ can be separated as a product of a space-dependent factor $\phi(\boldsymbol{r})$ and a **time-dependent** factor that takes the form of a pure harmonic wave $e^{-i\omega t}$ [i.e., there is only one Fourier component (one frequency) in the Fourier decomposition of the time dependence], where the frequency ω is given by $\omega = E/\hbar$ (the **Einstein relation** $E = \hbar\omega$), in which $\hbar$ ($= 1.055 \times 10^{-34}$ Js) is Planck's constant h divided by 2π.

Thus, the time dependence of the wave function of a particle (or, indeed, a system of many particles) with definite energy E is **always** of the form $e^{-iEt/\hbar}$.

If a particle is in a state of **definite momentum** $\boldsymbol{p} = \boldsymbol{i}p_x + \boldsymbol{j}p_y + \boldsymbol{k}p_z$ (so that the energy is also well defined and the above-mentioned factorization occurs), the **space-dependent** part $\phi(\boldsymbol{r})$ of the corresponding wave function $\psi(\boldsymbol{r}, t)$ is a harmonic plane wave $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ [i.e., there is only one Fourier component (one wave vector $\boldsymbol{k}$) in its Fourier decomposition], where the wave vector $\boldsymbol{k} = \boldsymbol{i}k_x + \boldsymbol{j}k_y + \boldsymbol{k}k_z$ is related to the momentum $\boldsymbol{p}$ by $\boldsymbol{k} = \boldsymbol{p}/\hbar$, i.e., $k_x = p_x/\hbar$, and so on. [This is the well known **de Broglie relation** $\boldsymbol{p} = \hbar\boldsymbol{k}$, which, for the magnitudes, gives

$$p \equiv |\boldsymbol{p}| = \hbar\,|\boldsymbol{k}| \equiv \hbar k = \frac{h}{2\pi} \cdot \frac{2\pi}{\lambda} = \frac{h}{\lambda},$$

where λ is the wavelength associated with wave number k.]

Thus, the space dependence of the wave function of a particle with definite momentum $\boldsymbol{p}$ is the **plane wave**

$$\phi_{\boldsymbol{p}}(\boldsymbol{r}) = e^{i\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}. \tag{1.4}$$

[To see the reason for the designation "plane wave", we choose the z axis to lie along the direction of the momentum $\boldsymbol{p}$. Then (1.4) takes the form $e^{ipz/\hbar}$, which clearly has the same phase (and so takes the same value) over any surface $z = \text{const}$ (a plane).]

Successive planes on which the function (1.4) has the same phase are separated by a distance equal to the wavelength λ and are called **wave fronts**.

To understand the motion of wave fronts, consider a particle of definite energy E and definite momentum $\boldsymbol{p}$ (along the z axis). Then the corresponding wave function is

$$\psi_{E,\boldsymbol{p}}(t) \propto e^{i(pz-Et)/\hbar},$$

and the equation of motion of a wave front, that is, of a plane $z = z_{\mathrm{P}}$ with given phase of the wave function, is found by equating the phase of $\psi_{E,\boldsymbol{p}}(t)$ to a constant:

$$pz_{\mathrm{P}} - Et = \text{const},$$

that is,

$$z_{\mathrm{P}} = z_{\mathrm{P}}(t) = \frac{E}{p}t + \text{const} \equiv v_{\mathrm{P}}t + \text{const},$$

Thus, a wave front (surface of constant phase) moves with **phase velocity** $v_{\mathrm{P}} = E/p$ (> 0) along the positive z axis, that is, in the direction of the momentum $\boldsymbol{p}$.

1.3 Observables and operators

With every physically measurable quantity ("**observable**") A we can associate an operator $\hat{A}$, such that, if a particle is in a state with a well defined value a of A [we denote the wave function corresponding to this state by $\psi_a(\boldsymbol{r}, t)$], then

$$\hat{A}\psi_a(\boldsymbol{r}, t) = a\psi_a(\boldsymbol{r}, t). \tag{1.5}$$

Thus, the action of the operator $\hat{A}$ on the function $\psi_a(\boldsymbol{r}, t)$ is to reproduce precisely the same function of $\boldsymbol{r}$ and t but scaled by a constant factor a equal to the well defined value of A in this state. We say that $\psi_a(\boldsymbol{r}, t)$ is an **eigenfunction** of the operator $\hat{A}$, and a is the corresponding **eigenvalue** of the operator $\hat{A}$.

1.4 Examples of operators

As follows from section 1.2, the state of a particle with definite energy E and definite momentum $\boldsymbol{p}$ is described by a wave function of the form

$$\psi_{E,\boldsymbol{p}}(\boldsymbol{r}, t) \propto \mathrm{e}^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} = \mathrm{e}^{i(\boldsymbol{p}\cdot\boldsymbol{r}-Et)/\hbar}.$$

Accordingly [see (1.5)], there must exist an **energy operator** $\hat{E}$ and a **momentum operator** $\hat{\boldsymbol{p}}$ such that

$$\hat{E}\mathrm{e}^{i(\boldsymbol{p}\cdot\boldsymbol{r}-Et)/\hbar} = E\mathrm{e}^{i(\boldsymbol{p}\cdot\boldsymbol{r}-Et)/\hbar}$$

and

$$\hat{\boldsymbol{p}}\mathrm{e}^{i(\boldsymbol{p}\cdot\boldsymbol{r}-Et)/\hbar} = \boldsymbol{p}\mathrm{e}^{i(\boldsymbol{p}\cdot\boldsymbol{r}-Et)/\hbar}.$$

The differential operators

$$\hat{E} = i\hbar\frac{\partial}{\partial t} \tag{1.6}$$

and

$$\hat{\boldsymbol{p}} = -i\hbar\left(\boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z}\right) \equiv -i\hbar\boldsymbol{\nabla} \tag{1.7}$$

clearly have the required properties (if we bear in mind that $\boldsymbol{p}\cdot\boldsymbol{r} = p_x x + p_y y + p_z z$).

Just as we have found the energy and momentum operators by considering wave functions corresponding to definite energy and definite momentum, we can find the **position operator** by considering a wave function corresponding to definite position. We restrict ourselves, for the moment, to motion confined to one dimension. Then a particle with (at some given time) a definite position $x = x_1$ will be described by a wave function with the following spatial dependence (at that time):

$$\phi_{x_1}(x) = \delta(x - x_1),$$

where $\delta(x - x_1)$ is the **Dirac delta-function**, equal to zero when $x \neq x_1$ and with the properties

$$\int f(x)\delta(x - x_1)\mathrm{d}x = f(x_1) \quad \text{and} \quad \int \delta(x - x_1)\mathrm{d}x = 1,$$

(provided that the range of integration includes the point $x = x_1$). We need an operator $\hat{x}$ with the following effect:

$$\hat{x}\phi_{x_1}(x) = x_1\phi_{x_1}(x),$$

that is,

$$\hat{x}\delta(x - x_1) = x_1\delta(x - x_1). \tag{1.8}$$

The purely multiplicative operator

$$\hat{x} = x$$

has the required property, since (1.8) becomes

$$x\delta(x - x_1) = x_1\delta(x - x_1),$$

which is clearly true both for $x \neq x_1$ and for $x = x_1$. Similarly, we have $\hat{y} = y$ and $\hat{z} = z$, and so the operator of the three-dimensional position vector $\boldsymbol{r}$ is $\hat{\boldsymbol{r}} = \boldsymbol{r}$.

The operator corresponding to a general observable $A(\boldsymbol{r}, \boldsymbol{p})$ is obtained by the prescription

$$A \rightarrow \hat{A} = A(\hat{\boldsymbol{r}}, \hat{\boldsymbol{p}}) \tag{1.9}$$

with

$$\hat{\boldsymbol{r}} = \boldsymbol{r} \quad \text{and} \quad \hat{\boldsymbol{p}} = -i\hbar\boldsymbol{\nabla}. \tag{1.10}$$

For example, the angular momentum, about the coordinate origin, of a particle at a vector distance $\boldsymbol{r}$ from the origin and with momentum $\boldsymbol{p}$ is given classically by the vector product $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$, and the corresponding operator $\hat{\boldsymbol{L}}$ is given by:

$$\hat{\boldsymbol{L}} = \hat{\boldsymbol{r}} \times \hat{\boldsymbol{p}} = -i\hbar\boldsymbol{r} \times \boldsymbol{\nabla}. \tag{1.11}$$

1.5 The time-dependent Schrödinger equation

The total energy E of a particle at a given point $\boldsymbol{r}$ at time t is given by the sum of the kinetic energy $T(\boldsymbol{r}, t)$ and potential energy $V(\boldsymbol{r}, t)$:

$$\begin{aligned} E &= E(\boldsymbol{r}, t) = T(\boldsymbol{r}, t) + V(\boldsymbol{r}, t) \\ &= \frac{\boldsymbol{p} \cdot \boldsymbol{p}}{2m} + V(\boldsymbol{r}, t) \equiv \frac{p^2}{2m} + V(\boldsymbol{r}, t), \end{aligned} \tag{1.12}$$

where $\boldsymbol{p} = \boldsymbol{p}(\boldsymbol{r}, t)$ is the momentum of the particle (of mass m) at the point $\boldsymbol{r}$ at time t.

We make the operator replacements (section 1.4):

$$\begin{aligned} E &\to \hat{E} = i\hbar\frac{\partial}{\partial t}, \\ \boldsymbol{p} &\to \hat{\boldsymbol{p}} = -i\hbar\boldsymbol{\nabla} \quad \text{(so that } p^2 \equiv \boldsymbol{p}\cdot\boldsymbol{p} \to -\hbar^2\boldsymbol{\nabla}\cdot\boldsymbol{\nabla} \equiv -\hbar^2\nabla^2\text{)}, \\ V(\boldsymbol{r},t) &\to \hat{V} = V(\hat{\boldsymbol{r}},t) = V(\boldsymbol{r},t). \end{aligned} \tag{1.13}$$

Since the relation $E = T + V$ is a physical requirement, a physically acceptable wave function $\psi(\boldsymbol{r},t)$ must satisfy the equation $\hat{E}\psi = \hat{T}\psi + \hat{V}\psi$, that is,

$$i\hbar\frac{\partial\psi(\boldsymbol{r},t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi(\boldsymbol{r},t) + V(\boldsymbol{r},t)\psi(\boldsymbol{r},t). \tag{1.14}$$

This is the **time-dependent Schrödinger equation**, and is (for the nonrelativistic case and in three dimensions) a **completely general** differential equation (first order in time and second order in space) for the evolution of the wave function $\psi(\boldsymbol{r},t)$ of a particle whose energy may or may not be well defined, moving in a potential that may or may not be varying in time.

1.6 Stationary states and the time-independent Schrödinger equation

We now specialize to the case when the particle is in a state with definite energy E. For this case (section 1.2) the wave function is always of the form

$$\psi(\boldsymbol{r},t) = \phi(\boldsymbol{r})\mathrm{e}^{-iEt/\hbar}, \tag{1.15}$$

where $\phi(\boldsymbol{r})$ specifies the space-dependent part of the function. Because, from (1.15), $|\psi(\boldsymbol{r},t)|^2 = |\phi(\boldsymbol{r})|^2$, which is independent of the time t, such a state is called a **stationary state**. When (1.15) is substituted into the time-dependent Schrödinger equation (1.14) the left-hand side becomes $E\psi(\boldsymbol{r},t)$, and so, after cancellation of the factor $\mathrm{e}^{-iEt/\hbar}$, we obtain

$$-\frac{\hbar^2}{2m}\nabla^2\phi(\boldsymbol{r}) + V(\boldsymbol{r})\phi(\boldsymbol{r}) = E\phi(\boldsymbol{r}), \tag{1.16}$$

where we have also replaced $V(\boldsymbol{r},t)$ by $V(\boldsymbol{r})$, since for a particle to be in a state of well-defined energy E it is a necessary (though not sufficient) condition that the potential-energy function be time independent. [This is clearly seen from the fact that equation (1.16) with different potentials $V_t(\boldsymbol{r}) \equiv V(\boldsymbol{r},t)$ at different times t will have different solutions $\phi_t(\boldsymbol{r})$ with different energy eigenvalues E_t, contradicting the assumption of a definite energy E.] The equation (1.16) is the **time-independent Schrödinger equation**, and is an eigenvalue equation [see (1.5)] of the form

$$\hat{H}\phi(\boldsymbol{r}) = E\phi(\boldsymbol{r}) \tag{1.17}$$

with

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r}) \tag{1.18}$$

(the hamiltonian operator, or **hamiltonian**).

1.7 Eigenvalue spectra and the results of measurements

The set of all possible eigenvalues a of an operator $\hat{A}$ is called the **spectrum** of the operator $\hat{A}$. **A measurement of A always yields a value belonging to this spectrum**.

Two cases are possible:

(a) The wave function $\psi(\boldsymbol{r}, t)$ describing the state of the particle is an eigenfunction $\psi_i(\boldsymbol{r}, t)$ of $\hat{A}$, that is,

$$\hat{A}\psi_i(\boldsymbol{r}, t) = a_i\psi_i(\boldsymbol{r}, t). \tag{1.19}$$

Then the result of measuring A will certainly be a_i.

(b) The wave function $\psi(\boldsymbol{r}, t)$ describing the state of the particle is not an eigenfunction of $\hat{A}$; that is, the action of $\hat{A}$ on $\psi(\boldsymbol{r}, t)$ gives a function that is not simply a scaled version of $\psi(\boldsymbol{r}, t)$. But since the eigenfunctions $\psi_i(\boldsymbol{r}, t)$ of $\hat{A}$ form a complete set, in the sense that any normalized function can be expanded in terms of them, we may write $\psi(\boldsymbol{r}, t)$ as such an expansion:

$$\psi(\boldsymbol{r}, t) = \sum_i c_i\psi_i(\boldsymbol{r}, t). \tag{1.20}$$

Then a measurement of A can yield any eigenvalue a_i for which the corresponding eigenfunction appears in the sum (1.20) with nonzero weight c_i. For example, a measurement of A yields the result a_j with probability $|c_j|^2$.

1.8 Hermitian operators

The operators corresponding to all physical observables are **hermitian**.

Definition. An operator $\hat{A}$ is said to be hermitian if, for any pair of normalizable wave functions $\psi(\boldsymbol{r}, t)$ and $\chi(\boldsymbol{r}, t)$, the relation

$$\int \chi^*\hat{A}\psi\,\mathrm{d}^3\boldsymbol{r} = \int (\hat{A}\chi)^*\psi\,\mathrm{d}^3\boldsymbol{r} \tag{1.21}$$

always holds.

The eigenvalues of hermitian operators are real. This is proved as follows. Choose χ and ψ to be the same eigenfunction ψ_i of the operator $\hat{A}$, with corresponding eigenvalue a_i; that is, choose

$$\chi = \psi = \psi_i \qquad (\hat{A}\psi_i = a_i\psi_i).$$

Then (1.21) becomes

$$a_i \int \psi_i^*\psi_i\,\mathrm{d}^3\boldsymbol{r} = a_i^* \int \psi_i^*\psi_i\,\mathrm{d}^3\boldsymbol{r},$$

that is, $a_i = a_i^*$, or, in other words, the eigenvalue a_i (and hence any eigenvalue of $\hat{A}$) is real.

Now choose χ and ψ to be two different eigenfunctions of $\hat{A}$: let χ be ψ_i with eigenvalue a_i, and let ψ be ψ_j with eigenvalue a_j. Then (1.21) becomes

$$\int \psi_i^* \hat{A} \psi_j \mathrm{d}^3 \boldsymbol{r} = \int (\hat{A}\psi_i)^* \psi_j \mathrm{d}^3 \boldsymbol{r},$$

whence

$$a_j \int \psi_i^* \psi_j \mathrm{d}^3 \boldsymbol{r} = a_i^* \int \psi_i^* \psi_j \mathrm{d}^3 \boldsymbol{r}.$$

But $a_i^* = a_i$ and, since $a_j \neq a_i$, we have:

$$\int \psi_i^* \psi_j \mathrm{d}^3 \boldsymbol{r} = 0. \tag{1.22}$$

This result states that eigenfunctions of a hermitian operator that correspond to different eigenvalues of that operator are **orthogonal**.

We now use the idea of orthogonality to prove an inequality

$$\left| \int \phi^* \chi \mathrm{d}^3 \boldsymbol{r} \right|^2 \leq \left(\int \phi^* \phi \mathrm{d}^3 \boldsymbol{r} \right) \left(\int \chi^* \chi \mathrm{d}^3 \boldsymbol{r} \right) \tag{1.23}$$

that will be used in section 1.11 to prove the **general uncertainty relation**. The case when the functions ϕ and χ are the same function clearly corresponds to the equality in (1.23). The inequality (1.23), if true, will be true for any normalization of ϕ and χ, since when we multiply ϕ and χ by arbitrary factors these factors cancel in the inequality (1.23). Therefore, for convenience in examining the left-hand side of (1.23), we choose unit normalization for ϕ and χ. But if in the left-hand side $|\int \phi^* \chi \mathrm{d}^3 \boldsymbol{r}|^2$ the function χ contains not only a part proportional to ϕ but also a part proportional to a normalized function ρ **orthogonal** to ϕ, so that $\chi = a\phi + b\rho$, with $|a|^2 + |b|^2 = 1$ (note that then, in particular, $|a| \leq 1$), we have:

$$\left| \int \phi^* \chi \mathrm{d}^3 \boldsymbol{r} \right| = \left| \int \phi^* (a\phi + b\rho) \mathrm{d}^3 \boldsymbol{r} \right| = |a| \int \phi^* \phi \mathrm{d}^3 \boldsymbol{r} \leq \int \phi^* \phi \mathrm{d}^3 \boldsymbol{r}. \tag{1.24}$$

Similarly, expressing ϕ in this case as $\phi = c\chi + d\eta$, with the function η orthogonal to χ and with both χ and η normalized to unity (so that $|c|^2 + |d|^2 = 1$ and, in particular, $|c| \leq 1$), we have

$$\left| \int \phi^* \chi \mathrm{d}^3 \boldsymbol{r} \right| = \left| \int \chi^* \phi \mathrm{d}^3 \boldsymbol{r} \right| = \left| \int \chi^* (c\chi + d\eta) \mathrm{d}^3 \boldsymbol{r} \right| = |c| \int \chi^* \chi \mathrm{d}^3 \boldsymbol{r} \leq \int \chi^* \chi \mathrm{d}^3 \boldsymbol{r}, \tag{1.25}$$

so that, taking the product of (1.24) and (1.25), we obtain (1.23).

1.9 Expectation values of observables

The **expectation value** of an observable A in a particle state described by a wave function $\psi(\boldsymbol{r}, t)$ is defined as the average result of $N (\to \infty)$ measurements of A, all carried out on a particle in the same state $\psi(\boldsymbol{r}, t)$. The prescription for calculating the expectation value when the (normalized) function $\psi(\boldsymbol{r}, t)$ is known is

$$\langle A \rangle_{\psi(\boldsymbol{r},t)} \equiv \int \psi^*(\boldsymbol{r}, t) \hat{A} \psi(\boldsymbol{r}, t) \mathrm{d}^3\boldsymbol{r}. \tag{1.26}$$

To show that this prescription is equivalent to the definition of the expectation value as an average, we substitute the expansion (1.20) into (1.26) and use the fact that the eigenfunctions $\psi_i(\boldsymbol{r}, t)$ of $\hat{A}$ are normalized and mutually orthogonal (an **orthonormal** set). (We showed in section 1.8 that eigenfunctions belonging to different eigenvalues of a hermitian operator are always mutually orthogonal. Also, although eigenfunctions belonging to the same eigenvalue need not be mutually orthogonal, it is always possible to form linear combinations of them that are.) The result is

$$\langle A \rangle_{\psi(\boldsymbol{r},t)} = \sum_i |c_i|^2 a_i. \tag{1.27}$$

This relation asserts (in precise accord with its definition as an average) that the expectation value of A in the state $\psi(\boldsymbol{r}, t)$ is the sum of all possible outcome a_i, each weighted by the probability $|c_i|^2$ that in this state a measurement of A will yield the result a_i.

1.10 Commuting observables and simultaneous observability

If the operators corresponding to two observables A and B have a common, complete set of eigenfunctions ψ_{ij} (i labels the corresponding eigenvalues a_i of $\hat{A}$, and j labels the corresponding eigenvalues b_j of $\hat{B}$), we say that A and B can be **simultaneously well defined**, or, equivalently, that they are simultaneously observable.

The **commutator** $[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}$ of any pair of simultaneously well-definable observables A and B is identically zero, in the sense that its action on an arbitrary normalizable function ψ gives zero. To show this, we use the fact that any such function ψ can be written as a linear combination of the common eigenfunctions ψ_{ij} of $\hat{A}$ and $\hat{B}$, so that

$$[\hat{A}, \hat{B}]\psi = [\hat{A}, \hat{B}] \sum_{i,j} c_{ij} \psi_{ij} = \sum_{i,j} c_{ij} [\hat{A}, \hat{B}] \psi_{ij} = 0, \tag{1.28}$$

which follows from the fact that

$$\begin{aligned} [\hat{A}, \hat{B}]\psi_{ij} &\equiv \hat{A}\hat{B}\psi_{ij} - \hat{B}\hat{A}\psi_{ij} = \hat{A} b_j \psi_{ij} - \hat{B} a_i \psi_{ij} \\ &= b_j \hat{A}\psi_{ij} - a_i \hat{B}\psi_{ij} = b_j a_i \psi_{ij} - a_i b_j \psi_{ij} = 0. \end{aligned}$$

The converse is clearly also true, in that if two operators $\hat{A}$ and $\hat{B}$ commute, that is, $[\hat{A}, \hat{B}] = 0$, they can be simultaneously well defined.

1.11 Noncommuting observables and the uncertainty principle

If $[\hat{A}, \hat{B}] \neq 0$, the observables A and B cannot be simultaneously well defined. We define the **uncertainty** in A in the state ψ by

$$(\Delta A)_\psi = \sqrt{\left\langle \left(\hat{A} - \langle \hat{A} \rangle_\psi\right)^2 \right\rangle_\psi} \tag{1.29}$$

(which is clearly the standard deviation of A, i.e., the root mean square deviation of A from the mean of A, where the "mean" here is the expectation value in the state ψ). Then, as is proved below, the product of the uncertainties in two observables A and B is given by the inequality (**general uncertainty relation**)

$$(\Delta A)_\psi (\Delta B)_\psi \geq \frac{1}{2} \left| \langle [\hat{A}, \hat{B}] \rangle_\psi \right|. \tag{1.30}$$

***Proof*:** With the definition (1.29) in mind, we set $\phi \equiv (\hat{A} - \langle \hat{A} \rangle_\psi)\psi$ and $\chi \equiv (\hat{B} - \langle \hat{B} \rangle_\psi)\psi$, so that

$$\begin{aligned} \int \phi^* \phi \, \mathrm{d}^3 \boldsymbol{r} &= \int \left[\left(\hat{A} - \langle \hat{A} \rangle_\psi\right) \psi \right]^* \left(\hat{A} - \langle \hat{A} \rangle_\psi\right) \psi \, \mathrm{d}^3 \boldsymbol{r} \\ &= \int \psi^* \left(\hat{A} - \langle \hat{A} \rangle_\psi\right) \left(\hat{A} - \langle \hat{A} \rangle_\psi\right) \psi \, \mathrm{d}^3 \boldsymbol{r} = (\Delta A)_\psi{}^2, \end{aligned}$$

where we have used the fact that $\hat{A}$ is a hermitian operator [see (1.21)]. Similarly, $\int \chi^* \chi \, \mathrm{d}^3 \boldsymbol{r} = (\Delta B)_\psi{}^2$. Then the inequality (1.23) becomes:

$$\begin{aligned} (\Delta A)_\psi{}^2 (\Delta B)_\psi{}^2 &= \left(\int \phi^* \phi \, \mathrm{d}^3 \boldsymbol{r} \right) \left(\int \chi^* \chi \, \mathrm{d}^3 \boldsymbol{r} \right) \geq \left| \int \phi^* \chi \, \mathrm{d}^3 \boldsymbol{r} \right|^2 = \left| \mathrm{Re} \left(\int \phi^* \chi \, \mathrm{d}^3 \boldsymbol{r} \right) \right|^2 + \left| \mathrm{Im} \left(\int \phi^* \chi \, \mathrm{d}^3 \boldsymbol{r} \right) \right|^2 \\ &\geq \left| \mathrm{Im} \left(\int \phi^* \chi \, \mathrm{d}^3 \boldsymbol{r} \right) \right|^2 = \left| \frac{\int \phi^* \chi \, \mathrm{d}^3 \boldsymbol{r} - \int \chi^* \phi \mathrm{d}^3 \boldsymbol{r}}{2i} \right|^2 \\ &= \frac{1}{4} \left| \int \left(\left(\hat{A} - \langle \hat{A} \rangle_\psi\right) \psi \right)^* \left(\hat{B} - \langle \hat{B} \rangle_\psi\right) \psi \, \mathrm{d}^3 \boldsymbol{r} - \int \left(\left(\hat{B} - \langle \hat{B} \rangle_\psi\right) \psi \right)^* \left(\hat{A} - \langle \hat{A} \rangle_\psi\right) \psi \, \mathrm{d}^3 \boldsymbol{r} \right|^2 \\ &= \frac{1}{4} \left| \int \psi^* \left(\hat{A} - \langle \hat{A} \rangle_\psi\right) \left(\hat{B} - \langle \hat{B} \rangle_\psi\right) \psi \, \mathrm{d}^3 \boldsymbol{r} - \int \psi^* \left(\hat{B} - \langle \hat{B} \rangle_\psi\right) \left(\hat{A} - \langle \hat{A} \rangle_\psi\right) \psi \, \mathrm{d}^3 \boldsymbol{r} \right|^2, \end{aligned}$$

where we have again used the fact that that $\hat{A}$ and $\hat{B}$ are hermitian operators. Thus, we have

$$\begin{aligned} (\Delta A)_\psi{}^2 (\Delta B)_\psi{}^2 &\geq \frac{1}{4} \left| \langle \hat{A}\hat{B} \rangle_\psi - 2 \langle \hat{A} \rangle_\psi \langle \hat{B} \rangle_\psi + \langle \hat{A} \rangle_\psi \langle \hat{B} \rangle_\psi - \left(\langle \hat{B}\hat{A} \rangle_\psi - 2 \langle \hat{B} \rangle_\psi \langle \hat{A} \rangle_\psi + \langle \hat{B} \rangle_\psi \langle \hat{A} \rangle_\psi \right) \right|^2 \\ &= \frac{1}{4} \left| \langle \hat{A}\hat{B} - \hat{B}\hat{A} \rangle_\psi \right|^2 \equiv \frac{1}{4} \left| \langle [\hat{A}, \hat{B}] \rangle_\psi \right|^2, \end{aligned}$$

from which (1.30 follows). For example, when $\hat{A} = \hat{x} = x$ and $\hat{B} = \hat{p}_x = -i\hbar\partial/\partial x$, we have $[\hat{x}, \hat{p}_x]\psi = -i\hbar \left(x \frac{\partial \psi}{\partial x} - \frac{\partial}{\partial x}(x\psi) \right) = i\hbar\psi$. From this we see that $[\hat{x}, \hat{p}_x] = i\hbar$, and so (1.30) yields $(\Delta x)_\psi (\Delta p_x)_\psi \geq \frac{\hbar}{2}$, which asserts that for a particle in one-dimensional motion no state ψ exists for which the product of the

uncertainties in the position and momentum is smaller than $\hbar/2$. In fact, the equality $(\Delta x)_\psi (\Delta p_x)_\psi = \frac{\hbar}{2}$ (corresponding to the smallest possible product of the position and momentum uncertainties) is found in the case of a one-dimensional harmonic oscillator in its ground state, as can be seen by substituting the gaussian space-dependent part $\phi(x) \propto \exp(-\alpha x^2)$ of the latter [see (3.36)] into the expressions for $(\Delta x)_\psi$ and $(\Delta p_x)_\psi$.

1.12 Time dependence of expectation values

Using the expression (1.26) for the expectation value of an observable A, and also the time-dependent Schrödinger equation (1.14) in the form

$$i\hbar \frac{\partial \psi(\boldsymbol{r}, t)}{\partial t} = \hat{H}\psi(\boldsymbol{r}, t) \tag{1.31}$$

(together with the complex conjugate of this equation), we can immediately obtain an expression for the rate of change of the expectation value with time:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \langle A \rangle_\psi = \left\langle [\hat{A}, \hat{H}] \right\rangle_\psi + i\hbar \left\langle \partial \hat{A}/\partial t \right\rangle_\psi , \tag{1.32}$$

which is known as **Ehrenfest's theorem**.

1.13 The probability-current density

We now use the time-dependent Schrödinger equation (1.14) to find the time rate of change of the probability density $P(\boldsymbol{r}, t) = \psi^*(\boldsymbol{r}, t)\psi(\boldsymbol{r}, t)$. We find

$$\frac{\partial P(\boldsymbol{r}, t)}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{j}(\boldsymbol{r}, t) = 0 \tag{1.33}$$

(the **continuity equation** for the probability density). Here,

$$\boldsymbol{j} = -\frac{i\hbar}{2m} \left(\psi^* \boldsymbol{\nabla}\psi - (\boldsymbol{\nabla}\psi^*)\psi \right) \tag{1.34}$$

is the **probability-current density**.

1.14 The general form of wave functions

In this section, for simplicity, we consider the case of **one-dimensional** motion of a particle of mass m and potential energy $V(x)$.

The space part $\phi_E(x)$ of the wave function of a stationary state with energy E satisfies the time-independent Schrödinger equation:

$$-\frac{\hbar^2}{2m} \frac{\mathrm{d}^2 \phi_E}{\mathrm{d}x^2} + V(x)\phi_E = E\phi_E, \tag{1.35}$$