

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
© 2013 John Wiley & Sons, Ltd.

Registered office

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

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

The right of the author to be identified as the author of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

The publisher and the author make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of fitness for a particular purpose. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for every situation. In view of ongoing research, equipment modifications, changes in governmental regulations, and the constant flow of information relating to the use of experimental reagents, equipment, and devices, the reader is urged to review and evaluate the information provided in the package insert or instructions for each chemical, piece of equipment, reagent, or device for, among other things, any changes in the instructions or indication of usage and for added warnings and precautions. The fact that an organization or Website is referred to in this work as a citation and/or a potential source of further information does not mean that the author or the publisher endorses the information the organization or Website may provide or recommendations it may make. Further, readers should be aware that Internet Websites listed in this work may have changed or disappeared between when this work was written and when it is read. No warranty may be created or extended by any promotional statements for this work. Neither the publisher nor the author shall be liable for any damages arising herefrom.

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	xiii
Preface	xv
I NONRELATIVISTIC QUANTUM MECHANICS	1
1 Basic Concepts of Quantum Mechanics	3
1.1 Probability interpretation of the wave function	3
1.2 States of definite energy and states of definite momentum	4
1.3 Observables and operators	5
1.4 Examples of operators	5
1.5 The time-dependent Schrödinger equation	6
1.6 Stationary states and the time-independent Schrödinger equation	7
1.7 Eigenvalue spectra and the results of measurements	8
1.8 Hermitian operators	8
1.9 Expectation values of observables	10
1.10 Commuting observables and simultaneous observability	10
1.11 Noncommuting observables and the uncertainty principle	11
1.12 Time dependence of expectation values	12
1.13 The probability-current density	12
1.14 The general form of wave functions	12
1.15 Angular momentum	15
1.16 Particle in a three-dimensional spherically symmetric potential	17
1.17 The hydrogen-like atom	18
2 Representation Theory	23
2.1 Dirac representation of quantum mechanical states	23
2.2 Completeness and closure	27
2.3 Changes of representation	28
2.4 Representation of operators	29
2.5 Hermitian operators	31
2.6 Products of operators	31
2.7 Formal theory of angular momentum	32
3 Approximation Methods	39
3.1 Time-independent perturbation theory for nondegenerate states	39
3.2 Time-independent perturbation theory for degenerate states	44
3.3 The variational method	50
3.4 Time-dependent perturbation theory	54

4	Scattering Theory	63
4.1	Evolution operators and Møller operators	63
4.2	The scattering operator and scattering matrix	66
4.3	The Green operator and T operator	70
4.4	The stationary scattering states	76
4.5	The optical theorem	83
4.6	The Born series and Born approximation	85
4.7	Spherically symmetric potentials and the method of partial waves	87
4.8	The partial-wave scattering states	92
II	THERMAL AND STATISTICAL PHYSICS	97
5	Fundamentals of Thermodynamics	99
5.1	The nature of thermodynamics	99
5.2	Walls and constraints	99
5.3	Energy	100
5.4	Microstates	100
5.5	Thermodynamic observables and thermal fluctuations	100
5.6	Thermodynamic degrees of freedom	102
5.7	Thermal contact and thermal equilibrium	103
5.8	The zeroth law of thermodynamics	104
5.9	Temperature	104
5.10	The International Practical Temperature Scale	107
5.11	Equations of state	107
5.12	Isotherms	108
5.13	Processes	109
	5.13.1 Nondissipative work	109
	5.13.2 Dissipative work	111
	5.13.3 Heat flow	112
5.14	Internal energy and heat	112
	5.14.1 Joule's experiments and internal energy	112
	5.14.2 Heat	113
5.15	Partial derivatives	115
5.16	Heat capacity and specific heat	116
	5.16.1 Constant-volume heat capacity	117
	5.16.2 Constant-pressure heat capacity	117
5.17	Applications of the first law to ideal gases	118
5.18	Difference of constant-pressure and constant-volume heat capacities	119
5.19	Nondissipative-compression/expansion adiabat of an ideal gas	120
6	Quantum States and Temperature	125
6.1	Quantum states	125
6.2	Effects of interactions	128
6.3	Statistical meaning of temperature	130
6.4	The Boltzmann distribution	134

7	Microstate Probabilities and Entropy	141
7.1	Definition of general entropy	141
7.2	Law of increase of entropy	142
7.3	Equilibrium entropy S	144
7.4	Additivity of the entropy	146
7.5	Statistical–mechanical description of the three types of energy transfer	147
8	The Ideal Monatomic Gas	151
8.1	Quantum states of a particle in a three-dimensional box	151
8.2	The velocity-component distribution and internal energy	153
8.3	The speed distribution	156
8.4	The equation of state	158
8.5	Mean free path and thermal conductivity	160
9	Applications of Classical Thermodynamics	163
9.1	Entropy statement of the second law of thermodynamics	163
9.2	Temperature statement of the second law of thermodynamics	164
9.3	Summary of the basic relations	166
9.4	Heat engines and the heat-engine statement of the second law of thermodynamics	167
9.5	Refrigerators and heat pumps	169
9.6	Example of a Carnot cycle	170
9.7	The third law of thermodynamics	172
9.8	Entropy-change calculations	174
10	Thermodynamic Potentials and Derivatives	177
10.1	Thermodynamic potentials	177
10.2	The Maxwell relations	179
10.3	Calculation of thermodynamic derivatives	180
11	Matter Transfer and Phase Diagrams	183
11.1	The chemical potential	183
11.2	Direction of matter flow	184
11.3	Isotherms and phase diagrams	184
11.4	The Euler relation	187
11.5	The Gibbs–Duhem relation	188
11.6	Slopes of coexistence lines in phase diagrams	188
12	Fermi–Dirac and Bose–Einstein Statistics	191
12.1	The Gibbs grand canonical probability distribution	191
12.2	Systems of noninteracting particles	193
12.3	Indistinguishability of identical particles	194
12.4	The Fermi–Dirac and Bose–Einstein distributions	195
12.5	The entropies of noninteracting fermions and bosons	197

III MANY-BODY THEORY	199
13 Quantum Mechanics and Low-Temperature Thermodynamics of Many-Particle Systems	201
13.1 Introduction	201
13.2 Systems of noninteracting particles	201
13.2.1 Bose systems	202
13.2.2 Fermi systems	204
13.3 Systems of interacting particles	209
13.4 Systems of interacting fermions (the Fermi liquid)	211
13.5 The Landau theory of the normal Fermi liquid	214
13.6 Collective excitations of a Fermi liquid	221
13.6.1 Zero sound in a neutral Fermi gas with repulsive interactions	221
13.6.2 Plasma oscillations in a charged Fermi liquid	221
13.7 Phonons and other excitations	223
13.7.1 Phonons in crystals	223
13.7.2 Phonons in liquid helium-4	232
13.7.3 Magnons in solids	233
13.7.4 Polarons and excitons	233
14 Second Quantization	235
14.1 The occupation-number representation	235
14.2 Particle-field operators	246
15 Gas of Interacting Electrons	251
15.1 Hamiltonian of an electron gas	251
16 Superconductivity	261
16.1 Superconductors	261
16.2 The theory of Bardeen, Cooper and Schrieffer	262
16.2.1 Cooper pairs	267
16.2.2 Calculation of the ground-state energy	269
16.2.3 First excited states	277
16.2.4 Thermodynamics of superconductors	280
IV CLASSICAL FIELD THEORY AND RELATIVITY	287
17 The Classical Theory of Fields	289
17.1 Mathematical preliminaries	289
17.1.1 Behavior of fields under coordinate transformations	289
17.1.2 Properties of the rotation matrix	293
17.1.3 Proof that a “dot product” is a scalar	295
17.1.4 A lemma on determinants	297
17.1.5 Proof that the “cross product” of two vectors is a “pseudovector”	298
17.1.6 Useful index relations	299
17.1.7 Use of index relations to prove vector identities	300
17.1.8 General definition of tensors of arbitrary rank	301

17.2	Introduction to Einsteinian relativity	302
17.2.1	Intervals	302
17.2.2	Timelike and spacelike intervals	304
17.2.3	The light cone	304
17.2.4	Variational principle for free motion	305
17.2.5	The Lorentz transformation	305
17.2.6	Length contraction and time dilation	307
17.2.7	Transformation of velocities	308
17.2.8	Four-tensors	308
17.2.9	Integration in four-space	314
17.2.10	Integral theorems	316
17.2.11	Four-velocity and four-acceleration	317
17.3	Principle of least action	318
17.3.1	Free particle	318
17.3.2	Three-space formulation	318
17.3.3	Momentum and energy of a free particle	319
17.3.4	Four-space formulation	321
17.4	Motion of a particle in a given electromagnetic field	325
17.4.1	Equations of motion of a charge in an electromagnetic field	326
17.4.2	Gauge invariance	328
17.4.3	Four-space derivation of the equations of motion	329
17.4.4	Lorentz transformation of the electromagnetic field	332
17.4.5	Lorentz invariants constructed from the electromagnetic field	334
17.4.6	The first pair of Maxwell equations	335
17.5	Dynamics of the electromagnetic field	337
17.5.1	The four-current and the second pair of Maxwell equations	338
17.5.2	Energy density and energy flux density of the electromagnetic field	342
17.6	The energy–momentum tensor	345
17.6.1	Energy–momentum tensor of the electromagnetic field	350
17.6.2	Energy–momentum tensor of particles	353
17.6.3	Energy–momentum tensor of continuous media	355
18	General Relativity	361
18.1	Introduction	361
18.2	Space–time metrics	362
18.3	Curvilinear coordinates	364
18.4	Products of tensors	365
18.5	Contraction of tensors	366
18.6	The unit tensor	366
18.7	Line element	366
18.8	Tensor inverses	366
18.9	Raising and lowering of indices	367
18.10	Integration in curved space–time	367
18.11	Covariant differentiation	369
18.12	Parallel transport of vectors	370
18.13	Curvature	374
18.14	The Einstein field equations	376

18.15	Equation of motion of a particle in a gravitational field	381
18.16	Newton's law of gravity	383
V	RELATIVISTIC QUANTUM MECHANICS AND GAUGE THEORIES	385
19	Relativistic Quantum Mechanics	387
19.1	The Dirac equation	387
19.2	Lorentz and rotational covariance of the Dirac equation	391
19.3	The current four-vector	398
19.4	Compact form of the Dirac equation	400
19.5	Dirac wave function of a free particle	401
19.6	Motion of an electron in an electromagnetic field	405
19.7	Behavior of spinors under spatial inversion	408
19.8	Unitarity properties of the spinor-transformation matrices	409
19.9	Proof that the four-current is a four-vector	411
19.10	Interpretation of the negative-energy states	412
19.11	Charge conjugation	413
19.12	Time reversal	414
19.13	PCT symmetry	417
19.14	Models of the weak interaction	422
20	Gauge Theories of Quark and Lepton Interactions	427
20.1	Global phase invariance	427
20.2	Local phase invariance?	427
20.3	Other global phase invariances	429
20.4	SU(2) local phase invariance (a non-abelian gauge theory)	433
20.5	The "gauging" of color SU(3) (quantum chromodynamics)	436
20.6	The weak interaction	436
20.7	The Higgs mechanism	439
20.8	The fermion masses	448
Appendices		451
A.1	Proof that the scattering states $ \phi+\rangle \equiv \Omega_+ \phi\rangle$ exist for all states $ \phi\rangle$ in the Hilbert space \mathcal{H}	451
A.2	The scattering matrix in momentum space	452
A.3	Calculation of the free Green function $\langle \mathbf{r} G^0(z) \mathbf{r}'\rangle$	454
Supplementary Reading		457
Index		459

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 (\mathbf{r}), three-momentum (\mathbf{p}), magnetic induction (\mathbf{B}), and angular momentum (\mathbf{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 λ , (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 1 and 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** $\psi(\mathbf{r}, t)$ of the space coordinate $\mathbf{r} (= ix + jy + kz)$ and time coordinate t . The state of the particle is “described” by $\psi(\mathbf{r}, t)$ 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 $dxdydz \equiv d^3 \mathbf{r}$ with centre at \mathbf{r} at time t is

$$P(\mathbf{r}, t)dxdydz = \psi^*(\mathbf{r}, t)\psi(\mathbf{r}, t)dxdydz, \quad (1.1)$$

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

$$P(\mathbf{r}, t)d^3 \mathbf{r} = |\psi(\mathbf{r}, t)|^2 d^3 \mathbf{r}. \quad (1.2)$$

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

The wave function $\psi(\mathbf{r}, t)$ 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:

$$\int_{\text{all space}} |\psi|^2 d^3 \mathbf{r} = 1. \quad (1.3)$$

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

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 $\psi(\mathbf{r}, t)$ can be separated as a product of a space-dependent factor $\phi(\mathbf{r})$ 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** $\mathbf{p} = ip_x + jp_y + kp_z$ (so that the energy is also well defined and the above-mentioned factorization occurs), the **space-dependent** part $\phi(\mathbf{r})$ of the corresponding wave function $\psi(\mathbf{r}, t)$ is a harmonic plane wave $e^{i\mathbf{k}\cdot\mathbf{r}}$ [i.e., there is only one Fourier component (one wave vector \mathbf{k}) in its Fourier decomposition], where the wave vector $\mathbf{k} = ik_x + jk_y + kk_z$ is related to the momentum \mathbf{p} by $\mathbf{k} = \mathbf{p}/\hbar$, i.e., $k_x = p_x/\hbar$, and so on. [This is the well known **de Broglie relation** $\mathbf{p} = \hbar\mathbf{k}$, which, for the magnitudes, gives

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

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

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

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

[To see the reason for the designation “plane wave”, we choose the z axis to lie along the direction of the momentum \mathbf{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 \mathbf{p} (along the z axis). Then the corresponding wave function is

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

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

$$pz_{\mathbf{p}} - Et = \text{const},$$

that is,

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

Thus, a wave front (surface of constant phase) moves with **phase velocity** $v_{\mathbf{p}} = E/p$ (> 0) along the positive z axis, that is, in the direction of the momentum \mathbf{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 $\psi_a(\mathbf{r}, t)$], then

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

Thus, the action of the operator \hat{A} on the function $\psi_a(\mathbf{r}, t)$ is to reproduce precisely the same function of \mathbf{r} and t but scaled by a constant factor a equal to the well defined value of A in this state. We say that $\psi_a(\mathbf{r}, t)$ 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 \mathbf{p} is described by a wave function of the form

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

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

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

and

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

The differential operators

$$\hat{E} = i\hbar\frac{\partial}{\partial t} \quad (1.6)$$

and

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

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

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

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

(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,

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

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 \mathbf{r} is $\hat{\mathbf{r}} = \mathbf{r}$.

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

$$A \rightarrow \hat{A} = A(\hat{\mathbf{r}}, \hat{\mathbf{p}}) \quad (1.9)$$

with

$$\hat{\mathbf{r}} = \mathbf{r} \quad \text{and} \quad \hat{\mathbf{p}} = -i\hbar\nabla. \quad (1.10)$$

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

$$\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}} = -i\hbar\mathbf{r} \times \nabla. \quad (1.11)$$

1.5 The time-dependent Schrödinger equation

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

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

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

We make the operator replacements (section 1.4):

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

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

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

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 $\psi(\mathbf{r}, t)$ 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

$$\psi(\mathbf{r}, t) = \phi(\mathbf{r})e^{-iEt/\hbar}, \quad (1.15)$$

where $\phi(\mathbf{r})$ specifies the space-dependent part of the function. Because, from (1.15), $|\psi(\mathbf{r}, t)|^2 = |\phi(\mathbf{r})|^2$, 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 $E\psi(\mathbf{r}, t)$, and so, after cancellation of the factor $e^{-iEt/\hbar}$, we obtain

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

where we have also replaced $V(\mathbf{r}, t)$ by $V(\mathbf{r})$, 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 $V_i(\mathbf{r}) \equiv V(\mathbf{r}, t)$ at different times t will have different solutions $\phi_i(\mathbf{r})$ with different energy eigenvalues E_i , 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

$$\hat{H}\phi(\mathbf{r}) = E\phi(\mathbf{r}) \quad (1.17)$$

with

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

(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 $\psi(\mathbf{r}, t)$ describing the state of the particle is an eigenfunction $\psi_i(\mathbf{r}, t)$ of \hat{A} , that is,

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

Then the result of measuring A will certainly be a_i .

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

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

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 $\psi(\mathbf{r}, t)$ and $\chi(\mathbf{r}, t)$, the relation

$$\int \chi^* \hat{A} \psi d^3 \mathbf{r} = \int (\hat{A} \chi)^* \psi d^3 \mathbf{r} \quad (1.21)$$

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

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

Then (1.21) becomes

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

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

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

whence

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

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

$$\int \psi_i^* \psi_j d^3 \mathbf{r} = 0. \quad (1.22)$$

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

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

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 $|\int \phi^* \chi d^3 \mathbf{r}|^2$ 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:

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

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

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

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

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

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 $\psi_i(\mathbf{r}, t)$ 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

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

This relation asserts (in precise accord with its definition as an average) that the expectation value of A in the state $\psi(\mathbf{r}, t)$ 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

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

which follows from the fact that

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

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

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

(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**)

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

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

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

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

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

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

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

(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:

$$i\hbar \frac{d}{dt} \langle A \rangle_\psi = \langle [\hat{A}, \hat{H}] \rangle_\psi + i\hbar \langle \partial \hat{A} / \partial t \rangle_\psi, \quad (1.32)$$

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 $P(\mathbf{r}, t) = \psi^*(\mathbf{r}, t)\psi(\mathbf{r}, t)$. We find

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

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

$$\mathbf{j} = -\frac{i\hbar}{2m} (\psi^* \nabla \psi - (\nabla \psi^*) \psi) \quad (1.34)$$

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:

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